RESEARCH ARTICLE

Ecology and Evolution
Open Access
WILEY

Pieces in a global puzzle: Population genetics at two whale shark aggregations in the western Indian Ocean

Royale S. Hardenstine[1] | Song He[1] | Jesse E. M. Cochran[1] | Camrin D. Braun[2] |
Edgar Fernando Cagua[3,4] | Simon J. Pierce[5] | Clare E. M. Prebble[5,6] |
Christoph A. Rohner[5] | Pablo Saenz-Agudelo[7] | Tane H. Sinclair-Taylor[8] |
Gregory B. Skomal[9] | Simon R. Thorrold[2] | Alexandra M. Watts[5,10] |
Casey J. Zakroff[1] | Michael L. Berumen[1]

[1]Division of Biological and Environmental Science and Engineering, Red Sea Research Center, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia

[2]Biology Department, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts, USA

[3]School of Biological Sciences, Centre for Integrative Ecology, University of Canterbury, Christchurch, New Zealand

[4]WorldFish, Bayan Lepas, Malaysia

[5]Marine Megafauna Foundation, Truckee, California, USA

[6]National Oceanography Centre, University of South Hampton, South Hamton, UK

[7]Facultad de Ciencias, Instituo de Ciencias Ambientales y Evolutivas, Universidad Austral de Chile, Valdivia, Chile

[8]Australian Institute of Marine Science, Townsville, Qld, Australia

[9]Massachusetts Division of Marine Fisheries, New Bedford, Massachusetts, USA

[10]Ecological Genetics and Conservation Laboratory, Manchester Metropolitan University, Manchester, UK

Correspondence
Royale S. Hardenstine, Division of Biological and Environmental Science and Engineering, Red Sea Research Center, King Abdullah University of Science and Technology, 23955-6900 Thuwal, Saudi Arabia.
Email: royale.hardenstine@kaust.edu.sa

[Corrections added on 04 March 2022 after first online publication: The surname of Pablo Saenz-Agudelo has been updated in this version.]

Abstract

The whale shark *Rhincodon typus* is found throughout the world's tropical and warm-temperate ocean basins. Despite their broad physical distribution, research on the species has been concentrated at a few aggregation sites. Comparing DNA sequences from sharks at different sites can provide a demographically neutral understanding of the whale shark's global ecology. Here, we created genetic profiles for 84 whale sharks from the Saudi Arabian Red Sea and 72 individuals from the coast of Tanzania using a combination of microsatellite and mitochondrial sequences. These two sites, separated by approximately 4500 km (shortest over-water distance), exhibit markedly different population demographics and behavioral ecologies. Eleven microsatellite DNA markers revealed that the two aggregation sites have similar levels of allelic richness and appear to be derived from the same source population. We sequenced the mitochondrial control region to produce multiple global haplotype networks (based on different alignment methodologies) that were broadly similar to each other in terms of population structure but suggested different demographic histories. Data from both microsatellite and mitochondrial markers demonstrated the stability of

genetic diversity within the Saudi Arabian aggregation site throughout the sampling period. These results contrast previously measured declines in diversity at Ningaloo Reef, Western Australia. Mapping the geographic distribution of whale shark lineages provides insight into the species' connectivity and can be used to direct management efforts at both local and global scales. Similarly, understanding historical fluctuations in whale shark abundance provides a baseline by which to assess current trends. Continued development of new sequencing methods and the incorporation of genomic data could lead to considerable advances in the scientific understanding of whale shark population ecology and corresponding improvements to conservation policy.

KEYWORDS
genetic diversity, global population structure, microsatellites, mtDNA, *Rhincodon typus*

JEL CLASSIFICATION
Population genetics

1 | INTRODUCTION

The whale shark, *Rhincodon typus* Smith 1828, has often been described as an enigmatic species (Rowat & Brooks, 2012). Despite its size (up to 20 m) (Chen et al., 1997) and circumglobal distribution, encounters with these ocean giants were once rare, with only 320 records documented between 1801 to 1985 (Wolfson & Sciara, 1981). Smith's initial description of the species and virtually all early whale shark research was based on chance encounters. The discovery of sites with predictable whale shark aggregations resulted in a rapid increase in data available for the species (Norman et al., 2017). Unfortunately, these data have shown sustained population declines, leading the International Union for Conservation of Natural Resources (IUCN) Red List to change the species' global status from Vulnerable (Norman, 2005) to Endangered (Pierce & Norman, 2016). An improved understanding of the whale shark's global and regional population structure and the species' patterns of dispersal can help determine the scale of management units. This is critical to long-term conservation efforts and planning because management strategies for a collection of independent aggregation sites may not be effective when applied to a wide-ranging species that exhibits functional connectivity across basins.

Most whale shark population and movement ecology studies are based on photo-identification (Araujo et al., 2014; Cochran et al., 2016; McKinney et al., 2017; Norman, Holmberg, et al., 2017; Robinson et al., 2016; Rohner et al., 2015) and/or telemetry (Berumen et al., 2014; Cagua et al., 2015; Cochran et al., 2019; Hueter et al., 2013; Norman, Whitty, et al., 2017; Robinson et al., 2016). While these methods have been successfully used to describe local population structure and to track movements within, between, and away from aggregation sites, they are limited by their focus on known aggregations (Sequeira et al., 2012). Most of these sites are dominated by either subadult males (Araujo et al., 2014;

Diamant et al., 2018; Graham & Roberts, 2007; Ketchum et al., 2013; McKinney et al., 2017; Meekan et al., 2006; Ramírez-Macías et al., 2012; Riley et al., 2010; Robinson et al., 2016; Rohner et al., 2015; Rowat et al., 2011), or more rarely, mature females (Acuña-Marrero et al., 2014; Ketchum et al., 2013). Both cases imply that much of the current data are demographically skewed. Mature sharks (particularly males), subadult females, and neonates of both sexes are underrepresented (Norman, Holmberg, et al., 2017). Genetic comparison among aggregations can overcome this issue because patterns in molecular data reflect the distribution and movement of previous generations, as well as those of sampled individuals, effectively providing a more holistic view of the species' demography across an evolutionary timescale.

Three independent global investigations of whale shark population genetics have been conducted: one using the mitochondrial control region (Castro et al., 2007), one using microsatellite loci (Schmidt et al., 2009), and one using both marker types (Vignaud et al., 2014). These studies show that genetic differentiation exists between whale sharks from the Indo-Pacific and Atlantic Oceans, and additional samples from more recent research (Meekan et al., 2017; Sigsgaard et al., 2017; Yagishita et al., 2020) have largely corroborated this result. The Indo-Pacific and Atlantic regions were subsequently treated as large-scale genetic "subpopulations" for the IUCN global assessment (Pierce & Norman, 2016). In addition to the delineation of basin-scale populations, one of the global comparison studies demonstrated a 6-year decline in genetic diversity of whale sharks that were sampled from Western Australia (Vignaud et al., 2014). However, the same study also found evidence of a recent population expansion within the broader Indo-Pacific (Vignaud et al., 2014), which has been confirmed by more recent work (Yagishita et al., 2020). Regardless, local assessments of temporal trends in genetic diversity have not been conducted at any aggregation sites outside of Western Australia.

Limited connectivity has been observed among the whale shark aggregation sites identified within the broader western Indian Ocean region. Photo-identification (Andrzejaczek et al., 2016; Brooks et al., 2010; Norman, Holmberg, et al., 2017; Robinson et al., 2016), telemetry (Berumen et al., 2014; Brunnschweiler et al., 2009; Cagua et al., 2015; Norman, Whitty, et al., 2017; Robinson et al., 2016; Rohner et al., 2020), and biochemical (Prebble et al., 2018) studies have not shown any direct evidence of large-scale connectivity between distant sites within this basin. Here, whale shark tissue samples collected from aggregation sites in the Saudi Arabian Red Sea (Berumen et al., 2014) and Mafia Island, Tanzania (Cagua et al., 2015) were analyzed. Genetic similarity between the two sites was assessed using both microsatellite markers (including novel primers developed for this study) and the whale shark mitochondrial control region. The same markers were also used to track genetic diversity within the Saudi Arabian Red Sea, providing the first direct comparison to similar studies from Western Australia (Vignaud et al., 2014). The novel sequences from both Saudi Arabia and Tanzania were combined with previously published data and used to generate updated global haplotype networks. These networks map the genetic similarity of sampled sharks based on mitochondrial sequences and can be used to identify geographic patterns in the molecular data.

2 | MATERIAL AND METHODS

2.1 | Ethics statement

This research was undertaken in accordance with the policies and procedures of the King Abdullah University of Science and Technology (KAUST) and approved by KAUST's Institutional Biosafety and Bioethics Committee and Institutional Animal Care and Use Committee (protocol #20IACUC07). Permissions relevant for KAUST to undertake research have been obtained from the applicable governmental agencies in the Kingdom of Saudi Arabia. All works in Tanzania were conducted with the approval from the Tanzanian Commission for Science and Technology (COSTECH #2018-545-NA-2015-161).

2.2 | Field sites

Shib Habil is a submerged reef platform in the Red Sea approximately 5 km from the Saudi Arabian port city of Al Lith (Figure 1). This site attracts juvenile whale sharks of both sexes, and shark presence is highly seasonal. Most whale shark encounters at Shib Habil occur in March, April, or May (Cochran et al., 2016), after which the animals disperse into the wider Red Sea and Indian Ocean (Berumen et al., 2014). The Tanzanian aggregation site is located in Kilindoni Bay to the southwest of Mafia Island (Figure 1). Here, most of the sharks are juvenile males, and are commonly seen feeding on sergestid shrimp from November through February with sightings declining through March (Rohner,

Richardson, et al., 2015; Rohner, Armstrong, et al., 2015). Acoustic telemetry has revealed that many of the sharks remain cryptically present within Kilindoni Bay or close by year-round, with two to four consecutive years of high residence recorded for some individuals (Cagua et al., 2015; Rohner et al., 2020).

2.3 | Sample collection

Whale shark tissue was collected over six seasons in Saudi Arabia (2010–2015) and over two seasons in Tanzania (late 2012 to early 2014). Sampling at both sites occurred during their respective whale shark tourism seasons and followed the same procedure. Free-swimming whale sharks were approached by snorkelers and tissue samples were taken using a Hawaiian sling pole-spear fitted with a biopsy tip. When possible, researchers estimated size, determined sex, and collected photos for individual identification. Identification photos were analyzed using both the Groth (Arzoumanian et al., 2005) and the Interactive Individual Identification System (I³S) (Van Tienhoven et al., 2007) algorithms to identify individuals and flag potential duplicate samples (Cochran et al., 2016; Rohner, Richardson, et al., 2015). Suspected duplicates were retained and eventually sequenced to confirm photo identification and to ensure that the highest quality sample was used for further analysis. Collected samples were preserved immediately in 70%–90% ethanol, or, in cases where ethanol was not available, samples were put on ice and transferred to a freezer. All samples were eventually placed in 90% ethanol and kept at −20°C for long-term storage.

