

Low worldwide genetic diversity in the basking shark (*Cetorhinus maximus*)

A. Rus Hoelzel^{1,*}, Mahmood S. Shivji², Jennifer Magnussen² and Malcolm P. Francis³

¹School of Biological and Biomedical Sciences, Durham University, South Road, Durham DH13LE, UK

²Guy Harvey Research Institute, Nova Southeastern University, 8000 North Ocean Drive, Dania Beach, FL 33004, USA

³National Institute of Water and Atmospheric Research, PO Box 14-901, Kilbirnie, Wellington 6003, New Zealand

*Author for correspondence (a.r.hoelzel@durham.ac.uk).

The basking shark (*Cetorhinus maximus*) is found in temperate waters throughout the world's oceans, and has been subjected to extensive exploitation in some regions. However, little is known about its current abundance and genetic status. Here, we investigate the diversity of the mitochondrial DNA control region among samples from the western North Atlantic, eastern North Atlantic, Mediterranean Sea, Indian Ocean and western Pacific. We find just six haplotypes defined by five variable sites, a comparatively low genetic diversity of $\pi=0.0013$ and no significant differentiation between ocean basins. We provide evidence for a bottleneck event within the Holocene, estimate an effective population size (N_e) that is low for a globally distributed species, and discuss the implications.

Keywords: biodiversity; marine fish; sharks; mitochondrial DNA

1. INTRODUCTION

The basking (*Cetorhinus maximus*) shark is a planktivorous species seasonally found in shoals nearshore, feeding near the surface. It is the second largest fish, sometimes exceeding 10 m in length and inhabits temperate regions in both hemispheres. Basking sharks mature slowly, requiring approximately 12–20 years and females have long gestation periods (approx. 1–3 years) after which they give birth to few offspring. These characteristics make this species especially vulnerable to over-exploitation (Compagno 2001). Basking sharks have been exploited for meat, fins, liver oil (Compagno 2001) and cartilage (Hoelzel 2001). They were listed by the IUCN as vulnerable worldwide (IUCN 2004), and in 2002 on Appendix II of the convention on international trade in endangered species. Despite the conservation concern for this species, there are few data on regional abundance, no estimates for abundance worldwide and no good data on population trends. The classification as 'vulnerable' is based largely on the rapid depletion of some populations subject to coastal harpoon fisheries, especially in the North Atlantic (Compagno 2001).

Here, we assess genetic diversity at the non-coding mtDNA control region and test the hypothesis that the species has been and remains sufficiently abundant to maintain levels of diversity comparable to those seen in

other broadly distributed elasmobranch species. Expected levels of diversity depend in part on the substitution rate. Martin *et al.* (1992) investigating the mitochondrial *Cytb* and *COI* genes (coding loci) in elasmobranchs calculated a mutation rate that was 7–8 times slower than had been calculated for primates and ungulates.

However, recent studies have shown high levels of diversity at the mtDNA control region for some shark species. For example, a comparison of white sharks (*Carcharodon carcharias*) from Australian/New Zealand and South African waters showed a level of diversity that is comparable to other widely distributed, pelagic marine species (table 1), with substantial genetic differentiation between the two regions ($F_{ST}=0.81$, $p<0.0001$; Pardini *et al.* 2001). In contrast, we show that basking shark genetic diversity is exceptionally low worldwide.

2. MATERIAL AND METHODS

Tissue samples were acquired from bycatch and strandings. The numbers of samples per region are given in table 2. DNA was extracted by standard phenol chloroform methods, and the full mtDNA control region amplified using primers set in the flanking tRNA_{Thr} and tRNA_{Phe} genes, designed on the basis of aligned sequences available in GenBank. The PCR reaction mixture consisted of 0.20 μ M each for the forward (5'-TCTTAGCATCTTCAGTGC) and reverse (5'-TCTTAGCATCTTCAGTGC) primers, 100 μ M dNTPs, 1.5 mM MgCl₂, 10 mM Tris-HCl pH 8.4, 50 mM KCl and 0.02 U (μ l)⁻¹ Taq polymerase. The PCR cycling profile was 5 min at 95 °C, 35 cycles of 45 s at 94 °C, 1.5 min at 48 °C and 1.5 min at 72 °C, followed by 8 min at 72 °C. PCR products were purified with QIAGEN PCR purification columns and sequenced directly using the ABI dye-terminator method. All samples were sequenced in both directions using the amplification primers, and internal primers when necessary (forward: 5'-GCACATTACTCATCTCGACTACATCAC, 5'-GAAGCAATCGCTATCAATCGAA; reverse: 5'-CTGGTCAA TTGGTGGGGATCAACCG, 5'-CGTTTATTGCGAATTTGT CCCCAGGGG). Resulting sequences were aligned using CLUSTALX.

