DATA NOTE

The chromosomal genome sequence of the Florida corallimorph, *Ricordea florida* (Duchassaing & Michelotti, 1860) (Corallimorpharia: Ricordeidae)

[version 1; peer review: awaiting peer review]

Anthony J. Bellantuono[1], Mónica Medina[1], Viridiana Avila-Magaña[2], Bishoy Kamel[3], Jose V. Lopez[4], Nina Pruzinsky[5,6], Graeme Oatley[7], Elizabeth Sinclair[7], Eerik Aunin[7], Noah Gettle[7], Camilla Santos[7], Michael Paulini[7], Haoyu Niu[7], Victoria McKenna[7], Rebecca O'Brien[7], Wellcome Sanger Institute Tree of Life Management, Samples and Laboratory team, Wellcome Sanger Institute Scientific Operations: Sequencing Operations, Wellcome Sanger Institute Tree of Life Core Informatics team, EBI Aquatic Symbiosis Genomics Data Portal Team, Aquatic Symbiosis Genomics Project Leadership

[1]Department of Ecology and Evolutionary Biology, University of California Los Angeles, Los Angeles, California, USA
[2]Carnegie Science, Pasadena, California, USA
[3]Lawrence Berkeley National Laboratory, Joint Genome Institute, Berkeley, California, USA
[4]National Coral Reef Institute, Department of Biological Sciences, Nova Southeastern University, Dania Beach, Florida, USA
[5]University Corporation for Atmospheric Research (UCAR), Boulder, Colorado, USA
[6]NOAA Ocean Exploration Affiliate, Silver Spring, Maryland, USA
[7]Tree of Life Programme, Wellcome Sanger Institute, Hinxton, England, UK

v1 **First published:** 18 Aug 2026, **11**:529
https://doi.org/10.12688/wellcomeopenres.27190.1
Latest published: 18 Aug 2026, **11**:529
https://doi.org/10.12688/wellcomeopenres.27190.1

Abstract

We present a genome assembly from an individual *Ricordea florida* (Florida corallimorph; Cnidaria; Anthozoa; Corallimorpharia; Ricordeidae). The genome sequence has a total length of 537.84 megabases. Most of the assembly (99.81%) is scaffolded into 16 chromosomal pseudomolecules. The mitochondrial genome has also been assembled, with a length of 21.38 kilobases.

Keywords

Ricordea florida, Florida corallimorph, genome sequence, chromosomal, Corallimorpharia, microbial metagenome assembly

Open Peer Review

Approval Status *AWAITING PEER REVIEW*

Any reports and responses or comments on the article can be found at the end of the article.

This article is included in the Tree of Life gateway.

Corresponding author: Aquatic Symbiosis Genomics Project Leadership (mark.blaxter@sanger.ac.uk)

Author roles: Bellantuono AJ: Investigation, Resources, Writing – Original Draft Preparation; **Medina M**: Investigation, Resources, Writing – Original Draft Preparation; **Avila-Magaña V**: Writing – Original Draft Preparation; **Kamel B**: Writing – Original Draft Preparation; **Lopez JV**: Resources; **Pruzinsky N**: Resources; **Oatley G**: Investigation; **Sinclair E**: Investigation; **Aunin E**: Formal Analysis; **Gettle N**: Formal Analysis; **Santos C**: Formal Analysis; **Paulini M**: Formal Analysis; **Niu H**: Project Administration; **McKenna V**: Project Administration; **O'Brien R**: Project Administration;

Competing interests: No competing interests were disclosed.

Grant information: This work was funded by the Gordon and Betty Moore Foundation through a grant (GBMF8897) to the Wellcome Sanger Institute to support the Aquatic Symbiosis Genomics Project, and by Wellcome through core funding to the Wellcome Sanger Institute (220540). We acknowledge funding from NSF-IOS 2227068 for sample collection.
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

How to cite this article: Bellantuono AJ, Medina M, Avila-Magaña V *et al.* **The chromosomal genome sequence of the Florida corallimorph, *Ricordea florida* (Duchassaing & Michelotti, 1860) (Corallimorpharia: Ricordeidae) [version 1; peer review: awaiting peer review]** Wellcome Open Research 2026, **11**:529 https://doi.org/10.12688/wellcomeopenres.27190.1

First published: 18 Aug 2026, **11**:529 https://doi.org/10.12688/wellcomeopenres.27190.1

Species taxonomy

Eukaryota; Opisthokonta; Metazoa; Eumetazoa; Cnidaria; Anthozoa; Hexacorallia; Corallimorpharia; Ricordeidae; *Ricordea*; *Ricordea florida* (Duchassaing & Michelotti, 1860) (NCBI:txid165100).

Background

Ricordea florida (Cnidaria: Anthozoa: Corallimorpharia) is commonly known as the Florida False Coral, Florida Corallimorph or Mushroom Coral. *Ricordea florida* is distributed throughout the Caribbean, with reports as far north in the Atlantic as Bermuda (GBIF Secretariat, 2026), and inhabits shallow waters, with a maximum reported depth of 54 m (den Hartog, 1980). Corallimorpharians, though closely related to stony corals, lack the ability to form calcium carbonate skeletons (den Hartog, 1980). Corallimorpharians exist as solitary polyps or colonial aggregates that live attached to substrates such as rocks or corals (Muhando *et al.*, 2002; Parr, 2019; Torres-Pratts *et al.*, 2011), and are represented by at least six species in the Caribbean, with *R. florida* being one of the most commonly encountered (den Hartog, 1980). They have a simple body structure consisting of a disk-shaped base and a central mouth surrounded by tentacles. The tentacles are used for capturing prey and defending the polyp. *R. florida*, in particular, is known for its vibrant colours, which can range from bright greens to oranges and purples, making it an attractive aquarium trade organism. The corallimorpharian family Ricordeidae comprises just two species, *R. florida* and *R. yuma*, though they are geographically separated, with *R. yuma* located in the Indo-Pacific (Fautin, 2016).

R. florida has a simple digestive system and relies on both filter-feeding and a symbiotic relationship with photosynthetic intracellular algae (Dinophyceae: Symbiodiniaceae) for nutrition (den Hartog, 1980). The symbiosis is facilitated by specialised cells within the coral's gastrodermis that house the algal cells. These cells provide a stable environment for the algae to thrive, and in return, the algae produce sugars and other organic compounds that the coral can utilise (Lin *et al.*, 2019). This symbiotic relationship is important for host survival in nutrient-poor waters (Davy *et al.*, 2012).

As the non-calcifying sister group of scleractinian corals that establishes symbiosis with the same algal family, corallimorphs are a critical taxon in understanding the evolution of threatened reef-building corals. Importantly, this lineage pre-existed the evolutionary acquisition of calcification in scleractinian corals, making the genome of *R. florida* a crucial resource for investigating the genetic and evolutionary basis of calcification using comparative genomics (Lin *et al.*, 2016). The unique taxonomic position of corallimorphs, combined with the fact that they express most of the genes known to be involved in carbonate skeleton biogenesis in corals, raises questions about whether corallimorpharians lost the capacity to form a skeleton, and how mineralisation pathways are regulated in these non-calcifying species (Lin *et al.*, 2017; Medina *et al.*, 2006).

