GBE

Chromosome-Level Assembly and Annotation of the Grey Reef Shark (*Carcharhinus amblyrhynchos*) Genome

Carolin Dahms [1,*], Laurent Vigliola [2], Leni Hiu Tong Cheung[3], Jerome Ho Lam Hui [3], Paolo Momigliano [1,4,*]

[1]Area of Ecology and Biodiversity, School of Biological Sciences, Faculty of Science, The University of Hong Kong, Hong Kong SAR, China
[2]ENTROPIE, SANTECO, IRD, CNRS, Ifremer, Université de la Réunion, Université de la Nouvelle-Calédonie, Nouméa, New Caledonia, France
[3]School of Life Sciences, Simon F.S. Li Marine Science Laboratory, State Key Laboratory of Agrobiotechnology, Institute of Environment, Energy and Sustainability, The Chinese University of Hong Kong, Hong Kong SAR, China
[4]Department of Ecological and Biological Sciences, University of Tuscia, Viterbo, Italy

*Corresponding authors: E-mails: carolin.dahms.ac@gmail.com; paolo.momigliano@gmail.com.

Accepted: August 19, 2025

Abstract

To date less than 5% of shark species have nuclear reference genomes, despite next-generation sequencing advances. Particularly for threatened shark species, there is a lack of reliable genomes which are crucial in facilitating research and conservation applications. We assembled the first nuclear reference genome of the endangered grey reef shark (*Carcharhinus amblyrhynchos*) using long-read PacBio HiFi and Omni-C sequencing to reach chromosome-level contiguity (36 pseudochromosomes; 2.9 Gbp) and high completeness (94% complete BUSCOs). BRAKER3 annotated 16,505 protein-coding genes after masking repetitive elements which accounted for 59% of the genome. We identified potential X and Y sex chromosomes on pseudochromosomes 36 and 57, respectively. The quality and completeness of the draft genome of *C. amblyrhynchos* will enable researchers to investigate genetic variations and adaptations specific to this species as well as across other *Carcharhinus* spp., opening new venues for comparative genomics and advancing conservation genetic applications.

Key words: reference genome, HiFi, Carcharhinidae, shark conservation, sex chromosome.

Significance

Stemming from an ancient vertebrate lineage, sharks present an interesting evolutionary study system. A third of shark species face extinction, yet critical genomic resources necessary for research and conservation remain scarce. To address this gap, we assembled and annotated the first chromosome-level nuclear reference genome of the threatened grey reef shark (*Carcharhinus amblyrhynchos*) at high completeness. This genome will help advance studies in evolution, phylogenetics, adaptation, and conservation, offering insights not only for this species but for wider elasmobranch and vertebrate research.

Introduction

Sharks belong to the cartilaginous fish (class Chondrichthyes), one of the oldest vertebrate lineages which emerged over 400 million years ago, encompassing a diverse array of taxa such as sharks, rays, skates, and chimaeras. Having lived through five mass extinction events, sharks now face unprecedented rates of population declines due to habitat loss and overfishing (Dulvy et al. 2014, 2021). According to

the IUCN Red List, nearly one-third of shark and ray species are currently classified as threatened (Critically Endangered, Endangered, or Vulnerable; Pacoureau et al. 2021), with some families, such as Carcharhinidae (requiem sharks), exhibiting even higher proportions—up to 68%—of threatened species.

The grey reef shark (*Carcharhinus amblyrhynchos*) is one of the most widespread requiem sharks in the Indo-Pacific. *C. amblyrhynchos* has high reef fidelity (Heupel et al. 2010; Papastamatiou et al. 2018), with rather rare long-range migrations (usually under 100 km), primarily undertaken by males (Bonnin et al. 2019). Their unique life history and habitat association has led to relatively high levels of genetic population structure compared with other shark species (Cortés 2000; Robbins 2006; Bernard et al. 2021). Ecologically, *C. amblyrhynchos* occupies a unique niche, representing up to 50% of higher-order predator biomass on some reefs (Friedlander et al. 2014) and preying on unusually large prey for its size (Barley et al. 2020). However, life history traits such as slow growth, late sexual maturity, and long lifespan make the species particularly susceptible to overexploitation (Stevens et al. 2000). Some populations in the Indian Ocean and western Pacific have declined by more than 90% (Graham et al. 2010), raising concerns about inbreeding and loss of genetic diversity, reducing their evolutionary potential (Frankham et al. 2002) and long-term viability of remaining populations (Sherman et al. 2023).

Genomic tools are increasingly being used in conservation to monitor genetic diversity, population connectivity, and inbreeding risk (Allendorf et al. 2013). In fisheries management, these approaches have informed conservation strategies and facilitated the restoration of many commercially important fish stocks, such as Pacific salmon (Waples 1995). The development of high-quality reference genomes is a critical step in enabling such analyses. Advances in long-read sequencing technologies, such as Pacific Biosciences (PacBio), producing lager sequence overlap and lower fragmentation (Lang et al. 2020) and chromatin conformation capture techniques, such as Hi-C and Omni-C, now allow for the generation of chromosome-level genome assemblies in a relatively affordable and rapid manner (Dudchenko et al. 2017).

Reliable genomes may be useful representatives of whole species to infer key conservation genomic statistics such as levels of heterozygosity, inbreeding, and demographic history. Reference genomes are key to most genetic studies commonly relying on low and mid coverage whole genome sequences, which need to be aligned to a species-specific reference genome (Lou et al. 2021). Reference assemblies significantly improve accuracy of various other analyses such as for local adaptations and provide a crucial baseline resource for nonmodel species, such as genetic data-deficient marine organisms (Oleksiak and Rajora 2020).

