Journal of Heredity, 2023, **114**, 279–285
https://doi.org/10.1093/jhered/esad014
Advance access publication 2 March 2023
Genome Resources

Genome Resources

A high-quality reference genome for the critically endangered Aeolian wall lizard, *Podarcis raffonei*

Maëva Gabrielli[1,*], Andrea Benazzo[1,*], Roberto Biello[1], Lorena Ancona[2], Silvia Fuselli[1], Alessio Iannucci[3], Jennifer Balacco[4], Jacqueline Mountcastle[4], Alan Tracey[5], Gentile Francesco Ficetola[6,7], Daniele Salvi[8], Marco Sollitto[9], Olivier Fedrigo[4], Giulio Formenti[3,4], Erich D. Jarvis[4,10], Marco Gerdol[9], Claudio Ciofi[3], Emiliano Trucchi[2], Giorgio Bertorelle[1]

[1]Department of Life Sciences and Biotechnology, University of Ferrara, Ferrara, Italy,
[2]Department of Life and Environmental Sciences, Marche Polytechnic University, Ancona, Italy,
[3]Department of Biology, University of Florence, Florence, Italy,
[4]Vertebrate Genome Laboratory, The Rockefeller University, New York, NY, United States,
[5]Tree of Life, Wellcome Sanger Institute, Cambridge, United Kingdom,
[6]Department of Environmental Sciences and Policy, University of Milan, Milan, Italy,
[7]Laboratoire d'Ecologie Alpine (LECA), CNRS, Université Grenoble Alpes and Université Savoie Mont Blanc, Grenoble, France,
[8]Department of Health, Life & Environmental Sciences—University of L'Aquila, L'Aquila, Italy,
[9]Department of Life Sciences, University of Trieste, Trieste, Italy,
[10]Howard Hughes Medical Institute, Chevy Chase, MD, United States

[*]These authors contributed equally.
Address correspondence to M. Gabrielli at the address above, or e-mail: maeva.gab@hotmail.fr.

Corresponding Editor: Christopher Blair

Abstract

The Aeolian wall lizard, *Podarcis raffonei*, is an endangered species endemic to the Aeolian archipelago, Italy, where it is present only in 3 tiny islets and a narrow promontory of a larger island. Because of the extremely limited area of occupancy, severe population fragmentation and observed decline, it has been classified as Critically Endangered by the International Union for the Conservation of Nature (IUCN). Using Pacific Biosciences (PacBio) High Fidelity (HiFi) long-read sequencing, Bionano optical mapping and Arima chromatin conformation capture sequencing (Hi-C), we produced a high-quality, chromosome-scale reference genome for the Aeolian wall lizard, including Z and W sexual chromosomes. The final assembly spans 1.51 Gb across 28 scaffolds with a contig N50 of 61.4 Mb, a scaffold N50 of 93.6 Mb, and a BUSCO completeness score of 97.3%. This genome constitutes a valuable resource for the species to guide potential conservation efforts and more generally for the squamate reptiles that are underrepresented in terms of available high-quality genomic resources.

Key words: conservation genetics, de novo assembly, Endemixit, Hi-C, Lacertids, PacBio HiFi

Introduction

The Aeolian wall lizard *Podarcis raffonei* (Fig. 1A) is one of the most endangered vertebrate in Europe (Gippoliti et al. 2017). It is endemic to 4 islands of the Aeolian archipelago, located North-East of Sicily, with an extremely restricted distribution range including 3 islets less than 0.01 km^2 (La Canna, Scoglio Faraglione, Strombolicchio) and a larger island (Vulcano, 21.2 km^2) where it currently occupies nonetheless a very limited area (Bonardi et al. 2022). The total area of occupancy has been estimated to be as small as 5,000 m^2 (Ficetola et al. 2021) and the total population size is estimated at about 2,000 individuals (Capula et al. 2002; Lo Cascio et al. 2014; Gippoliti et al. 2017; Ficetola et al. 2018, 2021). As a result, the Aeolian wall lizard has been listed as Critically Endangered in the Red List of Endangered Species of the IUCN (2009).

