DATA NOTE

The genome sequence of the Maltese wall lizard, *Podarcis filfolensis* (Bedriaga, 1876)

[version 1; peer review: 2 approved with reservations]

Nathalie Feiner[1,2], Tobias Uller[2], Javier Abalos[3], Joana Meier[4],
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,
Tree of Life Core Informatics collective

[1]Max Planck Institute for Evolutionary Biology, Plön, Germany
[2]Department of Biology, Lund University, Lund, Sweden
[3]Ethology Lab, Cavanilles Institute of Biodiversity and Evolutionary Biology, University of Valencia, Valencia, Spain
[4]Wellcome Sanger Institute, Hinxton, England, UK

v1 **First published:** 23 Apr 2025, **10**:213
https://doi.org/10.12688/wellcomeopenres.24042.1
Latest published: 23 Apr 2025, **10**:213
https://doi.org/10.12688/wellcomeopenres.24042.1

Abstract

We present a genome assembly from a female specimen of *Podarcis filfolensis* (Maltese wall lizard; Chordata; Lepidosauria; Squamata; Lacertidae). The assembly contains two haplotypes with total lengths of 1,506.95 megabases and 1,404.02 megabases. Most of haplotype 1 (98.73%) is scaffolded into 20 chromosomal pseudomolecules, including the W and Z sex chromosomes. Haplotype 2 was assembled to scaffold level. The mitochondrial genome has also been assembled, with a length of 17.23 kilobases.

Keywords

Podarcis filfolensis, Maltese wall lizard, genome sequence, chromosomal, Squamata

This article is included in the Tree of Life gateway.

Open Peer Review

Approval Status ? ?

	1	2
version 1 23 Apr 2025	? view	? view

1. **Terry Bertozzi**, South Australian Museum, Adelaide, Australia

2. **Hardip R. Patel**, Australian National University, Canberra, Australia
Arthur Georges, University of Canberra, Canberra, Australia
Paul D. Waters, UNSW Sydney, Sydney, Australia

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

Corresponding author: Wellcome Sanger Institute Tree of Life Management, Samples and Laboratory team (Mark.Blaxter@sanger.ac.uk)

Author roles: Feiner N: Investigation, Resources, Writing – Original Draft Preparation, Writing – Review & Editing; **Uller T**: Investigation, Resources, Writing – Original Draft Preparation; **Abalos J**: Resources, Writing – Review & Editing; **Meier J**: Investigation, Resources, Writing – Review & Editing;

Competing interests: No competing interests were disclosed.

Grant information: This work was supported by Wellcome through core funding to the Wellcome Sanger Institute (220540). N.F. was funded by Starting Grants from the European Research Council (948126) and the Swedish Research Council (2020-03650). J.A. was funded by a Margarita Salas contract from the Spanish Ministry of Science (MS21-053).

How to cite this article: Feiner N, Uller T, Abalos J *et al.* **The genome sequence of the Maltese wall lizard, *Podarcis filfolensis* (Bedriaga, 1876) [version 1; peer review: 2 approved with reservations]** Wellcome Open Research 2025, **10**:213 https://doi.org/10.12688/wellcomeopenres.24042.1

First published: 23 Apr 2025, **10**:213 https://doi.org/10.12688/wellcomeopenres.24042.1

Species taxonomy

Eukaryota; Opisthokonta; Metazoa; Eumetazoa; Bilateria; Deuterostomia; Chordata; Craniata; Vertebrata; Gnathostomata; Teleostomi; Euteleostomi; Sarcopterygii; Dipnotetrapodomorpha; Tetrapoda; Amniota; Sauropsida; Sauria; Lepidosauria; Squamata; Bifurcata; Unidentata; Episquamata; Laterata; Lacertibaenia; Lacertidae; Lacertinae; *Podarcis*; *Podarcis filfolensis* (Bedriaga, 1876) (NCBI:txid65481)

Background

The Maltese wall lizard *Podarcis filfolensis* (Bedriaga, 1876; Figure 1) is a diurnal lacertid lizard endemic to the Maltese and Pelagian islands in the Mediterranean Sea. The nominal subspecies *P. f. filfolensis* inhabits the island of Filfla, which is ca. 5 km southwest of Malta. The subspecies *P. f. laurentiimuelleri* inhabits the Pelagian islands of Linosa and Lampione, situated roughly 500 km west of Malta. *P. filfolensis* is classified as Least Concern (LC) by the IUCN (Schembri & Salvi, 2024). Nevertheless, the species is locally threatened, as demonstrated by the recent (in the early 2000s) eradication by rats of the population inhabiting St Paul's Island (also known as Selmunett), which was classified as subspecies *P. f. kieselbachi* (Sciberras & Schembri, 2008).

The Maltese wall lizard is one of 28 currently described species of the genus *Podarcis*. *Podarcis filfolensis* is the sister species to *P. waglerianus* and *P. raffonei*, which are endemic to Sicily and surrounding islands (Salvi *et al.*, 2014; Yang *et al.*, 2021). The divergence between *P. filfolensis* and this sister group is estimated to have occurred over 7 MYA (Yang *et al.*, 2021), before the Messinian Salinity Crisis.

The ecology of this species is typical for *Podarcis*. The species is found in a variety of habitats, including Mediterranean shrubland, karstland vegetation and grassland. The species is common in urban environments on both Malta and Gozo, but appears to be absent from some regions of the main island of Malta. It is generally highly abundant on Gozo and smaller islets such as Comino and Cominotto. Individuals can be active all year, but reproduction is restricted to spring and early summer. The species is oviparous, and Meiri *et al.* (2020) report an average clutch size of 2.5.

Maltese wall lizards are highly variable in terms of colouration and body size (Böhme, 1986). As in most species of *Podarcis*, the sexes differ in appearance, with females being generally smaller and with dorsal colour patterns dominated by longitudinal stripes. The subspecies *P. f. filfolensis* on Filfla is phenotypically divergent from other con-specifics and shows an overall darker body coloration than the subspecies found on the main islands.

After *P. raffonei* (Gabrielli *et al.*, 2023), *P. filfolensis* is the second species in the Sicilian-Maltese group of *Podarcis* (currently consisting of three species) for which a reference genome is constructed. This will be a valuable resource both for studies within *P. filfolensis*, for example as a support for identifying conservation priorities, and for comparative studies above the species level.

Genome sequence report

Sequencing data

The genome of a specimen of *Podarcis filfolensis* was sequenced using Pacific Biosciences single-molecule HiFi long reads, generating 49.37 Gb (gigabases) from 7.33 million reads. GenomeScope analysis of the PacBio HiFi data estimated the haploid genome size at 1,421.13 Mb, with a heterozygosity of 0.79% and repeat content of 20.30%. These values provide an initial assessment of genome complexity and the challenges anticipated during assembly. Based on this estimated genome size, the sequencing data provided approximately 33.0x coverage of the genome. Chromosome conformation Hi-C sequencing produced 205.77 Gb from 1,362.69 million reads. Table 1 summarises the specimen and sequencing information.