Despite potentially severe genetic consequences of critical shark population declines, only a few fisheries include assessment of population genetics (Ovenden et al. 2010), with genetic diversity and demographic history having been investigated in only about a tenth of shark species to date (Domingues et al. 2018). Contiguous high-quality genomes, allowing population and species-specific genetic assessments and genomic comparisons (Shafer et al. 2015), are currently available for less than 5% of shark species (23 species as of July 2025), due to the lack of quality samples and their large, repetitive genomes (Pearce et al. 2021). Within the Carcharhinidae, only three nuclear reference genomes exist to date: the Oceanic whitetip shark (*Carcharhinus longimanus*; Feldheim and Pirro 2023, unpublished), the lemon shark (*Negaprion brevirostris*; Baeza et al. 2024) and the great blue shark (*Prionace glauca*; Li 2024, unpublished).

Nevertheless, high-throughput sequencing and genome assembly projects like the Squalomix project (https://github.com/Squalomix/info/) have produced increasing numbers of shark and ray genomes (Stanhope et al. 2023; Yamaguchi et al. 2023; Baeza et al. 2024; Mayeur et al. 2024; Lee et al. 2025) and drastically expanded comparative genomics (Kuraku 2021). Due to their antiquity, Chondrichthyes are an interesting study system on the origins and evolution of genes (Marra et al. 2019), as well as a range of morphologies and life history traits such as jaws (Yu et al. 2008), the cerebellum (Sugahara et al. 2017), and oviparity (Nakaya et al. 2020). Sex determination in elasmobranchs has been relatively poorly understood compared with other clades (Yamaguchi et al. 2023). Recent findings suggest that sharks and rays have the oldest sex chromosomes among vertebrates, originating around 300 My ago, and express a unique dosage-dependent sex determination mechanism involving distinct molecules from other vertebrates (Niwa et al. 2025). Most studied elasmobranchs have highly differentiated XY chromosomes, although other systems such as XX/XO in *Potamotrygon* sp. (De Souza Valentim et al. 2013) and possibly ZZ/ZW in *Hypanus americanus* (Schwartz and Maddock 2002) might be confirmed as more Chondrichthyan karyotypes are investigated. To address the pressing need for shark genomic resources, we present the first assembled and annotated nuclear reference genome of the grey reef shark and identify potential sex chromosomes, the first within Carcharhinidae, which will hopefully aid conservation applications and wider evolutionary research.

Results and Discussion

Assembly

Three PacBio Revio Cells produced 90 Gb of sequence in 7,864,402 reads. The primary contig assembly generated by hifiasm contained 2,276 contigs with an N50 6.3 Mb,

an L50 of 126 contigs with 44% GC content after purging of duplicates (Table S1). The genome size estimate based on 21 kmers by Genomescope of 2.67 Gb from the contig assembly (Fig. S1) was similar to the final scaffolded assembly size of 2.9 Gb.

The Omni-C library used for scaffolding the contigs produced 511,721,951 read pairs of which 68.13% were high-quality pairs, with 31.66% of pairs mapping >10 kb apart. 23.57% were low quality, and 8.3% unmapped. After duplicate (24.14%) removal, Omni-C sequences amounted to a coverage of 23.5X. The proximity ligation libraries greatly improved the assembly quality, as YaHS joined the contigs into 1,520 scaffolds (Table S1), with an increased N50 of 90.6 Mb spanning 12 scaffolds, while 90% of the genome was covered in 36 pseudochromosomes (N95 = 276 Mb; L95 = 94; Fig. 1). Although the chromosome number of *C. amblyrhynchos* is currently not known, a haploid chromosome number of at least 36 is close to the expected range ($n = 38$ to 45) for *Carcharhinidae* according to karyotype studies (Stingo and Rocco 2001) and the same as in *Triakis scyllia* belonging to the *Carcharhiniformes* (Table S2; Asahida and Ida 1989; Uno et al. 2020). Assembly integrity was further evaluated through presence of the canonical telomeric repeat sequence AACCCT, a reverse complimented and string rotated equivalent to the chordate repeat TTAGGG, identified by the tidk "explore" function (v.0.2.65, Brown et al. 2025). The telomeric repeat was found at the ends in 32 out of the 36 pseudochromosomes, with 19 showing repeats on both ends (Table S3).

BUSCO results indicated a high degree of complete BUSCOS when compared with the vertebrata lineage dataset, finding 93.6% of the 3,354 genes, while 2.2% were complete and duplicated, 3.8% were missing, and 2.6% fragmented (Table 1, Table S2, Fig. S2). A generated Omni-C density map of the final scaffolded assembly indicated high contiguity (Fig. S3).

Annotation

The repeat content identified with RepeatMasker amounted to 1.73 Gb (59.33% of the genome) and was rich in long interspersed nuclear element (LINE) retrotransposons, comprising about 45% of the genome, the majority of which were of the L2, CR1 and Rex type (Fig. S3, Table S4). According to sequence divergence, LINEs experienced recent expansion waves at around 7% to 8% and 10% to 13% divergence (Fig. S4b). LTRs covered 5.88%, while only 1.58% were DNA transposons, of which the Tourist and Harbinger types were the most common (Fig. S4). The proportion of repetitive regions is within the expected range from 33% to 68% among sharks (Table S2). Overall, our results are in agreement with previously published studies showing that the repeat content of sharks is rich in LINE retrotransposons, which have the highest transposable element (TE) activity in more recent times (Fig. S4b; Hara et al. 2018; Marra et al. 2019; Zhang et al. 2020; Tan et al. 2021; Stanhope et al. 2023).

BRAKER3 inferred 21,372 proteins encoded by 16,505 genes with an average length of 52,698 bp (ranging 126 to 2,321,391 bp; Table S5). BLAST hits against the Uniprot SwissProt database functionally annotated 88% (14,602) of the dataset. This annotation yielded a relatively low number of predicted genes compared with related species annotations, which varied from 24,000 to 55,000 genes (Table S2), due to the suboptimal amount of RNA-seq evidence provided for model training. As obtaining tissue for endangered sharks is difficult, we had to rely on limited published RNA-seq from a different grey reef shark individual. Accordingly, the annotated assembly yielded a satisfactory, although slightly lower BUSCO completeness score of 86.5% (67.7% single, 18.8% duplicated, Table 1). An increase in duplicated copies is expected for annotated assemblies due to multiple protein sequences included per gene. Our score is similar (Wagner et al. 2023) and lower than some recent high-quality assemblies which generated transcriptome evidence (e.g. Mayeur et al. 2024; Niwa et al. 2025). We hope this annotation will provide a starting reference point for future studies.