Shallow, warm-water dwelling species of corallimorphs including *Ricordea* generally host endosymbiotic microalgal Symbiodiniaceae (den Hartog, 1980; LaJeunesse, 2002), while the deeper and cold-water species do not engage in photosymbiosis (Fautin *et al.*, 2002). This provides additional opportunity for comparative studies across taxa, allowing comparison of the genetic toolkits in symbiotic and non-symbiotic corallimorphs.

Like many cnidarians, *R. florida* exhibits extensive capacity for regeneration and has been examined using proteomics (Horricks *et al.*, 2023); genomic resources expand the potential to understand the molecular mechanisms underlying regeneration, symbiosis, and stress responses. Prior population genetic examination of *R. florida* by Torres-Pratts *et al.* (2011) identified two distinct, geographically overlapping genetic lineages within the species, potentially representing cryptic species. The availability of the *R. florida* genome provides a valuable resource for further studies into the genetic basis of these lineages, their evolutionary significance, and implications for conservation.

Little is known regarding the ecology of *R. florida*, but one study identified these corallimorphs as the dominant prey species of hawksbill turtles in the Dominican Republic (León & Bjorndal, 2002).

We present a chromosome-level genome sequence for *Ricordea florida*, generated via the Tree of Life pipeline using a specimen collected from the Florida Keys, Florida, USA (Figure 1). This assembly was generated as part of the Aquatic Symbiosis Genomics project.

Methods

Sample acquisition

The specimen used for genome sequencing was an adult *Ricordea florida* (specimen ID NSU0021801, ToLID jaRicFlor1; Figure 1). Another specimen was used for RNA sequencing (specimen ID NSU0021803, ToLID jaRicFlor3). This samples were acquired from Colin Foord (Coral Morphologic, Miami, Florida), with the sample originally collected in the middle Florida Keys.

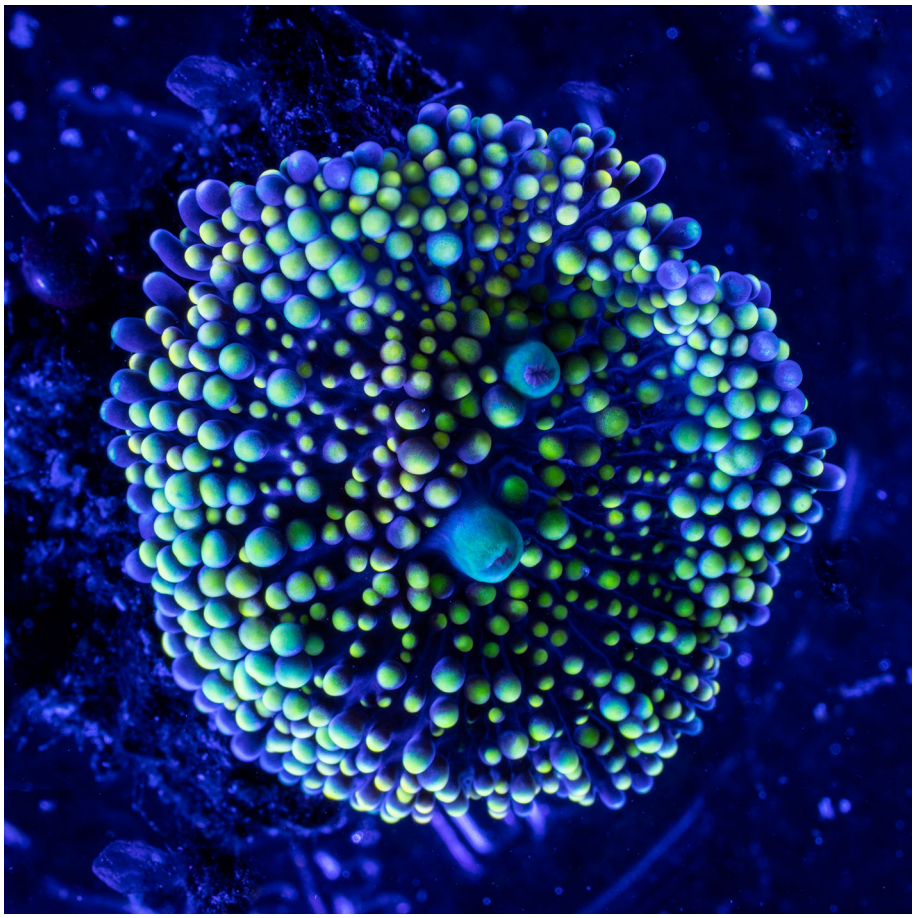

Figure 1. Photograph of the *Ricordea florida* (jaRicFlor1) specimen used for genome sequencing (approx. 2 cm diameter).

Nucleic acid extraction

Protocols for high molecular weight (HMW) DNA extraction developed at the Wellcome Sanger Institute (WSI) Tree of Life Core Laboratory are available on protocols.io (Howard *et al.*, 2025). The jaRicFlor1 sample was weighed and triaged to determine the appropriate extraction protocol. Tissue was homogenised by cryogenic disruption using the Covaris cryoPREP® Automated Dry Pulverizer. HMW DNA was extracted using the Manual MagAttract protocol. DNA was sheared into an average fragment size of 12–20 kb following the Megaruptor®3 for LI PacBio protocol. Sheared DNA was purified by manual SPRI (solid-phase reversible immobilisation). The concentration of the sheared and purified DNA was assessed using a Nanodrop spectrophotometer and Qubit Fluorometer using the Qubit dsDNA High Sensitivity Assay kit. Fragment size distribution was evaluated by running the sample on the FemtoPulse system.

RNA was extracted from tissue of jaRicFlor3 in the Tree of Life Laboratory at the WSI using the RNA Extraction: Automated MagMax™ *mir*Vana protocol. The RNA concentration was assessed using a Nanodrop spectrophotometer and a Qubit Fluorometer using the Qubit RNA Broad-Range Assay kit. Analysis of the integrity of the RNA was done using the Agilent RNA 6000 Pico Kit and Eukaryotic Total RNA assay.

PacBio HiFi library preparation and sequencing

Library preparation and sequencing were performed at the WSI Scientific Operations core. Libraries were prepared using the SMRTbell Prep Kit 3.0 (Pacific Biosciences, California, USA), following the manufacturer's instructions. The kit includes reagents for end repair/A-tailing, adapter ligation, post-ligation SMRTbell bead clean-up, and nuclease treatment. Size selection and clean-up were performed using diluted AMPure PB beads (Pacific Biosciences). DNA concentration was quantified using a Qubit Fluorometer v4.0 (ThermoFisher Scientific) and the Qubit 1X dsDNA HS assay kit. Final library fragment size was assessed with the Agilent Femto Pulse Automated Pulsed Field CE Instrument (Agilent Technologies) using the gDNA 55 kb BAC analysis kit.

