Gene xxx (2014) xxx–xxx

Contents lists available at ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

The complete mitochondrial genome sequence of the world's largest fish, the whale shark (*Rhincodon typus*), and its comparison with those of related shark species

Md Tauqeer Alam [a], Robert A. Petit III [a], Timothy D. Read [a,b], Alistair Dove [c,*]

[a] Emory University School of Medicine, Department of Medicine, Division of Infectious Diseases, Atlanta, GA, USA
[b] Emory University School of Medicine, Department of Human Genetics, Atlanta, GA, USA
[c] Georgia Aquarium Research Center, Atlanta, GA, USA

ARTICLE INFO

Article history:
Received 17 October 2013
Received in revised form 16 January 2014
Accepted 26 January 2014
Available online xxxx

Keywords:
Whale shark
Mitochondria
Genome
Mitogenome
Phylogeny

ABSTRACT

The whale shark (*Rhincodon typus*) is the largest extant species of fish, belonging to the order Orectolobiformes. It is listed as a "vulnerable" species on the International Union for Conservation of Nature (IUCN)'s Red List of Threatened Species, which makes it an important species for conservation efforts. We report here the first complete sequence of the mitochondrial genome (mitogenome) of the whale shark obtained by next-generation sequencing methods. The assembled mitogenome is a 16,875 bp circle, comprising of 13 protein-coding genes, two rRNA genes, 22 tRNA genes and a control region. We also performed comparative analysis of the whale shark mitogenome to the available mitogenome sequences of 17 other shark species, four from the order Orectolobiformes, five from Lamniformes and eight from Carcharhiniformes. The nucleotide composition, number and arrangement of the genes in whale shark mitogenome are the same as found in the mitogenomes of the other members of the order Orectolobiformes and its closest orders Lamniformes and Carcharhiniformes, although the whale shark mitogenome had a slightly longer control region. The availability of mitogenome sequence of whale shark will aid studies of molecular systematics, biogeography, genetic differentiation, and conservation genetics in this species.

© 2014 Published by Elsevier B.V.

1. Introduction

The whale shark *Rhincodon typus* Smith 1828, is the largest extant species of shark, indeed the largest extant species of fish. It belongs to the family Rhincodontidae within the order Orectolobiformes of class Chondrichthyes. There are 41 other species within this order, but *R. typus* is the only one that exclusively inhabits the pelagic zone. It is found mostly in tropical and warm temperate seas between latitudes 30°N and 35°S (Compagno, 2001). Whale sharks predominantly feed on macrozooplankton that they capture by filtering seawater through 20 specialised pad-like apparatuses situated in the oropharyngeal cavity (Motta et al., 2010; Rowat and Brooks, 2012; Taylor, 2007). The filter feeding habit is shared with only two other shark species, the lamniform basking shark *Cetorhinus maximus* and megamouth shark *Megachasma pelagios*, but the whale shark mechanism is unique and likely evolved separately (Martin and Naylor, 1997). Unlike the whale shark, the filtering structures in *C. maximus* and *M. pelagios* consist of bristle- or finger-like gill rakers, which they use to trap plankton as water filters through the mouth and over the gills (Rowat and Brooks, 2012). The whale shark is a charismatic and placid species and a popular target for ecotourism in places where it is found reliably, however it is also subject to several threats and listed as a Vulnerable species on the International Union for Conservation of Nature (IUCN) 's Red List of Threatened Species (IUCN, 2013). Whale sharks face several threats including targeted artisanal fisheries, incidental take associated with tuna fisheries (the two species are often co-located), ship strike and ecotourism. Unfortunately, due to their late maturation and presumed low fecundity, whale sharks are expected to have low population resilience in the face of these threats.

Despite being an iconic animal with great conservation importance, little research has been done on the genetics and genomics of this species. There is no comprehensive estimate of the inherent genetic diversity within the global population of *R. typus*. Also, the degree of genetic differentiation between *R. typus* sub-populations in different ocean basins is not known. Three genetic studies conducted to date seem to suggest that *R. typus* constitute a single panmictic population with little or no genetic differentiation between oceans (Castro et al.,

Abbreviations: bp, base pairs; IUCN, International Union for the Conservation of Nature; rRNA, Ribosomal ribose nucleic acid; tRNA, Transfer ribose nucleic acid; π, nucleotide diversity; Cytb, Cytochrome oxidase b gene; COI and COII, Cytochrome oxidase gene subunits 1 and 2, respectively; DNA, Deoxy-ribose nucleic acid; BLAST, Basic Local Alignment Search Tool; PCA, Polymerase chain reaction; NCBI, National Center for Biotechnology Information; ML, Maximum likelihood; AIC, Akaike Information Criterion; GC, The Guanine–Cytosine content of DNA, expressed as a percentage; EGC, Emory Genome Center.

* Corresponding author at: Georgia Aquarium Research Center, 225 Baker Street NW, Atlanta, GA 30313, USA. Tel.: +1 404 581 4364.
E-mail address: adove@georgiaaquarium.org (A. Dove).