Sex Chromosome Identification

Cytogenetic studies have identified heteromorphic XY sex chromosomes in sharks, with the Y being around one-third of the size of the X (Uno et al. 2020). In line with the expectation that the X chromosome should have reduced male sequencing coverage, while the Y chromosome should have low female coverage, we identified a 2-fold female to male depth ratio on pseudochromosome 36 (17.3 Mb) consistent with its hemizygous state in males and diploid state in females (Fig. 1b, Fig. S5). The smaller pseudochromosome 57 (1.5 Mb) exhibited both near 0 f/m coverage and depth ratios (Fig. 1b). Thus, we can assume a XY sex determination system in the grey reef shark with 36 to be the X chromosome (Fig. 1c) and 57 the Y chromosome (Fig. 1d). While no telomere sequences were identified on pseudochromosome 36, both ends of pseudochromosome 57 entailed high numbers of AACCCT repeats (Table S3). The first 4.7 Mb of the X chromosome is likely a pseudoautosomal region (PAR), which was assembled from a singular contig and showed significantly higher heterozygosity of quality-filtered SNPs (not shown) compared with the rest of the X chromosome in males (*t*-test: $t_{291820} = -155.6$, $P < 2.2e{-}16$).

Materials and Methods

Library Preparation and Sequencing

PacBio and Omni-C genomic sequences were generated from dorsal fin tissue of two different male grey reef sharks,

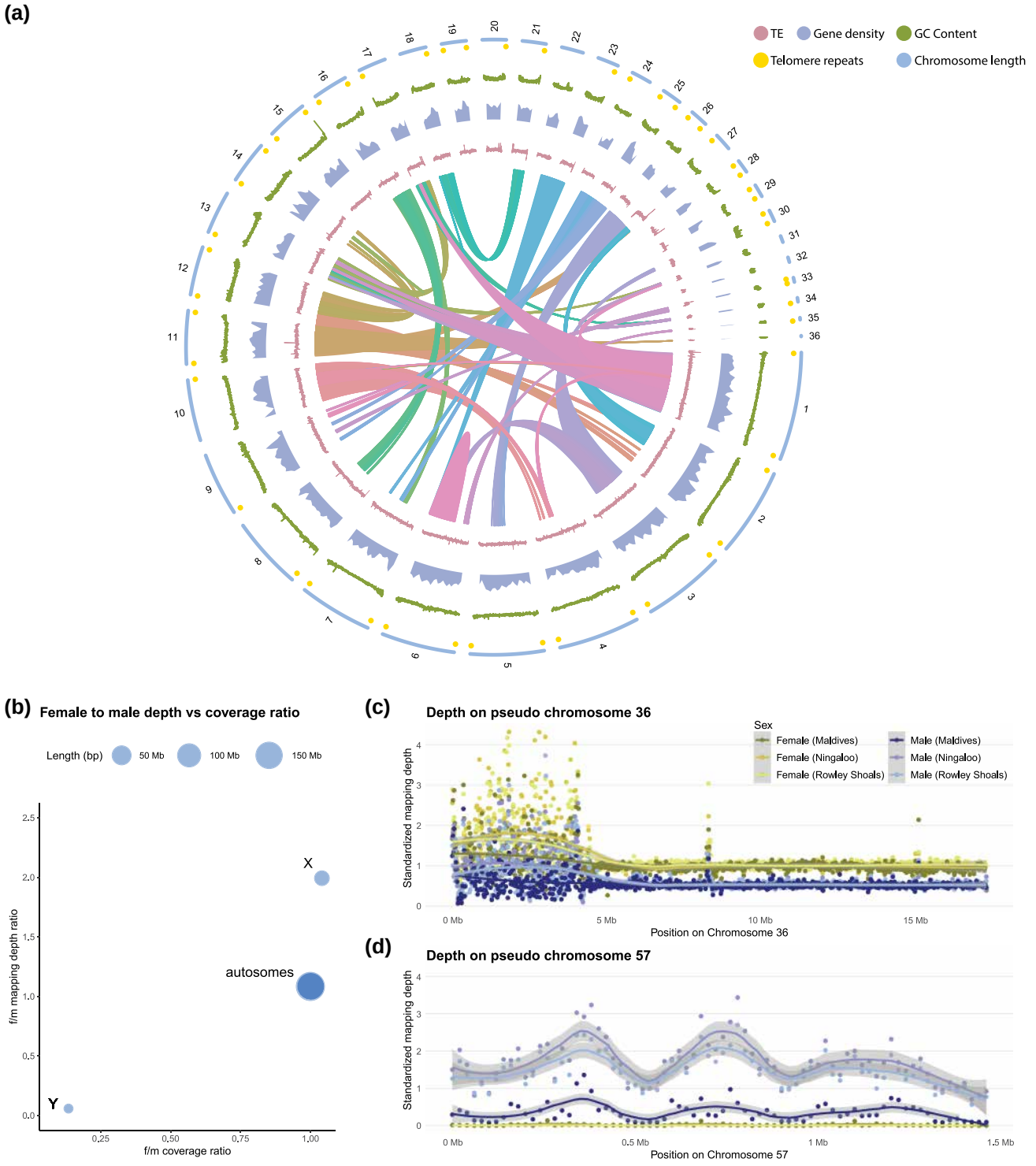

Fig. 1. Genome assembly summary and sex chromosome identification. (a) Genome assembly statistics, from the outer to inner circles: lengths of 36 pseudochromosomes are shown as numbered bars, dotted with locations of identified telomere repeats; line graphs of guanine-cytosine (GC) content in 100 kb nonoverlapping windows; gene density plots in 100 kb nonoverlapping windows; and transposable elements (TE). The inner circle shows genes with conserved order between pseudochromosomes. (b) Female to male sequencing coverage and mapping depth ratios for six males and females from the Maldives, Ningaloo, and Rowley Shoals, with near 0 coverage on the putative Y chromosome and two-fold depth ration on the X chromosome expected in a XY system. Mapping depths, standardized by whole genome depth in 20 kb nonoverlapping windows, are shown for the 18 individuals on pseudochromosome 36 (c) and pseudochromosome 57 (d).