2.4 | DNA extraction and PCR

Each sample consisted of white, subcutaneous tissue and a black dermal-cap. The dermal-cap was separated with a scalpel and used for all further analyses. DNA was extracted using one of two kits, the DNeasy Blood and Tissue Kit (Qiagen Inc.) or the NucleoSpin Tissue Kit (Macherey-Nagel), following the respective kit instructions. The quality and quantity of extracted DNA were measured using a NanoDrop 8000 spectrophotometer (Thermo Scientific; www.thermoscientific.com). Next-generation sequencing was performed for one specimen using a Roche 454 GS FLX (titanium) sequencer and genomic library was constructed following the manufacturer's protocol. Raw unassembled reads from this library were mined for putative microsatellite loci using the msat-commander v 1.0.8 software (Faircloth, 2008). Default settings were used to screen perfect dinucleotide and tetranucleotide repeats that were at least 20 bp long, resulting in 1588 putative microsatellite loci. Primer 3 (Rosen & Skaletsky, 2000) software was then used to design 353 primer pairs for all reads that contained suitable microsatellite repeat motifs. From these 14 novel primer pairs were selected and synthesized along with 8 primers from earlier publications (three from Ramirez-Macias et al., 2009, five from Schmidt et al., 2009), producing 22 candidates for PCR trials.
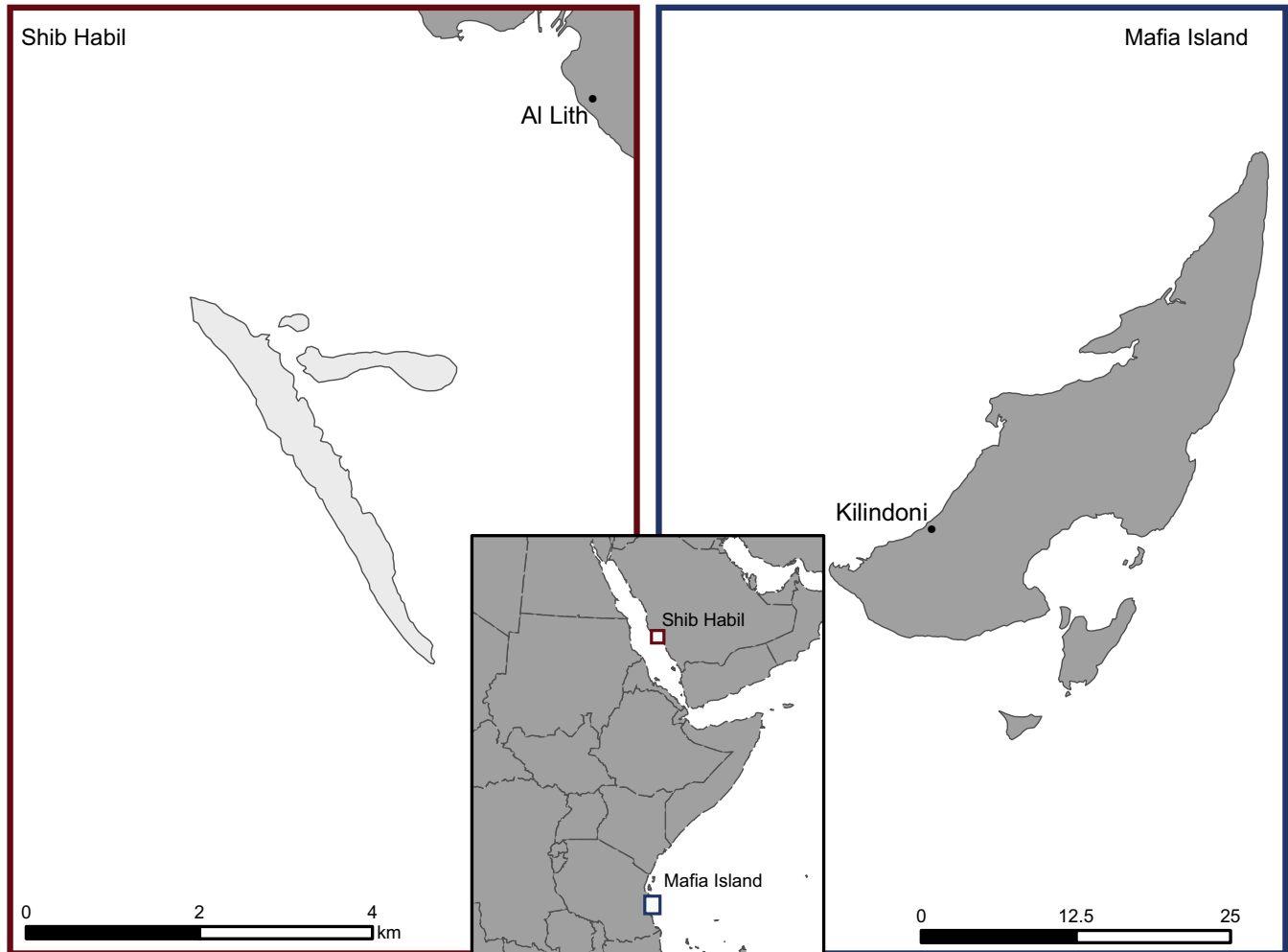
FIGURE 1 Two main sampling locations for this study. Center: Both locations shown in regional context. Shib Habil, Saudi Arabia, highlighted in red box and expanded to the left. Mafia Island, Tanzania, highlighted in blue box expanded on the right. Maps were composed in ArcPro (Esri Inc.) using layers sourced from MF Campbell Jr

(Appendix S1: SI 1) PCR reactions were run with 5 µl of Master Mix (Qiagen Inc.), 3 µl of RNA free water, 1 µl of primer mix (each primer concentration 1 µM), and 1 µl of template DNA. The PCR reactions were initiated at 95°C for 15 min, followed by 35 cycles of 94°C for 30 s, 55°C for 90 s, and 72°C for 60 s, and concluded at 60°C for 30 min. PCR products were then diluted to 1:50 with water before being sent for fragment analysis.

For mitochondrial analysis, the whale shark control region (primers WSCR1-F and WSCR2-R) from Castro et al. (2007) was chosen for comparison with previous studies and publicly available sequences. A 12.5 µl solution of Master Mix (Qiagen Inc.) (6.25 µl), water (4.25 µl), forward and reverse primers (0.5 µl each at 10 µM concentration), and template DNA (1 µl) was added to each well. The thermocycler program was as follows: 95°C for 15 min; 35 cycles of 94°C for 60 s, 58°C for 60 s, and 72°C for 45 s; and a final elongation step at 72°C for 10 min. A subsample of PCR product was checked using Qiaxcel (Qiagen Inc.). Following this, PCR products were cleaned using ExoSap-IT (Affymetrix). Purified products were sequenced using an ABI 3730xl platform.

2.5 | Microsatellite marker analyses

Microsatellite markers were used to assess local patterns in genetic diversity within and between the Tanzanian and Saudi Arabian aggregations. These markers are derived from nuclear DNA, so they reflect both the maternal and paternal histories of the sampled animals. Microsatellite allele size was read using Geneious 8.1.6 software (Biomatters Ltd.). After scoring, any duplicate genotypes were identified using the Microsatellite Toolkit (Park, 2001). A subset of duplicate sequences was then used to calculate genotyping error before being removed from further analysis. Primers were screened using Genepop 4.2 (Rousset, 2008) to exclude markers containing either null alleles or linked loci and to ensure Hardy–Weinburg equilibrium. The remaining loci were used for all further microsatellite analyses. In order to assess genetic diversity within the two aggregations, each site's expected and observed heterozygosity were compared using GenAlEx version 6.4 (Peakall & Smouse, 2006) and local genetic diversity was determined using the Allelic Diversity Analyzer (ADZE) version 1.0 (Szpiech et al., 2008) via the rarefaction method.

This technique analyzes private alleles and allelic richness within populations and is designed to be robust when sampling sizes vary. ADZE was used to determine overall genetic diversity for both the Tanzanian and Saudi Arabian populations individually. Genetic diversity was also calculated independently for each of the Saudi Arabian aggregation's five sampling seasons and compared using a Mann–Kendall trend test (Kendall, 1955; Mann, 1945) to determine if any substantial changes in local diversity occurred from 2010 to 2015.

Differentiation between the two aggregations was quantitated using a pairwise fixation index (F_{ST}) and Shannon's mutual information, both of which were calculated using on GenAlEx version 6.4 (Peakall & Smouse, 2006). The same program was also used to identify any alleles found at only one of the two sampling sites (private alleles) and to quantify the prevalence of such alleles within each area. Genetic population structure was modeled using STRUCTURE v2.3.4 (Pritchard et al., 2000). The model was run using locprior with one, two, and three possible populations ($K = 1$, $K = 2$, and $K = 3$). Selecting a model with a K value of one would indicate our sampling locations were from one panmictic population, two would mean that each site hosted a single distinct population, and three would imply that one or both aggregations were composed of multiple distinct populations. For each value of K, the model was run 10 times with a burn-in of 10,000 iterations, followed by 1,000,000 iterations, and 10 iterations of each K scenario. STRUCTURE results were uploaded to Structure Harvester (Earl, 2012) where both Evanno's delta K and averaged maximum likelihood scores were calculated and used to determine the most likely value of K (Appendix S1: SI 2). After selecting the optimal K value, results were uploaded to CLUMPAK (Kopelman et al., 2015) to generate the consolidated plot.

In order to assess the likelihood of recent population expansions in Saudi Arabia or Tanzania, microsatellite loci were analyzed using the Microsoft Excel macro KGTESTS (Bilgin, 2007). A within-locus k-test was used to compare observed microsatellite allelic distributions with those expected under mutation–drift equilibrium. A negative value in the k-test is indicative of population expansion while positive values suggest population stagnation (Reich et al., 1999). The significance of k was determined using a one-tailed binomial test. In addition, an interlocus g-test was used to compare observed versus expected allele size variances across all loci (Bilgin, 2007). This ratio is expected to be small in a recently expanded population in which allele genealogies show recent coalescence, but large in a population of constant size because of longer histories of variable mutation rates among loci. To determine the significance of the test, g values were compared to the expected 5% cut off of g under constant population size (Reich et al., 1999).

2.6 | Mitochondrial marker analysis

Microsatellite makers are useful, but differences in the chosen loci and scoring biases make results difficult to compare among studies. Using the mitochondrial control region allows for the incorporation of sequences from published sources to build a global comparison.

Mitochondrial sequences were aligned, edited, and trimmed using Geneious 8.1.6 (Biomatters Ltd.) and the original sequences were uploaded to GenBank (accession numbers: OL782199–OL782316). Whale shark control region sequences from previous studies (Djibouti, Qatar, Mozambique, Seychelles, Maldives, Western Australia, Philippines, Taiwan, Japan, Mexican Pacific, Mexican Atlantic; Castro et al., 2007; Meekan et al., 2017; Ramírez-Macías et al., 2007; Schmidt et al., 2010; Sigsgaard et al., 2017; Vignaud et al., 2014; Yagishita et al., 2020) were incorporated into the alignment for subsequent analyses (Appendix S1: SI 3).

Most prior studies (Castro et al., 2007; Vignaud et al., 2014; Yagishita et al., 2020) have retained gaps in the alignment on the assumption that these regions represented bioinformative deletion/insertion mutations. These papers followed a dual approach using a fully gap inclusive "raw alignment" to build their haplotype networks and a gap-reduced "modified alignment" for demographic analyses. The modified alignment reduced contiguous, nonvariable gap regions to a single transition, effectively assuming they were caused by a single indel. In contrast, at least one paper has elected to simplify the alignment methodology by excluding gaps entirely (Meekan et al., 2017). To facilitate comparison to all previous works and investigate the bioinformatic value of these regions, the Roehl data files for the haplotype network constructions were generated with the full gaps (raw alignment), reduced gaps (modified alignment), and without gaps. Haplotype networks were constructed for both the fully gap-inclusive and -exclusive data, while demographic analyses were calculated for all three alignments. Haplotype statistics and Roehl data were generated using DnaSP version 5.0 (Librado & Rozas, 2009) for subsequent haplotype network analysis. Within the Roehl data file, all the nonvariable sites in the alignment were discarded. Evolutionary relationships among whale shark mitochondrial haplotypes were assessed with a median-joining network, constructed with the program NETWORK version 4.5.1.0 (www.fluxus-engineering.com/network_terms.htm) using default settings.