0378-1119/$ – see front matter © 2014 Published by Elsevier B.V.
http://dx.doi.org/10.1016/j.gene.2014.01.064

Please cite this article as: Alam, M.T., et al., The complete mitochondrial genome sequence of the world's largest fish, the whale shark (*Rhincodon typus*), and its comparison with those of related shark species, Gene (2014), http://dx.doi.org/10.1016/j.gene.2014.01.064

ARTICLE IN PRESS

2007; Ramírez-Macías et al., 2007; Schmidt et al., 2009). These studies utilised either mitochondrial control region (Castro et al., 2007; Ramírez-Macías et al., 2007) or microsatellite loci (Schmidt et al., 2009) as markers to investigate genetic diversity and phylogeography of the *R. typus* populations. Castro et al (2007) observed 44 unique haplotypes of the control region sequence in the 70 *R. typus* individuals analysed from six locations around the world (haplotype diversity = 0.974 ± 0.008, π nucleotide diversity = 0.011 ± 0.006), however, none of the haplotypes showed a clear association with any specific geographic location. Ramírez-Macías et al. (2007) analysed control region sequence in 36 *R. typus* individuals from three areas in the Gulf of California, Mexico, and found no evidence of geographic clustering of any of the 14 control region haplotypes (haplotype diversity = 0.90 and π nucleotide diversity = 0.005). Schmidt et al. (2009) surveyed variation at 8 microsatellite loci in 68 *R. typus* individuals collected from Pacific Ocean, Indian Ocean and Caribbean Ocean and observed no significant genetic differentiation between any of these populations.

Short mitochondrial regions such as *Cytb*, *COI*, *COII* are widely used markers for population genetic studies, however phylogenetic inference based on the variations at such short loci is not always robust and may lack sufficient resolution compared to the phylogenies based on longer genomic sequences such as the complete mitogenome or nuclear genome (Cao et al., 1998a; Galtier et al., 2009; Powell et al., 2013; Velez-Zuazo and Agnarsson, 2011; Yu et al., 2007). Also, analysis of longer sequences from large number of samples provides more insightful view of population structure compared to single and shorter genetic loci. In this study we report the first complete sequence of the *R. typus* mitogenome using next-generation sequencing methods. We also compared *R. typus* mitogenome with mitogenomes of 17 other shark species belonging to the order Orectolobiformes and its closest orders Lamniformes and Carcharhiniformes.

2. Materials and methods

The genomic DNA used for this study was isolated from the liver tissue of a deceased male *R. typus* species from Taiwan, using a standard phenol–chloroform extraction method. The specimen was collected by the Georgia Aquarium Research Center, Atlanta as a part of the whale shark whole genome-sequencing project. Whole genome shotgun sequencing was performed on the Illumina Hiseq 2000 (Illumina Inc., San Diego, CA) and GS-FLX (454 Life Sciences, Branford, CT). The mitochondrial genome-specific sequences were retrieved from the whole genome shotgun reads in two sequential steps. We first performed *de novo* assembly of the whole genome shotgun reads using ABySS (Simpson et al., 2009), and then identified the mitochondrial genome-specific contigs using BLAST taking brown-banded bamboo shark *Chiloscyllium punctatum* complete mitogenome (GenBank ID JQ082337) as a query sequence (Chen et al., 2013g). These contigs were then used as a reference against which entire shotgun reads were remapped using BWA short read aligner (Li and Durbin, 2009, 2010). Approximately 299,687 mapped reads (1072 reads from 454, mean read length 500 bp; and 298,615 reads from Illumina, mean read length 100 bp) were extracted at this step, which resulted into a single contig of 16,875 bp upon *de novo* assembly by Velvet (Zerbino and Birney, 2008). Since, like any other vertebrate, the *R. typus* mitochondrial control region contains repeat sequence and is highly polymorphic, we also confirmed its size and sequence by PCR and Sanger sequencing methods.

The assembled 16,875 bp contig was annotated using MitoAnnotator, a web-based tool developed specifically for fish mitochondrial genome annotation (Wataru et al., 2013). The stand-alone programmes RNAmmer (Lagesen et al., 2007) and tRNAscan-SE (Lowe and Eddy, 1997) were used to confirm rRNA and tRNA annotation results, respectively. The boundaries of the predicted genes were also confirmed by sequence comparisons with the annotated mitogenome of *C. punctatum* (Chen et al., 2013g). The *R. typus* mitogenome genome has been deposited in NCBI GenBank database under accession number KF679782.

Comparison of *R. typus* mitogenome with other available mitogenomes of the members of the orders Orectolobiformes (*C. punctatum*, *C. griseum*, *C. plagiosum* and *Orectolobus japonicas*), Lamniformes (*M. pelagios*, *Mitsukurina owstoni*, *Alopias superciliosus*, *Carcharodon carcharias* and *Isurus oxyrinchus*) and Carcharhiniformes (*Pseudotriakis microdon*, *Mustelus manazo*, *Scyliorhinus canicula*, *Scoliodon macrorhynchos*, *Carcharhinus obscurus*, *Glyphis glyphis*, *Galeocerdo cuvier* and *Sphyrna lewini*) were made using CGView Comparison Tool (CCT). The NCBI GenBank ID and sizes of these mitogenomes are listed in Table 1. CCT uses BLAST to compare a query genome with all other genomes and then presents the results as a circular map (Grant et al., 2012). We also reconstructed protein-based maximum likelihood (ML) and Bayesian phylogenies to investigate the evolutionary relationships amongst these 18 shark species. All 13 protein sequences from each species were concatenated and aligned using MUSCLE programme (Edgar, 2004). For ML analysis we used RAxML version 7.0.4 and inference was made assuming MtMAM + G amino acid substitution model based on the Akaike information criterion (AIC) suggested by PROTTEST, a best-fit evolutionary model predictor for protein alignment data (Abascal et al., 2005; Stamatakis, 2006). The branch support of the ML tree was assessed by non-parametric bootstrapping method with 500 pseudo-replicates (Felsenstein, 1985). Bayesian analysis was performed in MrBayes version 3.2, under the same evolutionary model used for ML analysis (Ronquist et al., 2012). Two independent runs of 50,000 generations (until average standard deviation of split frequencies < 0.01) each were conducted simultaneously, sampling every 500 generations and discarding 25% of the initial trees as burn-in. Convergence of the likelihood values of the trees from both runs (stationarity of likelihood values) and burn-in frequency were determined by visualising the trace files and also by plotting a graph of likelihood values against number of trees sampled. ML and Bayesian analyses were also carried out under MtREV + G evolutionary model, whilst keeping all other parameters same as describe above.