Assembly statistics

The genome was assembled into two haplotypes using Hi-C phasing. Haplotype 1 was curated to chromosome level, while haplotype 2 was assembled to scaffold level. The assembly was improved by manual curation, which corrected 130 misjoins or missing joins. These interventions increased the total assembly length by 1.68%, decreased the scaffold count by 11.42%, and increased the scaffold N50 by 2.11%. The final assembly has a total length of 1,506.95 Mb in 348 scaffolds, with 872 gaps, and a scaffold N50 of 91.56 Mb (Table 2).

Figure 1. A) Female and **B**) male *Podarcis filfolensis filfolensis* photographed on the islet of Filfla (May 2022). Note that the specimen selected for genome sequencing is from the same locality, but not shown on the photographs. Photographs by Nathalie Feiner (**A**) and Javier Abalos (**B**).

Table 1. **Specimen and sequencing data for *Podarcis filfolensis*.**

Project information			
ENA Study title	Podarcis filfolensis (Filfola wall lizard)		
Umbrella BioProject	PRJEB73695		
Species	*Podarcis filfolensis*		
BioSpecimen	SAMEA113403362		
NCBI taxonomy ID	65481		
Specimen information			
Technology	**ToLID**	**BioSample accession**	**Organism part**
PacBio long read sequencing	rPodFil1	SAMEA113403363	terminal body
Hi-C sequencing	rPodFil1	SAMEA113403363	terminal body
Sequencing information			
Platform	**Run accession**	**Read count**	**Base count (Gb)**
Hi-C Illumina NovaSeq X	ERR12743667	1.36e+09	205.77
PacBio Revio	ERR12736920	7.33e+06	49.37

Table 2. **Genome assembly data for *Podarcis filfolensis*.**

Genome assembly	**Haplotype 1**	**Haplotype 2**
Assembly name	rPodFil1.hap1.1	rPodFil1.hap2.1
Assembly accession	GCA_964270895.1	GCA_964270875.1
Assembly level	chromosome	scaffold
Span (Mb)	1,506.95	1,404.02
Number of contigs	1,220	937
Number of scaffolds	348	142
Assembly metrics (benchmark)	**Haplotype 1**	**Haplotype 2**
Contig N50 length (≥ 1 Mb)	3.04 Mb	3.22 Mb
Scaffold N50 length (= chromosome N50)	91.56 Mb	91.99 Mb
Consensus quality (QV) (≥ 40)	60.7	60.9
k-mer completeness	83.62%	80.00%
Combined *k*-mer completeness (≥ 95%)	99.43%	
BUSCO* (S > 90%; D < 5%)	C:95.0%[S:93.1%,D:2.0%], F:0.8%,M:4.1%,n:7,480	C:92.2%[S:90.6%,D:1.6%], F:0.9%,M:6.9%,n:7,480
Percentage of assembly mapped to chromosomes (≥ 90%)	98.73%	-
Sex chromosomes (localised homologous pairs)	W and Z	-
Organelles (one complete allele)	Mitochondrial genome: 17.23 kb	-

* BUSCO scores based on the sauropsida_odb10 BUSCO set using version 5.5.0. C = complete [S = single copy, D = duplicated], F = fragmented, M = missing, n = number of orthologues in comparison.

The snail plot in Figure 2 provides a summary of the assembly statistics, indicating the distribution of scaffold lengths and other assembly metrics. Figure 3 shows the distribution of scaffolds by GC proportion and coverage. Figure 4 presents a cumulative assembly plot, with separate curves representing different scaffold subsets assigned to various phyla, illustrating the completeness of the assembly.

Most of the assembly sequence (98.73%) was assigned to 20 chromosomal-level scaffolds, representing 18 autosomes and the W and Z sex chromosomes. These chromosome-level scaffolds, confirmed by Hi-C data, are named according to size (Figure 5; Table 3).

The mitochondrial genome was also assembled. This sequence is included as a contig in the multifasta file of the genome submission and as a standalone record.

Assembly quality metrics

The estimated Quality Value (QV) and *k*-mer completeness metrics, along with BUSCO completeness scores, were calculated for each haplotype and the combined assembly. The