In Arlequin v3.5 (Excoffier & Lischer, 2010), each sampling site was treated as a population and genetic structure was tested with analysis of molecular variance (Nei & Jin, 1989) with 10,000 permutations, using the T92+G substitution model (Tamura, 1992) with a gamma shape parameter of 0.48 (Appendix S1: SI 4). This model aims to estimate the genetic divergences between pairwise samples using F_{ST} based on haplotype frequencies and molecular divergence. The resulting p-values were compared to both a standard α (.05) and a multiple-comparisons-corrected α (.01) (Narum, 2006). Several metrics of genetic diversity, including the haplotype diversity (h) and nucleotide diversity (π), were calculated for each location, and all samples combined, using Arlequin v 3.5 (Excoffier & Lischer, 2010). Haplotype diversity within the Saudi Arabian Red Sea was also calculated and compared among years using a Mann–Kendall trend test (Kendall, 1955; Mann, 1945) to assess recent temporal fluctuation in the genetic makeup of the aggregation. Longer term population trends were visualized using Bayesian skyline plots (BSPs) generated for the overall dataset, for the Indo-Pacific, and for the Atlantic. Several neutrality statistics including Tajima's D, Fu's

F, and Harpending's raggedness index (HRI) were calculated (again using Arlequin v 3.5; Excoffier & Lischer, 2010) to detect evidence of historical population expansions within each location, and for all samples combined. Fu's *F* is known to be extremely sensitive and possibly confounded by species-specific recombination rates (Rozas & Calafell, 2008). These rates are unknown for the whale shark, making Fu's *F* somewhat unreliable for this species. It has been included to facilitate comparison with previous work.

3 | RESULTS

3.1 | Data filtering

Of the 22 candidate microsatellite markers initially considered for this study, six loci (Rhin_t_04, Rtyp1, Rtyp3, Rtyp4, Rtyp8, and Rty_15) could not be successfully amplified. Another three loci (Rty_16, Rhin_t_28, and Rhin_t_46) contained null alleles, one locus (Rtyp7) showed strong evidence of linkage with multiple other loci, and one locus (Rhin_t_13) violated the assumptions of the Hardy–Weinberg equilibrium. The remaining 11 markers consisted of 2 from Ramirez-Macias et al. (2009) and 9 that were designed for the present study (Table 1). The probability of identity for these markers ranged from 0.03 to 0.59 with a combined probability of 7.3×10^{-9} when all 11 loci were used together.

A total of 239 tissue samples (102 from Saudi Arabia, 137 from Tanzania) were genotyped, revealing 87 duplicates (18 from Saudi Arabia, 65 from Tanzania). Comparing the two highest quality samples from a subset of duplicated sharks (*n* = 35) produced a genotyping error rate of 1.1% per locus. The 156 highest quality, unique genotypes were retained for further analysis. This filtered dataset included 84 unique individuals from Saudi Arabia (40 females, 33 males, 11 unknown) and 72 from Tanzania (12 females, 58 males,

2 unknown). Sharks sampled in Saudi Arabia had an average estimated total length of 4.04 m (range: 3–7 m) while those in Tanzania averaged 5.99 m (range 3.5–8 m). As of 2015, there were 136 photo-identified sharks from Saudi Arabia (Cochran et al., 2016), including 81 of the sharks sampled in this study. Using the same 2015 end-date for Tanzania (Prebble et al., 2018) yields 139 photo-identified sharks, including 65 individuals from the present study.

3.2 | Comparing Saudi Arabia and Tanzania

The microsatellite datasets from Saudi Arabia and Tanzania had an allelic richness of 5.30 and 5.07, respectively. No distinct private alleles were found at either aggregation. Given the number of sampled individuals, private alleles at either site likely have low prevalence (mean expected frequency of private alleles is 0.0087), if they exist at all. Genetic differentiation between the two sampling sites was small (microsatellite $F_{ST} = 0.0028$) indicating connectivity sufficient to maintain population homogeneity. This was also supported by Shannon's Information Index, which showed that more than 98% of the total information derived from individual variation within the two aggregation sites, as opposed to less than 2% from difference between them. Observed heterozygosity estimates at each site (Saudi Arabia: 0.55, Tanzania: 0.56) were similar (χ^2 test, $p > .05$) to each other and to expected values (0.57 in both cases). Results from the program STRUCTURE suggested $K = 1$ (one source population for both sites) as the most likely configuration (Appendix S1: SI 2).

3.3 | Trends in genetic diversity

The KGTESTS analyses of the microsatellite data also indicated population stability for both the Saudi Arabian and Tanzanian

Locus	Size range (bp)	N_a	H_o	H_e	k Test
Rty_18*	175–185	4	0.475	0.451	−349.13
Rty_38*	172–185	7	0.740	0.728	63.82
Rhin_t_03	247–255	5	0.639	0.645	1183.79
Rhin_t_05	220–248	12	0.835	0.863	−562.74
Rhin_t_07	264–286	10	0.821	0.831	195828.79
Rhin_t_10	144–154	5	0.661	0.667	−30.96
Rhin_t_11	143–149	3	0.242	0.239	−12.12
Rhin_t_30	321–336	6	0.075	0.073	214.22
Rhin_t_31	152–164	4	0.412	0.427	−28.70
Rhin_t_32	115–123	5	0.631	0.702	3230.23
Rhin_t_47	117–139	7	0.590	0.620	−1768.07
11 Total loci					p-Value = .46

TABLE 1 Subset of 11 microsatellite markers used for the analysis of 156 individual whale sharks (Saudi Arabia: 84, Tanzania: 72), including the number of alleles per locus (N_a), the observed heterozygosity (H_o), the expected heterozygosity (H_e)

Note: Intralocus variance (*k*) values are provided for each marker, an interlocus variance (*g*) is provided for the study as a whole. Locus names with * are sourced from Ramirez-Macias et al. (2009).

aggregations. Within-locus k-tests produced a mixture of positive and negative values for different markers (Table 1) and did not reveal any significant evidence of historical population changes at either aggregation ($p = .46$). Similarly, the interlocus g-test produced a value of 3.46, far above the modeled 5% cutoff of 0.19. In the Red Sea, population stability was further supported by quantitating genetic diversity within the Saudi Arabian aggregation site over the 6-year sampling period. Allelic richness showed little year-to-year variation (range: 2.93–3.16) and no consistent directional trend (Mann–Kendall trend test, tau = 0.316, $p > .05$) (Table 2). This pattern of stability was also maintained when only the first (2010, allelic richness: 4.30) and final (2015, allelic richness: 4.32) sampling years were compared, confirming that genetic diversity remained constant over the 6-year period and that low sample sizes in specific years did not bias the overall pattern. Mitochondrial analysis yielded similar results with haplotype diversity remaining stable in Saudi Arabia from 2010 (haplotype diversity: 0.859) to 2015 (0.940) (Mann–Kendall trend test, tau = 0.600, $p > .05$) (Table 2).

3.4 | Global structure

An initial alignment of 816 mitochondrial sequences was created from 699 publicly available sequences and 117 sequences from this study (Tanzania $n = 57$, Saudi Arabia $n = 60$). Previous work included 26 sequences from the Saudi Arabian Red Sea (Vignaud et al., 2014; Yagishita et al., 2020); we have updated the sequencing for 16 of these individuals and sourced another 8 directly from online databases. In the main text we have largely focused on the gap-inclusive results while the gap-exclusive data are fully available in the appendix. A final whale shark control region fragment of 673 bp was used for subsequent analyses, encompassing 816 individuals from 13 locations (Appendix S1: SI 3). The alignment contained 328 variable sites, including sequence gaps, resulting in 192 identified

TABLE 2 Temporal genetic diversity of Saudi Arabian Red Sea whale sharks

	2010	2011	2012	2013	2015
Microsatellite N	30	26	9	7	19
Allelic Richness	3.04	3.01	2.93	3.04	3.16
*Allelic Richness	4.3				4.32
Mitochondrial N	19	17	8	8	16
Haplotype Diversity	0.859	0.844	0.900	0.867	0.940

Note: Microsatellite *N* is the number of individuals for which microsatellite data were available from each season. Sample size fluctuated among years, which can affect the calculations for allelic richness. To account for this, change in allelic richness was also assessed using only data from the first and final seasons (noted with an *). Mitochondrial *N* is number of individuals for which mitochondrial sequence data were available. All analyses indicated population stability at this site over the study period.

haplotypes, an average haplotype diversity of 0.944 (range: 0.85–1.00), and an average nucleotide diversity of 0.12 (range: 0.05–0.17).

The haplotype network was divided into four major lineages with no clear pattern in how haplotypes were distributed among aggregation sites (Figure 2). The one exception to this general lack of geographic structure was the disproportionate grouping of Atlantic samples within lineage 4 (Figure 3). This was also reflected in the statistical analysis of the mitochondrial sequences, which revealed relatively strong (mitochondrial $F_{ST} = 0.0532$–0.1456) and significant ($p < .05$) differences between the Atlantic samples and those from the Indo-Pacific sites (Table 3). The analysis also showed some smaller (mitochondrial $F_{ST} = -0.0042$–0.023), but still statistically significant ($p < .05$), differences among Indo-Pacific aggregations. The Maldives exhibited some stronger differences to other Indo-Pacific sites (mitochondrial $F_{ST} = 0.0325$–0.0482), but this could be attributed to the limited number of samples available from that region (Table 3). Removing sequence gaps did not substantially affect the results (Appendix S1: SI 5). Overall, the gap-inclusive and -exclusive networks exhibited similar geographic distributions of haplotypes and patterns of population structure (Table 3, Appendix S1: SI 6).

3.5 | Demographic history

The gap-inclusive network was characterized by a highly reticulated structure, which is an indication of population stability (Figure 2). This was also reflected in the neutrality statistics. Tajima's *D* was nonsignificant for most study sites (Table 4) and for the Indo-Pacific overall (Tajima's $D = -1.23$, $p > .05$). Fu's *Fs* was nonsignificant for every individual study site but was significant for the Indo-Pacific as a whole (Fu's $Fs = -23.31$, $p < .05$). Both tests were nonsignificant for the Mexican Atlantic, as were the HRI values for all sites and both ocean basins. Unlike population structure, the demographic history results changed when the gap data were reduced or removed. The different lineages within the haplotype network tended to form star-shaped clusters more indicative of recent population expansions (Ferreri et al., 2011). Tajima's *D* was negative for all Indo-Pacific study sites, and significantly so for most (Appendix S1: SI 7 and 8). More importantly, Tajima's *D* was both negative (−1.45) and significant ($p < .05$) when these sites were pooled together and tested as a whole. Fu's *Fs* results were more regionally varied, but only Tanzania showed significant evidence of a regional expansion ($Fs = -10.32$, $p < .01$) and only when gaps were removed entirely. Despite these varied regional results, Fu's *Fs* indicated significant historical expansion of the Indo-Pacific population overall for both the gap-reduced ($Fs = -23.10$, $p < .05$) and gap-exclusive ($Fs = -23.41$, $p < .05$) alignments. The Mexican Atlantic never showed significant deviation from population neutrality. The HRI was not consistently significant for any sampling site, and was never significant for the Indo-Pacific collectively.