Measures of haplotypic (h) and nucleotide (π) diversity and theta (θ_S) and differentiation (ϕ_{ST} using the Kimura 2-parameter model and F_{ST} using haplotype frequencies) were calculated using ARLEQUIN 2.000 (Schneider & Excoffier 1999). The Kimura 2-parameter model was chosen because the amount of variation was small, distributed across the locus and consisted of only transitions. Proportional haplotype frequencies were compared using a Fisher exact test (with a Markov chain of 10 000). ARLEQUIN was also used to construct a minimum spanning tree (after Rohlf 1973), estimate Tajima's D (Tajima 1989) and construct a mismatch distribution (Rogers & Harpending 1992). The latter two analyses can help determine if a population has undergone a rapid expansion (possibly as a result of a population bottleneck). Tajima's D -tests for departure from mutation-drift equilibrium. The mismatch distribution will be multimodal in stable populations (if the generation of new mutations is offset by random drift), and unimodal for expanding populations (if new mutations are accumulated faster than loss due to drift). The time of a possible population expansion (t) can be calculated using the formulation: $\tau=2ut$ (Rogers & Harpending 1992), where τ is the mode of the mismatch distribution and u is the mutation rate of the sequence (such that $u=2\mu k$, where μ is the mutation rate per nucleotide and k is the number of nucleotides). The time (t) is measured in generations.

For basking sharks, neither the generation time nor the mtDNA control region mutation rate is precisely known. However, data from Pauly (1978) suggests a generation time of about 16 years. Donaldson & Wilson (1999) estimate an average control region mutation rate of 3.6% per million years for a range of fish species, while the Martin *et al.* (1992) data would imply roughly 2.7% and Duncan *et al.* (2006) have recently suggested 0.8% based on population isolation in one shark species. We use an average of these three (2.4%).

3. RESULTS AND DISCUSSION

In a preliminary study, Hoelzel (2001) compared basking shark populations in the North Atlantic ($N=11$) and South Pacific ($N=6$) based on sequence

Table 1. Diversity at the mtDNA control region among pelagic marine vertebrate species. (Multiple geographical regions represented except as indicated; WNA, western North Atlantic; SA, South Africa; M, Mediterranean Sea. All studies based on the whole control region except those marked by *. Blank lines separate sharks, teleost fishes, loggerhead turtle and cetaceans. Species or populations proposed to have been through a population bottleneck marked by †.)

species	nucleotide diversity (π)	haplotypic diversity (h)	reference
<i>Cetorhinus maximus</i>	0.0013 ± 0.0009	0.720 ± 0.028	this study
<i>Carcharhinus limbatus</i> (WNA)†	0.0021 ± 0.0013	0.805 ± 0.018	Keeney <i>et al.</i> (2005)
<i>Carcharias taurus</i> (SA)*	0.003 ± 0.0001	0.717 ± 0.01	Stow <i>et al.</i> (2006)
<i>Carcharodon carcharias</i>	0.0203	—	Pardini <i>et al.</i> (2001)
<i>Sphyrna lewini</i> *	0.013 ± 0.0068	0.80 ± 0.02	Duncan <i>et al.</i> (2006)
<i>Thunnus obesus</i> *	0.054	0.98–1.0	Martinez <i>et al.</i> (2006)
<i>Xiphias gladius</i>	0.0148 ± 0.0005	0.997	Lu <i>et al.</i> (2006)
<i>Thunnus thynnus thynnus</i> (M)	0.015	0.991	Carlsson <i>et al.</i> (2004)
<i>Acanthocybium solandri</i>	0.053	0.999	Garber <i>et al.</i> (2005)
<i>Caretta caretta</i> (WNA)*	0.0236 ± 0.0121	0.579 ± 0.028	Bowen <i>et al.</i> (2004)
<i>Physeter macrocephalus</i> †	0.002 ± 0.0003	0.86	Lyrholm <i>et al.</i> (1996)
<i>Orcinus orca</i> †	0.0053 ± 0.0031	0.874 ± 0.013	Hoelzel <i>et al.</i> (2002)
<i>Tursiops truncatus</i> *	0.013–0.024	0.42–0.92	Natoli <i>et al.</i> (2004)
<i>Delphinus delphis</i> *	0.012–0.021	0.853–1.0	Natoli <i>et al.</i> (2006)