The main threats to its survival include interactions with the invasive Italian wall lizard *Podarcis siculus*, combined with habitat degradation (Capula et al. 2002). This is particularly visible on the island Vulcano, where the intense habitat change that occurred in the last 50 yr may have favored the spread of *P. siculus*, leading to a sharp decline in the *P. raffonei* population (Capula et al. 2002).

The production of highly contiguous genomes has greatly accelerated in the last decade, refining our understanding of the genomic basis of organismal traits, the chromosome evolution, and allowing the detection of natural selection through genomic scans (Geneva et al. 2022). Furthermore, reference genomes can

Received January 10, 2023; Accepted March 1, 2023

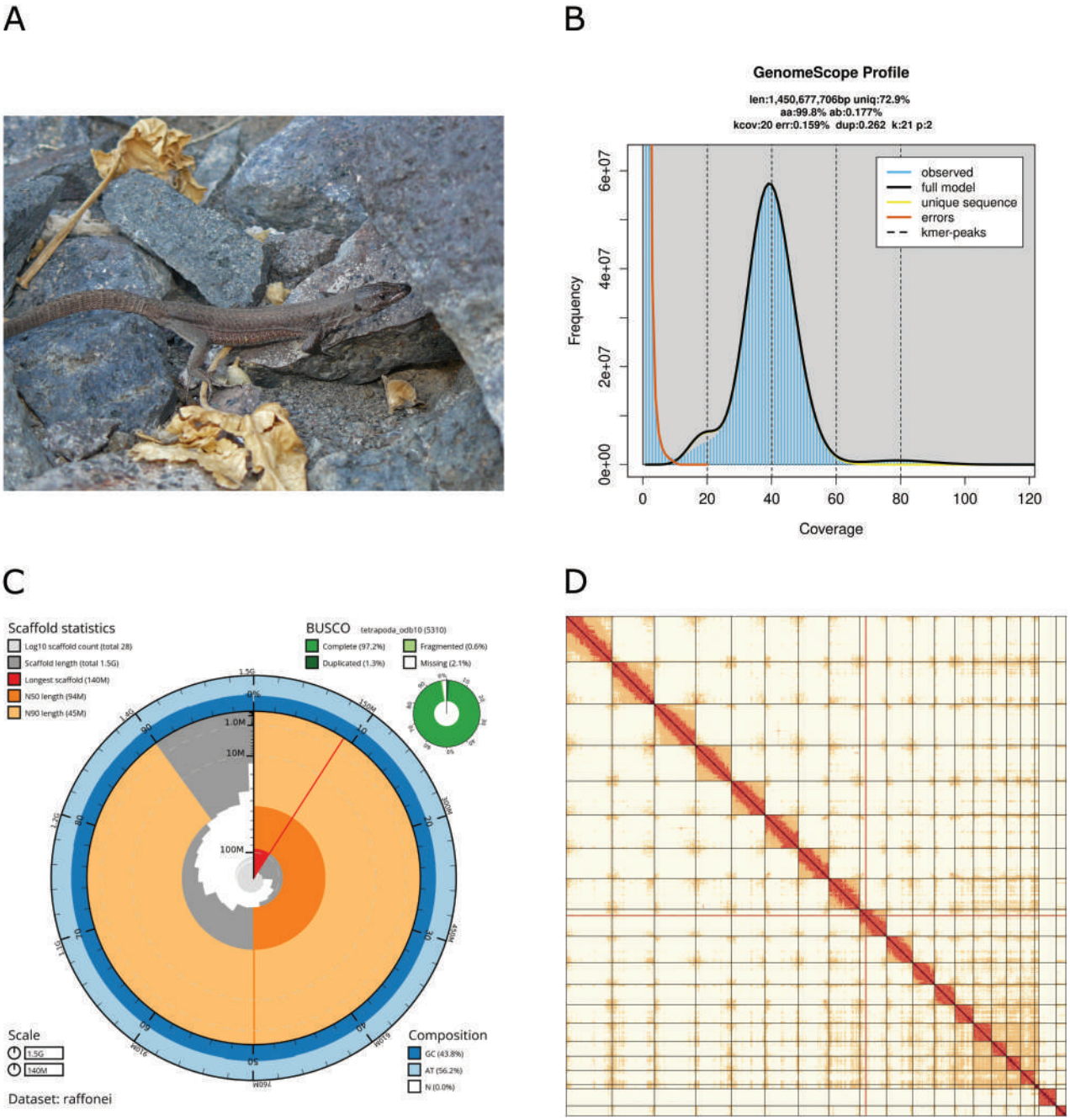

Fig. 1. A) Photography of an individual of *Podarcis raffonei*, on La Canna stack (Photo credit: Daniele Salvi), and visual overview of genome assembly metrics. B) K-mer spectra output and corresponding genome size and heterozygosity estimated with GenomeScope 2.0. C) BlobToolKit Snail plot showing a graphical representation of the quality metrics presented in Table 2 for the *Podarcis raffonei* primary assembly (rPodRaf1.pri). D) Hi-C contact map for the 20 scaffolds of the primary genome assembly generated with PretextSnapshot.

be key for conservation genomics as they may permit, in combination with whole-genome resequencing data, to assess genetic diversity, investigate inbreeding depression, or characterize deleterious mutations (Formenti et al. 2022b). High-quality reference genomes are unevenly distributed across the tree of life, and some clades, such as the squamate reptiles, are underrepresented (Pinto et al. 2022; Card et al. 2023).