The sample was sequenced using the Sequel IIe system (Pacific Biosciences, California, USA). The concentration of the library loaded onto the Sequel IIe was in the range 40–135 pM. The SMRT link software, a PacBio web-based end-to-end workflow manager, was used to set-up and monitor the run, and to perform primary and secondary analysis of the data upon completion.

Hi-C

Sample preparation and crosslinking

The Hi-C sample was prepared from 20–50 mg of frozen tissue from the jaRicFlor1 sample using the Arima-HiC v2 kit (Arima Genomics). Following the manufacturer's instructions, tissue was fixed and DNA crosslinked using TC buffer to a final formaldehyde concentration of 2%. The tissue was homogenised using the Diagnocine Power Masher-II. Crosslinked DNA was digested with a restriction enzyme master mix, biotinylated, and ligated. Clean-up was performed with SPRISelect beads before library preparation. DNA concentration was measured with the Qubit Fluorometer (Thermo Fisher Scientific) and Qubit HS Assay Kit. The biotinylation percentage was estimated using the Arima-HiC v2 QC beads.

Hi-C library preparation and sequencing

Biotinylated DNA constructs were fragmented using a Covaris E220 sonicator and size selected to 400–600 bp using SPRISelect beads. DNA was enriched with Arima-HiC v2 kit Enrichment beads. End repair, A-tailing, and adapter ligation were carried out with the NEBNext Ultra II DNA Library Prep Kit (New England Biolabs), following a modified protocol where library preparation occurs while DNA remains bound to the Enrichment beads. Library amplification was performed using KAPA HiFi HotStart mix and a custom Unique Dual Index (UDI) barcode set (Integrated DNA Technologies). Depending on sample concentration and biotinylation percentage determined at the crosslinking stage, libraries were amplified with 10–16 PCR cycles. Post-PCR clean-up was performed with SPRISelect beads. Libraries were quantified using the AccuClear Ultra High Sensitivity dsDNA Standards Assay Kit (Biotium) and a FLUOstar Omega plate reader (BMG Labtech).

Prior to sequencing, libraries were normalised to 10 ng/μL. Normalised libraries were quantified again to create equimolar and/or weighted 2.8 nM pools. Pool concentrations were checked using the Agilent 4200 TapeStation (Agilent) with High Sensitivity D500 reagents before sequencing. Sequencing was performed using paired-end 150 bp reads on the Illumina NovaSeq 6000.

RNA library preparation and sequencing

Libraries were prepared using the NEBNext® Ultra™ II Directional RNA Library Prep Kit for Illumina (New England Biolabs), following the manufacturer's instructions. Poly(A) mRNA in the total RNA solution was isolated using oligo (dT) beads, converted to cDNA, and uniquely indexed; 14 PCR cycles were performed. Libraries were size-selected to produce fragments between 100–300 bp. Libraries were quantified, normalised, pooled to a final concentration of 2.8 nM, and diluted to 150 pM for loading. Sequencing was carried out on the Illumina NovaSeq 6000, generating paired-end reads.

Genome assembly

Prior to assembly of the PacBio HiFi reads, a database of *k*-mer counts ($k = 31$) was generated from the filtered reads using FastK. GenomeScope2 (Ranallo-Benavidez *et al.*, 2020) was used to analyse the *k*-mer frequency distributions, providing estimates of genome size, heterozygosity, and repeat content.

The HiFi reads were assembled using Hifiasm (Cheng *et al.*, 2021) with the --primary option. Haplotypic duplications were identified and removed using purge_dups (Guan *et al.*, 2020). The Hi-C reads (Rao *et al.*, 2014) were mapped to the primary contigs using bwa-mem2 (Vasimuddin *et al.*, 2019), and the contigs were scaffolded in YaHS (Zhou *et al.*, 2023) with the --break option for handling potential misassemblies. The scaffolded assemblies were evaluated using Gfastats (Formenti *et al.*, 2022), BUSCO (Manni *et al.*, 2021) and MERQURY.FK (Rhie *et al.*, 2020).

The mitochondrial genome was assembled using MitoHiFi (Uliano-Silva *et al.*, 2023).

Assembly curation

The assembly was decontaminated using the Assembly Screen for Cobionts and Contaminants (ASCC) pipeline. TreeVal was used to generate the flat files and maps for use in curation. Manual curation was conducted primarily in PretextView and HiGlass (Kerpedjiev *et al.*, 2018). Scaffolds were visually inspected and corrected as described by Howe *et al.*

(2021). Manual corrections included 56 breaks, 76 joins, and removal of 45 haplotypic duplications. This reduced the scaffold count by 54.2%, increased the scaffold N50 by 4.3%, and reduced the total assembly length by 4.6%. The curation process is described at sanger-tol/curation-resources. PretextSnapshot was used to generate a Hi-C contact map of the final assembly.

Assembly quality assessment

The Merqury.FK tool (Rhie *et al.*, 2020) was run in a Singularity container (Kurtzer *et al.*, 2017) to evaluate *k*-mer completeness and assembly quality for the primary and alternate haplotypes using the *k*-mer databases ($k = 31$) computed prior to genome assembly. The analysis outputs included assembly QV scores and completeness statistics.

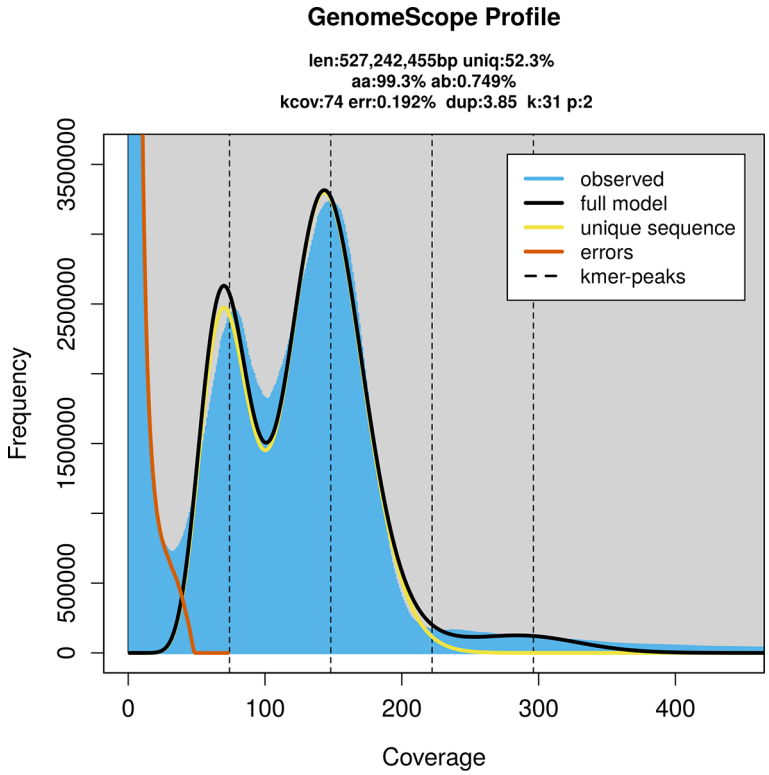

Figure 2. Frequency distribution of *k*-mers generated using GenomeScope2. The plot shows observed and modelled *k*-mer spectra, providing estimates of genome size, heterozygosity, and repeat content based on unassembled sequencing reads.