Table 1 Grey reef shark assembly statistics and comparison to published assemblies of the closest species

Statistic		Primary contig	Scaffolded assembly	Annotation	Other assemblies
Number of reads (Mb)		7.9	511.7	…	79.3[a], 109.5[b], 340.4[d]
Coverage (X)		19	23.5	…	10[a], 24[b], 35[c], 30 to 45[d]
Contigs/scaffolds		2,276	1,520	…	96,185[a], 68,774[b], 239[c], 1,658[d]
Total contig/scaffold length (Gb)		2.9	2.9	…	2.29 to 2.58[a], 2.68[b], 3.23[c], 2.77[d]
N50 (Mb)		6.2	90.6	…	92.9[a], 93.5[b], 109.8[c], 89.8[d]
L50		126	12	…	11[a,b,c,d]
N90 (Mb)		0.75	17.3	…	-
L90		644	36	…	-
GC content (%)		43.98	43.98	…	42.5[a,b,c,d]
BUSCO scores (%)	Complete	93.2	93.6	86.5	94.2[d]
Single-copy	Single copy	91.1	91.4	67.7	-
Duplicate	Duplicate	2.1	2.2	18.8	-
Fragmented	Fragmented	2.9	2.6	2.3	-
Missing	Missing	3.9	3.8	11.2	-
Protein-coding genes		-	-	16,505	26,110[d]
Proportion with functional annotation (%)		-	-	88	-
Repeat content (%)		-	-	59	50[a], 52[d]

Statistics of the primary contigs, scaffolded assembly, and annotation are compared with available Carcharhinidae reference genomes: [a]*Negaprion brevirostris* (lemon shark; Baeza et al. 2024), [b]*Carcharhinus longimanus* (oceanic whitetip shark; Feldheim and Pirro 2023, unpublished), [c]*Prionace glauca* (great blue shark; Li, 2024, unpublished), and closely related [d]*Sphyrna mokarran* (great hammerhead shark; Stanhope et al. 2023).

one from New Caledonia, Noumea, and one from the Southern Maldives, respectively, as tissue from the first individual did not provide enough high molecular weight DNA to cover both sequencing library preparations. The New New Caledonian individual was caught in 2016 as per Boussarie et al. (2022) under a permit from the Southern Province of New Caledonia (permits no. 479-2016/ARR/DENV and no. 2093-2016/ARR/DENV); the individual from the Maldives was obtained by this study in the Gaafu Dhaalu Atoll in 2024 under a permit from the Ministry of Fisheries, Marine Resources and Agriculture of the Maldives (permit no. NRP2023/44). Genomic DNA was extracted from alcohol-preserved tissue samples using a salting out protocol as per Momigliano et al. 2017. PacBio library preparation and sequencing was performed by the Centre for PanorOmic Sciences (Hong Kong). Briefly, quantity and integrity of the DNA sample was confirmed by Qubit and Pulse field gel electrophoresis before being fragmented to appropriate sizes. DNA fragments were damage-repaired, end-repaired, and A-tailed. The SMRTbell library was produced by ligating universal hairpin adapters onto double-stranded DNA fragments. The library was checked with Qubit for quantification on a Femto Pulse System (Agilent) for size distribution detection. Quantified libraries were pooled and sequenced on PacBio Revio system in 3 Revio Cells amounting to around 4 million paired-end reads containing 96 Gb of raw data.

Omni-C libraries with 570 bp inserts were prepared with the Dovetail Omni-C kit and Omni-C proximity Ligation Assay (v.1.0, Dovetail Genomics, Scotts Valley, USA) according to the manufacturer's protocol and sequenced on an Illumina NovaSeq 6000 PE150 (Novogene, Hong Kong) as pair-end 2×150 bp runs, generating 153.5 Gb of raw data.

Genome Assembly

We assembled the grey reef shark genome following the Vertebrate Genomes Project (VGP) assembly pipeline (Larivière et al. 2024). Remnant adapter sequences from the PacBio HiFi dataset were removed with BBTools (Bushnell 2014; BBDuk with the parameters *ktrim = r k = 23 mink = 11 hdist = 1 tpe tbo*). Reads were assembled into haplotype assemblies of primary and alternate contigs using hifiasm (v.0.19.8; Cheng et al. 2021) with light purging (*-l 1 -k 21*). Duplications, repeats and contig overlaps were removed with purge dups (v.1.2.5; Guan et al. 2020) and minimap2 (v.2.26; Li 2018) using presets for HiFi data (-x map-hifi) from the primary assembly and added to the alternate assembly.

Scaffolding

FastQC confirmed high quality of fastq reads which were mapped separately with bwa mem (Li 2013) according to the Omni-C pipeline by Dovetail (https://omni-c.readthedocs.io/en/latest/index.html). Valid Omni-C ligation junctions were identified using pairtools (v. 1.1.0; Open2C et al. 2024), parsed with a minimal MAPQ of "5unique" for reporting unrescuable walks, and a maximum inter alignment gap of 30. Pairs were sorted in scaffold order, and marked duplicates removed from downstream analyses. The assembly genome was scaffolded with YaHS (v.1.2.; Zhou et al. 2023) with presets. Generated Omni-C contacts

with Juicer (v.1.9.9; Lander and Lieberman Aiden 2016) were visualized in JuiceBox and checked for major misassemblies. None of the contigs or scaffolds were contaminated with either adapter or foreign species sequences according to an NCBI Foreign Contamination Screen (FCS; Astashyn et al. 2024).