3. Results and discussion

3.1. Characteristics of the R. typus mitogenome

In the present study, we determined the complete mitochondrial genome sequence of the whale shark *R. typus* by next-generation sequencing methods. A uniform coverage of ~1500× was obtained for most of the mitogenome, except at few places in the control region where the presence of the repeat sequences resulted in slightly lower coverage. The exact size (1225 bp) and sequence of the control region was verified by PCR followed by Sanger sequencing. The assembled mitogenome was 16,875 bp in length, which is in the range of the mitogenome size of its closely related species of the orders Orectolobiformes, Lamniformes and Carcharhiniformes (Table 1). Similar to a typical vertebrate mitogenome, it was comprised of 13 protein-coding genes, two rRNA genes (12S rRNA and 16S rRNA), 22 tRNA genes and a putative control region (or displacement loop) between $tRNA^{Pro}$ and $tRNA^{Phe}$ (Table 2). Eight of the 22 tRNAs ($tRNA^{Gln}$, $tRNA^{Ala}$, $tRNA^{Asn}$, $tRNA^{Cys}$, $tRNA^{Tyr}$, $tRNA^{Ser(UGA)}$, $tRNA^{Glu}$ and $tRNA^{Pro}$) and a protein-coding gene *ND6* were encoded by the light (L) strand, whilst the remaining genes were encoded by the heavy (H) strand (Table 2, Fig. 1). All 22 tRNAs were within the size range of 67 to 75 bp, and each of them folded into a typical cloverleaf secondary structure as predicted by tRNAscan-SE. The arrangements of the genes were also similar to a typical vertebrate mitogenome. With the exception of the *COI* gene, which starts with GTG codon, all other protein-coding genes had the usual ATG start codon. Seven (*ND1*, *COI*, *ATP8*, *COIII*, *ND4L*, *ND5* and *ND6*) of the 13 protein-coding genes have TAA stop codon, whilst six

Please cite this article as: Alam, M.T., et al., The complete mitochondrial genome sequence of the world's largest fish, the whale shark (*Rhincodon typus*), and its comparison with those of related shark species, Gene (2014), http://dx.doi.org/10.1016/j.gene.2014.01.064

UNCORRECTED PROOF

ARTICLE IN PRESS

Table 1
Shark species compared in this study.

Shark species	Order	GenBank ID	Mitogenome size (bp)	Reference
Rhincodon typus (whale shark)	Orectolobiformes	KF679782	16,875	This study
Orectolobus japonicus (Japanese wobbegong shark)	Orectolobiformes	KF111729	16,706	Chen et al. (2013c)
Chiloscyllium griseum	Orectolobiformes	NC_017882	16,755	Chen et al. (2013a)
Chiloscyllium plagiosum (whitespotted bamboo shark)	Orectolobiformes	NC_012570	16,726	Unpublished
Chiloscyllium punctatum (brown-banded bamboo shark)	Orectolobiformes	JQ082337	16,703	Chen et al. (2013g)
Alopias superciliosus (bigeye thresher shark)	Lamniformes	KC757415	16,719	Chang et al. (2013c)
Carcharodon carcharias (great white shark)	Lamniformes	KC914387	16,744	Chang et al. (2013b)
Isurus oxyrinchus (shortfin mako shark)	Lamniformes	KF361861	16,701	Chang et al. (2013d)
Megachasma pelagios (megamouth shark)	Lamniformes	KC702506	16,694	Chang et al. (2013a)
Mitsukurina owstoni (goblin shark)	Lamniformes	EU528659	17,743	Unpublished
Pseudotriakis microdon (false catshark)	Carcharhiniformes	AB560493	16,700	Tanaka et al. (2013)
Mustelus manazo (starspotted smooth-hound shark)	Carcharhiniformes	AB015962	16,707	Cao et al. (1998b)
Scyliorhinus canicula (small-spotted catshark),	Carcharhiniformes	Y16067	16,697	Delarbre et al. (1998)
Scoliodon macrorhynchos (Pacific spadenose shark)	Carcharhiniformes	JQ693102	16,693	Chen et al. (2013d)
Carcharhinus obscurus (dusky shark)	Carcharhiniformes	KC470543	16,706	(Blower et al., 2013)
Glyphis glyphis (speartooth shark)	Carcharhiniformes	KF006312	16,702	Chen et al. (2013b)
Sphyrna lewini (scalloped hammerhead shark)	Carcharhiniformes	JX827259	16,726	Chen et al. (2013e)
Galeocerdo cuvier (tiger shark)	Carcharhiniformes	KF111728	16,703	Chen et al. (2013f)

genes have incomplete stop codons either TA (*ND2*, *ATP6* and *Cytb*) or T (*COII*, *ND3* and *ND4*). The overall GC content of *R. typus* mitogenome is ~38%.