Here, we present a high-quality chromosome-scale reference genome for the Aeolian wall lizard, produced as part of the Endemixit project (www.endemixit.com). Our final genome assembly spans 1.51 Gb across 28 scaffolds, with a scaffold N50 of 93.6 Mb and a BUSCO completeness score of 97.3%. This high-quality reference genome is a valuable resource to assess the genetic diversity in the 4 extant populations of the Aeolian wall lizard and better develop the conservation strategy for this species.

Methods

Biological materials

An adult female was collected on the 31st of July 2020 by D. Salvi on the stack of La Canna (38°34′56.13″N to

14°31′16.61″E; see Supplementary Fig. 1), in the Aeolian archipelago, in a small terrace at 50 m a.s.l. on the eastern slope of the stack, reached by climbing with the technical assistance of the mountain guide Lorenzo Inzigneri. A piece of tail was cut and immediately frozen in liquid nitrogen until the final storage at −80 °C.

Nucleic acid extraction, library preparation, and sequencing

All the following steps were carried out at the Vertebrate Genomes Project (VGP, https://vertebrategenomesproject.org/) lab. High molecular weight (HMW) DNA was extracted from muscle with the Circulomics HMW DNA extraction standard TissueRuptor protocol with the Nanobind Tissue Big DNA Kit (PN NB-900-701-01). DNA absorbance was checked as quality and purity control with Nanodrop and average fragment length was verified with a Pulsed Field Gel Electrophoresis (PFGE).

Genomic data from 3 different sequencing technologies were used for the assembly: Pacific Biosciences (PacBio) High Fidelity (HiFi) reads, Bionano optical maps, and Hi-C reads from Arima Genomics.

PacBio HiFi libraries were prepared using the Pacific Biosciences Express Template Prep Kit 2.0. The library was then size selected (>10 kb) using the Circulomics Short Read Eliminator. The PacBio library was sequenced on 2 PacBio 8M v3 SMRT Cells on a PacBio Sequel II and 1 PacBio 8M SMRT Cell on a PacBio Sequel IIe using the sequencing kit 2.0 and a 30-h movie.

An aliquot of the HMW DNA was labeled for Bionano Genomics optical mapping using the Bionano Prep Direct Label and Stain (DLS) Protocol and run on 1 Saphyr instrument chip flowcell.