Telomere Identification

To identify telomeres, we used the TeloExplorer function of quarTeT (v1.2.5, Lin et al. 2023), which utilizes the "explore" and "search" functions of the Telomere Identification Toolkit (https://github.com/tolkit/telomeric-identifier) to identify repetitive telomere units and then search the unit string in the ends of pseudochromosomes with a minimum repeat number of 50.

Genome Annotation

De novo repeat element libraries were built with RepeatModeler (v.2.0.5; Flynn et al. 2020) and LTRStruct (v.2.9.5) for the scaffolded assembly (from contigs > 500 kb) and six other shark species (see Table S7 for accession numbers). The program repclassifier (v.1.1; https://github.com/darencard/GenomeAnnotation/blob/master/repclassifier) was run iteratively using newly annotated known elements from our reference assembly to classify more unknown elements in three rounds; one round with repeat families from the curated Dfam database (v.3.8; Storer et al. 2021) for ancestors of *C. amblyrhynchos* (72 elements), a second round against newly classified known elements (353 additional classified elements), a third round against known elements from a consensus library of the six related shark species, which added another 213 known elements.

The genome was serially annotated and masked with RepeatMasker (v.4.1.6, Smit et al. 2015) in four rounds: simple repeats (*-a -e ncbi -noint -xsmall*); known repeats from the curated *C. amblyrhynchos* Dfam database; *de novo* grey repeats from reference species identified by RepeatModeler; unknown grey reef shark specific repeats from RepeatModeler (*-a -e ncbi -nolow*). The combined repeat libraries were then used to mask the genome and repeat compositions from combined analysis of all RepeatMasker rounds were summarized with ProcessRepeats. For gene prediction, the parts of genome sequences detected as repeats are soft-masked with the options *-nolow -xsmall*.

Genes were predicted in BRAKER3 (version 3.0.8; Stanke et al. 2008) with GeneMark-ETP (Gabriel et al. 2024) using AUGUSTUS v.3.3.3 trained models with RNA and protein evidence. Since we did not perform RNA sequencing, we used published *C. amblyrchynhos* retina RNA-seq from NCBI (accession number SRR2146929). The protein evidence consisted of 374 Carcharhinidae proteins (NCBI search: (txid7805[organism:exp]) NOT mitochondrion), proteins from other shark species *Rhincodon typus* (36,827), *Chiloscyllium punctatum* (33,501), *Scyliorhinus torazame* (27,605), accessed from NCBI October 2024; the well annotated *Callorhinchus milii* (GenBank accession GCF_000165045.2) and the OrthoDB Vertebrata database (Kuznetsov et al. 2023). Genes were functionally annotated with AGAT (v.1.4.1; Dainat 2024) based on BLAST hits against the UnipotKB/SwissProt database (accessed February 2025) and filtered for a minimum length of 30 encoded amino acids, as per Weber et al. 2020, which removed 17 genes. Collinear gene blocks with conserved order between chromosomes were identified with MCScanX (v.1.0, Wang et al. 2012) using default parameters.

Genome Size Estimation and Quality Assessment

k-mer counts from the PacBio HiFi reads were generated with meryl for $k = 18,19,20,21,22,23,31$ (https://github.com/marbl/meryl). We then applied GenomeScope 2.0 (Ranallo-Benavidez et al. 2020) to the k-mer databases to estimate genome features including genome size, heterozygosity, and repeat content, with $k = 21$ producing the best assembly (Fig. S1). Genome quality and completeness were assessed with BUSCO (v.5.6.1; Manni et al. 2021) using the vertebrate and the actinopterygii ortholog databases, with the vertebrate database providing better results. We evaluated base level accuracy (QV) and k-mers using the previously generated meryl database and merqury (version 1.0; Rhie et al. 2020) generating spectral plots which confirmed successful removal of false duplicates (Fig. S6). Assembly statistics were calculated with gfastats (v1.3.6; Formenti et al. 2022) and QUAST (v.5.2.0; Mikheenko et al. 2023).

Identification of Sex Chromosomes

Sex chromosome identification was performed using whole genome sequences from six male and six female grey reef sharks from three populations, the Southern Maldives (Gaafu Dhaalu atoll, see previous methods), Ningaloo reef, and Rowley Shoals reef in Australia. The Australian samples were collected in 2013 and 2014 as described by Momigliano et al. (2015) under a permit from the Western Australia Department of Environment and Conservation (permit number: CE003632). DNA was extracted using a salting out protocol as per Momigliano et al. 2017 and sequenced at 10× by Novogene, Hong Kong SAR on a NovaSeq X Plus PE150 sequencing platform (see Table S6 for sample information). Reads were mapped to the reference assembly using bwa mem (v.0.7.17), samtools (v.1.16.1; Danecek et al. 2021), and GATK for indel realignment (v.3.8; McKenna et al. 2010). Read depths and coverages across chromosomes and average female-to-male ratios were calculated with bamdst (https://github.com/shiquan/bamdst) in 20 kb windows. A read depth ratio of 2 and coverage ratio of 1 was expected to

indicate the presence of a X chromosome, consistent with the hemizygous state of the X, while female-to-male coverage and depth ratios of near 0 would be expected for the Y chromosome. For the potential sex chromosomes read depths standardized by whole genome depth were plotted.

Supplementary Material

Supplementary material is available at *Genome Biology and Evolution* online.

Acknowledgments

We thank Paolo Franchini for his support with genome annotation and Sudeshna Chakraborty and Sean Law for assistance with DNA extractions and Omni-C library preparation. We would like to express our gratitude to Azim Musthag, Ahmed Mohamed, and Ahmed Shathir from the Small Island Research Group in the Maldives, as well as Ishmail Ashraf, William Robbins, Adam Stow, and Jarrah Stow for their invaluable support during fieldwork. This work was funded by the Hong Kong Research Grant Council Collaborative Research Fund (C4015-20EF) and the Hong Kong PhD Fellowship Scheme (HKPFS).