All of the BSPs, regardless of gap treatment or basin grouping, showed long-term stability of the targeted population followed by a
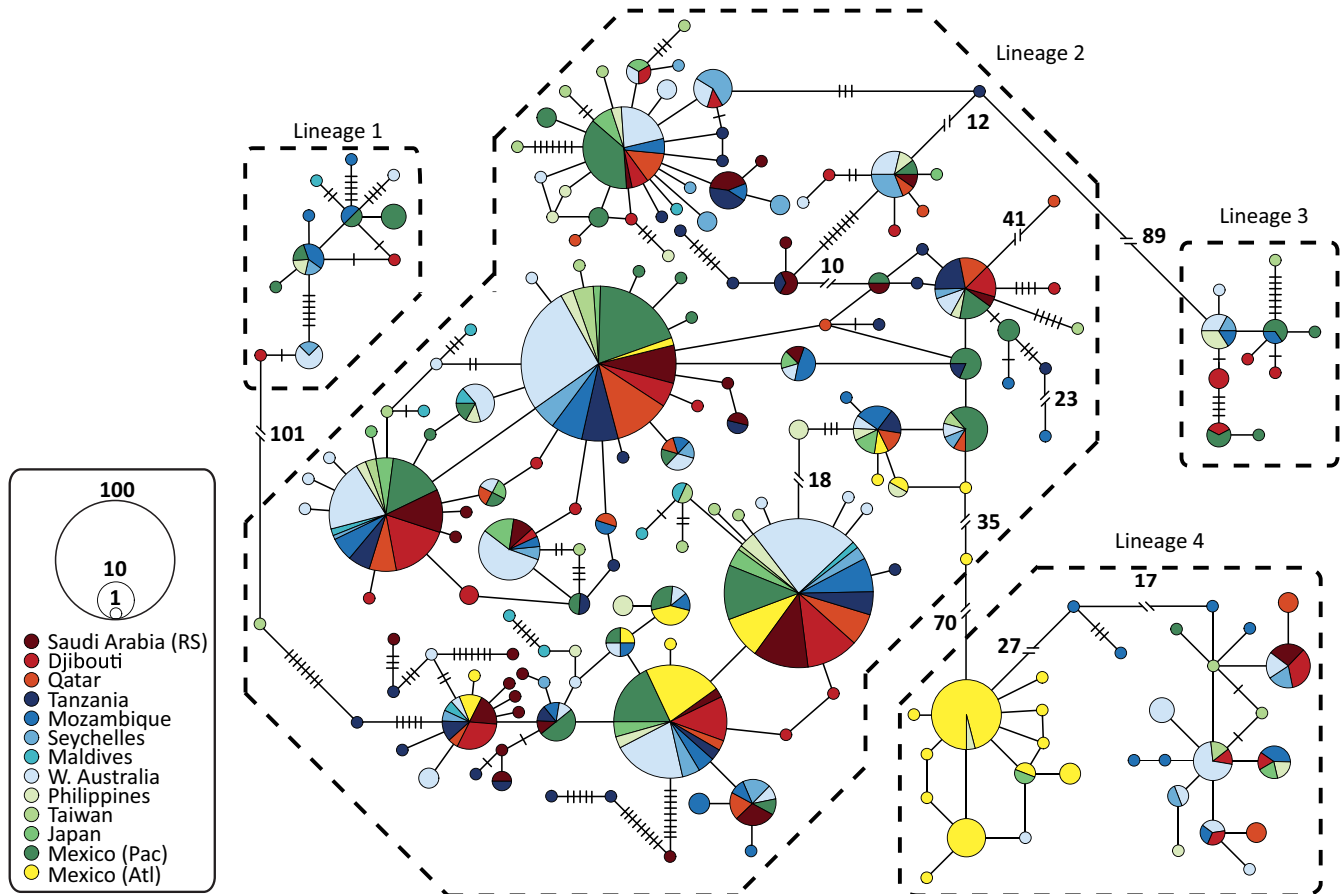
FIGURE 2 Relationships of *Rhincodon typus* haplotypes, from 13 different geographic locations (color in legend), in median-joining network created using mitochondrial whale shark control region sequences where gaps were considered informative. Each circle represents a unique haplotype and is proportional to total haplotype frequency. Branches connecting circles represents a single nucleotide substitution; black cross-bar represents an additional nucleotide substitution; black double slash bars represent more than 10 nucleotide substitutions (exact numbers noted). Areas encompassed by dashed lines represent four putative lineages. Atl, Atlantic; Pac, Pacific; RS, Red Sea

recent expansion (Appendix S1: SI 9). However, the broad trends proceeding the population expansion varied. For instance, the fully gap-inclusive plots for both the overall data and the Indo-Pacific showed a gradual, long-term population decline while the gap-reduced and gap-exclusive alignments showed a much more abrupt bottlenecking event (Appendix S1: SI 9). The Atlantic showed a bottleneck and subsequent recovery in all of the alignments, but in all cases the difference between the minimum and maximum population size was less pronounced than in the Indo-Pacific, indicating relative stability in the Atlantic population overall.

4 | DISCUSSION

This study provides the first mitochondrial and microsatellite sequences from the Mafia Island aggregation site in Tanzania and nearly quadruples the sample size from the Shib Habil aggregation site in the Saudi Arabian Red Sea. Both Tanzania and Saudi Arabia were shown to attract whale sharks from a single, basin-scale population that extends throughout the Indo-Pacific. Nucleotide and haplotype diversity were both stable at the Saudi Arabian aggregation

site over the 6-year sampling period, a direct contrast to the declines reported in Western Australia (Vignaud et al., 2014). Globally, our results corroborate the division of whale sharks into distinct Atlantic and Indo-Pacific populations, but contrast recent population expansions inferred by previous work (Vignaud et al., 2014; Yagishita et al., 2020).

4.1 | Comparing Saudi Arabia and Tanzania

Allelic richness was similar between Saudi Arabia and Tanzania, although the richness metrics for Saudi Arabia were slightly higher than those from previous work (Vignaud et al., 2014). This is likely due to differences in sample sizes, chosen microsatellite loci, or a combination of both among different studies. All analyses (F_{ST}, private allele assessment, both haplotype networks, and STRUCTURE analysis) indicate genetic exchange between the sharks from the Tanzanian site and from the Saudi Arabian site. This would imply that the sharks from both aggregation sites belong to the same population. Previous work has shown that sharks sampled at the Saudi Arabian site and aggregation sites along the east coast of Africa
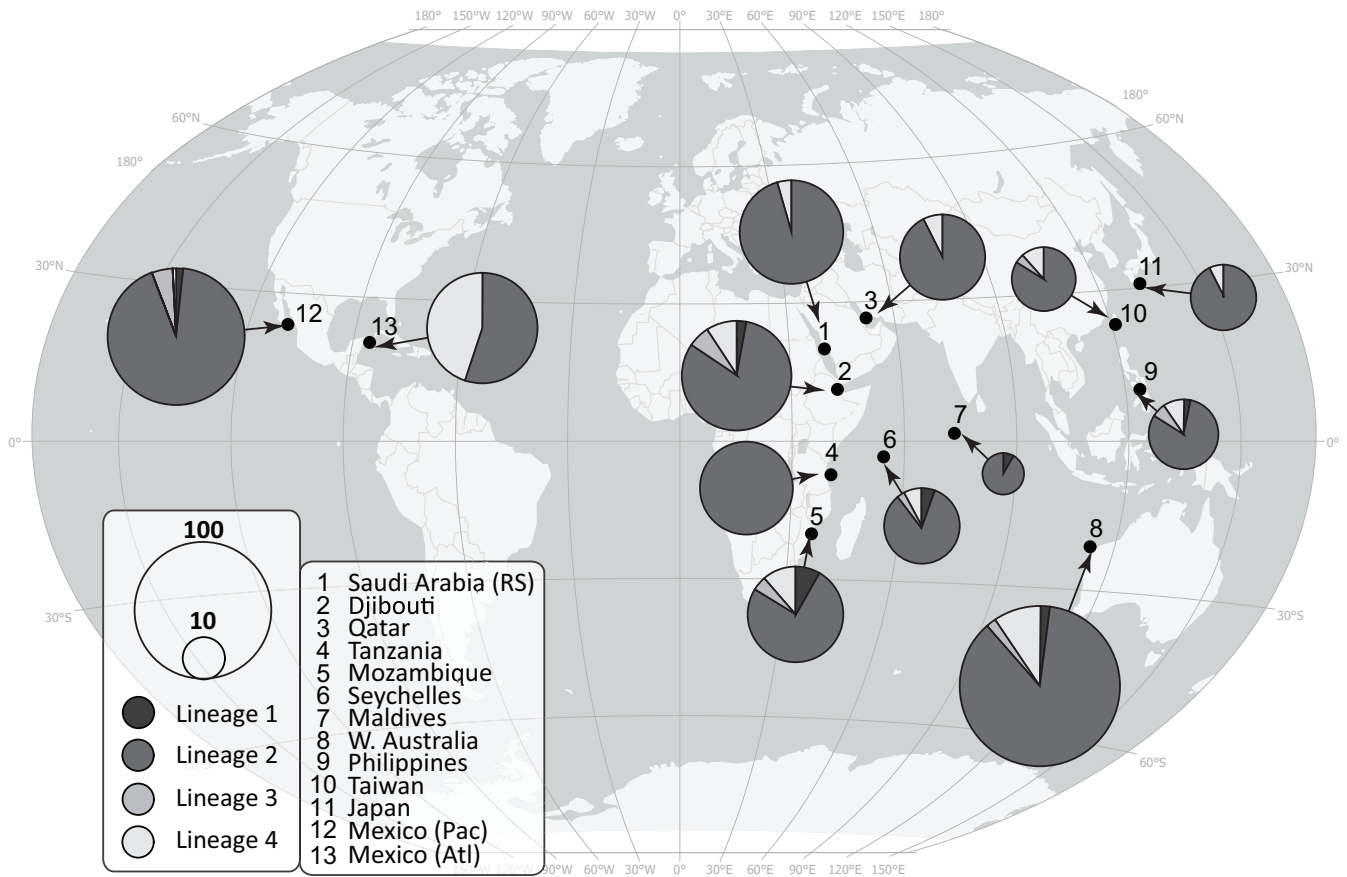
FIGURE 3 Locations from this study and mitochondrial sequences sourced from previous publications, and haplotype lineage composition of *Rhincodon typus* individuals at each aggregation site (numbers in legend). Circles display the lineage composition of individuals sequenced from that aggregation (colors representing four putative haplotype lineages in legend) and circle size is proportional to the number samples. Atl, Atlantic; Pac, Pacific; RS, Red Sea

(Djibouti and Mozambique) are all grouped within the Indo-Pacific population (Yagishita et al., 2020; Meekan et al., 2017; Vignaud et al., 2014). The results here demonstrate that the Tanzanian aggregation site, despite local residency of the sharks (Cagua et al., 2015; Rohner et al., 2020), is part of the same genetic population as neighboring aggregation sites along the east African coast.