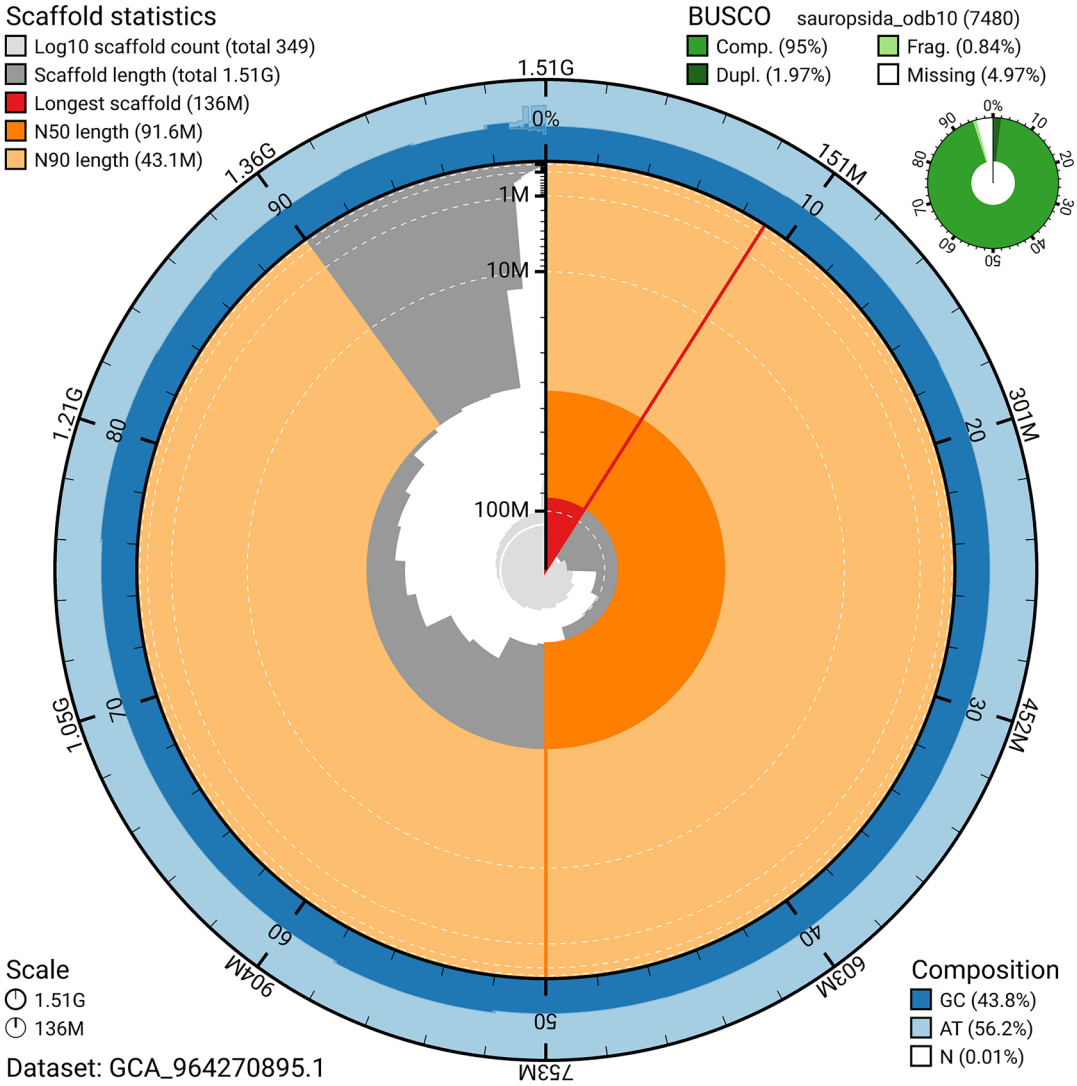

Figure 2. Genome assembly of *Podarcis filfolensis*, rPodFil1.hap1.1: metrics. The BlobToolKit snail plot provides an overview of assembly metrics and BUSCO gene completeness. The circumference represents the length of the whole genome sequence, and the main plot is divided into 1,000 bins around the 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 sauropsida_odb10 set is presented at the top right. An interactive version of this figure is available at https://blobtoolkit.genomehubs.org/view/GCA_964270895.1/dataset/GCA_964270895.1/snail.

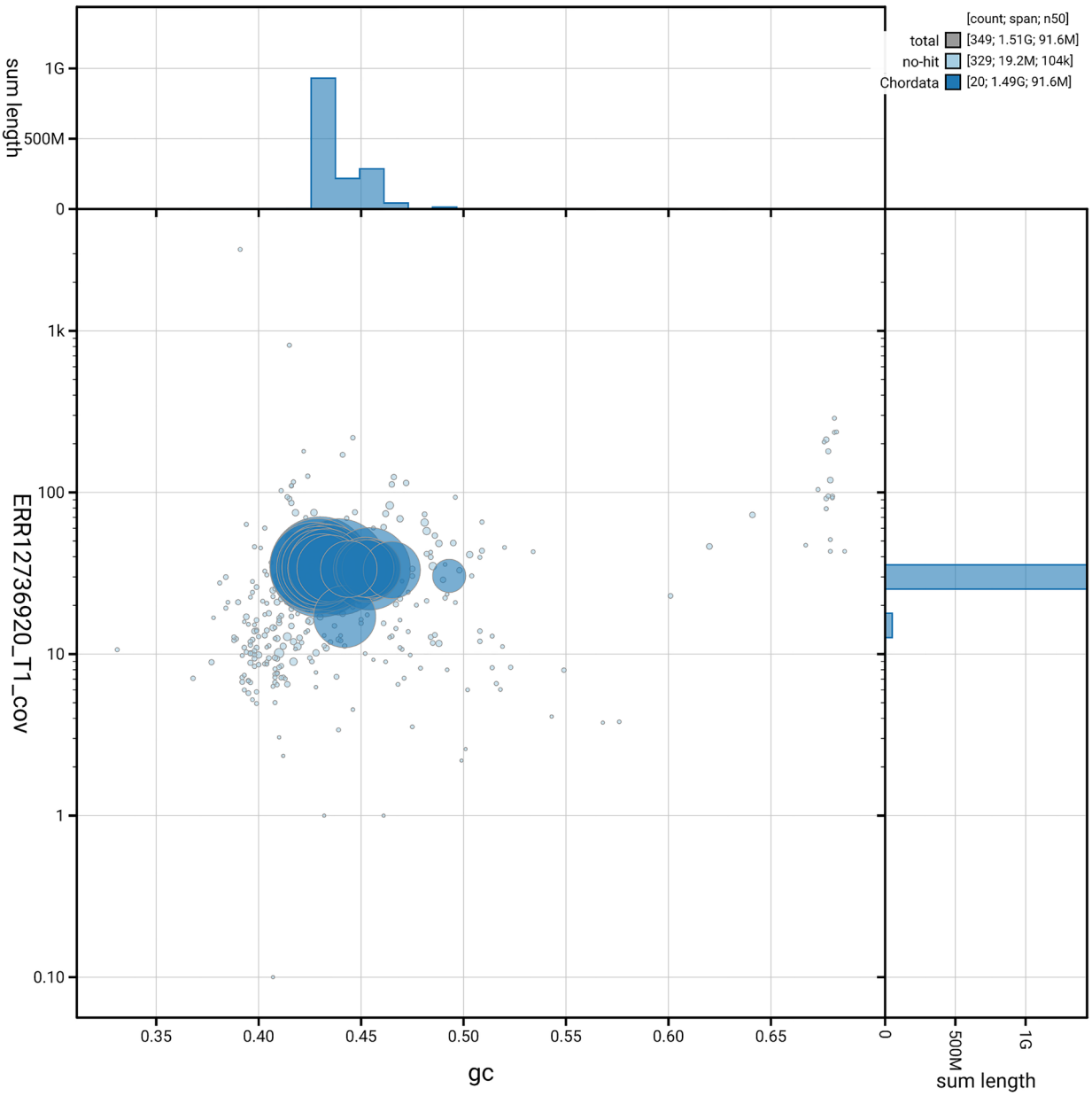

Figure 3. Genome assembly of *Podarcis filfolensis*, rPodFil1.hap1.1: BlobToolKit GC-coverage plot. Blob plot showing sequence coverage (vertical axis) and GC content (horizontal axis). The circles represent scaffolds, with the size proportional to scaffold length and the colour representing phylum membership. 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 at https://blobtoolkit.genomehubs.org/view/GCA_964270895.1/dataset/GCA_964270895.1/blob.

QV reflects the base-level accuracy of the assembly, while *k*-mer completeness indicates the proportion of expected *k*-mers identified in the assembly. BUSCO scores provide a measure of completeness based on benchmarking universal single-copy orthologues.

For haplotype 1, the estimated QV is 60.7, and for haplotype 2, 60.9. When the two haplotypes are combined, the assembly achieves an estimated QV of 60.8. The *k*-mer recovery for haplotype 1 is 83.62%, and for haplotype 2 80.00%, while the combined haplotypes have a *k*-mer recovery of 99.43%.