Table 2. Haplotypes and their frequency from sampling areas: NZ, New Zealand; TW, Taiwan; NOR, Norway; SCO, Scotland; WNA, western North Atlantic; MED, Mediterranean Sea; CAR, Caribbean; SA, East coast of South Africa. (Position numbers are with reference to light-strand sequence along the 1085 bp sequences submitted to GenBank.)

hap	nucleotide positions					sampling area							
	182	450	640	794	966	NZ	TW	NOR	SCO	WNA	MED	CAR	SA
BS1	T	A	G	G	—	13	—	2	1	5	—	—	—
BS2	.	.	.	A	.	9	1	1	1	6	4	—	—
BS3	C	G	A	.	.	5	—	1	—	3	—	1	—
BS4	C	4	—	—	1	1	—	—	1
BS5	C	.	.	A	A	1	—	—	—	—	—	—	—
BS6	.	.	.	A	A	1	—	—	—	—	—	—	—

data from 550 bp of the mtDNA *Cytb* locus. There were two *Cytb* haplotypes, and their frequency did not differ between regions. Here, we investigate the complete mtDNA control region from a much larger set of globally distributed animals. Although the results show more haplotypes (6) compared to *Cytb*, the overall level of diversity remains remarkably low. The haplotypic and nucleotide diversity values are given for the full sample ($N=62$) in table 1 and compared with values for various other globally widespread, pelagic marine species (including elasmobranchs, teleosts, a sea turtle and mammals). Genetic diversity values for basking sharks in each ocean basin considered separately were very similar to the worldwide values (Pacific: $h=0.7344 \pm 0.0418$, $\pi=0.0013 \pm 0.0009$, $N=34$; Atlantic: $h=0.7169 \pm 0.0495$, $\pi=0.0014 \pm 0.0009$, $N=27$).

Despite the global distribution of the samples, there were only five variable sites among the six haplotypes and 1085 bp in the sequence. The minimum spanning tree (figure 1) showed no evidence for structure and was dominated by two haplotypes (found in similar frequencies in each ocean basin; table 2). By contrast, the white shark (*C. carcharias*) had 77 variable sites defining 29 haplotypes for the 1149 bp control region sequence among a sample of 88 sharks (Pardini *et al.* 2001). Similarly, a study of young blacktip sharks (*Carcharhinus limbatus*) over a much smaller

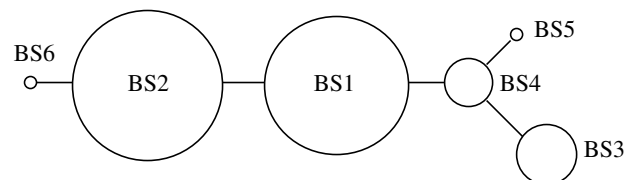


Figure 1. Minimum spanning network for the six haplotypes. Size of circle reflects relative frequency.

geographical range (part of the WNA) had 15 variable sites defining 23 haplotypes ($N=323$; 1067–1070 bp), with high genetic differentiation ($\phi_{ST}=0.35$, $p<0.001$) between sampled areas (Keeney *et al.* 2005).

There was no difference between Pacific and Atlantic basking shark samples grouped as putative populations (non-significant and negative ϕ_{ST} and F_{ST} and non-significant Fisher's exact test: $p=0.85$). Local sample sizes were small in the North Atlantic, Mediterranean, and Indian Ocean and possible further population substructure could not be assessed. However, the differences among all haplotypes were very small (table 2). The lack of structure may suggest that a bottleneck event preceded expansion into the current distributional range. Alternatively, it could suggest female mediated gene flow over a wide geographical range, but this would not account for the very low diversity levels.

