

RESEARCH ARTICLE

Open Access

Mitogenomics of the Spartooth Shark challenges ten years of control region sequencing

Pierre Feutry^{1*}, Peter M Kyne¹, Richard D Pillans², Xiao Chen³, Gavin JP Naylor⁴ and Peter M Grewe⁵

Abstract

Background: Mitochondrial DNA markers have long been used to identify population boundaries and are now a standard tool in conservation biology. In elasmobranchs, evolutionary rates of mitochondrial genes are low and variation between distinct populations can be hard to detect with commonly used control region sequencing or other single gene approaches. In this study we sequenced the whole mitogenome of 93 Critically Endangered Spartooth Shark *Glyphis glyphis* from the last three river drainages they inhabit in northern Australia.

Results: Genetic diversity was extremely low ($\pi = 0.00019$) but sufficient to demonstrate the existence of barriers to gene flow among river drainages (AMOVA $\Phi_{ST} = 0.28283$, $P < 0.00001$). Surprisingly, the comparison with single gene sub-datasets revealed that ND5 and 12S were the only ones carrying enough information to detect similar levels of genetic structure. The control region exhibited only one mutation, which was not sufficient to detect any structure among river drainages.

Conclusions: This study strongly supports the use of single river drainages as discrete management units for the conservation of *G. glyphis*. Furthermore when genetic diversity is low, as is often the case in elasmobranchs, our results demonstrate a clear advantage of using the whole mitogenome to inform population structure compared to single gene approaches. More specifically, this study questions the extensive use of the control region as the preferential marker for elasmobranch population genetic studies and whole mitogenome sequencing will probably uncover a large amount of cryptic population structure in future studies.

Keywords: Population genetics, Elasmobranchs, Philopatry, IUCN Red List, Dispersal, D-loop, Low genetic diversity

Background

Genetic markers have long been used to discriminate between populations and define evolutionary significant units [1] or management units [2], including for marine or euryhaline species for which barriers to migration are rarely easy to identify [3,4]. The choice of marker employed is critical as it has a strong influence on the ability to identify discrete populations. Markers with low mutation rates could prevent or reduce the probability of detecting population differentiation, even in truly non-panmictic populations [5]. Nearly three decades ago, Avise & Ellis [6] advocated using mitochondrial DNA (mtDNA) genes as preferential markers in phylogeographic genetic studies. Mitochondrial genes have since proved to be useful

markers to infer species or population boundaries, and have become a standard tool in conservation biology [7,8].

Progeny generally inherit their mtDNA from the female parent. As such, the analysis of mtDNA reflects matrilineal lineage processes only, providing information unobtainable from any nuclear marker [9]. This is particularly important in groups such as elasmobranchs (sharks and rays), where there is accumulating evidence of reproductive philopatry in females but not in males [10-12].

Mitochondrial DNA markers have been widely used in studies of teleost and elasmobranch fish populations over the past decade with much attention focused on the D-loop or control region (CR). This region has been shown to be extremely variable in many vertebrates including humans [13], teleosts [14] and birds [15]. In elasmobranchs however, evolutionary rates of mitochondrial genes are low compared to other vertebrates [16], and variation between distinct populations can sometimes be hard to detect with commonly used CR sequencing

* Correspondence: Pierre.Feutry@cdu.edu.au

¹Charles Darwin University, Research Institute for the Environment and Livelihoods, Ellengowan Drive, Darwin 0909, Northern Territory, Australia
Full list of author information is available at the end of the article

approaches [17,18]. However, a thorough examination of intraspecific rates of evolution of the different mtDNA regions in sharks and rays has not yet been carried out, and there may be alternative regions exhibiting useful polymorphisms that could be exploited for population genetic analysis. Interestingly, the highest single gene mitochondrial nucleotide diversity reported in any elasmobranch species so far was in the Dark Shyshark *Haploblepharus pictus* using the COXI gene [19] followed by the Discus Ray *Paratrygon aiereba* using the ND6 gene [20]. Also, Naylor et al. [21] collected 4,283 ND2 sequences and looked at intraspecific divergences in 595 different elasmobranch species. Many were found to show substantial differences among populations. In order to increase the ability to detect population structure, a few studies have included the analysis of more than one mtDNA marker with variable success [17,22,23]. One study included a preliminary screening of four different regions to find the most variable one, which happened to be the ND4 gene followed by the ATP gene and then the CR [24]. These results suggest that in elasmobranchs the CR could be under higher evolutionary constraints (compared to other mitochondrial regions) than in other vertebrates.

Recent advances in next generation sequencing have significantly reduced costs in a way that could easily facilitate routine investigation of intraspecific variation of whole elasmobranch mtDNA genomes at typical sampling levels used for standard population analyses. In the past, whole mitogenome sequencing has only been used in phylogenetic studies, where the number of individuals sequenced was generally low. Compared to single gene approaches, whole mitogenome sequencing can sometimes provide higher resolution of phylogenies and increased precision of divergence time estimates [25]. More recently, whole mitogenome sequencing has been used to investigate diversity at the intra-species level. For example, it provided improved resolution of population structure in the Green Turtle *Chelonia mydas* [26] and the North American Fisher *Martes pennanti* (a weasel) [27]. It has also been used in species delineation [28] and the detection of genes under selection in the Killer Whale *Orcinus orca* [29]. There are currently no publications using this method in elasmobranchs.

The Spartooth Shark *Glyphis glyphis* is a rare, habitat- and range-restricted, river shark of northern Australia and southern Papua New Guinea [30]. It is confirmed from only eight rivers of northern Australia, with the majority of records coming from Van Diemen Gulf drainages (Adelaide, South Alligator, and East Alligator Rivers) in the Northern Territory, and the Wenlock River, Western Cape York Peninsula, Queensland. There is an apparent large gap (>1,000 km) in its distribution between the Van Diemen Gulf drainages and the Wenlock River [30]. Its

rarity, limited range, and the suspected impacts of fishing activities resulted in a listing of Critically Endangered on the Australian *Environment Protection and Biodiversity Conservation Act* 1999. Little is known of its biology: juveniles and sub-adults have been recorded across a large salinity range in rivers, however, adults of this species have never been recorded and their range and habitat requirements are unknown [30]. An understanding of population structure is required to inform management and conservation measures for this species. Wynen et al. [17] developed a DNA barcoding approach for confirming species delineation using portions of the COX1 gene and the CR. This approach proved to be successful in distinguishing between *G. glyphis* and the sympatric Northern River Shark *G. garricki* and Bull Shark *Carcharhinus leucas*, but no intraspecific variation was found across 12 and 17 *G. glyphis* individuals from the Northern Territory and Queensland, respectively, thus preventing any assessment of potential population structure.