The control region of mitogenome is a widely used molecular marker in population genetic studies (Ahonen et al., 2009; Castro et al., 2007). It consists of tandem repeat sequences and has been

Table 2
Location and arrangement of genes on the 16,875 bp *R. typus* mitogenome.

Gene	Strand[a]	Gene					Intergenic spacer[d]
		From (b)	To (bp)	Size (bp)[b]	Start codon	Stop codon[c]	
$tRNA^{Phe}$	H	1	70	70			
12S rRNA	H	71	1025	955			
$tRNA^{Val}$	H	1026	1097	72			
16S rRNA	H	1098	2784	1687			
$tRNA^{Leu\ (UAA)}$	H	2785	2859	75			
ND1	H	2860	3834	975	ATG	TAA	3
$tRNA^{Ile}$	H	3838	3907	70			
$tRNA^{Gln}$	L	3908	3979	72			
$tRNA^{Met}$	H	3980	4048	69			
ND2	H	4049	5094	1046	ATG	TA—	
$tRNA^{Trp}$	H	5095	5163	69			1
$tRNA^{Ala}$	L	5165	5233	69			
$tRNA^{Asn}$	L	5234	5306	73			
OL[e]	-	5307	5339	33			
$tRNA^{Cys}$	L	5340	5406	67			1
$tRNA^{Tyr}$	L	5408	5477	70			1
COI	H	5479	7035	1557	GTG	TAA	3
$tRNA^{Ser\ (UGA)}$	L	7039	7109	71			5
$tRNA^{Asp}$	H	7115	7184	70			2
COII	H	7187	7877	691	ATG	T-	
$tRNA^{Lys}$	H	7878	7951	74			1
ATP8	H	7953	8120	168	ATG	TAA	−10
ATP6	H	8111	8793	683	ATG	TA-	
COIII	H	8794	9579	786	ATG	TAA	2
$tRNA^{Gly}$	H	9582	9651	70			
ND3	H	9652	10,000	349	ATG	T–	
$tRNA^{Arg}$	H	10,001	10,069	69			
ND4L	H	10,070	10,366	297	ATG	TAA	−7
ND4	H	10,360	11,737	1378	ATG	T–	3
$tRNA^{His}$	H	11,741	11,809	69			
$tRNA^{Ser\ (GCU)}$	H	11,810	11,876	67			
$tRNA^{Leu\ (UAG)}$	H	11,877	11,948	72			
ND5	H	11,949	13,787	1839	ATG	TAA	−18
ND6	L	13,770	14,294	525	ATG	TAA	
$tRNA^{Glu}$	L	14,295	14,364	70			2
Cytb	H	14,367	15,511	1145	ATG	TA-	
$tRNA^{Thr}$	H	15,512	15,581	70			
$tRNA^{Pro}$	L	15,582	15,650	69			
D-loop	–	15,651	16,875	1225			

[a] H, heavy strand; L, light strand.
[b] Includes stop codon also.
[c] T or TA indicates incomplete stop codon.
[d] Numbers indicate nucleotides separating two adjacent genes. Negative numbers indicate overlapping nucleotides.
[e] Origin of light strand replication.

Please cite this article as: Alam, M.T., et al., The complete mitochondrial genome sequence of the world's largest fish, the whale shark (*Rhincodon typus*), and its comparison with those of related shark species, Gene (2014), http://dx.doi.org/10.1016/j.gene.2014.01.064