The observation that whale sharks sampled in Saudi Arabia and Tanzania are members of the same population is particularly interesting as the movement ecologies and demography at the two aggregation sites are quite distinct. Sharks at the two sites exhibit markedly different residency behaviors (Cagua et al., 2015; Cochran et al.,2019), movement patterns (Berumen et al., 2014; Rohner et al., 2020), and size/sex demographics (Cochran et al., 2016; Rohner et al., 2015). Saudi Arabia hosts a seasonal aggregation (Cagua et al., 2015) that attracts relatively small (average total length: 4.04 m), juvenile sharks of both sexes in roughly equal numbers (Cochran et al., 2016). Satellite telemetry from this site has shown that the sharks disperse hundreds of kilometers across the wider Red Sea and into the Indian Ocean after the aggregation season ends (Berumen et al., 2014). Conversely, Tanzania hosts a year-round resident aggregation (Cagua et al., 2015) of comparatively large (average total length: 5.99 m) juvenile males (Rohner et al., 2015) that seasonally move

only a few tens of kilometers between near-shore and off-shore habitats (Rohner et al., 2020). The results here suggest that these differences are not driven by population-level genetic variation between the two sites but may instead be examples of behavioral plasticity in response to each aggregation's local environment. Protection legislation and other conservation actions should account for these differences because, despite genetic homogeneity within the Indo-Pacific, there is unlikely to be a "one-size-fits-all" approach to managing whale shark aggregations.

4.2 | Trends in genetic diversity

Based on the KGTEST results, both the Saudi Arabian and Tanzanian populations appear to be stable over relatively short time scales. Similarly, both allelic richness and haplotype diversity at the Saudi Arabian aggregation site appear to have been stable from 2010 through 2015. These results are counter to the declines in diversity reported in Western Australia from 2007 to 2012 (Vignaud et al., 2014). Based on present results and previous studies (Meekan et al., 2017; Vignaud et al., 2014; Yagishita et al., 2020), sharks sampled in Western Australia and Saudi Arabia are both part of a single

TABLE 3 Population pairwise F_{ST} and p-values based on haplotype frequencies (gaps considered informative)

	Saudi Arabia (RS)	Djibouti	Qatar	Tanzania	Mozambique	Seychelles	Maldives	W. Australia	Philippines	Taiwan	Japan	Mexico (Pac)	Mexico (Atl)
Saudi Arabia (RS)	–	–	–	–	–	–	–	–	–	–	–	++	++
Djibouti	0.0001	–	++	–	–	+	–	+	–	+	–	++	++
Qatar	0.0053	0.0183	–	–	–	–	–	–	–	–	–	–	++
Tanzania	–0.0012	0.0071	0.0013	–	–	–	–	–	–	–	–	–	++
Mozambique	0.0004	0.0068	0.0052	–0.0011	–	–	–	–	–	–	–	–	++
Seychelles	0.008	0.0133	0.0093	0.0046	0.0045	–	–	–	–	–	–	–	++
Maldives	0.0086	0.0051	0.0325	0.0057	0.0058	0.0142	–	+	–	–	–	–	++
W. Australia	0.0022	0.0092	0.0014	0.0055	0.0028	0.006	0.023	–	–	–	–	–	++
Philippines	0.0015	0.0016	0.0077	–0.0024	–0.0081	0.0004	–0.0031	0.0032	–	–	–	–	++
Taiwan	0.0088	0.019	0.003	–0.0012	0.0027	0.0042	0.005	0.0077	0.0035	–	+	–	++
Japan	0.0025	–0.0008	0.0121	0.0082	–0.0054	0.0118	0.0071	0.0014	–0.0048	0.0164	–	–	++
Mexico (Pac)	0.0106	0.0139	0.0007	0.006	0.0052	0.0106	0.0246	0.0041	0.0048	0.0062	0.0059	–	++
Mexico (Atl)	0.0791	0.068	0.0974	0.0747	0.0684	0.0742	0.0702	0.0791	0.0532	0.0873	0.0658	0.0797	

Note: Not significant; + $p < .05$; ++ $p < .01$ (corrected for multiple comparison).

Abbreviations: Atl, Atlantic; Pac, Pacific; RS, Red Sea.

Indo-Pacific population. The declines in Western Australia, and the absence of declines in Saudi Arabia, suggest that the former's losses of genetic diversity are the result of local processes as opposed to global- or population-scale phenomena (Vignaud et al., 2014). It is also possible, given the distance between the two sites, that there is a time lag between impacts on one portion of the population affecting the other. If this is the case, then continued monitoring in Saudi Arabia might eventually be expected to show declines similar to those reported from Western Australia.

Several potential threats may disproportionately affect whale sharks from Western Australia when compared to those from the Saudi Arabian Red Sea, including natural predation as well as anthropogenic pressures. Two predators are known to attack and occasionally kill whale sharks: the killer whale *Orcinus orca* (O'Sullivan, 2000) and the white shark *Carcharodon carcharias* (Fitzpatrick et al., 2006). Both are relatively common along the western coast of Australia (McAuley et al., 2017); Pitman et al., 2015), but rare (Notarbartolo di Sciara et al., 2017) or absent (Golani & Bogorodsky, 2010) within the Red Sea. This difference is reflected in the prevalence and types of scars reported from whale sharks in the Saudi Arabian Red Sea and Western Australia. Scarring consistent with boat strikes are commonly seen at both sites (Cochran et al., 2016; Speed et al., 2008), although scars directly attributed to boats are less frequent in Western Australia than in Saudi Arabia (Lester et al., 2020). Conversely, bite marks have not been documented in Saudi Arabia (Cochran et al., 2016) while in Western Australia 4.8%–11% of individuals had predator bite scars (Lester et al., 2020; Speed et al., 2008). In addition to natural predators, Western Australia is also in relatively close proximity to several historical whale shark fisheries (Alava et al., 1997; Chen et al., 1997; White & Cavanagh, 2007), and at least one that still actively targets these animals (Li et al., 2012). In contrast, there are no known current or former fisheries dedicated to landing whale sharks within the Red Sea.

Conversely, other threats (particularly boat strike) are arguably more prevalent in the Red Sea given its status as one of the world's busiest shipping lanes (Stevens, 2007). The use of spotter planes to direct tourism boats to whale shark locations within Ningaloo Marine Park may also decrease the threat of collisions in that immediate area. At Shib Habil, Saudi Arabia, tourism operations rely exclusively on boat-based surveys, increasing the potential risk of boat strike (Lester et al., 2020). Furthermore, the loss of genetic diversity reported from Western Australia does not necessarily correspond to declines in local whale shark abundance (Vignaud et al., 2014). While the different trends in genetic diversity in Saudi Arabia and Western Australia certainly warrant additional investigation, conclusions from both sites are based on only 6 years of data. These are relatively short time series when compared to the potential life span of the species (Ong et al., 2020). Continued sampling at both sites and at other aggregations could substantially improve this comparison and help to identify the exact causes for the declines in genetic diversity shown in Western Australia (Vignaud et al., 2014).

TABLE 4 Measures of genetic diversity at each location (gaps considered); number of sequences (*n*), number of haplotypes (N_{hp}), haplotype diversity (*h*), and nucleotide diversity (*π*), Tajima's *D*, Fu's *Fs*, and Harpending's raggedness index (HRI)

| Sampling locations | *n* | Genetic diversity | | | Neutrality tests | | Mismatch distribution (HRI) |
		N_{hp}	*h*	*π*	Tajima's *D*	Fu's *Fs*	
Saudi Arabia (RS)	68	31	0.93 ± 0.02	0.07 ± 0.04	−1.61[+]	0.16[NS]	0.02[NS]
Djibouti	77	31	0.93 ± 0.01	0.13 ± 0.06	−1.23[NS]	8.31[NS]	0.01[NS]
Qatar	54	20	0.90 ± 0.03	0.08 ± 0.04	−1.19[NS]	5.34[NS]	0.02[NS]
Tanzania	57	33	0.96 ± 0.01	0.05 ± 0.03	−0.62[NS]	−4.87[NS]	0.01[NS]
Mozambique	62	33	0.96 ± 0.01	0.17 ± 0.08	−0.51[NS]	5.18[NS]	0.01[NS]
Seychelles	38	21	0.95 ± 0.02	0.13 ± 0.06	−1.60[NS]	5.14[NS]	0.02[NS]
Maldives	12	12	1.00 ± 0.03	0.13 ± 0.06	−1.91[++]	−1.42[NS]	0.02[NS]
W. Australia	162	48	0.92 ± 0.01	0.10 ± 0.05	−1.34[NS]	4.08[NS]	0.01[NS]
Philippines	31	22	0.97 ± 0.02	0.15 ± 0.07	−1.60[+]	2.81[NS]	0.01[NS]
Taiwan	26	21	0.97 ± 0.03	0.12 ± 0.06	−1.13[NS]	−0.26[NS]	0.01[NS]
Japan	28	15	0.95 ± 0.02	0.09 ± 0.04	−1.44[NS]	−3.45[NS]	0.02[NS]
Mexico (Pac)	121	33	0.93 ± 0.01	0.11 ± 0.05	−1.50[+]	11.22[NS]	0.01[NS]
Mexico (Atl)	80	24	0.85 ± 0.03	0.17 ± 0.08	3.72[NS]	20.16[NS]	0.08[NS]
Indo-Pacific	736	178	0.94 ± 0.00	0.10 ± 0.05	−1.23[NS]	−23.31[+]	0.01[NS]
Overall	816	192	0.94 ± 0.00	0.12 ± 0.06	−0.89[NS]	−23.31[+]	0.01[NS]

Note: [+]$0.01 < p < .05$; [++]$p < .01$ (corrected for multiple comparisons).

Abbreviations: Atl, Atlantic; NS, not significant; Pac, Pacific; RS, Red Sea.

4.3 | Global structure

The haplotype networks all revealed relatively small, but statistically significant differences between Indo-Pacific aggregation sites (Table 3, Appendix S1: SI 6). This is consistent with previous research on whale shark population genetics (Vignaud et al., 2014; Yagishita et al., 2020). Given the low F_{ST} values, previous studies have simply dismissed these differences as either ecologically negligible or (in the case of the Maldives) artifacts of small sample sizes (Vignaud et al., 2014; Yagishita et al., 2020). This conclusion is also supported by more general guidelines for interpreting F_{ST} that describe values below 0.05 (as is the case for all Indo-Pacific comparisons in the present study) as relatively weak indicators of population structure (Hartl et al., 1997). While these differences might warrant further investigation, they are not sufficient evidence to justify further subdivision of the Indo-Pacific population.

Most importantly, the much larger differences recorded between the Atlantic and Indo-Pacific support the existence of two basin-scale whale shark populations. The first global haplotype network for whale sharks was built using 69 samples from 6 aggregations and split the species into distinct Atlantic and Indo-Pacific populations (Castro et al., 2007). This original network was updated in 2014 (Vignaud et al., 2014), 2017 (Meekan et al., 2017), (Yagishita et al., 2020), and again in the present study using 816 samples from 13 sites. Despite more than an order of magnitude increase in sample size and the use of three different alignment methodologies, the division of whale sharks into Atlantic and Indo-Pacific populations has been consistent across studies and is further confirmed by our results. Given the robustness of the two-population model for

global whale shark population dynamics, it seems unlikely that additional sampling from novel sites will significantly alter these results. Revealing finer scale differentiation among distant sites within each basin may require a fundamental shift in methodology, either toward single-nucleotide polymorphisms or population genomics.