The aim of this study was to explore the potential of whole mitogenome sequencing for phylogeographic studies in elasmobranchs, using *G. glyphis* as a case study. We anticipated that this approach would provide useful information for the conservation of this Critically Endangered species, as well as giving insights into the intraspecific genetic variability of the different mitogenomic regions in elasmobranchs.

Results

Data reliability

The complete mitogenomes of 93 *G. glyphis* were successfully sequenced and assembled (Genbank Accessions, KM100613 - KM100704). Individual mean coverage ranged from 152 to over 6,000 fold ensuring a high quality of mitogenome sequencing. The duplicated sequencing of the *G. glyphis* reference mitogenome [31] provided the exact same sequence twice, demonstrating a high degree of reliability in the dataset.

Less than 0.5% of the reads from *G. glyphis* samples were successfully mapped on the Largetooth Sawfish *Pristis pristis* mitogenome, the other elasmobranch species present in the Miseq run. Those reads were mostly mapped on the highly conserved 16S region, thus suggesting cross contamination did not occur during library preparation.

Mitogenomic diversity

Twelve haplotypes were observed among the 93 fish from the three river drainages sampled (Figure 1). No insertions or deletions were observed, and the *G. glyphis* mitogenome was calculated to be 16,701 bp in length. Interestingly, no transversions were observed but there were 19 transitions spread across 12 haplotypes. Nucleotide diversity (π) was 0.00019 and haplotype diversity (H_d) 0.76. Only one amino acid change was found in the

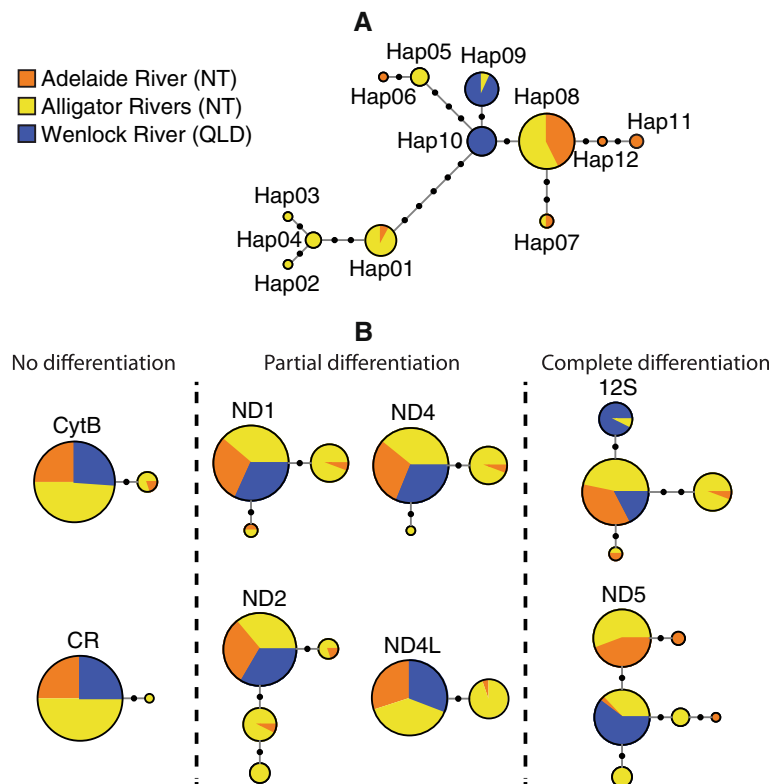


Figure 1 Haplotype networks inferred from (A) whole mitogenome dataset; (B) single region sub-datasets. Black circles on connecting lines between haplotypes represent single mutation steps for haplotypes not found in our current data.

complete dataset, at position 1348 of the ND5 alignment; this was not river drainage specific. The highest genetic diversity was found in the Alligator Rivers ($\pi = 0.00024$ and $H_d = 8$), whereas the Wenlock River exhibited the lowest genetic diversity ($\pi = 0.00003$ and $H_d = 2$). River drainages each contained from one to four drainage-specific haplotypes (Figure 1). Genetic diversity indices for each of the drainage locations derived from the whole genome dataset are given in Table 1.

The ‘all proteins’ sub-dataset recovered 13 out of the 19 variable sites. Among the single gene sub-datasets, ND5 was the most diverse with 5 variable sites and a nucleotide diversity of 0.00042, followed by ND2 with 3 variable sites and a nucleotide diversity of 0.00051. The 12S region harbored 4 variable sites for a nucleotide diversity of 0.00101. Several sub-datasets, including 16S, COX1, COX2, COX3, ND3, ND6, ATP6 and ATP8 did not exhibit any variation. Additional file 1: Table S1 summarizes the variable site positions per mitogenome region and per haplotype.

With the exception of Tajima’s D in the Adelaide River, neutrality tests calculated from the whole mitogenome were not significant (Table 1). Neutrality tests calculated from single region sub-dataset were generally concordant

with this result. A complete set of genetic diversity indices for the whole mitogenome and all sub-datasets is available in Additional file 2: Table S2.

Population structure

Significant genetic structure was detected among the three river drainages (AMOVA $\Phi_{ST} = 0.28283$, $P < 0.00001$).