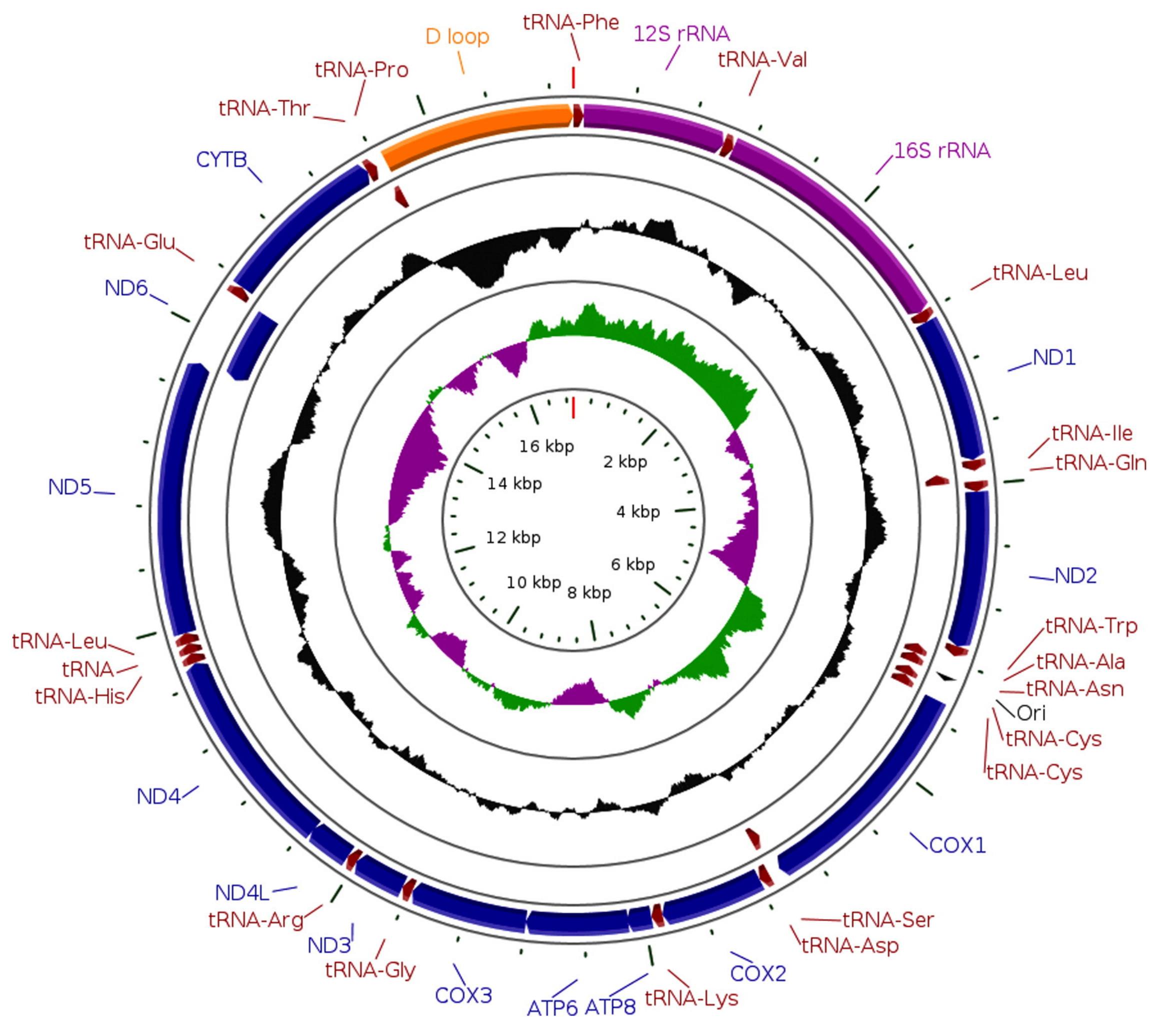


Fig. 1. Gene organisation map of *R. typus* mitogenome. The protein-coding genes, tRNAs, rRNAs, and two non-coding regions are shown in different colours. Direction of the arrows on the map indicates orientation of the genes on the heavy (H) and light (L) strand of the mitogenome. The black ring in the middle shows GC content (outer and inner peaks indicating above or below average GC content, respectively), whereas the innermost purple-green ring shows GC skew [(G − C/G + C), purple if between −1 and 0, green if between 0 and +1] of the mitogenome. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

found to exhibit extensive single-nucleotide and size polymorphism in several species of fish including *R. typus* (Castro et al., 2007; Ramírez-Macías et al., 2007). Castro et al (2007) observed 44 haplotypes of the control region with size ranging from 1143 to 1332 bp, in 70 *R. typus* individuals analysed from around the world, including 9 from Taiwan. The control region of the Taiwanese shark sequenced in our study was 1225 bp long, and is identical to haplotype H37 (a Western Indian Ocean shark) and nearly identical (1 bp difference) to the haplotypes H5 (a Western Indian Ocean shark) and H32 (a Northwest Pacific/Taiwanese Ocean shark) reported by Castro et al. (2007). Similarly, Ramírez-Macías et al. (2007) observed 14 different haplotypes of the control region amongst 36 individuals analysed from the Gulf of California, Mexico. Unlike Castro et al., however, that study analysed only 650 bp fragment of the control region (Ramírez-Macías et al., 2007).