Table 1. Specimen and sequencing data for *Ricordea florida* (BioProject PRJEB52859).

Platform	PacBio HiFi	Hi-C	RNA-seq
ToLID	jaRicFlor1	jaRicFlor1	jaRicFlor3
Specimen ID	NSU0021801	NSU0021801	NSU0021803
BioSample (source individual)	SAMEA9196950	SAMEA9196950	SAMEA9196952
BioSample (tissue)	SAMEA9197049	SAMEA9197048	SAMEA9197055
Instrument	Sequel II	Illumina NovaSeq 6000	Illumina NovaSeq 6000
Run accessions	ERR9742064; ERR9744410	ERR9767806	ERR10123701
Read count total	5.98 million reads	450.60 million read pairs	22.68 million read pairs
Base count total	78.80 Gb	136.08 Gb	6.85 Gb

Table 2. **Genome assembly data for *Ricordea florida*.**

Genome assembly	Primary assembly
Assembly name	jaRicFlor1.1
Assembly accession	GCA_949710005.1
Alternate haplotype accession	GCA_949710055.1
Assembly level	chromosome
Span (Mb)	537.84
Number of chromosomes	16
Number of contigs	363
Contig N50	2.97 Mb
Number of scaffolds	54
Scaffold N50	33.96 Mb
Organelles	Mitochondrial genome: 21.38 kb

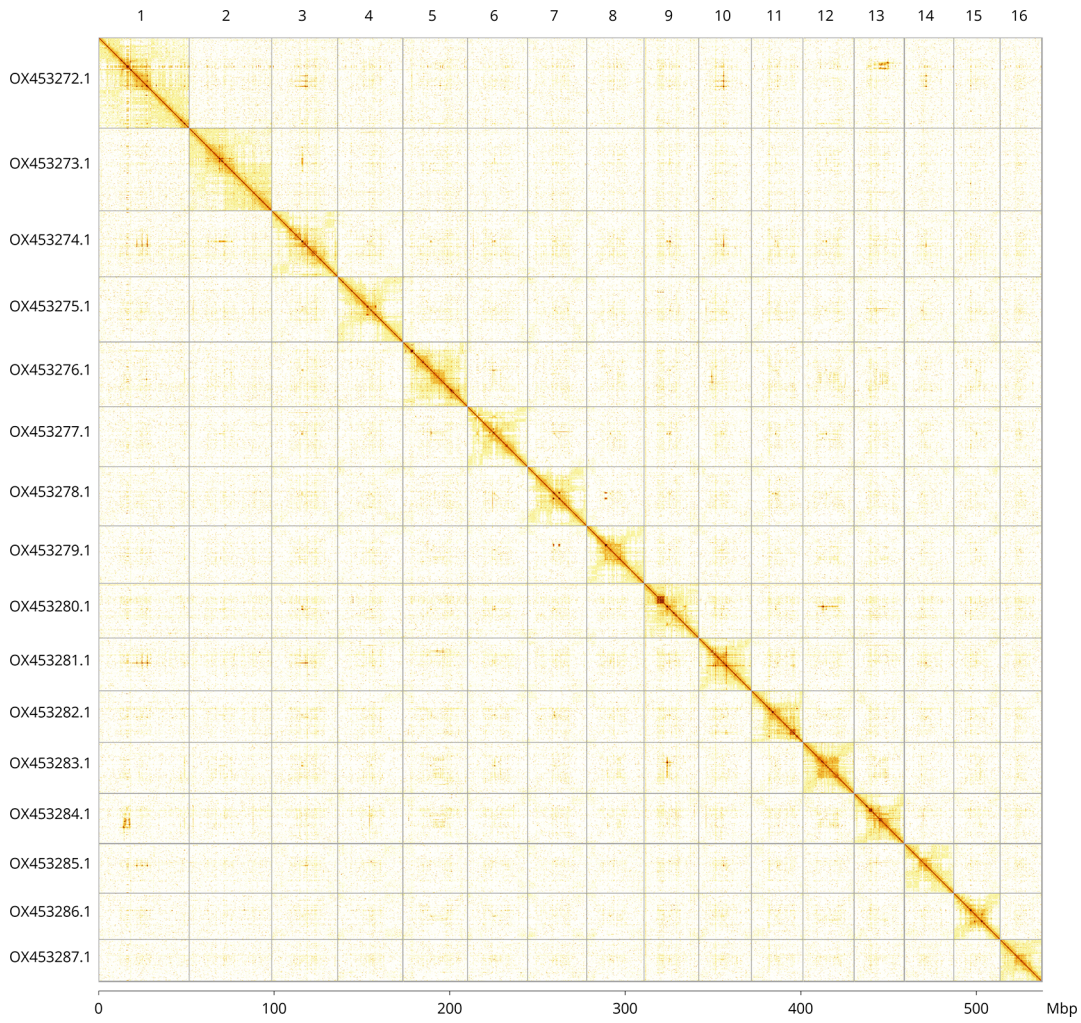

Figure 3. **Hi-C contact map of the *Ricordea florida* genome assembly.** Assembled chromosomes are shown in order of size and labelled along the axes, with a megabase scale shown below. The plot was generated using PretextSnapshot.

The genome was analysed using the BlobToolKit pipeline, a Nextflow implementation of the earlier Snakemake version (Challis *et al.*, 2020). PacBio reads were aligned to the assembly using minimap2 (Li, 2018) and processed with SAMtools (Danecek *et al.*, 2021) to generate coverage tracks. The pipeline runs BUSCO (Manni *et al.*, 2021) using lineages identified from the NCBI Taxonomy (Schoch *et al.*, 2020). For the three basal BUSCO lineages, eukaryota, bacteria and archaea, BUSCO genes are aligned to the UniProt Reference Proteomes database (Bateman *et al.*, 2023) using DIAMOND BLASTp (Buchfink *et al.*, 2021). The genome assembly is divided into chunks based on the density of

Table 3. Chromosomal pseudomolecules in the primary genome assembly of *ricordea florida* jaRicFlor1.