Table 1 Summary of genetic diversity indices inferred from whole mitogenomes across the study sites and for each of the three river drainages sampled

| Diversity indices | All samples | Adelaide River | Alligator Rivers | Wenlock River |
|--------------------|-------------|----------------|------------------|---------------|
| π | 0.00019 | 0.00009 | 0.00024 | 0.00003 |
| Variable sites - S | 19 | 14 | 16 | 1 |
| Singleton sites | 3 | 11 | 5 | 0 |
| Haplotypes | 12 | 6 | 8 | 2 |
| H_d | 0.760 | 0.458 | 0.685 | 0.514 |
| Tajima’s D | -0.42548 | -2.11498 | 0.32904 | 1.53267 |
| Significance | $P > 0.10$ | $P = 0.001$ | $P > 0.10$ | $P > 0.10$ |
| Fu’s F_s | -0.22900 | -0.79791 | 2.18415 | 1.52985 |
| Significance | $P > 0.10$ | $P > 0.10$ | $P > 0.10$ | $P > 0.10$ |

The pairwise differentiation Φ_{ST} values ranged from 0.17755 to 0.53088 and all were significant (Table 2). The genetic differentiation inferred from each sub-dataset varied greatly (Additional file 3: Table S3). Apart from the ‘all proteins’ sub-dataset, only two single region sub-datasets, 12S and ND5, recovered significant Φ_{ST} values for all population pairs. ND1, ND2, ND4 and ND4L suggested the existence of barriers to gene flow between the Adelaide River and the Alligator Rivers, and between the Alligator Rivers and the Wenlock River, but not between the Adelaide River and the Wenlock River. No signs of genetic differentiation were found in CytB and CR, the remaining sub-datasets that exhibited some degree of genetic variation. Interestingly, overall values of Φ_{st} appeared concordant with geographical distances among the three populations, with Adelaide-Alligators ($\Phi_{st} = 0.18$) having lower values than the Alligators-Wenlock and Adelaide-Wenlock ($\Phi_{st} = 0.30$ and 0.53 , respectively).

Discussion

The whole mitogenomic survey using next generation sequencing (MiSeq, Illumina) proved to be an efficient and accurate method for routine surveys of *G. glyphis* population structure in Australia. All mitogenomic sequences were easily amplified by PCR and sequenced with a high degree of reliability in a time-efficient manner. The data collection pipeline, from PCR amplification to the mitogenome alignment ready for analysis, could be completed in one week if all individuals are sequenced on the same Miseq run. Despite some discrepancies in coverage across samples, the high read coverage provided by the Miseq ensured sufficient coverage for each individual. Additionally, the attempt to map reads from *G. glyphis* barcodes on the *P. pristis* reference mitogenome demonstrated that cross contamination of sequence data was not significant. Following the method described here, whole mitogenome sequencing is currently 4-5 fold more expensive than single fragment analysis by Sanger sequencing. However, this method could be optimized in several ways to make the best use of the high number of reads offered by high throughput sequencers, including (i) pooling more libraries on a given sequencing run [32]; or (ii) pooling several species in each library [33]. Whole mitogenome

sequencing costs will also continue to fall as technology improves.

Mitogenomic sequencing in *G. glyphis* revealed particularly low levels of genetic diversity. Previous to this study, the species that exhibited the lowest recorded genetic diversity inferred from a whole mitogenome was the Giant Squid, *Architeuthis* sp., in a survey of 43 individuals from 10 sampling locations [34]. This species harbored a nucleotide diversity of 0.00066, more than three times higher than that of *G. glyphis*. The low diversity in the Giant Squid is believed to be due to a recent bottleneck followed by expansion, possibly coupled with a low mutation rate [34]. With the exception of the Adelaide River, there were no signs of recent population expansion in *G. glyphis*, suggesting that a bottleneck followed by population expansion is unlikely to be the best explanation for this species.

A low mutation rate, however, could partly explain such low diversity: rates of mitochondrial DNA evolution in elasmobranchs are slow compared to other taxa [16]. To the best of our knowledge, there is no other study investigating the intraspecific variability of whole mitogenomic sequences in elasmobranchs to compare with our data. However, our review of the current literature examining single gene analysis of mitochondrial diversity in sharks and rays (Additional file 4: Table S4) indicates that slow mutation rate alone cannot completely explain the low nucleotide diversity found in *G. glyphis*. All species included in this review had populations with higher nucleotide diversity indices; 42 out of 50 species exhibited nucleotide diversity indices at least one order of magnitude higher than *G. glyphis* (Additional file 4: Table S4). Some shark species however, exhibited populations without any or only very low levels of nucleotide diversity for the gene under study. Again, a recent bottleneck event was proposed to explain the low genetic diversity observed in the Basking Shark *Cetorhinus maximus* [35] and the Nurse Shark *Ginglymostoma cirratum* [36], two coastal species with large ranges. In the Grey Nurse Shark *Carcharias taurus*, demographic events in the deep history of the species (bottlenecks or founder effects) and especially low levels of molecular evolution were proposed to explain current levels of genetic diversity [37]. In addition to past demographic events, current effective population size can also affect levels of nucleotide diversity [38,39]. *Glyphis glyphis* has a naturally very limited range, being recorded in only a small number of tidal rivers and their estuaries in northern Australia and southern Papua New Guinea resulting in a low expected total number of individuals for this species [40]. We thus argue that the low genetic diversity observed in *G. glyphis* is probably due to the low evolutionary rate of mitochondrial genes (common in elasmobranch species, [16])

Table 2 Pairwise Φ_{st} (below) and associated p-values (above) for population comparisons derived from the whole mitogenome dataset

| Population | Adelaide River | Alligator Rivers | Wenlock River |
|------------------|----------------|------------------|---------------|
| Adelaide River | - | 0.00168 | < 0.00001 |
| Alligator Rivers | 0.17756 | - | < 0.00001 |
| Wenlock River | 0.53088 | 0.30185 | - |

combined with an extremely small population size leading to subsequent loss of diversity through genetic drift [41]. These factors also apply in *G. garricki*, which shares with *G. glyphis* its limited habitat, limited range and probably small localized populations. This could explain the absence of genetic variation in the CR of 15 *G. garricki* analysed by Wynen et al. [17]. Low levels of genetic diversity were also observed in *C. leucas* in northern Australia, another species sharing juvenile life cycle and habitat characteristics with *G. glyphis* [23].