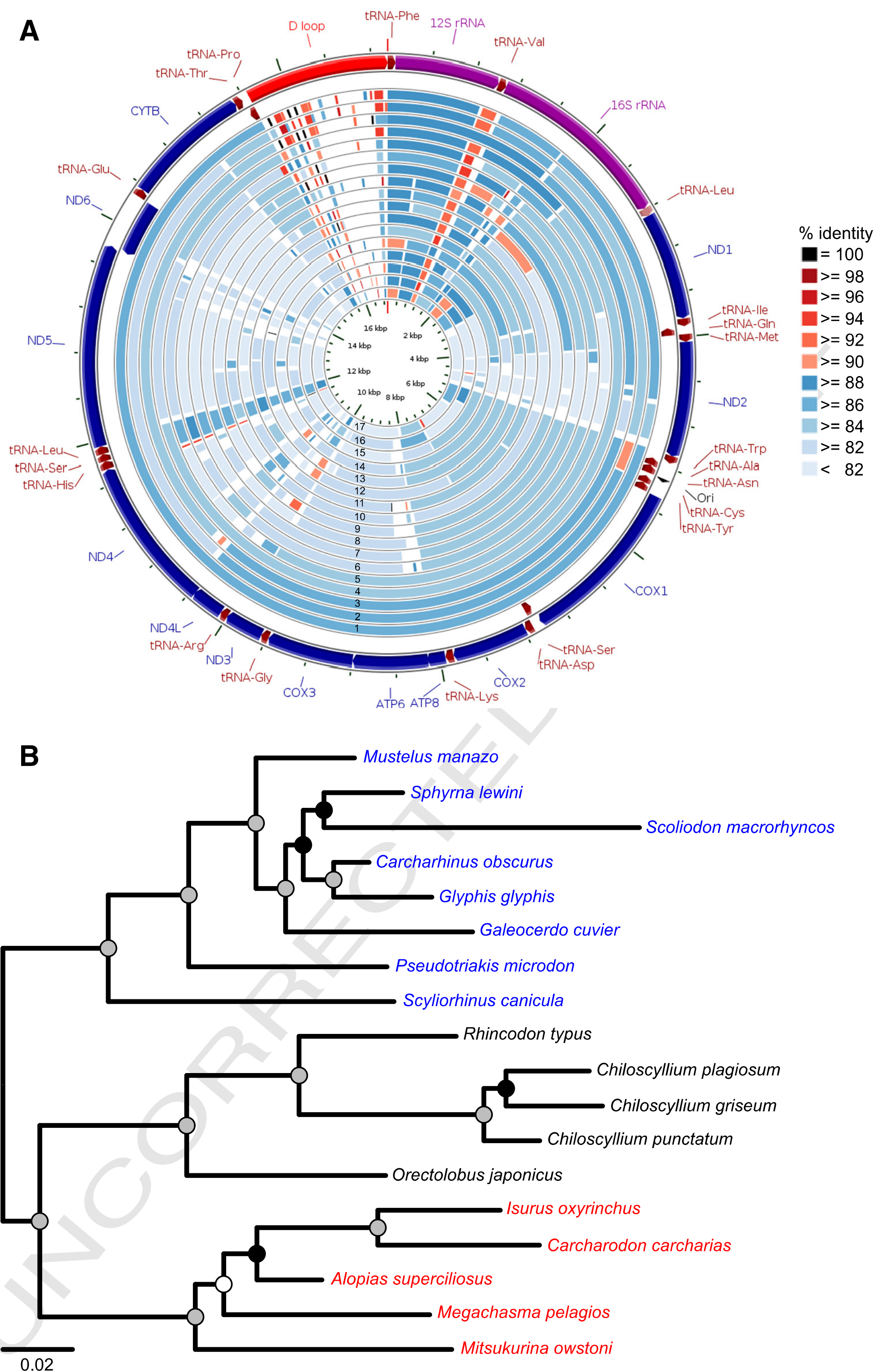
3.2. Comparison with closely related shark species

The mitogenomes of 17 shark species closely related to *R. typus* have been sequenced recently, four from the order Orectolobiformes, five from Lamniformes and eight from Carcharhiniformes (Table 1). The comparative mitogenome analysis revealed that the nucleotide composition, number and arrangement of the genes are same in all the shark species compared. However, the *R. typus* mitogenome was slightly longer than the mitogenomes of other shark species because of the longer control region. The sequence identity between *R. typus* and other shark species varied between 71 and 83% at nucleotide level and 87–93% at protein level (Fig. 2A). As can be seen in the CCT BLAST map, the control region sequences of these sharks are highly divergent (Fig. 2A). As expected, all five orectolobiform sharks, including *R. typus*, formed a monophyletic clade. Similarly, members of the orders Lamniformes and Carcharhiniformes grouped together in their independent clade on both ML and Bayesian phylogenetic trees (Fig. 2B). As shown, both ML and Bayesian methods produced identical tree topology with strong bootstrap (for ML tree) and posterior probability (for Bayesian tree)-supported branch nodes (Fig. 2B). The branch supports and topology of the tree were identical under both MtMAM + G (ML log likelihood score = −26,692.96) and MtREV + G (ML log likelihood score = −26,981.06) evolutionary models tested. The phylogenetic proximity of *R. typus* to other shark species compared here was consistent with the positions of these species on a traditional phylogeny and the phylogenies based on few mitochondrial genes (Velez-Zuazo and Agnarsson, 2011).

4. Conclusion

The mitogenome of whale shark *R. typus* was found to be a 16,875 bp circle with typical features of a vertebrate mitogenome, consisting of 13 protein-coding genes, two rRNA genes, 22 tRNA genes and a control

Please cite this article as: Alam, M.T., et al., The complete mitochondrial genome sequence of the world's largest fish, the whale shark (*Rhincodon typus*), and its comparison with those of related shark species, Gene (2014), http://dx.doi.org/10.1016/j.gene.2014.01.064



UNCORRECTED

Fig. 2. (A) Graphical map of the BLAST results showing nucleotide identity between *R. typus* mitogenome and 17 other shark species listed in Table 1, as generated by the CGView comparison tool (CCT). CCT arranges BLAST result in an order where sequence that is most similar to the reference (*R. typus*) is placed closer to the outer edge of the map. The rings labelled 1 to 17 indicate BLAST results of *R. typus* mitogenome against *C. punctatum*, *C. plagiosum*, *C. griseum*, *O. japonicus*, *A. superciliosus*, *C. obscurus*, *G. glyphis*, *G. cuvier*, *S. lewini*, *S. macrorhynchos*, *P. microdon*, *M. pelagios*, *M. owstoni*, *M. manazo*, *S. canicula*, *C. carcharias* and *I. oxyrinchus*, respectively. (B) Protein-based phylogenetic tree of 18 shark species. Both ML and Bayesian analyses produced identical tree topologies. The ML bootstrap and Bayesian posterior probability values for each node are indicated (grey circles: bootstrap value ≥ 90% and posterior probability of 1; black circles: bootstrap value < 90% and posterior probability of 1; white circle: bootstrap value < 90% and posterior probability < 1).