Hi-C libraries were generated by Arima Genomics (https://arimagenomics.com/) using muscle in vivo cross-linking with the Arima-HiC kit with 2-enzyme proximity ligation. Proximally ligated DNA was subjected to shearing, size selection (~200 to 600 bp) with SPRI beads, and enrichment with streptavidin beads for the biotin-labeled DNA. KAPA Hyper Prep kit was employed to generate libraries compatible with Illumina technologies. Libraries were amplified through PCR, purified with SPRI beads and sequenced on an Illumina HiSeq X (~60× coverage) after a quality check with Bioanalyzer and qPCR.

Nuclear genome assembly

The genome of the Aeolian wall lizard was assembled following the VGP assembly pipeline v2.0 (Rhie et al. 2021), as outlined in Table 1. Briefly, PacBio HiFi long reads were processed using hifiasm (Cheng et al. 2021, 2022) producing a set of primary contigs representing the initial haploid assembly and separating alternative haplotypic variants. Primary contigs were then processed with purge_dups (Guan et al. 2020) to identify residual haplotype duplication in the assembly. Such duplicated sequences were moved to the alternate assembly that was then exposed to a second round of purge_dups to obtain the final set of nonredundant haplotypic variants. Primary contigs were anchored to scaffolds using Bionano optical maps, adjusting the gap size according to the observed optical distance with the bionano_solve pipeline v3.6.1 (Chan et al. 2018). A second round of scaffolding was performed using Hi-C data. Paired-end reads were aligned to the primary assembly using the Arima genomics' pipeline (https://github.com/ArimaGenomics/mapping_pipeline) and the obtained contact data were used to guide the scaffolding procedure using salsa2 (Ghurye et al. 2017, 2019). Hi-C contact maps were generated and visually inspected using PretextSuite (https://github.com/wtsi-hpag/PretextView; https://github.com/wtsi-hpag/PretextMap; https://github.com/wtsi-hpag/PretextSnapshot) before and after the last scaffolding step. The resulting primary and alternate assemblies were screened for residual contaminations (Howe et al. 2021) and manual curation was performed on the primary assembly using the gEVAL browser release 73 (Howe et al. 2021), PretextView and HiGlass (Kerpedjiev et al. 2018) to anchor scaffolds to chromosomes and check their coherence.

Genome size estimation and quality assessment

We estimated the genome size from the PacBio HiFi reads using a k-mer-based approach. The distribution of k-mers of length 21 was generated using meryl v1.3 (Miller et al. 2008) and GenomeScope 2.0 (Ranallo-Benavidez et al. 2020) was subsequently used to infer the genome length, genome-wide heterozygosity, and error rate.

We assessed the quality of our genome assembly using 2 independent methods. First, we used the BUSCO quality control tool to check for genome completeness using a set of conserved single-copy orthologous genes. We ran BUSCO v5.3.2 (Manni et al. 2021) in the genome mode with default parameters on the tetrapod dataset (tetrapoda_odb10) that contains 5,310 orthologous genes. Second, we used Mercury v1.3 (Rhie et al. 2020) to estimate the base level accuracy (QV) and the assembly completeness comparing the k-mers in the assembly and those observed in the HiFi reads. All assembly metrics were computed using gfastats v1.2.3 (Formenti et al. 2022a).

Identification of repetitive elements and gene annotation

To identify repetitive elements, we first generated a de novo repeat library using the Extensive de novo TE Annotator (EDTA) v1.9.9 (Ou et al. 2019) and DeepTE (Yan et al. 2020) to refine classifications within this library. We then used the final library to mask the genome with RepeatMasker v4.1.2 (Smit et al. n.d.). We used the same pipeline to identify repeats in the genome of *Podarcis muralis* (assembly PodMur_1.0; Andrade et al. 2019).