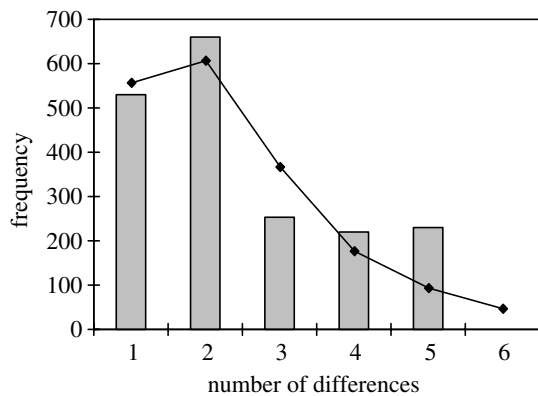


Figure 2. Mismatch distribution with the observed values as bars and a solid line representing the expected distribution according to the sudden expansion model.

Tajima's D was not significantly different from random expectations, but the mismatch distribution was unimodal (figure 2). Estimated $\tau = 0.56$, and given our assumed parameters for generation time and mutation rate, this would suggest a population expansion beginning approximately 86 000 years ago, well before the influence of modern fisheries. The lack of transversion mutations is consistent with a bottleneck, and with this timeframe. Uncertainty about parameter estimates means that correlation with any specific event is not possible. However, events such as the warming periods preceding Holocene interglacials can theoretically cause considerable disruption to oceanic systems, and therefore to the resources that basking sharks depend on, and plausibly lead to a population bottleneck. Selective sweeps can also reduce diversity, but it is less likely for this to result in the same common haplotypes occurring in distant populations.

The census number of basking sharks in the world's oceans is unknown. However, our data suggest an N_e of 8200, calculated from estimated Theta ($\theta_S = 0.852 \pm 0.467$; $\theta = 2N_e u$). Again, given the underlying uncertainties in the parameter estimates, it is necessarily a rough approximation. Even so, given the global distribution of the species, it is surprisingly low. In a meta-analysis, Frankham (1995) identified a median ratio of N_e to census population size (N_c) of 10%, though the range was very broad. However, historical bottlenecks can depress N_e/N_c , especially in rapidly expanding populations (e.g. this has been proposed to explain a human N_e of 10 000; e.g. Harpending *et al.* 1993).

Although diversity at this locus is low for some other elasmobranch species assessed in local populations, or following known bottlenecks, none have so far shown such low diversity worldwide (see table 1). We suggest that, the most parsimonious explanation is that the basking shark suffered a population bottleneck affecting global diversity (cf. similar data for the killer whale (*Orcinus orca*); Hoelzel *et al.* 2002), and propose that the low diversity revealed should be a key consideration for the development of conservation and management strategies for this species, on the assumption that further loss of diversity could affect evolutionary potential.

This study was supported in part by DETR, the Pew Institute for Ocean Science, the Florida Sea Grant Program and the Guy Harvey Research Institute. We thank M. Affronte, R. Baird, L. Boren, J. Cassin, G. Cliff, T. Knott, D. Mattila, J. Morrissey, L. Natanson, Mark O'Connell, M. Preide, S. Fowler, T. Thom, R. Torres, S. Wintner and New Zealand Ministry of Fisheries observers for help with the acquisition of basking shark samples.