Data Availability

The final genome assembly has been deposited at DDBJ/ENA/GenBank under accession numbers GCA_965287045 and GCA_965295155. HiFi PacBio and Omni-C Illumina reads have been deposited under ENA accessions SAMEA 118362358 and SAMEA118362359, respectively. Whole genome sequences are deposited under study accession number PRJEB90057. The genome annotation is available at Zenodo (https://doi.org/10.5281/zenodo.15599239). Code to generate the scaffolded assembly can be found at https://github.com/carolindahms/Assembly-with-Omni-C.

Literature Cited

Allendorf FW, Luikart GH, Aitken SN. Conservation and the genetics of populations. 2nd ed Wiley-Blackwell; 2013.

Asahida T, Ida H. Karyological notes on four sharks in the order Carcharhiniformes. Jpn J Ichthyol. 1989:36:275–280. https://doi.org/10.11369/jji1950.36.275.

Astashyn A, et al. Rapid and sensitive detection of genome contamination at scale with FCS-GX. Genome Biol. 2024:25:60. https://doi.org/10.1186/s13059-024-03198-7.

Baeza JA, et al. Insights into the nuclear and mitochondrial genome of the lemon shark *Negaprion brevirostris* using low-coverage sequencing: genome size, repetitive elements, mitochondrial genome, and phylogenetic placement. Gene. 2024:894:147939. https://doi.org/10.1016/j.gene.2024.147939.

Barley SC, Clark TD, Meeuwig JJ. Ecological redundancy between coral reef sharks and predatory teleosts. Rev Fish Biol Fish. 2020:30:153–172. https://doi.org/10.1007/s11160-019-09588-6.

Bernard AM, Finnegan KA, Pavinski Bitar P, Stanhope MJ, Shivji MS. Genomic assessment of global population structure in a highly migratory and habitat versatile apex predator, the tiger shark (*Galeocerdo cuvier*). J Hered. 2021:112:497–507. https://doi.org/10.1093/jhered/esab046.

Bonnin L, et al. Repeated long-range migrations of adult males in a common Indo-Pacific reef shark. Coral Reefs. 2019:38:1121–1132. https://doi.org/10.1007/s00338-019-01858-w.

Boussarie G, et al. Identifying barriers to gene flow and hierarchical conservation units from seascape genomics: a modelling framework applied to a marine predator. Ecography. 2022:2022:e06158. https://doi.org/10.1111/ecog.06158.

Brown MR, Gonzalez de la Rosa PM, Blaxter M. Tidk: a toolkit to rapidly identify telomeric repeats from genomic dataset. Bioinformatics. 2025:41, btaf049. https://doi.org/10.1093/bioinformatics/btaf049.

Bushnell B. BBMap. [Software]. Version 38.90. SourceForge; 2014. https://sourceforge.net/projects/bbmap/.

Cheng H, Concepcion GT, Feng X, Zhang H, Li H. Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. Nat Methods. 2021:18:170–175. https://doi.org/10.1038/s41592-020-01056-5.

Cortés E. Life history patterns and correlations in sharks. Rev Fish Sci. 2000:8:299–344. https://doi.org/10.1080/10408340308951115.

Dainat J. AGAT: Another GFF Analysis Toolkit to handle annotations in any GTF/GFF format; 2024. Version v1.4.1. Zenodo. https://doi.org/10.5281/zenodo.3552717.

Danecek P, et al. Twelve years of SAMtools and BCFtools. GigaScience. 2021:10:giab008. https://doi.org/10.1093/gigascience/giab008.

De Souza Valentim FC, Porto JIR, Bertollo LAC, Gross MC, Feldberg E. XX/XO, a rare sex chromosome system in Potamotrygon freshwater stingray from the Amazon Basin, Brazil. Genetica. 2013:141:381–387. https://doi.org/10.1007/s10709-013-9737-2.

Domingues RR, Hilsdorf AWS, Gadig OBF. The importance of considering genetic diversity in shark and ray conservation policies. Conserv Genet. 2018:19:501–525. https://doi.org/10.1007/s10592-017-1038-3.

Dudchenko O, et al. De novo assembly of the *Aedes aegypti* genome using Hi-C yields chromosome-length scaffolds. Science. 2017:356:92–95. https://doi.org/10.1126/science.aal3327.

Dulvy NK, et al. Extinction risk and conservation of the world's sharks and rays. Elife. 2014:3:e00590. https://doi.org/10.7554/eLife.00590.

Dulvy NK, et al. Overfishing drives over one-third of all sharks and rays toward a global extinction crisis. Curr Biol. 2021:31:4773–-4787.e8. https://doi.org/10.1016/j.cub.2021.08.062.

Feldheim K, Pirro S. Carcharhinus longimanus voucher IRGN:5ZTL7XCE87 (JASCQW000000000.1) [Unpublished genome]; 2023. NCBI. https://www.ncbi.nlm.nih.gov/nuccore/JASCQW000000000.1.

Flynn JM, et al. RepeatModeler2 for automated genomic discovery of transposable element families. Proc Natl Acad Sci U S A. 2020:117:9451–9457. https://doi.org/10.1073/pnas.1921047117.

Formenti G, et al. Gfastats: conversion, evaluation and manipulation of genome sequences using assembly graphs. Bioinformatics. 2022:38:4214–4216. https://doi.org/10.1093/bioinformatics/btac460.

Frankham R, Briscoe DA, Ballou JD. Introduction to conservation genetics. Cambridge University Press; 2002.

Friedlander AM, et al. The real bounty: marine biodiversity in the Pitcairn Islands. PLoS One. 2014:9:e100142. https://doi.org/10.1371/journal.pone.0100142.

Gabriel L, et al. BRAKER3: fully automated genome annotation using RNA-seq and protein evidence with GeneMark-ETP, AUGUSTUS, and TSEBRA. Genome Res. 2024:34:769–777. https://doi.org/10.1101/gr.278090.123.