Despite very low observed genetic diversity among river drainages, our whole mitogenome sequencing approach was able to detect sufficient variation to provide the first insight into the population structure of *G. glyphis*. Haplotype diversity appears to be partitioned in a way that suggests strong female philopatry as each of the three river drainages is home to a distinct genetic population. Prior to our study, no evidence of population structure had been observed, however this was inferred from single gene approaches [17]. Our result is critical for the implementation of conservation measures to manage and ultimately recover this Critically Endangered species. Tissue samples of *G. glyphis* used in this study were obtained from juvenile and sub-adult animals captured in the Adelaide, Alligators and Wenlock River drainages. In Australia, this represents the entire known extant distribution of *G. glyphis* with the population in the Bizant River, Eastern Cape York Peninsula presumed extinct [30]. Given the population structuring and low levels of female dispersal highlighted in this study, remaining population centers need to be managed as discrete population units. The life history of *G. glyphis* is not fully understood, although given the apparent absence of adults from river systems, adults presumably occur in marine habitats [30]. The presence of a boundary to gene flow in females indicates that females return to the river of their birth to undertake parturition, but whether adult populations are also segregated remains unknown. As mtDNA strictly reflects female-mediated gene flow, further investigation of nuclear markers is warranted to evaluate patterns of male-mediated gene flow among these populations. In *C. leucas*, population structuring occurred only at very large scales and not between adjacent rivers [23]. *Glyphis glyphis* either has more discrete reproductive habitat preferences than *C. leucas* (at least for females) or the two gene approach used by Tillett et al. [23] did not have the power to detect structure at a finer scale, as would have been the case in *G. glyphis* using a similar approach (Figure 1; Additional file 3: Table S3).

The analysis of each sub-section of the mitogenome further highlighted the advantage of whole mitogenome sequencing compared to single gene approaches. Several commonly used mitogenome regions, including CR and COX1 were invariant and hence failed to detect any

barrier to gene flow between the three river drainages under study. These results were consistent with previous work on *G. glyphis* by Wynen et al. [17]. In the current study, ND1, ND2, ND4 and ND4L genes all provided partial insights into the population structuring of *G. glyphis*. A similar finding was made in *Chelonia mydas*, where whole mitogenomic sequences improved the population structure resolution compared to single gene analysis [26]. Interestingly, the strongest barrier to gene flow inferred from the whole mitogenome dataset (between the Adelaide River and the Wenlock River), was the most difficult to identify with single region datasets. This is because the haplotypes harbored by sharks from the Wenlock River differ from the most common haplotype by 1 or 2 mutations only. The uniqueness of those haplotypes relies on a very low number of mutations (Figure 1), clearly demonstrating the difference that can be made by sequencing the whole mitogenome compared to a subset of the mitogenome only. Not sequencing the mitogenome sections containing those critical mutations would have missed this population structure pattern. This has particularly important ramifications for species with low mutation rates such as elasmobranchs or in cases where isolation between populations has only occurred recently [26].

In *G. glyphis*, 12S and ND5 were the most variable mitochondrial regions and the only ones providing comparable resolution to that of the whole mitogenome to detect barriers to gene flow between each river drainage pair. This result challenges the assumptions commonly made over the past decade on which marker should be used in elasmobranch phylogeographic studies. Our literature review revealed that the CR was the most commonly chosen marker with ND4 and CytB the two next most common (CR, 40 studies; ND4, 7 studies; CytB, 6 studies) (Additional file 4: Table S4). In the absence of any comprehensive study investigating intraspecific rates of evolution in the different mitogenomic regions in elasmobranchs, we argue that the wide use of the CR in elasmobranchs relies mostly on an assumption that this region will be the most variable in the mitogenome, as observed in other taxa [13]. In fact, the few times CR has been compared to other regions, it was not found to be most variable [22,24]. We do not argue that the CR is always a poor marker for phylogeographic studies in elasmobranch species and should not be used. For example, at least five studies analyzing the CR in elasmobranch species report nucleotide diversity indices over 0.01 thus providing a good capacity to detect population structure [42-46]. Also, many other studies were successful in detecting some degree of population structure despite lower levels of nucleotide diversity in the CR (Additional file 4: Table S4). However, our study demonstrates clearly that other regions of the mitogenome should be examined as they potentially improve population resolution.

The 12S and ND5 regions are the two most variable markers reported here, but they have not previously been used for intra-specific phylogeographic studies in elasmobranchs. The 12S region, however, is often used to infer phylogenetic relationships between elasmobranch species [47] or for species identification [48]. Similarly, the ND5 gene has been rarely used in phylogeographic studies of elasmobranchs, although see [49] for *Manta* species delineation. Analysis of other elasmobranch species is required to further investigate the potential of 12S and ND5 as a source of mitochondrial intraspecific diversity.

Conclusions

The current study clearly demonstrates the difference that whole mitogenome sequencing can make in population genetic studies that focus on species with low genetic diversity. Given the current state of technology and sequencing costs, we strongly recommend the use of whole mitogenome sequencing for future population genetic studies in elasmobranchs. Intraspecific mitogenomic surveys may expose a large amount of cryptic population structure in species that have only been examined for a single mitochondrial region such as the CR. At a minimum, we suggest preliminary whole mitogenome analyses on a small number of individuals should be performed to identify the more variable mitochondrial regions when population-level sequencing is cost-prohibitive. Forthcoming mitogenomic surveys might also help to identify more general patterns of variability permitting a better choice of specific mitochondrial genes containing variable regions of interest in elasmobranchs.