For gene prediction, we first downloaded RNA-seq reads available on NCBI from various tissues of closely related species (4 species of the genus *Podarcis*; see Supplementary Table 1). Quality control and trimming for adapters and low-quality bases (quality score <20) of the raw reads were performed using fastqc v0.11.8 (Andrews 2010) and TrimGalore v0.5.0 (https://github.com/FelixKrueger/TrimGalore), respectively. High-quality reads were then mapped to the soft-masked assembly with hisat2 v2.1.0 (Kim et al. 2015), and sorted with samtools v1.10 (Li et al. 2009). All the BAM files were filtered to remove invalid splice junctions with Portcullis v1.1.2 (Mapleson et al. 2018). Filtered RNA-seq alignments were passed to Braker v2.1.6 (Hoff et al. 2016, 2019), together with amino acid sequences of the whole exome of 22 closely related species from the order Squamata belonging to 11 families including 3 Lacertidae (*P. muralis*, *Lacerta agilis*, and *Zootoca vivipara*; see Supplementary Table 2). The Braker gene prediction pipeline was run with the options "--softmasking --prg=gth --gth2traingenes." The resulting

Table 1. Pipeline and software used for the genome assembly.

Assembly	Software	Version
K-mer counting	Meryl	1.3
Estimation of genome size and heterozygosity	GenomeScope2	2.0
De novo assembly (contigging)	HiFiasm	0.16.1-r375
Remove low-coverage, duplicated contigs	purge_dups	1.2.5
Scaffolding		
Bionano scaffolding	bionano_solve	3.6.1
Hi-C mapping for SALSA	Arima Genomics mapping pipeline	Commit 2e74ea4
Hi-C scaffolding	salsa2	2.3
Hi-C contact map generation		
Short-read alignment	Bwa	0.7.17
SAM/BAM processing	Samtools	1.10
Pairs processing	Bedtools	2.30
Contact map visualization	PretextView	0.2.2
	PretextMap	0.1.8
	PretextSnapshot	0.0.4
Genome assembly refinement		
Manual curation and contamination screening	gEVAL	Release 73
Genome quality assessment		
Basic assembly metrics	Gfastats	1.2.3
Assembly completeness	BUSCO	5.3.2
	Merqury	1.3
Repeat element identification		
Repeat identification	EDTA	1.9.9
	DeepTE	Commit babd65e
Repeat annotation	RepeatMasker	4.1.2
Gene annotation		
RNA-seq read quality control	Fastqc	0.11.8
	TrimGalore	0.5.0
Mapping RNA-seq reads genome	hisat2	2.1.0
Filtering splice junctions	Portcullis	1.1.2
Gene prediction	Braker	2.1.6
Comparison to *P. Muralis*		
Genome–genome alignment	minimap2	2.22
Synteny visualization	Circos	0.69-8

gene set was further filtered by evidence, keeping only gene predictions supported by RNA-seq or protein evidence using a BRAKER2 script (selectSupportedSubsets.py). The completeness of the final gene set was checked with BUSCO v5.3.2 (Manni et al. 2021) using the longest transcript of each gene as the representative transcript.

Mitochondrial genome sequencing and assembly

To characterize the entire sequence of the mitochondrial DNA via Sanger sequencing, we designed 4 different, and partially overlapping, amplicons of expected length between 4 and 7.3 kb. Primers were designed based on mitochondrial DNA sequences of congeneric species (*P. siculus* NC_011609.1, *P. muralis* NC_011607 and NC_011609). Amplifications were carried out starting from 50 ng of extracted DNA, in a 50 µL reaction with 0.2 µM primers and 1.25 u of PrimeSTAR GXL DNA Polymerase. Amplification primers and additional internal primers were used for Sanger sequencing reactions (see Supplementary Table 3). Fragments were visually inspected and manually assembled to reconstruct the mitochondrial sequence.

Comparative analyses with *P. muralis*

We performed a synteny comparison with the *P. muralis* assembly (PodMur_1.0; Andrade et al. 2019), the only chromosome-scale assembly presently available for the *Podarcis* genus. Phylogenetic reconstructions based on whole-genome data suggest that the 2 species diverged ~18 Mya during Miocene (Yang et al. 2021). We used minimap2 (Li 2018) to map the genome assembly of *P. raffonei* to the genome reference of *P. muralis* allowing a maximum sequence divergence of 5% (parameter -x asm20). We then filtered the alignment by mapping quality (>60) and length of the mapped fragments (>1 Mb) and plotted the alignment between the 18 autosomes and Z sexual chromosome (the W chromosome being absent from the *P. muralis* assembly) using Circos v0.69-8 (Krzywinski et al. 2009). Synteny between the 2 species was finally used to annotate the scaffolds of the *P. raffonei* assembly as chromosomes.