Graham NAJ, Spalding MD, Sheppard CRC. Reef shark declines in remote atolls highlight the need for multi-faceted conservation

action. Aquat Conserv Mar Freshw Ecosyst. 2010:20:543–548. https://doi.org/10.1002/aqc.1116.
Guan D, et al. Identifying and removing haplotypic duplication in primary genome assemblies. Bioinformatics. 2020:36:2896–2898. https://doi.org/10.1093/bioinformatics/btaa157.
Hara Y, et al. Shark genomes provide insights into elasmobranch evolution and the origin of vertebrates. Nat Ecol Evol. 2018:2:1761–1771. https://doi.org/10.1038/s41559-018-0673-5.
Heupel MR, Simpfendorfer CA, Fitzpatrick R. Large-scale movement and reef fidelity of grey reef sharks. PLoS One. 2010:5:e9650. https://doi.org/10.1371/journal.pone.0009650.
Kuraku S. Shark and ray genomics for disentangling their morphological diversity and vertebrate evolution. Dev Biol. 2021:477:262–272. https://doi.org/10.1016/j.ydbio.2021.06.001.
Kuznetsov D, et al. OrthoDB v11: annotation of orthologs in the widest sampling of organismal diversity. Nucleic Acids Res. 2023:51:D445–D451. https://doi.org/10.1093/nar/gkac998.
Lander E, Lieberman Aiden E. Juicer provides a one-click system for analyzing loop-resolution Hi-C experiments. Cell Syst. 2016:3:1–8. https://doi.org/10.1016/j.cels.2016.07.002.
Lang D, et al. Comparison of the two up-to-date sequencing technologies for genome assembly: HiFi reads of Pacific Biosciences Sequel II system and ultralong reads of Oxford Nanopore. GigaScience. 2020:9. https://doi.org/10.1093/gigascience/giaa123.
Larivière D, et al. Scalable, accessible and reproducible reference genome assembly and evaluation in Galaxy. Nat Biotechnol. 2024:42:367–370. https://doi.org/10.1038/s41587-023-02100-3.
Lee S-H, et al. Insights into the Evolution of Ancient Shark and Ray Sex Chromosomes. bioRxiv. 2025. https://doi.org/10.1101/2025.02.26.637739.
Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv preprint. 2013. https://arxiv.org/abs/1303.3997.
Li H. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics. 2018:34:3094–3100. https://doi.org/10.1093/bioinformatics/bty191.
Li YY. Genome assembly ASM3797433v1 (*Prionace glauca*); 2024. National Center for Biotechnology Information (NCBI). Available from: https://www.ncbi.nlm.nih.gov/datasets/genome/GCA_037974335.1/. Accessed July 26, 2025.
Lin Y, et al. Quartet: a telomere-to-telomere toolkit for gap-free genome assembly and centromeric repeat identification. Hortic Res 2023:10:uhad127. https://doi.org/10.1093/hr/uhad127.
Lou RN, Jacobs A, Wilder AP, Therkildsen NO. A beginner's guide to low-coverage whole genome sequencing for population genomics. Mol Ecol. 2021:30:5966–5993. https://doi.org/10.1111/mec.16077.
Manni M, Berkeley MR, Seppey M, Zdobnov EM. BUSCO: Assessing Genomic Data Quality and Beyond. Curr Protoc. 2021:1. https://doi.org/10.1002/cpz1.323.
Marra NJ, et al. White shark genome reveals ancient elasmobranch adaptations associated with wound healing and the maintenance of genome stability. Proc Natl Acad Sci U S A. 2019:116:4446–4455. https://doi.org/10.1073/pnas.1819778116.
Mayeur H, et al. The sensory shark: high-quality morphological, genomic and transcriptomic data for the small-spotted catshark *Scyliorhinus canicula* reveal the molecular bases of sensory organ evolution in jawed vertebrates. Mol Biol Evol. 2024:41:msae246. https://doi.org/10.1093/molbev/msae246.
Mayeur H, et al. The Sensory Shark: High-quality Morphological, Genomic and Transcriptomic Data for the Small-spotted Catshark *Scyliorhinus Canicula* Reveal the Molecular Bases of Sensory Organ Evolution in Jawed Vertebrates. Mol Biol Evol. 2024:41. https://doi.org/10.1093/molbev/msae246.
McKenna A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010:20:1297–1303. https://doi.org/10.1101/gr.107524.110.
Mikheenko A, Saveliev V, Hirsch P, Gurevich A. WebQUAST: online evaluation of genome assemblies. Nucleic Acids Res. 2023:51:W601–W606. https://doi.org/10.1093/nar/gkad406.
Momigliano P, et al. Genetic structure and signatures of selection in grey reef sharks (*Carcharhinus amblyrhynchos*). Heredity. 2017:119:142–153. https://doi.org/10.1038/hdy.2017.21.
Momigliano P, Harcourt R, Robbins WD, Stow A. Connectivity in grey reef sharks (*Carcharhinus amblyrhynchos*) determined using empirical and simulated genetic data. Sci Rep. 2015:5:13229. https://doi.org/10.1038/srep13229.
Nakaya K, White WT, Ho HC. Discovery of a new mode of oviparous reproduction in sharks and its evolutionary implications. Sci Rep. 2020:10:1–12. https://doi.org/10.1038/s41598-020-68923-1.
Niwa T, et al. Sharks and rays have the oldest vertebrate sex chromosome with unique sex determination mechanisms. Proc Natl Acad Sci U S A. 2025:122:e2513676122. https://doi.org/10.1073/pnas.2513676122.
Oleksiak MF, Rajora OP. Marine population genomics: challenges and opportunities. Population genomics, Vol. 2020. Springer; 2020. p. 3–5
Open2C AN, et al. Pairtools: from sequencing data to chromosome contacts. PLoS Comput Biol. 2024:20:e1012164. https://doi.org/10.1371/journal.pcbi.1012164.
Ovenden JR, Morgan JA, Kashiwagi T, Broderick D, Salini J. Towards better management of Australia's shark fishery: genetic analyses reveal unexpected ratios of cryptic blacktip species *Carcharhinus tilstoni* and *C. limbatus*. Mar Freshw Res. 2010:61:253–262. https://doi.org/10.1071/MF09151.
Pacoureau N, et al. Half a century of global decline in oceanic sharks and rays. Nature. 2021:589:567–571. https://doi.org/10.1038/s41586-020-03173-9.
Papastamatiou YP, et al. Activity seascapes highlight central place foraging strategies in marine predators that never stop swimming. Mov Ecol. 2018:6:1–15. https://doi.org/10.1186/s40462-018-0127-3.
Pearce J, Fraser MW, Sequeira AMM, Kaur P. State of shark and ray genomics in an era of extinction. Front Mar Sci. 2021:8. https://doi.org/10.3389/fmars.2021.744986.
Ranallo-Benavidez TR, Jaron KS, Schatz MC. GenomeScope 2.0 and Smudgeplot for reference-free profiling of polyploid genomes. Nat Commun. 2020:11. https://doi.org/10.1038/s41467-020-14998-3.
Rhie A, Walenz BP, Koren S, Phillippy AM. Merqury: reference-free quality, completeness, and phasing assessment for genome assemblies. Genome Biol. 2020:21. https://doi.org/10.1186/s13059-020-02134-9.
Robbins WD. 2006. Abundance, demography and population structure of the grey reef shark (*Carcharhinus amblyrhynchos*) and the white tip reef shark (*Triaenodon obesus*) (Fam. Charcharhinidae) [Doctoral dissertation]. Australia: James Cook University.
Schwartz FJ, Maddock MB. Cytogenetics of the elasmobranchs: genome evolution and phylogenetic implications. Mar Freshw Res. 2002:53:491–502. https://doi.org/10.1071/MF01139.
Shafer ABA, et al. Genomics and the challenging translation into conservation practice. Trends Ecol Evol. 2015:30:78–87. https://doi.org/10.1016/j.tree.2014.11.009.
Sherman CS, et al. Half a century of rising extinction risk of coral reef sharks and rays. Nat Commun. 2023:14:15. https://doi.org/10.1038/s41467-022-35091-x
Smit AFA, Hubley R, Green P. 2015. RepeatMasker Open-4.0. http://www.repeatmasker.org (Accessed June 2024).
Stanhope MJ, et al. Genomes of endangered great hammerhead and shortfin mako sharks reveal historic population declines and high