INSDC accession	Molecule	Length (Mb)	GC%
OX453272.1	1	48.64	38.50
OX453273.1	2	47.90	38.50
OX453274.1	3	39.05	38.50
OX453275.1	4	36.60	38.50
OX453276.1	5	35.76	38.50
OX453277.1	6	34.70	38.50
OX453278.1	7	33.96	38.50
OX453279.1	8	32.72	38.50
OX453280.1	9	30.99	38.50
OX453281.1	10	30.42	38.50
OX453282.1	11	28.96	38
OX453283.1	12	28.85	38.50
OX453284.1	13	28.79	38.50
OX453285.1	14	27.72	38.50
OX453286.1	15	27.09	38.50
OX453287.1	16	24.66	38.50

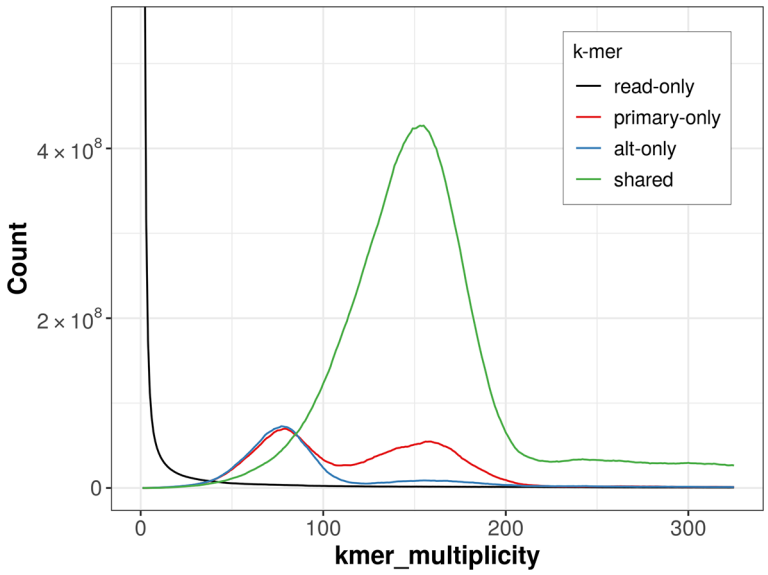

Figure 4. Evaluation of *k*-mer completeness using MerquryFK. This plot illustrates the recovery of *k*-mers from the original read data in the final assemblies. The horizontal axis represents *k*-mer multiplicity, and the vertical axis shows the number of *k*-mers. The black curve represents *k*-mers that appear in the reads but are not assembled. The green curve corresponds to *k*-mers shared by both haplotypes, and the red and blue curves show *k*-mers found only in one of the haplotypes.

BUSCO genes from the closest taxonomic lineage, and each chunk is aligned to the UniProt Reference Proteomes database using DIAMOND BLASTx. Sequences without DIAMOND BLASTx hits are extracted using seqtk and aligned to the NT database using BLASTn (Altschul *et al.*, 1990). The outputs are consolidated into a BlobDir for visualisation in BlobToolKit. The BlobToolKit pipeline was developed using nf-core tooling (Ewels *et al.*, 2020) and MultiQC (Ewels *et al.*, 2016), with containerisation through Docker (Merkel, 2014) and Singularity (Kurtzer *et al.*, 2017).

Genome sequence report

Sequence data

PacBio sequencing of the *Ricordea florida* specimen generated 78.80 Gb (gigabases) from 5.98 million reads, which were used to assemble the genome. GenomeScope2.0 analysis estimated the haploid genome size at 528.54 Mb, with a heterozygosity of 0.75% and repeat content of 47.73% (Figure 2). These estimates guided expectations for the assembly.

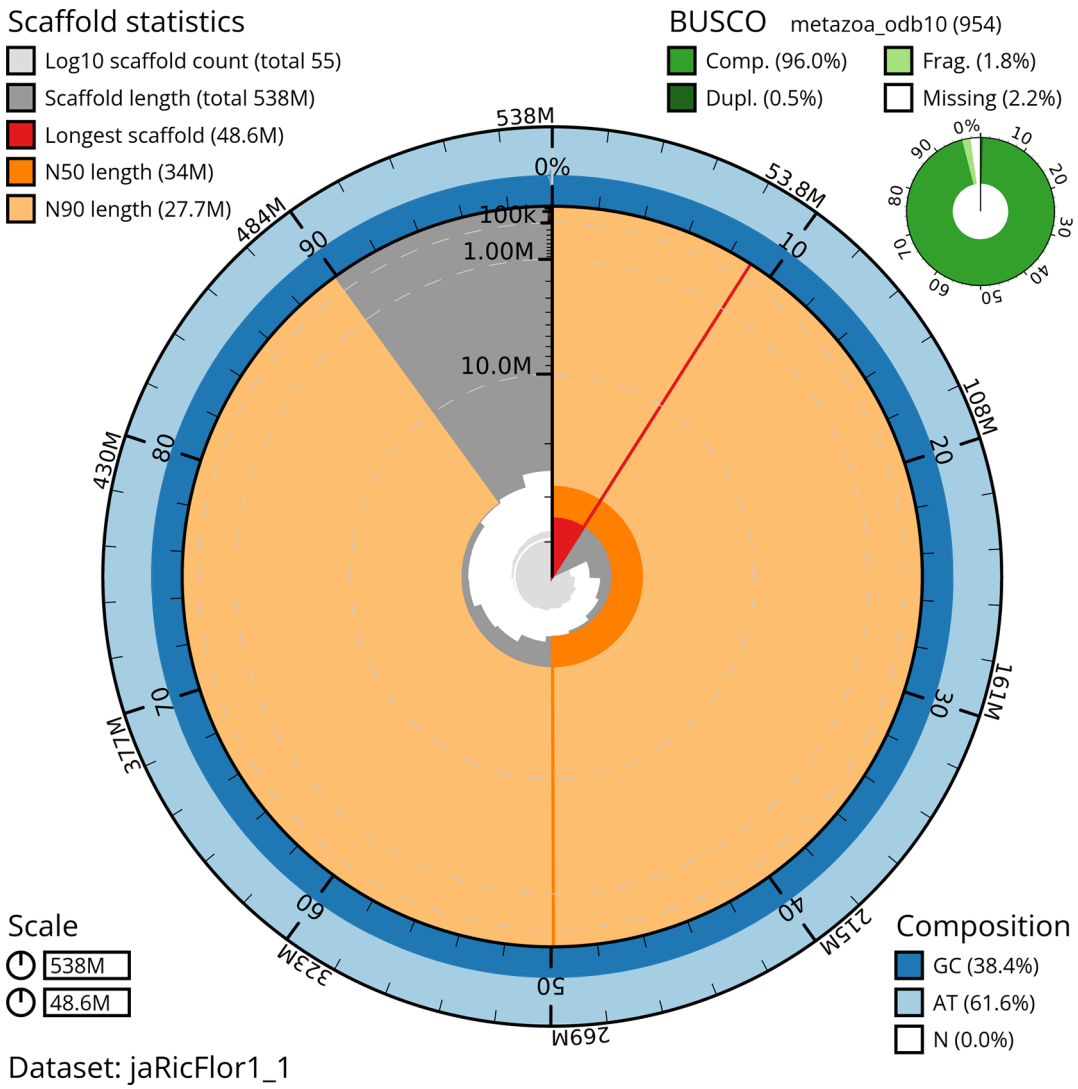

Figure 5. Assembly metrics for jaRicFlor1.1. The BlobToolKit snail plot provides an overview of assembly metrics and BUSCO gene completeness. The circumference represents the length of the whole genome sequence, and the main plot is divided into 1 000 bins around the circumference. The outermost blue tracks display the distribution of GC, AT, and N percentages across the bins. Scaffolds are arranged clockwise from longest to shortest and are depicted in dark grey. The longest scaffold is indicated by the red arc, and the deeper orange and pale orange arcs represent the N50 and N90 lengths. A light grey spiral at the centre shows the cumulative scaffold count on a logarithmic scale. A summary of complete, fragmented, duplicated, and missing BUSCO genes in the metazoa_odb10 set is presented at the top right. An interactive version of this figure can be accessed on the BlobToolKit viewer.