- Bowen, B. W. *et al.* 2004 Natal homing in juvenile loggerhead turtles (*Caretta caretta*). *Mol. Ecol.* **13**, 3797–3808. (doi:10.1111/j.1365-294X.2004.02356.x)
- Carlsson, J., McDowell, J. R., Diaz-Jaimes, P., Carlsson, J. E. L., Boles, S. B., Gold, J. R. & Graves, J. E. 2004 Microsatellite and mitochondrial DNA analyses of Atlantic bluefin tuna (*Thunnus thynnus thynnus*) population structure in the Mediterranean sea. *Mol. Ecol.* **13**, 3345–3356. (doi:10.1111/j.1365-294X.2004.02336.x)
- Compagno, L. J. V. 2001 *Sharks of the world: an annotated and illustrated catalogue of sharks species known to date. Bullhead, mackerel and carpet sharks (Heterodontiformes, Lamniformes and Orectolobiformes)*, vol. 2. Rome, Italy: FAO.
- Donaldson, K. A. & Wilson, R. R. 1999 Amphipanamic geminates of snook (Percoidei: Centropomidae) provide a calibration of the divergence rate in the mitochondrial control region of fishes. *Mol. Phylog. Evol.* **13**, 208–213. (doi:10.1006/mpev.1999.0625)
- Duncan, K. M., Martin, A. P., Bowen, B. W. & de Couet, H. G. 2006 Global phylogeography of the scalloped hammerhead shark (*Sphyrna lewini*). *Mol. Ecol.* **15**, 2239. (doi:10.1111/j.1365-294X.2006.02933.x)
- Frankham, R. 1995 Effective population size and population size ratios in wildlife—a review. *Genet. Res.* **66**, 95–107.
- Garber, A. F., Tringali, M. D. & Franks, J. S. 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)
- Harpending, H. C., Sherry, S. T., Rogers, A. R. & Stoneking, M. 1993 The genetic structure of ancient human populations. *Curr. Anthropol.* **34**, 483–496. (doi:10.1086/204195)
- Hoelzel, A. R. 2001 Shark fishing in fin soup. *Conserv. Genet.* **2**, 69–72. (doi:10.1023/A:1011590517389)
- Hoelzel, A. R., Natoli, A., Dahlheim, M., Olavarria, C., Baird, R. W. & Black, N. 2002 Low world-wide genetic diversity in the killer whale (*Orcinus orca*); implications for demographic history. *Proc. R. Soc. B* **269**, 1467–1475. (doi:10.1098/rspb.2002.2033)
- IUCN 2004 <http://www.redlist.org/>.
- Keeney, D. B., Heupel, M. R., Hueter, R. E. & Heist, E. J. 2005 Microsatellite and mitochondrial DNA analyses of the genetic structure of blacktip shark (*Carcharhinus limbatus*) nurseries in the northwestern Atlantic, Gulf of Mexico, and Caribbean Sea. *Mol. Ecol.* **14**, 1911–1923. (doi:10.1111/j.1365-294X.2005.02549.x)
- Lu, C.-L., Chen, C. A., Hui, C.-F., Tzeng, T.-D. & Yeh, S.-Y. 2006 Population genetic structure of the swordfish, *Xiphias gladius* (Linnaeus, 1758), in the Indian Ocean and west Pacific inferred from the complete DNA sequence of the mitochondrial control region. *Zool. Stud.* **45**.
- Lyrholm, T., Leimar, O. & Gyllenstein, U. 1996 Low diversity and biased substitution patterns in the mitochondrial DNA control region of sperm whales: implications for estimates of time since common ancestry. *Mol. Biol. Evol.* **13**, 1318–1326.
- Martin, A. P., Naylor, G. J. P. & Palumbi, S. R. 1992 Rates of mtDNA evolution in sharks are slow compared to humans. *Nature* **357**, 153–155. (doi:10.1038/357153a0)

- Martinez, P., Gonzalez, E. G., Castilho, R. & Zardoya, R. 2006 Genetic diversity and historical demography of Atlantic bigeye tuna (*Thunnus obesus*). *Mol. Phylogen. Evol.* **39**, 404–416. (doi:10.1016/j.ympev.2005.07.022)
- Natoli, A., Peddemors, V. & Hoelzel, A. R. 2004 Population structure and speciation in the genus *Tursiops* based on microsatellite and mitochondrial DNA analyses. *J. Evol. Biol.* **17**, 363–375. (doi:10.1046/j.1420-9101.2003.00672.x)
- Natoli, A., Cañadas, A., Peddemors, V. M., Aguilar, A., Vaquero, C., Fernández-Piqueras, P. & Hoelzel, A. R. 2006 Phylogeography and alpha taxonomy of the common dolphin (*Delphinus* sp.). *J. Evol. Biol.* **19**, 943–954. (doi:10.1111/j.1420-9101.2005.01033.x)
- Pardini, A. T. *et al.* 2001 Sex-biased dispersal of great white sharks. *Nature* **412**, 139–140. (doi:10.1038/35084125)
- Pauly, D. 1978 A preliminary compilation of fish length growth parameters. *Ber. Inst. Meereskd. Christian-Albrechts-Univ. Kiel* **55**, 1–200.
- Rogers, A. R. & Harpending, H. 1992 Population growth makes waves in the distribution of pairwise genetic differences. *Mol. Biol. Evol.* **9**, 552–569.
- Rohlf, F. J. 1973 Algorithm 76. Hierarchical clustering using the minimum spanning tree. *Comput. J.* **16**, 93–95.
- Schneider, S. & Excoffier, L. 1999 Estimation of past demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: application to human mitochondrial DNA. *Genetics* **152**, 1079–1089.
- Stow, A., Zenger, K., Briscoe, D., Gillings, 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)
- Tajima, F. 1989 Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**, 585–595.