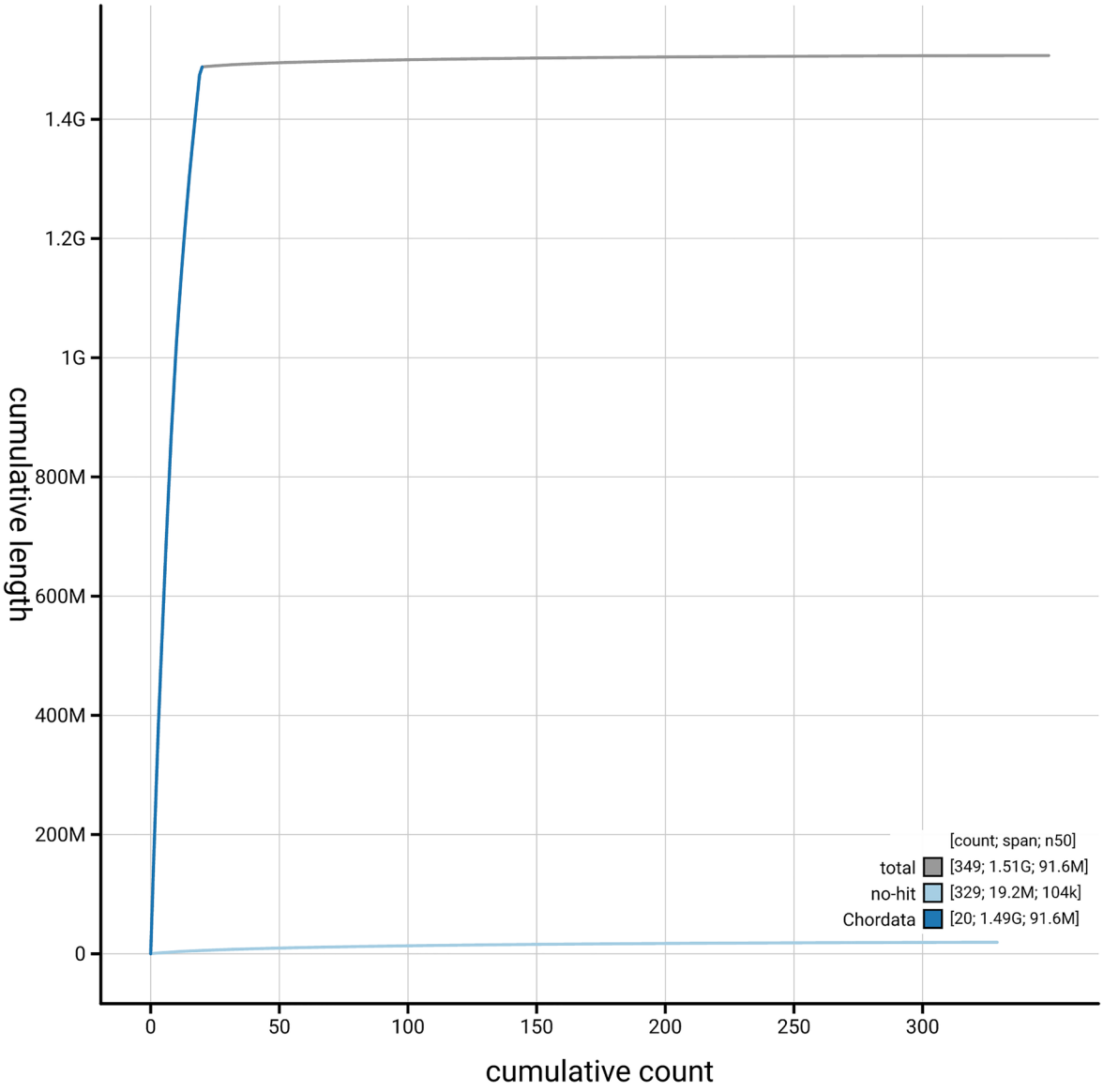

Figure 4. Genome assembly of *Podarcis filfolensis*, rPodFil1.hap1.1: BlobToolKit cumulative sequence plot. The grey line shows cumulative length for all scaffolds. Coloured lines show cumulative lengths of scaffolds assigned to each phylum using the buscogenes taxrule. An interactive version of this figure is available at https://blobtoolkit.genomehubs.org/view/GCA_964270895.1/dataset/GCA_964270895.1/cumulative.

BUSCO 5.5.0 analysis using the sauropsida_odb10 reference set (n = 7,480) identified 95.0% of the expected gene set (single = 93.1%, duplicated = 2.0%) for haplotype 1.

Table 2 provides assembly metric benchmarks adapted from Rhie *et al.* (2021) and the Earth BioGenome Project Report on Assembly Standards September 2024. The haplotype 1 assembly achieves the EBP reference standard of **6.C.Q60**.

Methods

Sample acquisition

The specimen, an adult female *P. filfolensis* lizard (specimen ID SAN25001765, ToLID rPodFil1) was collected on 2022-05-04 from Mediterranean garigue habitat on the island of Filfla (latitude: 35.787928; longitude: 14.409401). The specimen was caught by noosing, standard morphometric measurements were taken, and the tip of the tail (ca. 2 cm)

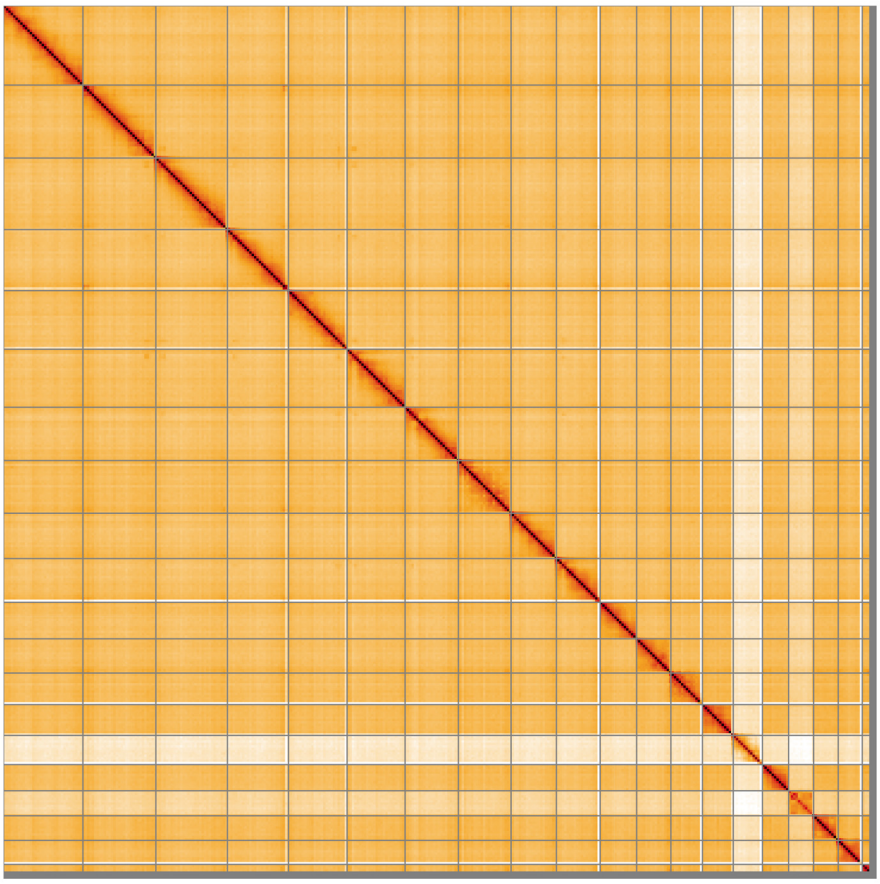

Figure 5. Genome assembly of *Podarcis filfolensis*: Hi-C contact map of the rPodFil1.hap1.1 assembly, visualised using HiGlass. Chromosomes are shown in order of size from left to right and top to bottom. An interactive version of this figure may be viewed at https://genome-note-higlass.tol.sanger.ac.uk/l/?d=NjJ735eWTxKirO6cd9ZaQw.