Results

The final genome size (1.51 Gb) is in agreement with the size estimated from the k-mer analysis with GenomeScope 2.0

(Fig. 1B) and very close to the genome size of *P. muralis* (1.51 Gb, Andrade et al. 2019). The k-mer spectrum shows a bimodal distribution with 2 major peaks, at ~20- and ~40-fold coverage, corresponding to heterozygous and homozygous states, respectively. Based on PacBio HiFi reads, we estimated a 0.159% sequencing error rate and a 0.177% nucleotide heterozygosity rate (Fig. 1B). The mitochondrial genome size is 17,038 bp, in agreement with the mitochondrial genome size of other species of *Podarcis* (17,311 bp for *P. muralis* and 17,297 bp for *P. siculus*; Podnar et al. 2009). The primary assembly contains 28 scaffolds for a total length of 1.51 Gb, with a contig N50 of 61.4 Mb, a scaffold N50 of 93.6 Mb, a longest contig size of 104.8 Mb, and a longest scaffold size of 139.1 Mb (Table 2; Fig. 1C). The alternate assembly contains 4,811 scaffolds spanning 182 Mb, having a N50 of 38.4 kb.

Table 2. Genome assembly statistics.

Measure	rPodRaf1
Total length	1.513 Gb
Number of scaffolds	28
Scaffold L50/N50	7 scaffolds; 93.6 Mb
Longest scaffold	139.1 Mb
Number of contigs	53
Contig L50/N50	10 contigs; 61.4 Mb
Longest contig	104.8 Mb
BUSCO completeness	97.3%
Single copy	5,095
Duplicated	67
Fragmented	34
Missing	114
Total	5,310

Fig. 2. Comparison of the chromosomal structure between the 18 autosomal chromosomes and Z chromosome between *P. raffonei* (right) and *P. muralis* (left). The different colors correspond to the different chromosomes of *P. raffonei*. The chromosomes were aligned using minimap2 and the resulting alignment between fragments longer than 1 Mb is represented with a ribbon plot using Circos.

This assembly is highly contiguous, as shown in the Hi-C contact map (Fig. 1D), with the 20 first scaffolds being of chromosome length and corresponding to the 18 autosomes and the 2 sexual chromosomes Z and W (see Supplementary Table 4). The sequencing depth of the HiFi reads along chromosomes is approximately uniform and does not reveal discrepancies in the assembly (see Supplementary Fig. 2). The completeness of the assembly is very high, with a BUSCO completeness score of 97.3% ([Single copy: 96.0%, Duplicated: 1.3%], Fragmented: 0.6%, Missing: 2.1%) using the tetrapod gene set and a k-mer completeness of 99.5%. Per base quality (QV) as estimated by Merqury is 62, corresponding to less than 1 incorrect nucleotide per megabase.

In total, 22,463 protein-coding genes were predicted. The BUSCO completeness of the gene annotation using the same tetrapod gene set was 92.1% ([Single copy: 91.1%, Duplicated: 1.0%], Fragmented: 3.9%, Missing: 4.0%). The identification of repetitive elements resulted in a 48.2% repeat content, falling within the range of repeat contents for other squamate species (24.4% to 73.0%; Pasquesi et al. 2018). In Lacertidae and Teiidae, the repeat content was estimated to be 45.1% and 44.5%for *P. muralis* and *Salvator merianae* (Roscito et al. 2018), respectively (see Supplementary Tables 5 and 6). The major class of repetitive elements was constituted by LTR elements and DNA transposons (see Supplementary Table 5).