levels of inbreeding in great hammerhead. iScience. 2023:26, 105815. https://doi.org/10.1016/j.isci.2023.105106.
Stanke M, Diekhans M, Baertsch R, Haussler D. Using native and syntenically mapped cDNA alignments to improve de novo gene finding. Bioinformatics. 2008:24:637–644. https://doi.org/10.1093/bioinformatics/btn013.
Stevens JD, Bonfil R, Dulvy NK, Walker P. The effects of fishing on sharks, rays, and chimaeras (chondrichthyans), and the implications for marine ecosystems. ICES J Mar Sci. 2000:57:476–494. https://doi.org/10.1006/jmsc.2000.0724.
Stingo V, Rocco L. Selachian cytogenetics: a review. Genetica. 2001:111:329–347. https://doi.org/10.1023/A:1013747215866.
Storer J, Hubley R, Rosen J, Wheeler TJ, Smit AF. The Dfam community resource of transposable element families, sequence models, and genome annotations. Mob DNA. 2021:12, 2. https://doi.org/10.1186/s13100-020-00203-8.
Sugahara F, Murakami Y, Pascual-Anaya J, Kuratani S. Reconstructing the ancestral vertebrate brain. Dev Growth Differ. 2017:59:163–174. https://doi.org/10.1111/dgd.12347.
Tan M, et al. The whale shark genome reveals patterns of vertebrate gene family evolution. eLife. 2021:10. https://doi.org/10.7554/eLife.65394.
Uno Y, et al. Cell culture-based karyotyping of orectolobiform sharks for chromosome-scale genome analysis. Commun Biol. 2020:3:652. https://doi.org/10.1038/s42003-020-01373-7.
Wagner CI, et al. Characteristics of the spiny dogfish (*Squalus acanthias*) nuclear genome. G3. 2023:13:jkad146. https://doi.org/10.1093/g3journal/jkad146.
Wang Y, et al. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucleic Acids Res. 2012:40:e49. https://doi.org/10.1093/nar/gkr1293.
Waples RS. Evolutionary significant units and the conservation of biological diversity under the endangered species act. Am Fish Soc Symp. 1995:17:8–27. https://www.webapps.nwfsc.noaa.gov/assets/4/6878_09172014_172219_Waples.1995.pdf (Accessed May 2025).
Weber JA, et al. The whale shark genome reveals how genomic and physiological properties scale with body size. Proc Natl Acad Sci U S A. 2020:117:20662–20671. https://doi.org/10.1073/pnas.1922576117.
Yamaguchi K, et al. Elasmobranch genome sequencing reveals evolutionary trends of vertebrate karyotype organization. Genome Res. 2023:33:1527–1540. https://doi.org/10.1101/gr.276840.122.
Yu WP, et al. Elephant shark sequence reveals unique insights into the evolutionary history of vertebrate genes: a comparative analysis of the protocadherin cluster. Proc Natl Acad Sci U S A.2008:105:3819–3824. https://doi.org/10.1073/pnas.0800398105.
Zhang Y, et al. The white-spotted bamboo shark genome reveals chromosome rearrangements and fast-evolving immune genes of cartilaginous fish. iScience. 2020:23, 101754. https://doi.org/10.1016/j.isci.2020.101754.
Zhou C, McCarthy SA, Durbin R. YaHS: yet another Hi-C scaffolding tool. Bioinformatics. 2023:39:btac808. https://doi.org/10.1093/bioinformatics/btac808.

Associate editor: Diego Cortez