Methods

Sampling, DNA extraction and long-range PCR amplification

Ninety-three *G. glyphis* were sampled from the Adelaide River (n = 23), the combined (South, East and West) Alligator Rivers (n = 47) (referred to as the Alligator Rivers) in the Northern Territory, and the Wenlock River (n = 23) in Queensland, Australia (Figure 2). This represents the entire extant northern Australian distribution of the species, with the Adelaide River the westernmost occurrence and the Wenlock River the easternmost occurrence [30]. Most sharks were sampled in 2012 and 2013, except six from the Adelaide River, five from the Alligator Rivers and thirteen from the Wenlock River that were sampled as part as a previous study (Wynen et al. [17]). Sharks ranged from neonates with open umbilical scars to sub-adults up to 156 cm TL, indicating that our samples comprised multiple year classes.

DNA was extracted from fin clips using DNeasy Blood & Tissue kits (Qiagen, Venlo, Netherlands). Whole mitogenomes were amplified in two overlapping fragments using Takara LA Taq, a proofreading Taq polymerase mixture

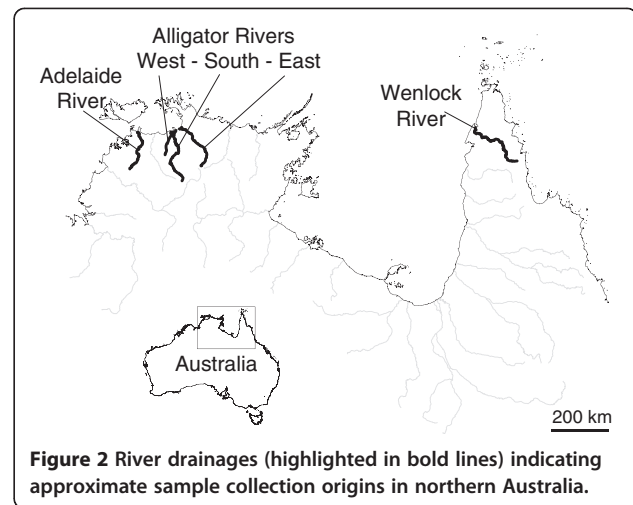


Figure 2 River drainages (highlighted in bold lines) indicating approximate sample collection origins in northern Australia.

(Takara, Otsu, Shiga, Japan). The PCR reaction mix consisted of 1 unit of Takara LA Taq, 2 μ l of 10x LA PCR Buffer II (Mg²⁺ free), 2 μ l of 25 mM MgCl₂, 3.2 μ l dNTP (2.5 mM each), 0.5 μ l of each primer at 10 mM, and 10.6 μ l sterilized distilled water. The primer pairs GL965F - 5' TATTTCTCCAACAAGAGGAGGCAAGTCGTAAC 3'/GL10307R - 5' GTAGTTGGTCAGAACCGTGT 3' and GL10014F - 5' GTCCAAATCAAGACCGCTAA 3'/GL1403R - 5' CTTTTGCCACAGAGACGG 3' were designed using previously acquired conserved areas of whole mitogenomic sequences of *G. glyphis* [31] and *G. garricki* [50]. PCR conditions consisted of an initial denaturation at 94°C for 3 min followed by 35 cycles of 94°C for 35 sec and 68°C for 10 min, with a final extension step of 72°C for 10 min. PCR amplicons for each individual were purified with Agencourt AMPure XP magnetic beads (Beckman Coulter Inc., Indianapolis, Indiana, USA), quantified with NanoDrop 8000 Spectrophotometer (Thermo Fisher Scientific, Waltham, Massachusetts, USA) and pooled at equimolar concentrations for each individual previous to the preparation of libraries.

Nextera library preparation and Miseq sequencing

The two purified fragments obtained for each individual were pooled at equimolar concentrations and subjected to library preparation using Nextera XT DNA Sample Preparation kits (Illumina, San Diego, California, USA). Individual Nextera XT libraries were simultaneously fragmented and barcoded using the 96 sample Nextera Index kit (Illumina, San Diego, California, USA). All libraries were quantified by a fluorometric method, using the Qubit dsDNA HS assay kit (Life Technologies, Carlsbad, California, USA), and concentrations were normalized to give equal mitogenome coverage from each individual fish. Libraries were then pooled and sequenced on a Miseq desktop sequencer using 2 × 250 bp paired-end reads Miseq reagent

kit v2 (Illumina, San Diego, California, USA). The sequencing was done in two different runs, approximately half of the samples for each run being from *P. pristis*, another elasmobranch species (data not shown).

Mitogenome assembly and data analysis

Fastq files were imported into GENEIOUS PRO (v. 7.0.6) (Biomatters Ltd., Auckland, New Zealand). Before mitogenome assembly, reads were paired; 5' and 3' ends as well as regions with more than 5% chance of an error per base were trimmed. Reads with more than 15 low quality bases and/or shorter than 150 bp after trimming were discarded from subsequent analyses. Reads for each individual were then mapped onto a previously published reference sequence [31] using 'Map to Reference' tool in GENEIOUS PRO without fine-tuning and default parameters. PCR duplicates were removed using the rmdup SAMTools toolkit (v. 1.0.0) [51]. The Majority rule consensus (>50% of reads for any single SNP, insertion, deletion) for each fish was exported. Consensus sequences were then aligned using MUSCLE alignment tool with default parameters in GENEIOUS PRO. The mitogenome of the exact same fish sequenced by Chen et al. [31] was amplified, sequenced once in each of the Miseq runs and assembled twice. Replicates were compared to the reference sequence to check for data reliability. As a second quality check, reads from each *G. glyphis* sample were mapped on the reference mitogenome of the second elasmobranch species included in the Miseq sequencing runs.

Preliminary reconstruction of a maximum likelihood haplotype tree indicated that the level of nucleotide variation among haplotypes was too low to provide a well-supported phylogenetic tree (Additional file 5: Figure S1). As an alternative to represent haplotype relationships, we drew a median-joining network [52] with NETWORK (v 4.6.1.2) (Fluxus Technology Ltd, Clare, Suffolk, England) using default parameters (Figure 1A). The same method was employed to draw haplotype networks for each single region sub-dataset (Figure 1B).