region or D-loop. Given that the whale shark is the largest extant fish and also a vulnerable species; we hope that the availability of its complete mitogenome sequence will prove an important resource that will help to better understand its biology. These studies should include more complete analyses of phylogenetic relationships with other shark species, conservation genetic studies that more thoroughly address the sub-structure of the global whale shark population, and studies that explore the energy metabolism of this species and how it differs from

Please cite this article as: Alam, M.T., et al., The complete mitochondrial genome sequence of the world's largest fish, the whale shark (*Rhincodon typus*), and its comparison with those of related shark species, Gene (2014), http://dx.doi.org/10.1016/j.gene.2014.01.064

ARTICLE IN PRESS

other sharks, since it has been proposed that filter feeding sharks exist on a metabolic knife-edge relative to food abundance. These studies should also aid in conservation efforts for this species and the habitats where it occurs, all of which contributes to developing the whale shark as a charismatic ambassador for the pelagic oceans.

Acknowledgments

The authors thank those current and former staff at Georgia Aquarium whose efforts supported this study, particularly Chris Coco, Chris Schreiber, Tonya Clauss, Tim Binder, Ray Davis, Greg Bossart, Harry Webb, and the husbandry team of the Ocean Voyager exhibit. We would also like to acknowledge the support of the 454 service facility in Branford, CT and the Emory Genome Center (EGC). DNA was prepared at the EGC by Chad Haase. Equipment used by the EGC was funded by the Atlanta Clinical and Translational Sciences Institute, the Department of Human Genetics, and the School of Medicine. Funding for sequencing and analysis was provided by the Georgia Aquarium and Emory University.