The relative lack of genetic structure among global whale shark aggregations is unusual when compared to many other pelagic megaplanktivores. There is evidence for genetic differentiation among of several cetaceans species (Jackson et al., 2014; Sremba et al., 2012) and both manta species (*Mobula birostris* and *Mobula alfredi*) (Stewart et al., 2016; Venables et al., 2021) within ocean basins. However, the lack of basin-scale differentiation appears to be a shared trait among the filter-feeding sharks. The basking shark *Cetorhinus maximus* (Rus Hoelzel et al., 2006) and the megamouth shark *Megachasma pelagios* (Liu et al., 2018) are both characterized by a single, globally distributed population making both of these species less divergent than the whale shark. While recent research has suggested some level of genetic structure for the basking shark in the northeast Atlantic, a single panmictic population remains the prevailing view (Lieber et al., 2020). If correct, the single population model for both the basking and megamouth sharks suggests a clear difference in the recent dispersal patterns of the three planktivorous shark species. This difference may derive from the whale shark's clear preference for the tropical and warm temperate seas (Rowat & Brooks, 2012), where several landmasses act as barriers to migration and gene flow. The basking and megamouth sharks are more tolerant of colder waters and are known to exhibit extended periods of residency in the deep ocean which may alleviate some barriers to movement and genetic exchange across basins. Alternatively, the global homogeneity found

in these two species could be due to historical, transarctic connectivity during the late Pleistocene (Lieber et al., 2020).

4.4 | Demographic history

Two previous studies used neutrality statistics and other methods to detect fluctuations in the Indo-Pacific whale shark population (Vignaud et al., 2014; Yagishita et al., 2020). Both of these studies reduced gap regions within their mitochondrial alignments for these analyses, effectively assuming that contiguous gaps resulted from singular events rather than independent mutations (Simmons & Ochoterena, 2000), and both found strong evidence for population expansion (Vignaud et al., 2014; Yagishita et al., 2020). These previous results are consistent with both the gap-exclusive and gap-reduced mitochondrial networks from the present study (Appendix S1: SI 5). The Indo-Pacific Tajima's D and Fu's Fs were significantly negative indicating recent population expansion in both cases (Appendix S1: SI 7 and 8) The evidence for population expansion became much weaker when contiguous gaps were analyzed. The Indo-Pacific Tajima's D was not significant. Fu's Fs remained significantly negative, but this may be attributed to the metric's oversensitivity or to the potential confounding effects of species-specific recombination rates (Rozas & Calafell, 2008). Similarly, HRI values were not significant for all sites. This is consistent with Yagishita et al. (2020), who found nonsignificant HRI values for whale sharks in the Atlantic, where most other metrics indicated a stable population. The alignment differences in the BSPs were less pronounced, with all groupings showing some evidence of population expansion. However, the abruptness and magnitude of the population declines preceding the recent expansions differed between gap treatments and between ocean basins. These mixed results raise some doubt on the recent population expansions proposed by earlier studies (Vignaud et al., 2014; Yagishita et al., 2020).

The differences between the gap-inclusive, -reduced, and -exclusive networks also highlight the potential effects that relatively minor methodological changes can have on final results. The gaps themselves were highly varied and their positions within the alignment were consistent across samples from different studies and geographic regions. This suggests that they are unlikely to be artifacts of sequencing or analytical methodologies and are most likely bioinformative collections of deletion mutations. Regardless, disentangling whale shark demographic history will require additional sampling and more advanced sequencing methods, but an improved understanding of past fluctuations in whale shark abundance may help conservation efforts to gauge and manage current or future population changes in the face of anthropogenic pressures.

5 | CONCLUSIONS

This paper adds samples from whale sharks in two understudied regions, describes new microsatellite markers, and provides a point of comparison to previous results from other aggregations (Vignaud

et al., 2014). In addition, most of the new mitochondrial sequences from this study have been directly linked to identification photos for sampled sharks (Appendix S1: SI 10 and 11), with accession numbers added to the online profile of each individual (sharkbook.ai). Linking the genetic and photographic datasets in this way can simplify future analysis and facilitate multi-method studies. The results here strongly suggest that the Saudi Arabian and Tanzanian aggregation sites are part of a larger Indo-Pacific population despite the two sites being over 4500 km apart, showing no direct evidence of connectivity (through photo-ID or tagging studies), and exhibiting markedly different local behavioral ecologies. This would seem to indicate that basin-scale genetic exchange occurs later in the whale shark's life history or that aggregations are "seeded" with the offspring of a more interconnected, adult population. In terms of conservation, this would mean that local protections at aggregation sites are important but probably insufficient as the sole means of managing a largely pelagic species. Effective high seas protections will likely be crucial for the long-term survival of whale shark and other species that frequently move through international waters (Queiroz et al., 2019). However, the development of such protections is currently hindered by the inability to reliably find mature sharks and the relatively short-term deployment capabilities of most current satellite transmitters. Until direct, long-term monitoring of large whale sharks becomes readily available, genetic inference remains the best method for mapping the species' global patterns of dispersal.

The broad division of whale sharks into two basin-scale populations (i.e., Indo-Pacific and Atlantic) appears to be robust and resolving finer scale differences among aggregating subpopulations (e.g., within the Indo-Pacific) may require researchers to move beyond microsatellite and mitochondrial markers and on to more comprehensive sequencing methods. The recent publication of complete mitochondrial-genome (Alam et al., 2014) and whole-genome (Hara et al., 2018; Read et al., 2017; Weber et al., 2020) assemblies for the whale shark represents an important first step away from population genetics and toward population genomics. Improved scientific understanding of the genetic structure and long-term patterns of connectivity among whale shark aggregation sites will be vital to the effective global conservation of this endangered species.

CONFLICT OF INTEREST

The authors declare no competing interests.

AUTHOR CONTRIBUTIONS

Royale S. Hardenstine: Conceptualization (equal); Data curation (equal); Formal analysis (supporting); Methodology (equal); Resources (supporting); Visualization (equal); Writing – original draft (equal). **Song He:** Conceptualization (supporting); Data curation (equal); Formal analysis (lead); Methodology (equal); Visualization (equal); Writing – original draft (supporting). **Jesse E. M. Cochran:** Conceptualization (supporting); Methodology (equal); Resources (equal); Writing – original draft (equal); Writing – review & editing (lead). **Camrin D. Braun:** Resources (equal); Writing – review & editing (equal). **E. Fernando Cagua:** Resources (equal); Writing – review & editing (equal). **Simon J. Pierce:** Funding acquisition (equal); Project administration (equal); Resources (equal); Writing – review & editing (equal). **Clare E. M. Prebble:** Data curation (supporting); Resources (equal); Writing – review & editing (equal). **Christoph A. Rohner:** Data curation (supporting); Resources (equal); Writing – review & editing (equal). **Pablo Saenz-Angudelo:** Methodology (equal); Resources (equal); Writing – review & editing (equal). **Tane H. Sinclair-Taylor:** Resources (equal); Writing – review & editing (equal). **Gregory B. Skomal:** Funding acquisition (equal); Writing – review & editing (equal). **Simon R. Thorrold:** Funding acquisition (equal); Writing – review & editing (equal). **Alexandra M. Watts:** Formal analysis (supporting); Writing – review & editing (equal). **Casey J. Zakroff:** Methodology (supporting); Writing – review & editing (equal). **Michael L. Berumen:** Conceptualization (equal); Funding acquisition (equal); Project administration (equal); Supervision (lead); Writing – review & editing (equal).

DATA AVAILABILITY STATEMENT

All microsatellite genotypes and primers have been deposited onto Dryad (https://doi.org/10.5061/dryad.sqv9s4n58). All mitochondrial sequences are available on GenBank (accession numbers: OL782199–OL782316). If available, the accession number for an individual shark will be included on their respective Wildbook for Whale Sharks profile.

ORCID

Royale S. Hardenstine 🔟 https://orcid.org/0000-0003-4693-7045
Song He 🔟 https://orcid.org/0000-0002-1293-6514
Jesse E. M. Cochran 🔟 http://orcid.org/0000-0002-6027-5052
Camrin D. Braun 🔟 https://orcid.org/0000-0002-9317-9489
Simon J. Pierce 🔟 https://orcid.org/0000-0002-9375-5175
Christoph A. Rohner 🔟 https://orcid.org/0000-0001-8760-8972
Pablo Saenz-Agudelo 🔟 https://orcid.org/0000-0001-8197-2861
Tane H. Sinclair-Taylor 🔟 https://orcid.org/0000-0003-4240-0435
Simon R. Thorrold 🔟 https://orcid.org/0000-0002-1533-7517
Alexandra M. Watts 🔟 https://orcid.org/0000-0002-1251-1735
Michael L. Berumen 🔟 https://orcid.org/0000-0003-2463-2742

REFERENCES

Acuña-Marrero, D., Jiménez, J., Smith, F., Doherty, P. F., Hearn, A., Green, J. R., Paredes-Jarrín, J., & Salinas-de-León, P. (2014). Whale Shark (*Rhincodon typus*) seasonal presence, residence time and habitat use at Darwin Island, Galapagos Marine Reserve. *PLoS One*, *9*, e115946. https://doi.org/10.1371/journal.pone.0115946

Alam, M. T., Petit, R. A. III, Read, T. D., & Dove, A. D. (2014). 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*, *539*, 44–49. https://doi.org/10.1016/j.gene.2014.01.064

Alava, M., Yaptinchay, A., Trono, R., & Dolumbal, E. (1997). Fishery and trade of whale sharks and manta rays in the Bohol Sea, Philippines. *International seminar and workshop on shark and ray biodiversity Conservation and Management*. Sabah Malaysia: Department of Fisheries; Sabahand IUCN Species Survival Commission's Shark Specialist Group.

Andrzejaczek, S., Meeuwig, J., Rowat, D., Pierce, S., Davies, T., Fisher, R., & Meekan, M. (2016). The ecological connectivity of whale shark aggregations in the Indian Ocean: A photo-identification approach. *Royal Society Open Science*, *3*, 160455. https://doi.org/10.1098/rsos.160455

Araujo, G., Lucey, A., Labaja, J., So, C. L., Snow, S., & Ponzo, A. (2014). Population structure and residency patterns of whale sharks, *Rhincodon typus*, at a provisioning site in Cebu, Philippines. *PeerJ*, *2*, e543.

Arzoumanian, Z., Holmberg, J., & Norman, B. (2005). An astronomical pattern-matching algorithm for computer-aided identification of whale sharks *Rhincodon typus*. *Journal of Applied Ecology*, *42*, 999–1011.

Berumen, M. L., Braun, C. D., Cochran, J. E., Skomal, G. B., & Thorrold, S. R. (2014). Movement patterns of juvenile whale sharks tagged at an aggregation site in the Red Sea. *PLoS One*, *9*, e103536. https://doi.org/10.1371/journal.pone.0103536

Bilgin, R. (2007). Kgtests: A simple Excel Macro program to detect signatures of population expansion using microsatellites. *Molecular Ecology Notes*, *7*, 416–417.

Brooks, K., Rowat, D., Pierce, S. J., Jouannet, D., & Vely, M. (2010). Seeing spots: photo-identification as a regional tool for whale shark identification. *Western Indian Ocean Journal of Marine Science*, *9*, 185–194.

Brunnschweiler, J. M., Baensch, H., Pierce, S. J., & Sims, D. W. (2009). Deep-diving behaviour of a whale shark *Rhincodon typus* during long-distance movement in the western Indian Ocean. *Journal of Fish Biology*, *74*, 706–714. https://doi.org/10.1111/j.1095-8649.2008.02155.x

Cagua, E. F., Cochran, J. E. M., Rohner, C. A., Prebble, C. E. M., Sinclair-Taylor, T. H., Pierce, S. J., & Berumen, M. L. (2015). Acoustic telemetry reveals cryptic residency of whale sharks. *Biology Letters*, *11*, 20150092. https://doi.org/10.1098/rsbl.2015.0092

Castro, A., Stewart, B. S., Wilson, S. G., Hueter, R. E., Meekan, M. G., Motta, P. J., Bowen, B. W., & Karl, S. A. (2007). Population genetic structure of Earth's largest fish, the whale shark (*Rhincodon typus*). *Molecular Ecology*, *16*, 5183–5192.