Table 3. Chromosomal pseudomolecules in the genome assembly of *Podarcis filfolensis*, rPodFil1.

INSDC accession	Name	Length (Mb)	GC%
OZ185435.1	1	136.39	43
OZ185436.1	2	125.39	44
OZ185437.1	3	122.62	43
OZ185438.1	4	104.86	42.5
OZ185439.1	5	100.65	43
OZ185440.1	6	99.61	43.5
OZ185441.1	7	91.56	43
OZ185442.1	8	90.41	45.5
OZ185443.1	9	77.8	43
OZ185444.1	10	75.34	43.5
OZ185445.1	11	62.55	43
OZ185446.1	12	58.7	43.5
OZ185447.1	13	54.33	45.5
OZ185448.1	14	52.8	45
OZ185450.1	15	44.74	45.5
OZ185452.1	16	42.1	46.5
OZ185453.1	17	41.38	44.5
OZ185454.1	18	13.14	49.5
OZ185451.1	W	43.07	45
OZ185449.1	Z	50.29	44
OZ185455.1	MT	0.02	40

was collected and preserved in ethanol. The specimen was released again at the site of capture. Field work was conducted under the permit ID EP 1042/22 from the Environment and Resources Authority of Malta.

Nucleic acid extraction

The workflow for high molecular weight (HMW) DNA extraction at the Wellcome Sanger Institute (WSI) Tree of Life Core Laboratory includes a sequence of procedures: sample preparation and homogenisation, DNA extraction, fragmentation and purification. Detailed protocols are available on protocols.io (Denton *et al.*, 2023b). The rPodFil1 sample was prepared for DNA extraction by weighing and dissecting it on dry ice (Jay *et al.*, 2023). Tissue from the terminal body was homogenised using a PowerMasher II tissue disruptor (Denton *et al.*, 2023a). HMW DNA was extracted using the Manual MagAttract v1 protocol (Strickland *et al.*, 2023b). DNA was sheared into an average fragment size of 12–20 kb in a Megaruptor 3 system (Todorovic *et al.*, 2023). Sheared DNA was purified by solid-phase reversible immobilisation, using AMPure PB beads to eliminate shorter fragments and concentrate the DNA (Strickland *et al.*, 2023a). 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.

Hi-C sample preparation and crosslinking

Tissue from the terminal body of the rPodFil1 sample was processed for Hi-C sequencing at the WSI Scientific Operations core, using the Arima-HiC v2 kit. In brief, 20–50 mg of frozen tissue (stored at –80 °C) was fixed, and the DNA crosslinked using a TC buffer with 22% formaldehyde concentration. After crosslinking, the tissue was homogenised using the Diagnocine Power Masher-II and BioMasher-II tubes and pestles. Following the Arima-HiC v2 kit manufacturer's instructions, crosslinked DNA was digested using a restriction enzyme master mix. The 5'-overhangs were filled in and labelled with biotinylated nucleotides and proximally ligated. An overnight incubation was carried out for enzymes to digest remaining proteins and for crosslinks to reverse. A clean up was performed with SPRIselect beads prior to library preparation. Additionally, the biotinylation percentage was estimated using the Qubit Fluorometer v4.0 (Thermo Fisher Scientific) and Qubit HS Assay Kit and Arima-HiC v2 QC beads.

Library preparation and sequencing

Library preparation and sequencing were performed at the WSI Scientific Operations core.

PacBio HiFi

At a minimum, samples were required to have an average fragment size exceeding 8 kb and a total mass over 400 ng to proceed to the low input SMRTbell Prep Kit 3.0 protocol (Pacific Biosciences, California, USA), depending on genome size and sequencing depth required. Libraries were prepared using the SMRTbell Prep Kit 3.0 (Pacific Biosciences, California, USA) as per the manufacturer's instructions. The kit includes the reagents required for end repair/A-tailing, adapter ligation, post-ligation SMRTbell bead cleanup, and nuclease treatment. Following the manufacturer's instructions, size selection and clean up was carried out using diluted AMPure PB beads (Pacific Biosciences, California, USA). DNA concentration was quantified using the Qubit Fluorometer v4.0 (Thermo Fisher Scientific) with Qubit 1X dsDNA HS assay kit and the final library fragment size analysis was carried out using the Agilent Femto Pulse Automated Pulsed Field CE Instrument (Agilent Technologies) and gDNA 55kb BAC analysis kit.

Samples were sequenced on a Revio instrument (Pacific Biosciences, California, USA). Prepared libraries were normalised to 2 nM, and 15 µL was used for making complexes. Primers were annealed and polymerases were hybridised to create circularised complexes according to manufacturer's instructions. The complexes were purified with the 1.2X clean up with SMRTbell beads. The purified complexes were then diluted to the Revio loading concentration (in the range 200–300 pM), and spiked with a Revio sequencing internal control. Samples were sequenced on Revio 25M SMRT cells (Pacific Biosciences, California, USA). The SMRT link software, a PacBio web-based end-to-end workflow manager, was used to set-up and monitor the run, as well as perform primary and secondary analysis of the data upon completion.

Hi-C

For Hi-C library preparation, DNA was fragmented using the Covaris E220 sonicator (Covaris) and size selected using SPRISelect beads to 400 to 600 bp. The DNA was then enriched using the Arima-HiC v2 kit Enrichment beads. Using the NEBNext Ultra II DNA Library Prep Kit (New England Biolabs) for end repair, A-tailing, and adapter ligation. This uses a custom protocol which resembles the standard NEBNext Ultra II DNA Library Prep protocol but where library preparation occurs while DNA is bound to the Enrichment beads. For library amplification, 10 to 16 PCR cycles were required, determined by the sample biotinylation percentage. The Hi-C sequencing was performed using paired-end sequencing with a read length of 150 bp on an Illumina NovaSeq X instrument.

Genome assembly, curation and evaluation

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 in Hi-C phasing mode (Cheng *et al.*, 2021; Cheng *et al.*, 2022), resulting in a pair of haplotype-resolved assemblies. The Hi-C reads were mapped to the primary contigs using bwa-mem2

(Vasimuddin *et al.*, 2019). The contigs were further scaffolded using the provided Hi-C data (Rao *et al.*, 2014) in YaHS (Zhou *et al.*, 2023) using 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), which runs MitoFinder (Allio *et al.*, 2020) and uses these annotations to select the final mitochondrial contig and to ensure the general quality of the sequence.