A range of genetic diversity indices was calculated for all sites and also each river separately using ARLEQUIN (v 3.5.1.2) [53]. This includes the nucleotide diversity index π [54], the number of variable sites S [55], the number of singleton sites, the number of haplotypes, and the haplotype diversity index H_d [54]. Tajima's D [56] and Fu's F_s [57] statistics were also calculated to infer the demographic history of *G. glyphis*.

Analyses of molecular variance (AMOVA) were conducted with ARLEQUIN (v 3.5.1.2), to test for the existence of genetic structure among animals from different river drainages, and we calculated the fixation index (Φ_{ST}) to estimate the degree of genetic differentiation among the three sampling locations. This was done for the whole mitogenome dataset and for each sub-dataset

containing variable sites. The model of nucleotide evolution used for the AMOVA and to calculate Φ_{ST} values was Tamura-Nei [58].

A literature review was conducted to evaluate preferences in mitochondrial markers and to estimate their degree of genetic diversity (π) in elasmobranchs (Additional file 4: Table S4). Scopus and Google Scholar were used as the search engines with the following keywords: phylogeography, elasmobranch(s), mitochondrial DNA, shark(s), ray(s), population structure, genetic diversity, genetic(s). This review was not intended to be comprehensive, but to provide a good representation of previous research. For this reason we avoided using marker name (e.g. COXI, 16S or control region) as keywords in order not to bias the review towards a particular marker.

Availability of supporting data

Mitogenomic sequences: Genbank accessions KM100613 - KM100704.

Additional files

Additional file 1: Table S1. Variable site positions in *G. glyphis* mitogenome alignment.

Additional file 2: Table S2. Genetic diversity indices per river and per mitogenomic region.

Additional file 3: Table S3. Pairwise Φ_{it} values derived from each mitogenomic region exhibiting some degree of genetic variation.

Additional file 4: Table S4. Mitochondrial nucleotide diversity π in elasmobranchs, literature review summary.

Additional file 5: Figure S1. Maximum likelihood haplotype tree.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

PF designed the research, carried out the molecular genetic studies and data analysis and drafted the manuscript. PMK and RDP carried out the sample collection and contributed to the draft of the manuscript. XC helped to perform the molecular studies. GJPN contributed to the draft of the manuscript. PMG contributed to the design of the study and helped to draft the manuscript. All authors read and approved the final manuscript.

Acknowledgements

Thanks to James Marthick at Menzies Research Institute Tasmania and Gareth Elvidge for their help using the Illumina Miseq, and Rasanthi Gunasekera for help in the laboratory. We also thank Ehsan Kayal for early discussions on DNA amplification and Jenny Ovenden for useful comments on the manuscript. This study was supported by the Marine Biodiversity Hub, a collaborative partnership supported through funding from the Australian Government's National Environmental Research Program (NERP). PF was partly supported by the North Australia Marine Research Alliance (NAMRA). PMK was partly supported by the NERP Northern Australia Hub. Thanks to Karen Gibb for access to samples collected previous to this study. Thanks also to the Traditional Owners and staff of Kakadu National Park, especially Anne O'Dea and Khan Spokes, and to Northern Territory Fisheries, in particular Grant Johnson, Mark Grubert and Thor Saunders, for field assistance. Samples were collected under NT Fisheries Special Permit S17/3252, Kakadu National Park Research Permit RK805, Queensland Fisheries General Research Permit 163582, and Charles Darwin University Animal Ethics Committee A11041. Finally we would like to thank the two anonymous reviewers for useful comments made on earlier versions of this manuscript.

Author details

¹Charles Darwin University, Research Institute for the Environment and Livelihoods, Ellengowan Drive, Darwin 0909, Northern Territory, Australia. ²CSIRO Oceans and Atmosphere Flagship, 41 Boggo Road, Dutton Park 4102, Queensland, Australia. ³Guangxi Key Lab for Mangrove Conservation and Utilization, Guangxi Mangrove Research Center, Guangxi Academy of Sciences, Beihai 536000, PR China. ⁴Department of Biology, College of Charleston, Charleston, SC 29412, USA. ⁵CSIRO Oceans and Atmosphere Flagship, Castray Esplanade, Hobart 7000, Tasmania, Australia.