Chen, C.-T., Liu, K. M., & Joung, S. J. (1997). Preliminary report on Taiwan's whale shark fishery. *Traffic Bulletin*, *17*, 53–57.

Cochran, J. E. M., Braun, C. D., Cagua, E. F., Campbell, M. F., Hardenstine, R. S., Kattan, A., Priest, M. A., Sinclair-Taylor, T. H., Skomal, G. B., Sultan, S., Sun, L. U., Thorrold, S. R., & Berumen, M. L. (2019). Multi-method assessment of whale shark (*Rhincodon typus*) residency, distribution, and dispersal behavior at an aggregation site in the Red Sea. *PLoS One*, *14*, e0222285. https://doi.org/10.1371/journal.pone.0222285

Cochran, J., Hardenstine, R. S., Braun, C. D., Skomal, G. B., Thorrold, S. R., Xu, K., Genton, M. G., & Berumen, M. L. (2016). Population structure of a whale shark *Rhincodon typus* aggregation in the Red Sea. *Journal of Fish Biology*, *89*, 1570–1582.

Diamant, S., Rohner, C. A., Kiszka, J. J., Guillemain d'Echon, A., Guillemain d'Echon, T., Sourisseau, E., & Pierce, S. J. (2018). Movements and habitat use of satellite-tagged whale sharks off western Madagascar. *Endangered Species Research*, *36*, 49–58. https://doi.org/10.3354/esr00889

Earl, D. A. (2012). STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, *4*, 359–361. https://doi.org/10.1007/s12686-011-9548-7

Excoffier, L., & Lischer, H. E. (2010). Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, *10*, 564–567. https://doi.org/10.1111/j.1755-0998.2010.02847.x

Faircloth, B. C. (2008). MSATCOMMANDER: Detection of microsatellite repeat arrays and automated, locus-specific primer design. *Molecular Ecology Resources*, *8*, 92–94. https://doi.org/10.1111/j.1471-8286.2007.01884.x

Ferreri, M., Qu, W., & Han, B. O. (2011). Phylogenetic networks: A tool to display character conflict and demographic history. *African Journal of Biotechnology*, *10*(60), 12799–12803. https://doi.org/10.5897/AJB11.010

Fitzpatrick, B., Meekan, M., & Richards, A. (2006). Shark attacks on a whale shark (*Rhincodon typus*) at Ningaloo Reef, western Australia. *Bulletin of Marine Science*, *78*, 397–402.

Golani, D., & Bogorodsky, S. V. (2010). The fishes of the Red Sea—Reappraisal and updated checklist. *Zootaxa*, *2463*, 1–135. https://doi.org/10.11646/zootaxa.2463.1.1

Graham, R. T., & Roberts, C. M. (2007). Assessing the size, growth rate and structure of a seasonal population of whale sharks (*Rhincodon typus* Smith 1828) using conventional tagging and photo identification. *Fisheries Research*, *84*, 71–80. https://doi.org/10.1016/j.fishres.2006.11.026

Hara, Y., Yamaguchi, K., Onimaru, K., Kadota, M., Koyanagi, M., Keeley, S. D., Tatsumi, K., Tanaka, K., Motone, F., Kageyama, Y., Nozu, R., Adachi, N., Nishimura, O., Nakagawa, R., Tanegashima, C., Kiyatake, I., Matsumoto, R., Murakumo, K., Nishida, K., … Kuraku, S. (2018). Shark genomes provide insights into elasmobranch evolution and the origin of vertebrates. *Nature Ecology & Evolution*, *2*, 1761–1771. https://doi.org/10.1038/s41559-018-0673-5

Hartl, D. L., Clark, A. G., & Clark, A. G. (1997). *Principles of population genetics* (vol. *116*). Sinauer Associates.

Hueter, R. E., Tyminski, J. P., & de la Parra, R. (2013). Horizontal movements, migration patterns, and population structure of whale sharks in the Gulf of Mexico and northwestern Caribbean Sea. *PLoS One*, *8*, e71883. https://doi.org/10.1371/journal.pone.0071883

Jackson, J. A., Steel, D. J., Beerli, P., Congdon, B. C., Olavarría, C., Leslie, M. S., Pomilla, C., Rosenbaum, H., & Baker, C. S. (2014). Global diversity and oceanic divergence of humpback whales (*Megaptera novaeangliae*). *Proceedings of the Royal Society B: Biological Sciences*, *281*, 20133222.

Kendall, M. G. (1955). Further contributions to the theory of paired comparisons. *Biometrics*, *11*, 43–62. https://doi.org/10.2307/3001479

Ketchum, J. T., Galván-Magaña, F., & Klimley, A. P. (2013). Segregation and foraging ecology of whale sharks, *Rhincodon typus*, in the southwestern Gulf of California. *Environmental Biology of Fishes*, *96*, 779–795. https://doi.org/10.1007/s10641-012-0071-9

Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A., & Mayrose, I. (2015). Clumpak: A program for identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology Resources*, *15*, 1179–1191.

Lester, E., Meekan, M. G., Barnes, P., Raudino, H., Rob, D., Waples, K., & Speed, C. W. (2020). Multi-year patterns in scarring, survival and residency of whale sharks in Ningaloo Marine Park, Western Australia. *Marine Ecology Progress Series*, *634*, 115–125. https://doi.org/10.3354/meps13173

Li, W., Wang, Y., & Norman, B. (2012). A preliminary survey of whale shark *Rhincodon typus* catch and trade in China: An emerging crisis. *Journal of Fish Biology*, *80*, 1608–1618. https://doi.org/10.1111/j.1095-8649.2012.03250.x

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

Lieber, L., Hall, G., Hall, J., Berrow, S., Johnston, E., Gubili, C., Sarginson, J., Francis, M., Duffy, C., Wintner, S. P., & Doherty, P. D. (2020). Spatio-temporal genetic tagging of a cosmopolitan planktivorous shark provides insight to gene flow, temporal variation and site-specific re-encounters. *Scientific Reports*, *10*(1), 1–17. https://doi.org/10.1038/s41598-020-58086-4

Liu, S. Y. V., Joung, S. J., Yu, C.-J., Hsu, H.-H., Tsai, W.-P., & Liu, K. M. (2018). Genetic diversity and connectivity of the megamouth shark (*Megachasma pelagios*). *PeerJ*, *6*, e4432.

Mann, H. B. (1945). Nonparametric tests against trend. *Econometrica*, *13*(3), 245–259. https://doi.org/10.2307/1907187

McAuley, R., Bruce, B., Keay, I., Mountford, S., Pinnell, T., & Whoriskey, F. (2017). Broad-scale coastal movements of white sharks off Western Australia described by passive acoustic telemetry data. *Marine and Freshwater Research*, *68*, 1518–1531. https://doi.org/10.1071/MF16222

McKinney, J. A., Hoffmayer, E. R., Holmberg, J., Graham, R. T., Driggers, W. B., de la Parra-Venegas, R., Galván-Pastoriza, B. E., Fox, S., Pierce, S. J., & Dove, A. D. M. (2017). Long-term assessment of whale shark population demography and connectivity using photo-identification in the Western Atlantic Ocean. *PLoS One*, *12*, e0180495. https://doi.org/10.1371/journal.pone.0180495

Meekan, M., Austin, C. M., Tan, M. H., Wei, N.-W., Miller, A., Pierce, S. J., Rowat, D., Stevens, G., Davies, T. K., Ponzo, A., & Gan, H. M. (2017). iDNA at sea: Recovery of whale shark (*Rhincodon typus*) mitochondrial DNA sequences from the whale shark copepod (*Pandarus rhincodonicus*) confirms global population structure. *Frontiers in Marine Science*, *4*, 420. https://doi.org/10.3389/fmars.2017.00420

Meekan, M. G., Bradshaw, C., Press, M., McLean, C., Richards, A., Quasnichka, S., & Taylor, J. G. (2006). Population size and structure of whale sharks (*Rhincodon typus*) at Ningaloo Reef Western Australia. *Marine Ecology-Progress Series*, *319*, 275–285. https://doi.org/10.3354/meps319275

Narum, S. R. (2006). Beyond Bonferroni: Less conservative analyses for conservation genetics. *Conservation Genetics*, *7*, 783–787. https://doi.org/10.1007/s10592-005-9056-y

Nei, M., & Jin, L. (1989). Variances of the average numbers of nucleotide substitutions within and between populations. *Molecular Biology and Evolution*, *6*, 290–300.

Norman, B. (2005). *Rhincodon typus*. IUCN 2014 IUCN list of threatened species version 20141.

Norman, B. M., Holmberg, J. A., Arzoumanian, Z., Reynolds, S. D., Wilson, R. P., Rob, D., Pierce, S. J., Gleiss, A. C., de la Parra, R., Galvan, B., Ramirez-Macias, D., Robinson, D., Fox, S., Graham, R., Rowat, D., Potenski, M., Levine, M., Mckinney, J. A., Hoffmayer, E., … Morgan,

D. L. (2017). Undersea constellations: the global biology of an endangered marine megavertebrate further informed through citizen science. *BioScience*, 67, 1029–1043. https://doi.org/10.1093/biosci/bix127

Norman, B., Whitty, J., Beatty, S., Reynolds, S., & Morgan, D. (2017). Do they stay or do they go? Acoustic monitoring of whale sharks at Ningaloo Marine Park, Western Australia. *Journal of Fish Biology*, 91, 1713–1720. https://doi.org/10.1111/jfb.13461

Notarbartolo di Sciara, G., Kerem, D., Smeenk, C., Rudolph, P., Cesario, A., Costa, M., Elasar, M., Feingold, D., Fumagalli, M., Goffman, O., Hadar, N., Mebrathu, Y.T., & Scheinin, A. (2017). Cetaceans of the red sea. CMS Technical Series, Vol 33, (p. 86).

Ong, J. J., Meekan, M. G., Hsu, H. H., Fanning, L. P., & Campana, S. E. (2020). Annual bands in vertebrae validated by bomb radiocarbon assays provide estimates of age and growth of whale sharks. *Frontiers in Marine Science*, 7, 188. https://doi.org/10.3389/fmars.2020.00188

O'Sullivan, J. B. (2000). A fatal attack on a whale shark *Rhincodon typus*, by killer whales *Orcinus orca* off Bahia de los Angeles Baja California. Abstracts of the American Elasmobranch Society 16th Annual Meeting, La Paz, Mexico. (p. 282).

Park, S. (2001). *The Excel microsatellite toolkit*.