Assembly curation

The assembly was decontaminated using the Assembly Screen for Cobionts and Contaminants (ASCC) pipeline. Flat files and maps used in curation were generated via the TreeVal pipeline (Pointon *et al.*, 2023). Manual curation was conducted primarily in PretextView (Harry, 2022) and HiGlass (Kerpedjiev *et al.*, 2018), with additional insights provided by JBrowse2 (Diesh *et al.*, 2023). Scaffolds were visually inspected and corrected as described by Howe *et al.* (2021). Any identified contamination, missed joins, and mis-joins were amended, and duplicate sequences were tagged and removed. The curation process is documented at https://gitlab.com/wtsi-grit/rapid-curation.

Assembly quality assessment

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

A Hi-C contact map was produced for the final version of the assembly. The Hi-C reads were aligned using bwa-mem2 (Vasimuddin *et al.*, 2019) and the alignment files were combined using SAMtools (Danecek *et al.*, 2021). The Hi-C alignments were converted into a contact map using BEDTools (Quinlan & Hall, 2010) and the Cooler tool suite (Abdennur & Mirny, 2020). The contact map was visualised in HiGlass (Kerpedjiev *et al.*, 2018).

The blobtoolkit pipeline is a Nextflow (Di Tommaso *et al.*, 2017) port of the previous Snakemake Blobtoolkit pipeline (Challis *et al.*, 2020). It aligns the PacBio reads in SAMtools and minimap2 (Li, 2018) and generates coverage tracks for regions of fixed size. In parallel, it queries the GoaT database (Challis *et al.*, 2023) to identify all matching BUSCO lineages to run BUSCO (Manni *et al.*, 2021). For the three domain-level BUSCO lineages, the pipeline aligns the BUSCO genes to the UniProt Reference Proteomes database (Bateman *et al.*, 2023) with DIAMOND blastp (Buchfink *et al.*, 2021). The genome is also divided into chunks according to the density of the BUSCO genes from the closest taxonomic lineage, and each chunk is aligned to the UniProt Reference Proteomes database using DIAMOND blastx. Genome sequences without a hit are chunked using seqtk and aligned to the NT database with blastn (Altschul *et al.*, 1990). The blobtools suite combines all these outputs into a blobdir for visualisation.

The blobtoolkit pipeline was developed using nf-core tooling (Ewels *et al.*, 2020) and MultiQC (Ewels *et al.*, 2016), relying on the Conda package manager, the Bioconda initiative (Grüning *et al.*, 2018), the Biocontainers infrastructure (da Veiga Leprevost *et al.*, 2017), as well as the Docker (Merkel, 2014) and Singularity (Kurtzer *et al.*, 2017) containerisation solutions.

Table 4 contains a list of relevant software tool versions and sources.

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

Table 4. Software tools: versions and sources.

Software tool	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	4.3.9	https://github.com/blobtoolkit/blobtoolkit
BUSCO	5.5.0	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
DIAMOND	2.1.8	https://github.com/bbuchfink/diamond
fasta_windows	0.2.4	https://github.com/tolkit/fasta_windows
FastK	666652151335353eef2fcd58880bcef5bc2928e1	https://github.com/thegenemyers/FASTK
Gfastats	1.3.6	https://github.com/vgl-hub/gfastats

Software tool	Version	Source
GoaT CLI	0.2.5	https://github.com/genomehubs/goat-cli
Hifiasm	0.19.8-r603	https://github.com/chhylp123/hifiasm
HiGlass	44086069ee7d4d3f6f3f0012569789ec138f42b84 aa44357826c0b6753eb28de	https://github.com/higlass/higlass
MerquryFK	d00d98157618f4e8d1a9190026b19b471055b22e	https://github.com/thegenemyers/MERQURY.FK
Minimap2	2.24-r1122	https://github.com/lh3/minimap2
MitoHiFi	3	https://github.com/marcelauliano/MitoHiFi
MultiQC	1.14, 1.17, and 1.18	https://github.com/MultiQC/MultiQC
Nextflow	23.10.0	https://github.com/nextflow-io/nextflow
PretextView	0.2.5	https://github.com/sanger-tol/PretextView
samtools	1.19.2	https://github.com/samtools/samtools
sanger-tol/ascc	-	https://github.com/sanger-tol/ascc
sanger-tol/blobtoolkit	0.5.1	https://github.com/sanger-tol/blobtoolkit
Seqtk	1.3	https://github.com/lh3/seqtk
Singularity	3.9.0	https://github.com/sylabs/singularity
TreeVal	1.2.0	https://github.com/sanger-tol/treeval
YaHS	1.2a.2	https://github.com/c-zhou/yahs

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: Podarcis filfolensis (Filfola wall lizard). Accession number PRJEB73695; https://identifiers.org/ena.embl/PRJEB73695. The genome sequence is released openly for reuse. The *Podarcis filfolensis* genome assembly is provided by the Wellcome Sanger Institute Tree of Life Programme (https://www.sanger.ac.uk/programme/tree-of-life/) and is part of the Vertebrate Genomes Project (PRJNA489243). 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 Table 1 and Table 2.

Author information

Members of the Wellcome Sanger Institute Tree of Life Management, Samples and Laboratory team are listed here: https://doi.org/10.5281/zenodo.12162482.

Members of Wellcome Sanger Institute Scientific Operations: Sequencing Operations are listed here: https://doi.org/10.5281/zenodo.12165051.

Members of the Wellcome Sanger Institute Tree of Life Core Informatics team are listed here: https://doi.org/10.5281/zenodo.12160324.

Members of the Tree of Life Core Informatics collective are listed here: https://doi.org/10.5281/zenodo.12205391.

References

Abdennur N, Mirny LA: **Cooler: scalable storage for Hi-C data and other genomically labeled arrays.** *Bioinformatics.* 2020; **36**(1): 311–316.
PubMed Abstract | Publisher Full Text | Free Full Text

Allio R, Schomaker-Bastos A, Romiguier J, *et al.*: **MitoFinder: efficient automated large-scale extraction of mitogenomic data in target enrichment phylogenomics.** *Mol Ecol Resour.* 2020; **20**(4): 892–905.
PubMed Abstract | Publisher Full Text | Free Full Text

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 MJ, 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

Böhme W: **Handbuch der Reptilien und Amphibien Europas**. Aula-Verlag, 1986; **2/II**.
Reference Source

Buchfink B, Reuter K, Drost HG: **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, Kumar S, Sotero-Caio C, *et al.*: **Genomes on a Tree (GoaT): a versatile, scalable search engine for genomic and sequencing project metadata across the eukaryotic Tree of Life [version 1; peer review: 2 approved].** *Wellcome Open Res.* 2023; **8**: 24.
PubMed Abstract | Publisher Full Text | Free Full Text

Challis R, Richards E, Rajan J, *et al.*: **BlobToolKit – interactive quality assessment of genome assemblies.** *G3 (Bethesda).* 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

Cheng H, Jarvis ED, Fedrigo O, *et al.*: **Haplotype-resolved assembly of diploid genomes without parental data.** *Nat Biotechnol.* 2022; **40**(9): 1332–1335.
PubMed Abstract | Publisher Full Text | Free Full Text

da Veiga Leprevost F, Grüning BA, Alves Aflitos S, *et al.*: **BioContainers: an open-source and community-driven framework for software standardization.** *Bioinformatics.* 2017; **33**(16): 2580–2582.
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