The alignment of the genomes of *P. muralis* and *P. raffonei* revealed a very high congruency in the chromosomal organization (Fig. 2). The only chromosomal segment that did not map to the homologous chromosome from the other species was a 1.5 Mb segment of the chromosome 2 of *P. raffonei* that mapped to the chromosome 18 of *P. muralis*. We analyzed the depth of coverage profile and the reads mapping in the edges of this segment of chromosome 2 in *P. raffonei* and did not find any discrepancies in the assembly (see Supplementary Fig. 3). The 2 species have a similar number of genes (24,656 protein-coding genes were predicted in *P. muralis*; Andrade et al. 2019).

Discussion

We present here the first chromosome-scale genome assembly for the Aeolian wall lizard (scaffold N50 of 93.6 Mb). Several metrics indicate that our genome assembly possesses a very high quality being chromosome-scale, accurate and complete. It constitutes a useful resource for squamates, a group composed of ~11,000 species for which only 29 high-quality genome assemblies are currently available (Card et al. 2023). In comparison to the other squamates, the *P. raffonei* assembly has a high scaffold N50 and the highest BUSCO completeness score (see Supplementary Table 6).

The alignment between the genomes of *P. raffonei* and *P. muralis* showed a very high synteny, suggesting that both assemblies are structurally accurate and that the 2 species share a very similar chromosomal organization. Only 1 segment of the chromosome 2 of *P. raffonei* mapped to the chromosome 18 of *P. muralis*. This finding could be a biological chromosomal rearrangement between these 2 species (that belong to distinct clades of the genus *Podarcis*; Salvi et al. 2021; Yang et al. 2021) or a disjunction in the genome assembly of *P. muralis*.

The genome assembly of the Aeolian wall lizard, one of the most endangered vertebrate species in Europe, is a useful resource to better plan conservation efforts. Previous studies have highlighted that the Aeolian wall lizard exhibits low levels of genetic diversity and that the populations inhabiting different islands show a very reduced gene flow, constituting additional threats to this species (Capula 2004). Accordingly, our genome assembly suggests a very low heterozygosity (0.177% as estimated by GenomeScope), the lowest value documented among 7 species belonging to distinct squamate families (see Supplementary Table 6). The genome resequencing of several individuals from different islands is in progress to comprehensively characterize the genetic diversity of this species and evaluate its extinction risk.

Supplementary material

Supplementary material is available at *Journal of Heredity* online.

Acknowledgments

We thank the mountain guide Lorenzo Inzigneri for its help in the sampling in La Canna. Sampling authorization was given under Prot. 45382, Rif. 35644/2019.

Funding

This work was supported by the University of Ferrara (Italy) and funded by the MIUR PRIN 2017 grant 201794ZXTL to GB.

Data availability

Raw sequencing data, primary genome assembly, and mitochondrial DNA sequence are available under NCBI BioProject PRJNA916649 and PRJNA839511. Gene annotations and RepeatMasker output are available on https://zenodo.org/record/7473296.

References

Andrade P, Pinho C, Pérez i de Lanuza G, Afonso S, Brejcha J, Rubin C-J, Wallerman O, Pereira P, Sabatino SJ, Bellati A, et al. Regulatory changes in pterin and carotenoid genes underlie balanced color polymorphisms in the wall lizard. *Proc Natl Acad Sci USA*. 2019;116(12):5633–5642.

Andrews S. FastQC: a quality control tool for high throughput sequence data. 2010. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/. [accessed 2022 January 8].

Bonardi A, Francesco Ficetola G, Razzetti E, Canedoli C, Falaschi M, Parrino EL, Rota N, Padoa-Schioppa E, Roberto S. ReptIslands: Mediterranean islands and the distribution of their reptile fauna. *Glob Ecol Biogeogr*. 2022;31(5):840–847.

Capula M. Low genetic variation in a critically endangered Mediterranean lizard: conservation concerns for *Podarcis raffonei* (Reptilia, Lacertidae). *Ital J Zool*. 2004;71(suppl 1):161–166.

Capula M, Luca L, Marco AB, Arianna C. The decline of the Aeolian wall lizard, *Podarcis raffonei*: causes and conservation proposals. *Oryx*. 2002;36(1):66–72.