Peakall, R., & Smouse, P. E. (2006). GENALEX 6: Genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, 6, 288–295. https://doi.org/10.1111/j.1471-8286.2005.01155.x

Pierce, S. J., & Norman, B. (2016). *Rhincodon typus. The IUCN Red List of Threatened Species 2016: e.T19488A2365291*. https://doi.org/10.2305/IUCN.UK.2016-1.RLTS.T19488A2365291.en

Pitman, R. L., Totterdell, J. A., Fearnbach, H., Ballance, L. T., Durban, J. W., & Kemps, H. (2015). Whale killers: Prevalence and ecological implications of killer whale predation on humpback whale calves off Western Australia. *Marine Mammal Science*, 31, 629–657. https://doi.org/10.1111/mms.12182

Prebble, C., Rohner, C. A., Pierce, S. J., Robinson, D. P., Jaidah, M. Y., Bach, S. S., & Trueman, C. N. (2018). Limited latitudinal ranging of juvenile whale sharks in the Western Indian Ocean suggests the existence of regional management units. *Marine Ecology Progress Series*, 601, 167–183. https://doi.org/10.3354/meps12667

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, 155, 945–959. https://doi.org/10.1093/genetics/155.2.945

Queiroz, N., Humphries, N. E., Couto, A., Vedor, M., da Costa, I., Sequeira, A. M. M., Mucientes, G., Santos, A. M., Abascal, F. J., Abercrombie, D. L., Abrantes, K., Acuña-Marrero, D., Afonso, A. S., Afonso, P., Anders, D., Araujo, G., Arauz, R., Bach, P., Barnett, A., ... Sims, D. W. (2019). Global spatial risk assessment of sharks under the footprint of fisheries. *Nature*, 572(7770), 461–466. https://doi.org/10.1038/s41586-019-1444-4

Ramírez-Macías, D. et al (2012). Patterns in composition, abundance and scarring of whale sharks *Rhincodon typus* near Holbox Island, Mexico. *Journal of Fish Biology*, 80, 1401–1416.

Ramirez-Macias, D., Shaw, K., Ward, R., Galvan-Magana, F., & Vazquez-Juarez, R. (2009). Isolation and characterization of microsatellite loci in the whale shark (*Rhincodon typus*). *Molecular Ecology Resources*, 9, 798–800.

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. *Fisheries Research*, 84, 87–95. https://doi.org/10.1016/j.fishres.2006.11.038

Read, T. D. et al (2017). Draft sequencing and assembly of the genome of the world's largest fish, the whale shark: *Rhincodon typus* Smith 1828. *BMC Genomics*, 18, 532.

Reich, D. E., Feldman, M. W., & Goldstein, D. B. (1999). Statistical properties of two tests that use multilocus data sets to detect population expansions. *Molecular Biology and Evolution*, 16, 453–466. https://doi.org/10.1093/oxfordjournals.molbev.a026127

Riley, M., Hale, M., Harman, A., & Rees, R. (2010). Analysis of whale shark *Rhincodon typus* aggregations near South Ari Atoll, Maldives Archipelago. *Aquatic Biology*, 8, 145–150. https://doi.org/10.3354/ab00215

Robinson, D. P., Jaidah, M. Y., Bach, S., Lee, K., Jabado, R. W., Rohner, C. A., March, A., Caprodossi, S., Henderson, A. C., Mair, J. M., Ormond, R., & Pierce, S. J. (2016). Population structure, abundance and movement of whale sharks in the Arabian Gulf and the Gulf of Oman. *PLoS One*, 11, e0158593. https://doi.org/10.1371/journal.pone.0158593

Rohner, C. A., Armstrong, A. J., Pierce, S. J., Prebble, C. E. M., Cagua, E. F., Cochran, J. E. M., Berumen, M. L., & Richardson, A. J. (2015). Whale sharks target dense prey patches of sergestid shrimp off Tanzania. *Journal of Plankton Research*, 37(2), 352–362. https://doi.org/10.1093/plankt/fbv010

Rohner, C. A., Cochran, J. E. M., Cagua, E. F., Prebble, C. E. M., Venables, S. K., Berumen, M. L., Kuguru, B. L., Rubens, J., Brunnschweiler, J. M., & Pierce, S. J. (2020). No place like home? High residency and predictable seasonal movement of whale sharks off Tanzania. *Frontiers in Marine Science*, 7, 423. https://doi.org/10.3389/fmars.2020.00423

Rohner, C. A., Richardson, A. J., Prebble, C. E. M., Marshall, A. D., Bennett, M. B., Weeks, S. J., Cliff, G., Wintner, S. P., & Pierce, S. J. (2015). Laser photogrammetry improves size and demographic estimates for whale sharks. *PeerJ*, 3, e886. https://doi.org/10.7717/peerj.886

Rosen, S., & Skaletsky, H. (2000). Primer3 on the WWW for general users and for biologist programmers. In S. Misener & S. A. Krawetz (Eds.), *Bioinformatics methods and protocols: Methods in Molecular Biology* (pp. 365–386). Humana Press Inc.

Rousset, F. (2008). genepop'007: A complete re-implementation of the genepop software for Windows and Linux. *Molecular Ecology Resources*, 8, 103–106. https://doi.org/10.1111/j.1471-8286.2007.01931.x

Rowat, D., & Brooks, K. S. (2012). A review of the biology, fisheries and conservation of the whale shark *Rhincodon typus*. *Journal of Fish Biology*, 80, 1019–1056. https://doi.org/10.1111/j.1095-8649.2012.03252.x

Rowat, D., Brooks, K., March, A., McCarten, C., Jouannet, D., Riley, L., Jeffreys, G., Perri, M., Vely, M., & Pardigon, B. (2011). Long-term membership of whale sharks (*Rhincodon typus*) in coastal aggregations in Seychelles and Djibouti. *Marine and Freshwater Research*, 62, 621–627. https://doi.org/10.1071/MF10135

Rozas, J., & Calafell, F. (2008). Statistical power analysis of neutrality tests under demographic expansions, contractions and bottlenecks with. *Genetics*, 179, 555–567.

Rus Hoelzel, A., Shivji, M. S., Magnussen, J., & Francis, M. P. (2006). Low worldwide genetic diversity in the basking shark (*Cetorhinus maximus*). *Biology Letters*, 2, 639–642.

Schmidt, J. V., Chen, C. C., Sheikh, S. I., Meekan, M. G., Norman, B. M., & Joung, S. J. (2010). Paternity analysis in a litter of whale shark embryos. *Endangered Species Research*, 12(2), 117–124. https://doi.org/10.3354/esr00300

Schmidt, J. V., Schmidt, C. L., Ozer, F., Ernst, R. E., Feldheim, K. A., Ashley, M. V., & Levine, M. (2009). Low genetic differentiation across three major ocean populations of the whale shark, *Rhincodon typus*. *PLoS One*, 4, e4988. https://doi.org/10.1371/journal.pone.0004988

Sequeira, A., Mellin, C., Rowat, D., Meekan, M. G., & Bradshaw, C. J. (2012). Ocean-scale prediction of whale shark distribution. *Diversity and Distributions*, 18, 504–518. https://doi.org/10.1111/j.1472-4642.2011.00853.x

Sigsgaard, E. E., Nielsen, I. B., Bach, S. S., Lorenzen, E. D., Robinson, D. P., Knudsen, S. W., Pedersen, M. W., Jaidah, M. A., Orlando, L., Willerslev, E., Møller, P. R., & Thomsen, P. F. (2017). Population characteristics of a large whale shark aggregation inferred from

seawater environmental DNA. *Nature Ecology & Evolution, 1,* 4. https://doi.org/10.1038/s41559-016-0004

Simmons, M. P., & Ochoterena, H. (2000). Gaps as characters in sequence-based phylogenetic analyses. *Systematic Biology, 49,* 369–381. https://doi.org/10.1093/sysbio/49.2.369

Speed, C., Meekan, M. G., Rowat, D., Pierce, S. J., Marshall, A. D., & Bradshaw, C. J. A. (2008). Scarring patterns and relative mortality rates of Indian Ocean whale sharks. *Journal of Fish Biology, 72,* 1488–1503. https://doi.org/10.1111/j.1095-8649.2008.01810.x

Sremba, A. L., Hancock-Hanser, B., Branch, T. A., LeDuc, R. L., & Baker, C. S. (2012). Circumpolar diversity and geographic differentiation of mtDNA in the critically endangered Antarctic blue whale (*Balaenoptera musculus* intermedia). *PLoS One, 7,* e32579. https://doi.org/10.1371/journal.pone.0032579

Stevens, J. D. (2007). Whale shark (*Rhincodon typus*) biology and ecology: A review of the primary literature. *Fisheries Research, 84,* 4–9. https://doi.org/10.1016/j.fishres.2006.11.008

Stewart, J. D., Beale, C. S., Fernando, D., Sianipar, A. B., Burton, R. S., Semmens, B. X., & Aburto-Oropeza, O. (2016). Spatial ecology and conservation of Manta birostris in the Indo-Pacific. *Biological Conservation, 200,* 178–183. https://doi.org/10.1016/j.biocon.2016.05.016

Szpiech, Z., Jakobsson, M., & Rosenberg, N. (2008). *ADZE: Allelic diversity analyzer, version 1.0.* Distributed by authors at http://rosenberglabbioinformatics.med.umich.edu/adze.html

Tamura, K. (1992). Estimation of the number of nucleotide substitutions when there are strong transition-transversion and G+ C-content biases. *Molecular Biology and Evolution, 9,* 678–687.

Van Tienhoven, A., Den Hartog, J., Reijns, R., & Peddemors, V. (2007). A computer-aided program for pattern-matching of natural marks on the spotted raggedtooth shark Carcharias taurus. *Journal of Applied Ecology, 44,* 273–280. https://doi.org/10.1111/j.1365-2664.2006.01273.x

Venables, S. K., Marshall, A. D., Armstrong, A. J., Tomkins, J. L., & Kennington, W. J. (2021). Genome-wide SNPs detect no evidence of genetic population structure for reef manta rays (*Mobula alfredi*) in southern Mozambique. *Heredity, 126,* 308–319. https://doi.org/10.1038/s41437-020-00373-x

Vignaud, T. M., Maynard, J. A., Leblois, R., Meekan, M. G., Vázquez-Juárez, R., Ramírez-Macías, D., Pierce, S. J., Rowat, D., Berumen, M. L., Beeravolu, C., Baksay, S., & Planes, S. (2014). Genetic structure

of populations of whale sharks among ocean basins and evidence for their historic rise and recent decline. *Molecular Ecology, 23,* 2590–2601. https://doi.org/10.1111/mec.12754

Weber, J. A., Park, S. G., Luria, V., Jeon, S., Kim, H.-M., Jeon, Y., Bhak, Y., Jun, J. H., Kim, S. W., Hong, W. H., Lee, S., Cho, Y. S., Karger, A., Cain, J. W., Manica, A., Kim, S., Kim, J.-H., Edwards, J. S., Bhak, J., & Church, G. M. (2020). The whale shark genome reveals how genomic and physiological properties scale with body size. *Proceedings of the National Academy of Sciences of the United States of America, 117,* 20662–20671. https://doi.org/10.1073/pnas.1922576117

White, W. T., & Cavanagh, R. D. (2007). Whale shark landings in Indonesian artisanal shark and ray fisheries. *Fisheries Research, 84,* 128–131. https://doi.org/10.1016/j.fishres.2006.11.022

Wolfson, F. H. & Sciara, G. N. (1981). The whaleshark, *Rhiniodon typus* Smith 1828: An annotated bibliography. *Atti Della Società Italiana Di Scienze Naturali E Del Museo Civico Di Storia Naturale Di Milano, 122,* 171–203.

Yagishita, N., Ikeguchi, S.-I., & Matsumoto, R. (2020). Re-estimation of genetic population structure and demographic history of the whale shark (*Rhincodon typus*) with additional japanese samples, inferred from mitochondrial DNA Sequences. *Pacific Science, 74,* 31–47. https://doi.org/10.2984/74.1.3

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.