Denton A, Oatley G, Cornwell C, *et al.*: **Sanger Tree of Life sample homogenisation: PowerMash.** *protocols.io.* 2023a.
Publisher Full Text

Denton A, Yatsenko H, Jay J, *et al.*: **Sanger Tree of Life wet laboratory protocol collection V.1.** *protocols.io.* 2023b.
Publisher Full Text

Di Tommaso P, Chatzou M, Floden EW, *et al.*: **Nextflow enables reproducible computational workflows.** *Nat Biotechnol.* 2017; **35**(4): 316–319.
PubMed Abstract | Publisher Full Text

Diesh C, Stevens GJ, Xie P, *et al.*: **JBrowse 2: a modular genome browser with views of synteny and structural variation.** *Genome Biol.* 2023; **24**(1): 74.
PubMed Abstract | Publisher Full Text | Free Full Text

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

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

Gabrielli M, Benazzo A, Biello R, *et al.*: **A high-quality reference genome for the critically endangered Aeolian wall lizard, *Podarcis raffonei*.** *J Hered.* 2023; **114**(3): 279–285.
PubMed Abstract | Publisher Full Text

Grüning B, Dale R, Sjödin A, *et al.*: **Bioconda: sustainable and comprehensive software distribution for the life sciences.** *Nat Methods.* 2018; **15**(7): 475–476.
PubMed Abstract | Publisher Full Text | Free Full Text

Harry E: **PretextView (Paired REad TEXTure Viewer): a desktop application for viewing pretext contact maps**. 2022.
Reference Source

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

Jay J, Yatsenko H, Narváez-Gómez JP, *et al.*: **Sanger Tree of Life sample preparation: triage and dissection.** *protocols.io.* 2023.
Publisher 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

Li H: **Minimap2: pairwise alignment for nucleotide sequences.** *Bioinformatics.* 2018; **34**(18): 3094–3100.
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

Meiri S, Avila L, Bauer AM, *et al.*: **The global diversity and distribution of lizard clutch sizes.** *Glob Ecol Biogeogr.* 2020; **29**(9): 1515–1530.
Publisher Full Text

Merkel D: **Docker: lightweight Linux containers for consistent development and deployment.** *Linux J.* 2014; **2014**(239): 2, [Accessed 2 April 2024].
Reference Source

Pointon DL, Eagles W, Sims Y, *et al.*: **sanger-tol/treeval v1.0.0 – Ancient Atlantis**. 2023.
Publisher Full Text

Quinlan AR, Hall IM: **BEDTools: a flexible suite of utilities for comparing genomic features.** *Bioinformatics.* 2010; **26**(6): 841–842.
PubMed Abstract | Publisher Full Text | Free Full Text

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

Salvi D, Schembri PJ, Sciberras A, *et al.*: **Evolutionary history of the Maltese wall lizard *Podarcis filfolensis*: insights on the 'Expansion–Contraction' model of Pleistocene biogeography.** *Mol Ecol.* 2014; **23**(5): 1167–1187.
PubMed Abstract | Publisher Full Text

Schembri PJ, Salvi D: ***Podarcis filfolensis*.** *The IUCN Red List of Threatened Species.* 2024; **2024**: e.T17793A137843120.
Publisher Full Text

Sciberras A, Schembri P: **Conservation status of the St Paul's Island wall lizard (*Podarcis filfolensis kieselbachi*).** *Herpetol Bull.* 2008; **105**: 28–34.
Reference Source

Strickland M, Cornwell C, Howard C: **Sanger Tree of Life fragmented DNA clean up: manual SPRI.** *protocols.io.* 2023a.
Publisher Full Text

Strickland M, Moll R, Cornwell C, *et al.*: **Sanger Tree of Life HMW DNA extraction: manual MagAttract.** *protocols.io.* 2023b.
Publisher Full Text

Todorovic M, Sampaio F, Howard C: **Sanger Tree of Life HMW DNA fragmentation: diagenode Megaruptor®3 for PacBio HiFi.** *protocols.io.* 2023.
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.** In: *2019 IEEE International Parallel and Distributed Processing Symposium (IPDPS).* IEEE, 2019; 314–324.
Publisher Full Text

Yang W, Feiner N, Pinho C, *et al.*: **Extensive introgression and mosaic genomes of Mediterranean endemic lizards.** *Nat Commun.* 2021; **12**(1): 2762.
PubMed Abstract | Publisher Full Text | Free 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

Open Peer Review

Current Peer Review Status: ? ?

Version 1

Reviewer Report 27 May 2025

https://doi.org/10.21956/wellcomeopenres.26526.r123199

?

Hardip R. Patel
Australian National University, Canberra, Canberra, Australia
Arthur Georges
University of Canberra, Canberra, Australia
Paul D. Waters
Faculty of Science, School of Biotechnology and Biomolecular Science, UNSW Sydney, Sydney, Australia

We found this genome note to be technically sound but found the communication to be wanting. Part of this deficiency arises from lack of clarity on what the genome note is designed to achieve. A genome note is an announcement of the availability of a new genome assembly made available concurrently with the assembly file becoming available in a public repository. In this the genome note succeeds. However, it should also describe the sequences and how they were generated and analysed in sufficient detail for one to be able to repeat the workflow or, if necessary, turn to the genome note for insight as to why subsequent use of the genome assembly has delivered unexpected results. It should highlight any novel approaches to the assembly for the benefit of others. In this the genome note does not succeed. More detail is required as in for example the description of the progressive gains from the hifiasm base assembly, the value add of the YAHS scaffolding, and then the manual curation required and the value add of that manual curation. Finally, the genome note should point to knowledge gaps where this assembly can be used to advance our knowledge and understanding. Pointers to the future.

We would like to see the genome note revised to bring out these sentient points. Otherwise, the value of the publication is not a lot greater than simply the submission of the assembly on NCBI.

Abstract
The abstract needs to be expanded to include (a) a brief sentence or two introducing the species, its karyotype and why it is important; (b) “a female ZW specimen” so that one can see immediately that the heterogametic sex has been sequenced; (c) a statement of the key benchmark statistics for the assembly against the standard set published by the Earth BioGenome Project; (d) a statement to the effect that the assembly is published without annotation; and (e) a final

concluding sentence or two on the value of having this assembly out. At the moment it is way too brief and skeletal.

Species Taxonomy
The NCBI abbreviated version would be more appropriate here. Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Lepidosauria; Squamata; Bifurcata; Unidentata; Episquamata; Laterata; Lacertibaenia; Lacertidae; *Podarcis filfolensis* (Bedriaga, 1876) [NCBI:txid65481]

Background
Notwithstanding the information being included in the abstract, the background needs to include a description of the karyotype and state clearly that the species has a ZZ/ZW system of sex determination, and that a female (ZW) animal has been sequenced.
The authors could consider being more specific in how the generation of this assembly will provide the basis for an acceleration of new knowledge in conservation and comparative studies, and provide opportunity to address challenges that would otherwise be intractable.