Received: 13 August 2014 Accepted: 4 November 2014

Published online: 19 November 2014

References

- Moritz C: Defining 'evolutionarily significant units' for conservation. *Trends Ecol Evol* 1994, **9**(10):373–375.
- Taylor BL, Dizon AE: First policy then science: why a management unit based solely on genetic criteria cannot work. *Mol Ecol* 1999, **8**:S11–S16.
- King TL, Kalinowski ST, Schill WB, Spidle AP, Lubinski BA: Population structure of Atlantic salmon (*Salmo salar* L.): a range-wide perspective from microsatellite DNA variation. *Mol Ecol* 2001, **10**(4):807–821.
- Palsbøll PJ, Berube M, Allendorf FW: Identification of management units using population genetic data. *Trends Ecol Evol* 2007, **22**(1):11–16.
- Waples RS, Gaggiotti O: What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Mol Ecol* 2006, **15**(6):1419–1439.
- Avise JC, Ellis D: Mitochondrial DNA and the evolutionary genetics of higher animals [and discussion]. *Philos Trans R Soc B* 1986, **312**(1154):325–342.
- Rubinoff D: Utility of mitochondrial DNA barcodes in species conservation. *Conserv Biol* 2006, **20**(4):1026–1033.
- Moritz C: Applications of mitochondrial DNA analysis in conservation: a critical review. *Mol Ecol* 1994, **3**(4):401–411.
- Avise JC, Arnold J, Ball RM, Bermingham E, Lamb T, Neigel JE, Reeb CA, Saunders NC: Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annu Rev Ecol Syst* 1987, **18**(1):489–522.
- Portnoy DS: Molecular insights into elasmobranch reproductive behavior for conservation and management. In *Sharks and Their Relatives II: Biodiversity, Adaptive Physiology, and Conservation*, Volume 2. Edited by Carrier JC, Musick JA, Heithaus MR. Boca Raton, FL: CRC Press; 2010:435–457.
- Feldheim KA, Gruber SH, DiBattista JD, Babcock EA, Kessel ST, Hendry AP, Pikitch EK, Ashley MV, Chapman DD: Two decades of genetic profiling yields first evidence of natal philopatry and long-term fidelity to parturition sites in sharks. *Mol Ecol* 2014, **23**(1):110–117.
- Dudgeon CL, Blower DC, Broderick D, Giles JL, Holmes BJ, Kashiwagi T, Krück NC, Morgan JAT, Tillett BJ, Ovenden JR: A review of the application of molecular genetics for fisheries management and conservation of sharks and rays. *J Fish Biol* 2012, **80**(5):1789–1843.
- Aquadro CF, Greenberg BD: Human mitochondrial DNA variation and evolution: analysis of nucleotide sequences from seven individuals. *Genetics* 1983, **103**(2):287–312.
- Lee W-J, Conroy J, Howell WH, Kocher TD: Structure and evolution of teleost mitochondrial control regions. *J Mol Evol* 1995, **41**(1):54–66.
- Wenink PW, Baker AJ, Tilanus M: Mitochondrial control-region sequences in two shorebird species, the Turnstone and the Dunlin, and their utility in population genetic studies. *Mol Biol Evol* 1994, **11**(1):22–31.
- Martin AP, Naylor GJP, Palumbi SR: Rates of mitochondrial DNA evolution in sharks are slow compared with mammals. *Nature* 1992, **357**(6374):153–155.
- Wynen L, Larson H, Thorburn D, Peverell S, Morgan D, Field I, Gibb K: Mitochondrial DNA supports the identification of two endangered river sharks (*Glyphis glyphis* and *Glyphis garricki*) across northern Australia. *Mar Freshw Res* 2009, **60**(6):554–562.
- Stoner DS, Grady JM, Priede KA, Quattro JM: Amplification primers for the mitochondrial control region and sixth intron of the nuclear-encoded lactate dehydrogenase A gene in elasmobranch fishes. *Conserv Genet* 2003, **4**(6):805–808.
- McLachlan A: *Comparative phylogeography of the catshark, Haploblepharus pictus and its nematode parasite, Proleptus obtusus*. Stellenbosch: University of Stellenbosch; 2011.
- Frederico RG, Farias IP, Araújo MLG, Charvet-Almeida P, Alves-Gomes JA: Phylogeography and conservation genetics of the Amazonian freshwater stingray *Paratrygon aiereba* Müller & Henle, 1841 (Chondrichthyes: Potamotrygonidae). *Neotrop Ichthyol* 2012, **10**:71–80.
- Naylor GJP, Caira JN, Jensen K, Rosana KAM, White WT, Last PR: A DNA sequence-based approach to the identification of shark and ray species and its implications for global elasmobranch diversity and parasitology. In *Bulletin of the American Museum of Natural History*. New York: Department of Library Services, American Museum of Natural History; 2012:1–262.
- Castillo-Páez A, Sosa-Nishizaki O, Sandoval-Castillo J, Galván-Magaña F, Rocha-Olivares A: Strong population structure and shallow mitochondrial phylogeny in the banded guitarfish, *Zapteryx exasperata* (Jordan y Gilbert, 1880), from the northern Mexican Pacific. *J Hered* 2014, **105**(1):91–100.
- Tillett B, Meekan M, Field I, Thorburn D, Ovenden J: Evidence for reproductive philopatry in the bull shark *Carcharhinus leucas*. *J Fish Biol* 2012, **80**(6):2140–2158.
- Dudgeon CL, Broderick D, Ovenden JR: IUCN classification zones concord with, but underestimate, the population genetic structure of the zebra shark *Stegostoma fasciatum* in the Indo-West Pacific. *Mol Ecol* 2009, **18**(2):248–261.
- Arnason U, Adegoke JA, Gullberg A, Harley EH, Janke A, Kullberg M: Mitogenomic relationships of placental mammals and molecular estimates of their divergences. *Gene* 2008, **421**(1):37–51.
- Shamblin BM, Bjørndal KA, Bolten AB, Hillis-Starr ZM, Lundgren IAN, Naro-Maciel E, Nairn CJ: Mitogenomic sequences better resolve stock structure of southern Greater Caribbean green turtle rookeries. *Mol Ecol* 2012, **21**(10):2330–2340.
- Knaus BJ, Cronn R, Liston A, Pilgrim K, Schwartz MK: Mitochondrial genome sequences illuminate maternal lineages of conservation concern in a rare carnivore. *BMC Ecol* 2011, **11**:10.
- Morin PA, Archer FI, Foote AD, Vilstrup J, Allen EE, Wade P, Durban J, Parsons K, Pitman R, Li L, Bouffard P, Abel Nielsen SC, Rasmussen M, Willerslev E, Gilbert MT, Harkins T: Complete mitochondrial genome phylogeographic analysis of killer whales (*Orcinus orca*) indicates multiple species. *Genome Res* 2010, **20**:908–916.
- Foote AD, Morin PA, Durban JW, Pitman RL, Wade P, Willerslev E, Gilbert MTP, da Fonseca RR: Positive selection on the killer whale mitogenome. *Biol Lett* 2011, **7**(1):116–118.
- Pillans RD, Stevens JD, Kyne PM, Salini J: Observations on the distribution, biology, short-term movements and habitat requirements of river sharks *Glyphis* spp. in northern Australia. *Endanger Species Res* 2009, **10**:321–332.
- Chen X, Liu M, Grewe PM, Kyne PM, Feutry P: Complete mitochondrial genome of the critically endangered Speartooth Shark *Glyphis glyphis* (Carcharhiniformes: Carcharhinidae). *Mitochondrial DNA* 2014, **25**(6):431–432.
- Quail MA, Smith M, Jackson D, Leonard S, Skelly T, Swerdlow HP, Gu Y, Ellis P: SASI-Seq: sample assurance Spike-Ins, and highly differentiating 384 barcoding for Illumina sequencing. *BMC Genomics* 2014, **15**(1):110.
- Timmermans MJTN, Dodsworth S, Culverwell CL, Bocak L, Ahrens D, Littlewood DTJ, Pons J, Vogler AP: Why barcode? High-throughput multiplex sequencing of mitochondrial genomes for molecular systematics. *Nucleic Acids Res* 2010, **38**(21):e197–e197.
- Winkelmann I, Campos PF, Strugnell J, Cheral Y, Smith PJ, Kubodera T, Allcock L, Kampmann M-L, Schroeder H, Guerra A: 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* 2013, **280**(1759):20130273.
- Hoelzel AR, Shivji MS, Magnussen J, Francis MP: Low worldwide genetic diversity in the basking shark (*Cetorhinus maximus*). *Biol Lett* 2006, **2**(4):639–642.
- Karl SA, Castro AL, Garla RC: Population genetics of the nurse shark (*Ginglymostoma cirratum*) in the western Atlantic. *Mar Biol* 2012, **159**(3):489–498.
- Ahonen H, Harcourt R, Stow A: Nuclear and mitochondrial DNA reveals isolation of imperilled grey nurse shark populations (*Carcharias taurus*). *Mol Ecol* 2009, **18**(21):4409–4421.
- Charlesworth B: Effective population size and patterns of molecular evolution and variation. *Nat Rev Genet* 2009, **10**(3):195–205.
- Frankham R: Relationship of genetic variation to population size in wildlife. *Conserv Biol* 1996, **10**(6):1500–1508.