Based on the estimated genome size, the sequencing data provided approximately 62× coverage. Hi-C sequencing produced 135.48 Gb from 897.22 million reads, which were used to scaffold the assembly. RNA sequencing data were also generated and are available in public sequence repositories. Table 1 summarises the specimen and sequencing details.

Assembly statistics

The primary haplotype was assembled, and contigs corresponding to an alternate haplotype were also deposited in INSDC databases. The final assembly has a total length of 537.84 Mb in 54 scaffolds, with a scaffold N50 of 33.96 Mb (Table 2).

Most of the assembly sequence (99.81%) was assigned to 16 chromosomal-level scaffolds. These chromosome-level scaffolds, confirmed by Hi-C data, are named according to size (Figure 3; Table 3).

The mitochondrial genome was also assembled (length 21.38 kb, OX453288.1). This sequence is included as a contig in the multifasta file of the genome submission and as a standalone record.

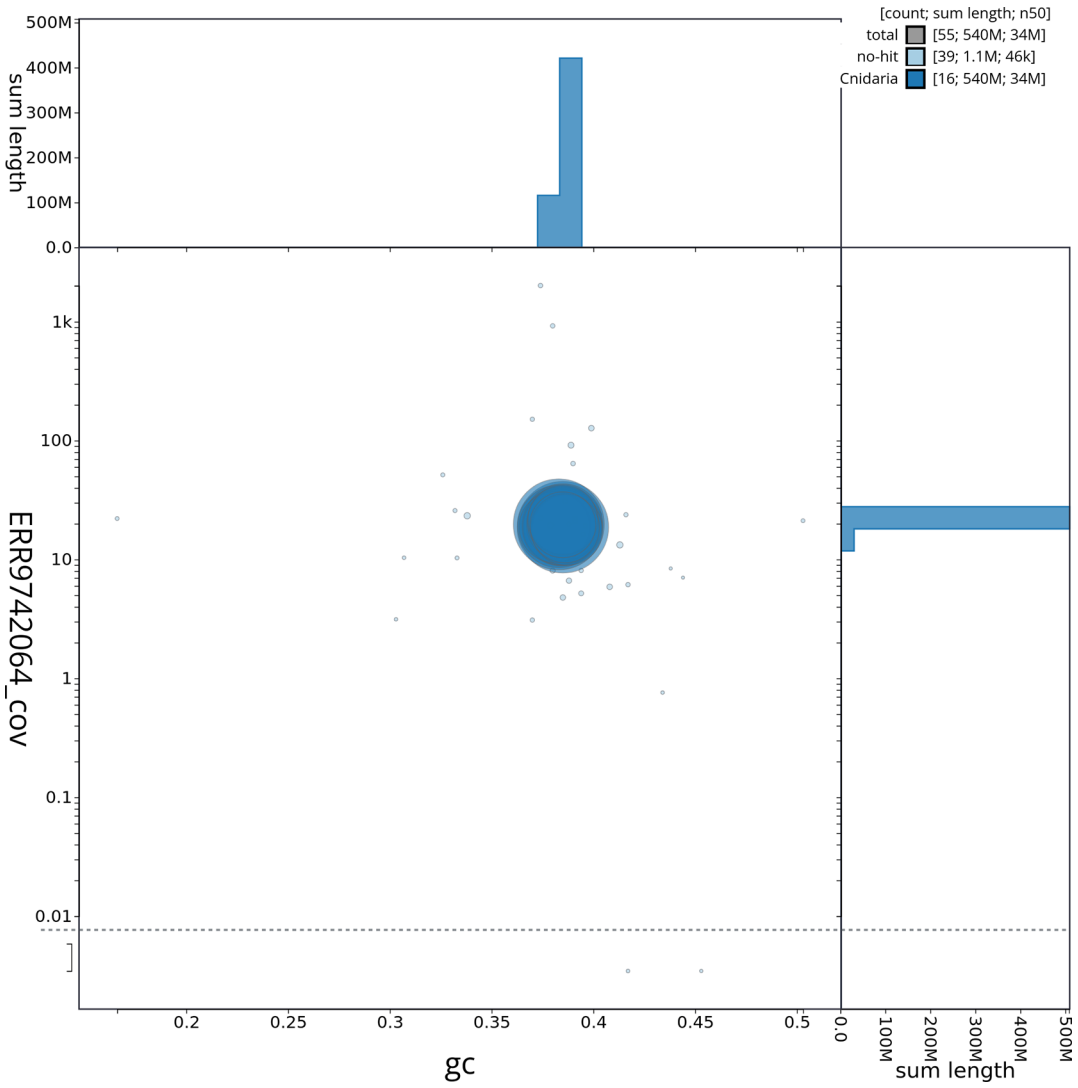

Figure 6. BlobToolKit GC-coverage plot for jaRicFlor1.1. Blob plot showing sequence coverage (vertical axis) and GC content (horizontal axis). The circles represent scaffolds, with the size proportional to scaffold length and the colour representing phylum membership. The histograms along the axes display the total length of sequences distributed across different levels of coverage and GC content. An interactive version of this figure is available on the BlobToolKit viewer.

Assembly quality metrics

The combined primary and alternate assemblies achieve an estimated QV of 58.6. The *k*-mer completeness is 85.47% for the primary assembly, 80.67% for the alternate haplotype, and 97.86% for the combined assemblies (Figure 4).

BUSCO v.5.3.2 analysis using the metazoa_odb10 reference set (n = 954) identified 96.0% of the expected gene set (single = 95.5%, duplicated = 0.5%). The snail plot in Figure 5 summarises the scaffold length distribution and other assembly statistics for the primary assembly. The blob plot in Figure 6 shows the distribution of scaffolds by GC proportion and coverage.

Table 4. Earth Biogenome Project summary metrics for the *ricordea florida* assembly.

Measure	Value	Benchmark
EBP summary (primary)	6.C.Q62	6.C.Q40
Contig N50 length	2.97 Mb	≥ 1 Mb
Scaffold N50 length	33.96 Mb	= chromosome N50
Consensus quality (QV)	Primary: 62.8; alternate: 57.6; combined: 58.6	≥ 40
k-mer completeness	Primary: 85.47%; alternate: 80.67%; combined: 97.86%	≥ 90%
BUSCO	C:96.0% [S:95.5%; D:0.5%]; F:1.8%; M:2.2%; n:954	S > 90%; D < 5%
Percentage of assembly assigned to chromosomes	99.81%	≥ 90%

Table 5. Software versions and sources.