Genome Sequence Report
Consider using read depth in place of coverage, as the latter is unambiguous [throughout].

Approximately 33.0x? Do you mean approximately 33x? And in any case 49.37Gb/1.42113 gives 34.74x.
As outlined earlier, it would be nice to see how the assembly progressed in terms of base statistics from the base hifiasm assembly, to the YAHS assembly, to the final manual curated assembly. Also, more detail is required on what the manual curation involved.
What do you mean exactly by curated to chromosome level. Need to specify what you mean by chromosome level.

It is not clear what the final assembly is. Is it Haplotype 1? Why not the hifiasm pseudohaplotype? Have the Z and the W been added to this final assembly? How? There is a discrepancy between the length of the final assembly and the genomescope estimate. This warrants a comment.

The authors report the percentage of the assembly mapped to chromosomes, but the assembly has not been physically mapped. Should these be referred to as putative chromosomes?

Most of the assembly sequence (98.73%) was assigned to 20 chromosomal-level scaffolds, representing 18 autosomes and the W and Z sex chromosomes. Is this consistent with the karyotype? Be sure to add citations for the karyotype.
Not really enamored with Figure 2. This information might be better presented in a table. “A light grey spiral at the centre shows the cumulative scaffold count on a logarithmic scale” – not sure what this means. Where is the log scale? Very hard to interpret.
Figure 3. I really struggled to see the different colours for the smaller dots in this graph. Suggest using more definitive colours.

The paragraph immediately following Figure 3 needs a substantial rewrite.

“The *k*-mer recovery for haplotype 1 is 83.62%, and for haplotype 2 80.00%, while the combined haplotypes have a *k*-mer recovery of 99.43%.” This likely arises from the high heterozygosity. Inclusion of a GenomeScope figure would assist the reader here.

Methods
I was disappointed to see that the specimen used for the assembly was released. It should have been vouchered and lodged with a reputable museum.
An assembly often includes mitochondria sequences as multiple contigs. Were they removed? More clarity on mitohifi would help here. Was mitohifi run on the reads or the assembly?

“Detailed protocols are available on protocols.io (Denton *et al.*, 2023b)” – a quick perusal of these two papers shows this claim is not true. Need to add the quantity and type of the lysis buffer, quantity of the tissued used after dissection, etc. The specific protocol within these papers needs to be identified. Also confused as to how tissue 20-50 mg was used for the HiC when the lizard was released. Was the HiC from a different individual from the focal assembly individual?
20-50 mg of frozen tissue? But only one individual. Which was it, 20 mg or 50 mg?

“At a minimum, samples were required to have an average fragment size exceeding 8 kb and a total mass over 400 ng to proceed to the low input SMRTbell Prep Kit 3.0 protocol (Pacific Biosciences, California, USA), depending on genome size and sequencing depth required”. This is a generic statement presumably taken from the manufacturers guidelines. What did you actually do. Needs to be specific to the current study.
“The Hi-C alignments were converted into a contact map using BEDTools” Further details required here for reproducibility.

The blobkit pipeline – what did you use it for. There is a statement on what it can be used for, but what did you specifically use it for in this assembly?
“Any identified contamination, missed joins, and mis-joins were amended, and duplicate sequences were tagged and removed”. Can you provide details of how these decisions were made. For example, how is a mis-join identified?

The Hi-C reads were aligned using bwa-mem2’. The parameters used would be useful here.

“The genome will be annotated using available RNA-Seq data and presented through the Ensembl pipeline at the European Bioinformatics Institute”. Hmmm. Why no annotation here? Greatly limits the utility of the assembly and thus the resource note.

Is the rationale for creating the dataset(s) clearly described?
Partly

Are the protocols appropriate and is the work technically sound?
Yes

Are sufficient details of methods and materials provided to allow replication by others?
Partly

Are the datasets clearly presented in a useable and accessible format?
Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Comparative Genomics; Genomics of sex determination

We confirm that we have read this submission and believe that we have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however we have significant reservations, as outlined above.

Reviewer Report 05 May 2025

https://doi.org/10.21956/wellcomeopenres.26526.r122664

?

Terry Bertozzi
South Australian Museum, Adelaide, Australia

The authors report on the sequencing and genome assembly of *Podarcis filfolensis*, collected from the Island of Filfla. The genome size is within the range expected for squamates and appears to be of excellent quality and completeness based on the metrics provided. The methods used are appropriate for the scope of work and the datasets and metadata are well laid out. However, there are some minor issues that should be addressed to improve the clarity and repeatability of the work. I have addressed these below titled by their relevant sections.

Background
A quick search indicates there are between 4-6 subspecies of *Podarcis filfolensis.* I am not familiar enough with the taxonomy of this group to know the current recognised taxa, however, it makes me wonder why the authors chose to highlight some of the subspecies of *Podarcis* that occur in the area but exclude others?

What was the rationale for choosing this particular subspecies (*P. filfolensis filfolensis*) for genome sequencing?

Genome sequencing report
The manuscript text indicates that that Figure 4 illustrates the “completeness of the assembly”. This is incorrect. This is a graph that shows “the length of the assembly as scaffolds are added from the longest to the shortest”, accounting for the likely taxonomic origin of the reads. The purpose of Blobtools is to detect contamination. The graph is also confusing as the grey line which has its origin at 0 is hidden by the dark blue line, which starts at the same place. The plot axes are not specific enough and should contain the word “scaffold”. The legend also does not describe what the axes are – compare with figure 3.

Material and methods
I dislike the term “terminal body”. It doesn’t actually explain what tissue was sequenced. “Muscle

tissue from the tail tip" is much easier for a reader to understand and more explanatory. This should also be fixed in the next paragraph which states "Tissue from the terminal body of the rPodFil1 sample…". This makes no sense. What is the "terminal body" of a sample? Please correct in Table 1 as well.

In the "Hi-C sample preparation and crosslinking" section, it states that "20–50 mg of frozen tissue (stored at –80 °C)" was used, however, it was noted that the original tissue collected was fixed in ethanol. This is not consistent with the "Sample acquisition" section!

Please indicate how many "25M SMRT cells" were used to produce the PacBio HiFi sequence data.

The Blobtoolkit pipeline does NOT align PacBio reads in SAMtools as stated in the manuscript. The reads are aligned with minimap2 and SAMtools is used process the resultant BAM file and index them.

The last two paragraphs in the "Assembly quality assessment" section suggest that the Nextflow implementation of the original Snakemake pipeline was done as part of this study. Given that none of the authors of this manuscript seem to have contributed to the blobtoolkit github repo, I assume this is not the case. Therefore, please just describe how you used the tool in your work. The whole last paragraph (and associated references) can be deleted.

Data availability
Italicise the genus and species name.

Is the rationale for creating the dataset(s) clearly described?
Partly

Are the protocols appropriate and is the work technically sound?
Partly

Are sufficient details of methods and materials provided to allow replication by others?
Partly

Are the datasets clearly presented in a useable and accessible format?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Bioinformatics, genome assembly and annotation, evolutionary biology

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.