40. Compagno LJV, Pogonoski J, Pollard D: *Glyphis glyphis*. In: *IUCN 2013 The IUCN Red List of Threatened Species Version 2013.2*. <http://www.iucnredlist.org/details/39379/0>; 2009.
41. Barrett SCH, Kohn JR, Falk DA, Holsinger KE: **Genetic and evolutionary consequences of small population size in plants: implications for conservation**. In *Genetics and Conservation of Rare Plants*. Edited by Falk DA, Holsinger KE. New York, USA: Oxford University Press; 1991:3–30.
42. Sandoval-Castillo J, Rocha-Olivares A, Villavicencio-Garayzar C, Balart E: **Cryptic isolation of Gulf of California shovelnose guitarfish evidenced by mitochondrial DNA**. *Mar Biol* 2004, **145**(5):983–988.
43. Griffiths AM, Sims DW, Cotterell SP, El Nagar A, Ellis JR, Lynghammar A, McHugh M, Neat FC, Pade NG, Queiroz N: **Molecular markers reveal spatially segregated cryptic species in a critically endangered fish, the common skate (*Dipturus batis*)**. *Proc R Soc Lond B Biol Sci* 2010, **277**(1687):1497–1503.
44. Castro A, Stewart B, Wilson S, Hueter R, Meekan M, Motta P, Bowen B, Karl S: **Population genetic structure of Earth's largest fish, the whale shark (*Rhincodon typus*)**. *Mol Ecol* 2007, **16**(24):5183–5192.
45. Benavides MT, Feldheim KA, Duffy CA, Wintner S, Braccini JM, Boomer J, Huveneres C, Rogers P, Mangel JC, Alfaro-Shigueto J: **Phylogeography of the copper shark (*Carcharhinus brachyurus*) in the southern hemisphere: implications for the conservation of a coastal apex predator**. *Mar Freshw Res* 2011, **62**(7):861–869.
46. Duncan KM, Martin AP, Bowen BW, De Couet HG: **Global phylogeography of the scalloped hammerhead shark (*Sphyrna lewini*)**. *Mol Ecol* 2006, **15**(8):2239–2251.
47. Douady CJ, Dosay M, Shivji MS, Stanhope MJ: **Molecular phylogenetic evidence refuting the hypothesis of Batoidea (rays and skates) as derived sharks**. *Mol Phylogenet Evol* 2003, **26**(2):215–221.
48. Greig TW, Moore MK, Woodley CM, Quattro JM: **Mitochondrial gene sequences useful for species identification of western North Atlantic Ocean sharks**. *Fish Bull* 2005, **103**(3):516–523.
49. Kashiwagi T, Marshall AD, Bennett MB, Ovenden JR: **The genetic signature of recent speciation in manta rays (*Manta alfredi* and *M. birostris*)**. *Mol Phylogenet Evol* 2012, **64**(1):212–218.
50. Feutry P, Grewe PM, Kyne PM, Chen X: **Complete mitogenomic sequence of the Critically Endangered Northern River Shark *Glyphis garricki* (Carcharhiniformes: Carcharhinidae)**. *Mitochondrial DNA*. In Press.
51. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R: **The sequence alignment/map format and SAMtools**. *Bioinformatics* 2009, **25**(16):2078–2079.
52. Bandelt HJ, Forster P, Röhl A: **Median-joining networks for inferring intraspecific phylogenies**. *Mol Biol Evol* 1999, **16**(1):37–48.
53. Excoffier L, Lischer HEL: **Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows**. *Mol Ecol Resour* 2010, **10**(3):564–567.
54. Nei M: *Molecular Evolutionary Genetics*. New York: Columbia University Press; 1987.
55. Watterson GA: **On the number of segregating sites in genetical models without recombination**. *Theor Popul Biol* 1975, **7**(2):256–276.
56. Tajima F: **Evolutionary relationship of DNA sequences in finite populations**. *Genetics* 1983, **105**(2):437–460.
57. Fu Y-X: **Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection**. *Genetics* 1997, **147**(2):915–925.
58. Tamura K, Nei M: **Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees**. *Mol Biol Evol* 1993, **10**(3):512–526.

doi:10.1186/s12862-014-0232-x

Cite this article as: Feutry et al.: Mitogenomics of the Speartooth Shark challenges ten years of control region sequencing. *BMC Evolutionary Biology* 2014 **14**:232.

Submit your next manuscript to BioMed Central and take full advantage of:

- **Convenient online submission**
- **Thorough peer review**
- **No space constraints or color figure charges**
- **Immediate publication on acceptance**
- **Inclusion in PubMed, CAS, Scopus and Google Scholar**
- **Research which is freely available for redistribution**

Submit your manuscript at
www.biomedcentral.com/submit