Software	Version	Source
BEDTools	2.30.0	https://github.com/arq5x/bedtools2
BLAST	2.14.0	ftp://ftp.ncbi.nlm.nih.gov/blast/executables/blast+/
BlobToolKit	4.2.1	https://github.com/blobtoolkit/blobtoolkit
BUSCO	5.3.2	https://gitlab.com/ezlab/busco
bwa-mem2	2.2.1	https://github.com/bwa-mem2/bwa-mem2
Cooler	0.8.11	https://github.com/open2c/cooler
fasta_windows	0.2.4	https://github.com/tolkit/fasta_windows
FastK	1.1	https://github.com/thegenemyers/FASTK
Gfastats	1.3.6	https://github.com/vgl-hub/gfastats
GenomeScope2.0	2.0.1	https://github.com/tbenavi1/genomescope2.0
Hifiasm	0.16.1	https://github.com/chhylp123/hifiasm
HiGlass	1.13.4	https://github.com/higlass/higlass
MerquryFK	1.1.2	https://github.com/thegenemyers/MERQURY.FK
MitoHiFi	2	https://github.com/marcelauliano/MitoHiFi
MultiQC	1.14; 1.17 and 1.18	https://github.com/MultiQC/MultiQC
PretextSnapshot	0.0.5	https://github.com/sanger-tol/PretextSnapshot
PretextView	1.0.3	https://github.com/sanger-tol/PretextView
Seqtk	1.3	https://github.com/lh3/seqtk
Singularity	3.9.0	https://github.com/sylabs/singularity
sanger-tol/ascc	0.1.0	https://github.com/sanger-tol/ascc
sanger-tol/curationpretext	1.4.2	https://github.com/sanger-tol/curationpretext
TreeVal	1.4.0	https://github.com/sanger-tol/treeval
YaHS	1.1a.2	https://github.com/c-zhou/yahs

Table 4 lists the assembly metric benchmarks adapted from Rhie *et al.* (2021) and the Earth BioGenome Project Report on Assembly Standards September 2024. The EBP metric, calculated for primary assembly, is **6.C.Q58**, meeting the recommended reference standard.

Author information

Contributors are listed at the following links:

- Members of the Wellcome Sanger Institute Tree of Life Management, Samples and Laboratory team
- Members of Wellcome Sanger Institute Scientific Operations – Sequencing Operations
- Members of the Wellcome Sanger Institute Tree of Life Core Informatics team
- Members of the EBI Aquatic Symbiosis Genomics Data Portal Team
- The Aquatic Symbiosis Genomics Project leadership

Wellcome sanger institute – Legal and governance

The materials that have contributed to this genome note have been supplied by a Tree of Life collaborator. The Wellcome Sanger Institute employs a process whereby due diligence is carried out proportionate to the nature of the materials themselves, and the circumstances under which they have been/are to be collected and provided for use. The purpose of this is to address and mitigate any potential legal and/or ethical implications of receipt and use of the materials as part of the research project, and to ensure that in doing so we align with best practice wherever possible.

The overarching areas of consideration are:

- Ethical review of provenance and sourcing of the material
- Legality of collection, transfer and use (national and international)

Each transfer of samples is undertaken according to a Research Collaboration Agreement or Material Transfer Agreement entered into by the Tree of Life collaborator, Genome Research Limited (operating as the Wellcome Sanger Institute) and in some circumstances other Tree of Life collaborators.

Data availability

European Nucleotide Archive: Ricordea florida (Florida false coral). Accession number PRJEB52859. The genome sequence is released openly for reuse. The *Ricordea florida* genome sequencing initiative is part of the Aquatic Symbiosis Genomics Project (PRJEB43743) and the Sanger Institute Tree of Life Programme (PRJEB43745). All raw sequence data and the assembly have been deposited in INSDC databases. The genome will be annotated using available RNA-Seq data and presented through the Ensembl pipeline at the European Bioinformatics Institute. Raw data and assembly accession identifiers are reported in Tables 1 and 2.

Production code used in genome assembly at the WSI Tree of Life is available at https://github.com/sanger-tol. Table **5** lists software versions used in this study.

References

Altschul SF, Gish W, Miller W, *et al.*: **Basic Local Alignment Search Tool.** *J. Mol. Biol.* 1990; **215**(3): 403–410.
PubMed Abstract | Publisher Full Text

Bateman A, Martin M-J, Orchard S, *et al.*: **UniProt: The Universal Protein Knowledgebase in 2023.** *Nucleic Acids Res.* 2023; **51**(D1): D523–D531.
PubMed Abstract | Publisher Full Text | Free Full Text

Buchfink B, Reuter K, Drost H-G: **Sensitive protein alignments at tree-of-life scale using DIAMOND.** *Nat. Methods.* 2021; **18**(4): 366–368.
PubMed Abstract | Publisher Full Text | Free Full Text

Challis R, Richards E, Rajan J, *et al.*: **BlobToolKit – interactive quality assessment of genome assemblies.** *G3: Genes, Genomes, Genetics.* 2020; **10**(4): 1361–1374.
PubMed Abstract | Publisher Full Text | Free Full Text

Cheng H, Concepcion GT, Feng X, *et al.*: **Haplotype-resolved *de novo* assembly using phased assembly graphs with Hifiasm.** *Nat. Methods.* 2021; **18**(2): 170–175.
PubMed Abstract | Publisher Full Text | Free Full Text

Danecek P, Bonfield JK, Liddle J, *et al.*: **Twelve years of SAMtools and BCFtools.** *GigaScience.* 2021; **10**(2): giab008.
PubMed Abstract | Publisher Full Text | Free Full Text

Davy SK, Allemand D, Weis VM: **Cell biology of cnidarian-dinoflagellate symbiosis.** *Microbiol. Mol. Biol. Rev.* 2012; **76**(2): 229–261.
PubMed Abstract | Publisher Full Text | Free Full Text

den Hartog JC: **Caribbean shallow water Corallimorpharia.** *Zool. Verh.* 1980; **176**(1): 1–83.
Reference Source

Ewels P, Magnusson M, Lundin S, *et al.*: **MultiQC: Summarize analysis results for multiple tools and samples in a single report.** *Bioinformatics.* 2016; **32**(19): 3047–3048.
PubMed Abstract | Publisher Full Text | Free Full Text

Ewels PA, Peltzer A, Fillinger S, *et al.*: **The nf-core framework for community-curated bioinformatics pipelines.** *Nat. Biotechnol.* 2020; **38**(3): 276–278.
PubMed Abstract | Publisher Full Text

Fautin D: **Catalog to families, genera, and species of orders Actiniaria and Corallimorpharia (Cnidaria: Anthozoa).** *Zootaxa.* 2016; **4145**(1): 1–449.
PubMed Abstract | Publisher Full Text

Fautin D, White TR, Pearson K: **Two new species of deep-water Corallimorpharia (Cnidaria: Anthozoa) from the Northeast Pacific, *Corallimorphus denhartogi* and *C. pilatus*.** *Pac. Sci.* 2002; **56**(2): 113–124.
Publisher Full Text