References

Abascal, F., Zardoya, R., Posada, D., 2005. ProtTest: selection of best-fit models of protein evolution. Bioinformatics 21, 2104–2105.
Ahonen, H., Harcourt, R.G., Stow, J., 2009. Nuclear and mitochondrial DNA reveals isolation of imperilled grey nurse shark populations (*Carcharias taurus*). Mol. Ecol. 18, 4409–4421.
Blower, D.C., Hereward, J.P., Ovenden, J.R., 2013. The complete mitochondrial genome of the dusky shark *Carcharhinus obscurus*. Mitochondrial DNA (130404005116000).
Cao, Y., et al., 1998a. Conflict among individual mitochondrial proteins in resolving the phylogeny of eutherian orders. J. Mol. Evol. 47, 307–322.
Cao, Y., Waddell, P.J., Okada, N., Hasegawa, M., 1998b. The complete mitochondrial DNA sequence of the shark *Mustelus manazo*: evaluating rooting contradictions to living bony vertebrates. Mol. Biol. Evol. 15, 1637–1646.
Castro, A.L.F., et al., 2007. Population genetic structure of earth's largest fish, the whale shark (*Rhincodon typus*). Mol. Ecol. 16, 5183–5192.
Chang, C.-H., Shao, K.-T., Lin, Y.-S., Chiang, W.-C., Jang-Liaw, N.-H., 2013a. Complete mitochondrial genome of the megamouth shark *Megachasma pelagios* (Chondrichthyes, Megachasmidae). Mitochondrial DNA 1–3.
Chang, C.-H., Shao, K.-T., Lin, Y.-S., Fang, Y.-C., Ho, H.-C., 2013b. The complete mitochondrial genome of the great white shark, *Carcharodon carcharias* (Chondrichthyes, Lamnidae). Mitochondrial DNA 1–2.
Chang, C.-H., Shao, K.-T., Lin, Y.-S., Ho, H.-C., Liao, Y.-C., 2013c. The complete mitochondrial genome of the big-eye thresher shark, *Alopias superciliosus* (Chondrichthyes, Alopiidae). Mitochondrial DNA 1–3.
Chang, C.-H., Shao, K.-T., Lin, Y.-S., Tsai, A.-Y., Su, P.-X., Ho, H.-C., 2013d. The complete mitochondrial genome of the shortfin mako, *Isurus oxyrinchus* (Chondrichthyes, Lamnidae). Mitochondrial DNA.
Chen, X., Ai, W., Ye, L., Wang, X., Lin, C., Yang, S., 2013a. The complete mitochondrial genome of the grey bamboo shark (*Chiloscyllium griseum*) (Orectolobiformes: Hemiscylliidae): genomic characterization and phylogenetic application. Acta Oceanol. Sin. 32, 59–65.
Chen, X., Liu, M., Grewe, P.M., Kyne, P.M., Feutry, P., 2013b. Complete mitochondrial genome of the Critically endangered speartooth shark *Glyphis glyphis* (Carcharhiniformes: Carcharhinidae). Mitochondrial DNA.
Chen, X., Liu, M., Xiang, D., Ai, W., 2013c. Complete mitochondrial genome of the Japanese wobbegong *Orectolobus japonicus* (Orectolobiformes: Orectolobidae). Mitochondrial DNA.
Chen, X., Peng, X., Zhang, P., Yang, S., Liu, M., 2013d. Complete mitochondrial genome of the spadenose shark (*Scoliodon macrorhynchos*). Mitochondrial DNA.
Chen, X., Xiang, D., Xu, Y., Shi, X., 2013e. Complete mitochondrial genome of the scalloped hammerhead *Sphyrna lewini* (Carcharhiniformes: Sphyrnidae). Mitochondrial DNA.
Chen, X., Yu, J., Zhang, S., Ding, W., Xiang, D., 2013f. Complete mitochondrial genome of the tiger shark *Galeocerdo cuvier* (Carcharhiniformes: Carcharhinidae). Mitochondrial DNA.
Chen, X., Zhou, Z., Pichai, S., Huang, X., Zhang, H., 2013g. Complete mitochondrial genome of the brownbanded bamboo shark *Chiloscyllium punctatum*. Mitochondrial DNA 1–2.
Compagno, L.J.V., 2001. Sharks of the world. An annotated and illustrated catalogue of shark species known to date. Bullhead, mackerel and carpet sharks (Heterodontiformes, Lamniformes and Orectolobiformes). FAO Species Catalogue for Fishery Purposes No. 1, vol. 2. FAO, Rome (269 pp.).
Delarbre, C., et al., 1998. The complete nucleotide sequence of the mitochondrial DNA of the dogfish, *Scyliorhinus canicula*. Genetics 150, 331–344.
Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32, 1792–1797.
Felsenstein, J., 1985. Confidence limits on phylogenies: an approach using bootstrap. Evolution 39, 783–791.
Galtier, N., Nabholz, B., Glémin, S., Hurst, G.D.D., 2009. Mitochondrial DNA as a marker of molecular diversity: a reappraisal. Mol. Ecol. 18, 4541–4550.
Grant, J.R., Arantes, A.S., Stothard, P., 2012. Comparing thousands of circular genomes using the CGView Comparison Tool. BMC Genomics 13, 1-1.
IUCN, 2013. The IUCN red list of threatened species. Version 2013.1. http://www.iucnredlist.org (Downloaded on 02 July 2013).
Lagesen, K., Hallin, P., Rødland, E.A., Staerfeldt, H.-H., Rognes, T., Ussery, D.W., 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res. 35, 3100–3108.
Li, H., Durbin, R., 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics 25, 1754–1760.
Li, H., Durbin, R., 2010. Fast and accurate long-read alignment with Burrows–Wheeler transform. Bioinformatics 26, 589–595.
Lowe, T.M., Eddy, S.R., 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. 25, 955–964.
Martin, A.P., Naylor, G.J.P., 1997. Independent origins of filter-feeding in Megamouth and Basking sharks (Order *Lamniformes*) inferred from phylogenetic analysis of Cytochrome *B* gene sequences. In: Yano, K., Morrissey, J.F., Yabumoto, Y., Nakaya, K. (Eds.), Biology of the Megamouth Shark. Tokai Univ. Press, Tokyo, pp. 39–50.
Motta, P.J., et al., 2010. Feeding anatomy, filter-feeding rate, and diet of whale sharks *Rhincodon typus* during surface ram filter feeding off the Yucatan Peninsula, Mexico. Zoology (Jena, Germany). 113, 199–212.
Powell, A.F.L.A., Barker, F.K., Lanyon, S.M., 2013. Empirical evaluation of partitioning schemes for phylogenetic analyses of mitogenomic data: an avian case study. Mol. Phylogenet. Evol. 66, 69–79.
Ramírez-Macías, D., Vázquez-Juárez, R., Galván-Magaña, F., Munguía-Vega, A., 2007. Variations of the mitochondrial control region sequence in whale sharks (*Rhincodon typus*) from the Gulf of California, Mexico. Fish. Res. 84, 87–95.
Ronquist, F., et al., 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. 61, 539–542.
Rowat, D., Brooks, K.S., 2012. A review of the biology, fisheries and conservation of the whale shark *Rhincodon typus*. J. Fish Biol. 80, 1019–1056.
Schmidt, J.V., et al., 2009. Low genetic differentiation across three major ocean populations of the whale shark, *Rhincodon typus*. PLoS ONE 4, e4988.
Simpson, J.T., Wong, K., Jackman, S.D., Schein, J.E., Jones, S.J.M., Birol, I., 2009. ABySS: a parallel assembler for short read sequence data. Genome Res. 19, 1117–1123.
Stamatakis, A., 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22, 2688–2690.
Tanaka, K., et al., 2013. Evolutionary relations of Hexanchiformes deep-sea sharks elucidated by whole mitochondrial genome sequences. Biomed. Res. Int. (147064).
Taylor, 2007. Ram filter-feeding and nocturnal feeding of whale sharks (*Rhincodon typus*) at Ningaloo Reef, Western Australia. Fish. Res. 84, 6-6.
Velez-Zuazo, X., Agnarsson, I., 2011. Shark tales: a molecular species-level phylogeny of sharks (Selachimorpha, Chondrichthyes). Mol. Phylogenet. Evol. 58, 207–217.
Wataru, I., et al., 2013. MitoFish and MitoAnnotator: a mitochondrial genome database of fish with an accurate and automatic annotation pipeline. Mol. Biol. Evol. 1–30.
Yu, L.L., Li, Y.-W.Y., Ryder, O.A.O., Zhang, Y.-P.Y., 2007. Analysis of complete mitochondrial genome sequences increases phylogenetic resolution of bears (Ursidae), a mammalian family that experienced rapid speciation. BMC Evol. Biol. 7, 198-198.
Zerbino, D.R., Birney, E., 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. 18, 821–829.

UNCORRECTED PROOF

Please cite this article as: Alam, M.T., et al., The complete mitochondrial genome sequence of the world's largest fish, the whale shark (*Rhincodon typus*), and its comparison with those of related shark species, Gene (2014), http://dx.doi.org/10.1016/j.gene.2014.01.064