Formenti G, Abueg L, Brajuka A, *et al.*: **Gfastats: Conversion, evaluation and manipulation of genome sequences using assembly graphs.** *Bioinformatics.* 2022; **38**(17): 4214–4216.
PubMed Abstract | Publisher Full Text | Free Full Text

GBIF Secretariat: ***Ricordea florida* Duchassaing & Michelotti, 1860 in GBIF Backbone Taxonomy.** 2026.
Reference Source

Guan D, McCarthy SA, Wood J, *et al.*: **Identifying and removing haplotypic duplication in primary genome assemblies.** *Bioinformatics.* 2020; **36**(9): 2896–2898.
PubMed Abstract | Publisher Full Text | Free Full Text

Horricks R, Millar Z, Sykes K, *et al.*: **Differential protein abundance in the regenerating Caribbean corallimorpharian *Ricordea florida*.** *Journal of Zoo and Aquarium Research.* 2023; **11**(1): 240–248.
Publisher Full Text

Howard C, Denton A, Jackson B, *et al.*: **On the path to reference genomes for all biodiversity: Lessons learned and laboratory protocols created in the Sanger Tree of Life core laboratory over the first 2000 species.** *bioRxiv.* 2025.
Publisher Full Text

Howe K, Chow W, Collins J, *et al.*: **Significantly improving the quality of genome assemblies through curation.** *GigaScience.* 2021; **10**(1): giaa153.
PubMed Abstract | Publisher Full Text | Free Full Text

Kerpedjiev P, Abdennur N, Lekschas F, *et al.*: **HiGlass: Web-based visual exploration and analysis of genome interaction maps.** *Genome Biol.* 2018; **19**(1): 125.
PubMed Abstract | Publisher Full Text | Free Full Text

Kurtzer GM, Sochat V, Bauer MW: **Singularity: Scientific containers for mobility of compute.** *PLOS ONE.* 2017; **12**(5): e0177459.
PubMed Abstract | Publisher Full Text | Free Full Text

LaJeunesse TC: **Diversity and community structure of symbiotic dinoflagellates from Caribbean coral reefs.** *Mar. Biol.* 2002; **141**(2): 387–400.
Publisher Full Text

León YM, Bjorndal KA: **Selective feeding in the hawksbill turtle, an important predator in coral reef ecosystems.** *Mar. Ecol. Prog. Ser.* 2002; **245**: 249–258.
Publisher Full Text

Li H: **Minimap2: Pairwise alignment for nucleotide sequences.** *Bioinformatics.* 2018; **34**(18): 3094–3100.
PubMed Abstract | Publisher Full Text | Free Full Text

Lin M-F, Chou WH, Kitahara MV, *et al.*: **Corallimorpharians are not "naked corals": Insights into relationships between Scleractinia and Corallimorpharia from phylogenomic analyses.** *PeerJ.* 2016; **4**: e2463.
PubMed Abstract | Publisher Full Text | Free Full Text

Lin M-F, Moya A, Ying H, *et al.*: **Analyses of corallimorpharian transcriptomes provide new perspectives on the evolution of calcification in the Scleractinia (corals).** *Genome Biol. Evol.* 2017; **9**(1): 150–160.
PubMed Abstract | Publisher Full Text | Free Full Text

Lin M-F, Takahashi S, Forêt S, *et al.*: **Transcriptomic analyses highlight the likely metabolic consequences of colonization of a cnidarian host by native or non-native *Symbiodinium* species.** *Biol. Open.* 2019; **8**(3): bio038281.
PubMed Abstract | Publisher Full Text | Free Full Text

Manni M, Berkeley MR, Seppey M, *et al.*: **BUSCO update: Novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes.** *Mol. Biol. Evol.* 2021; **38**(10): 4647–4654.
PubMed Abstract | Publisher Full Text | Free Full Text

Medina M, Collins AG, Takaoka TL, *et al.*: **Naked corals: Skeleton loss in Scleractinia.** *Proc. Natl. Acad. Sci. USA.* 2006; **103**(24): 9096–9100.
PubMed Abstract | Publisher Full Text | Free Full Text

Merkel D: **Docker: Lightweight Linux containers for consistent development and deployment.** *Linux J.* 2014; **2014**(239): 2.
Publisher Full Text

Muhando CA, Kuguru BL, Wagner GM, *et al.*: **Environmental effects on the distribution of corallimorpharians in Tanzania.** *Ambio.* 2002; **31**(7): 558–561.
PubMed Abstract | Publisher Full Text

Parr ND: ***Environmental tolerance and reproduction of Florida false corals Ricordea florida (Anthozoa: Corallimorpharia): Implications for ornamental fisheries management.*** 2019.
Reference Source

Ranallo-Benavidez TR, Jaron KS, Schatz MC: **GenomeScope 2.0 and Smudgeplot for reference-free profiling of polyploid genomes.** *Nat. Commun.* 2020; **11**(1): 1432.
PubMed Abstract | Publisher Full Text | Free Full Text

Rao SSP, Huntley MH, Durand NC, *et al.*: **A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping.** *Cell.* 2014; **159**(7): 1665–1680.
PubMed Abstract | Publisher Full Text | Free Full Text

Rhie A, McCarthy SA, Fedrigo O, *et al.*: **Towards complete and error-free genome assemblies of all vertebrate species.** *Nature.* 2021; **592**(7856): 737–746.
PubMed Abstract | Publisher Full Text | Free Full Text

Rhie A, Walenz BP, Koren S, *et al.*: **Merqury: Reference-free quality, completeness, and phasing assessment for genome assemblies.** *Genome Biol.* 2020; **21**(1): 245.
PubMed Abstract | Publisher Full Text | Free Full Text

Schoch CL, Ciufo S, Domrachev M, *et al.*: **NCBI taxonomy: A comprehensive update on curation, resources and tools.** *Database.* 2020; **2020**: baaa062.
PubMed Abstract | Publisher Full Text | Free Full Text

Torres-Pratts H, Lado-Insua T, Rhyne AL, *et al.*: **Two distinct, geographically overlapping lineages of the corallimorpharian *Ricordea florida* (Cnidaria: Hexacorallia: Ricordeidae).** *Coral Reefs.* 2011; **30**(2): 391–396.
Publisher Full Text

Uliano-Silva M, Ferreira JGRN, Krasheninnikova K, *et al.*: **MitoHiFi: A Python pipeline for mitochondrial genome assembly from PacBio high fidelity reads.** *BMC Bioinformatics.* 2023; **24**(1): 288.
PubMed Abstract | Publisher Full Text | Free Full Text

Vasimuddin M, Misra S, Li H, *et al.*: **Efficient architecture-aware acceleration of BWA-MEM for multicore systems.** *2019 IEEE International Parallel and Distributed Processing Symposium (IPDPS).* IEEE; 2019; 314–24.
Publisher Full Text

Zhou C, McCarthy SA, Durbin R: **YaHS: Yet another Hi-C scaffolding tool.** *Bioinformatics.* 2023; **39**(1): btac808.
PubMed Abstract | Publisher Full Text | Free Full Text