Card DC, Bryan Jennings W, Scott VE. Genome evolution and the future of phylogenomics of non-avian reptiles. *Animals*. 2023;13(3):471.

Chan S, Lam E, Saghbini M, Bocklandt S, Hastie A, Cao H, Holmlin E, Borodkin M. Structural variation detection and analysis using Bionano optical mapping. In: Bickhart DM, editor. *Copy number variants: methods and protocols*. Methods in Molecular Biology. New York (NY): Springer; 2018. p. 193–203.

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(2):170–175.

Cheng H, Jarvis ED, Fedrigo O, Koepfli K-P, Urban L, Gemmell NJ, Li H. Haplotype-resolved assembly of diploid genomes without parental data. *Nat Biotechnol*. 2022;40(9):1332–1335.

Ficetola GF, Barzaghi B, Melotto A, Muraro M, Lunghi E, Canedoli C, Lo Parrino E, Nanni V, Silva-Rocha I, Urso A, et al. N-mixture models reliably estimate the abundance of small vertebrates. *Sci Rep*. 2018;8(1):10357.

Ficetola GF, Silva-Rocha I, Carretero MA, Vignoli L, Sacchi R, Melotto A, Scali S, Salvi D. Status of the largest extant population of the critically endangered Aeolian lizard *Podarcis raffonei* (Capo Grosso, Vulcano Island). *PLoS One*. 2021;16(6):e0253631.

Formenti G, Abueg L, Brajuka A, Brajuka N, Gallardo-Alba C, Giani A, Fedrigo O, Jarvis ED. Gfastats: conversion, evaluation and manipulation of genome sequences using assembly graphs. *Bioinformatics*. 2022a:38(17):4214–4216.

Formenti G, Theissinger K, Fernandes C, Bista I, Bombarely A, Bleidorn C, Ciofi C, Crottini A, Godoy JA, Höglund J, et al.; European Reference Genome Atlas (ERGA) Consortium. The era of reference genomes in conservation genomics. *Trends Ecol Evol*. 2022b:37(3):197–202.

Geneva AJ, Park S, Bock DG, de Mello PLH, Sarigol F, Tollis M, Donihue CM, Reynolds RG, Feiner N, Rasys AM, et al. Chromosome-scale genome assembly of the brown Anole (*Anolis sagrei*), an emerging model species. *Commun Biol*. 2022;5(1):1–13.

Ghurye J, Pop M, Koren S, Bickhart D, Chin C-S. Scaffolding of long read assemblies using long range contact information. *BMC Genomics*. 2017;18(1):527.

Ghurye J, Rhie A, Walenz BP, Schmitt A, Selvaraj S, Pop M, Phillippy AM, Koren S. Integrating Hi-C links with assembly graphs for chromosome-scale assembly. *PLoS Comput Biol*. 2019;15(8):e1007273.

Gippoliti S, Capula M, Ficetola GF, Salvi D, Andreone F. Threatened by legislative conservationism? The case of the critically endangered Aeolian lizard. *Front Ecol Evol*. 2017;5:130. doi:10.3389/fevo.2017.00130

Guan D, McCarthy SA, Wood J, Howe K, Wang Y, Durbin R. Identifying and removing haplotypic duplication in primary genome assemblies. *Bioinformatics*. 2020;36(9):2896–2898.

Hoff KJ, Lange S, Lomsadze A, Borodovsky M, Stanke M. BRAKER1: unsupervised RNA-Seq-based genome annotation with GeneMark-ET and AUGUSTUS. *Bioinformatics*. 2016;32(5):767–769.

Hoff KJ, Lomsadze A, Borodovsky M, Stanke M. Whole-genome annotation with BRAKER. *Methods Mol Biol*. 2019;1962:65–95.

Howe K, Chow W, Collins J, Pelan S, Pointon D-L, Sims Y, Torrance J, Tracey A, Wood J. Significantly improving the quality of genome assemblies through curation. *GigaScience*. 2021;10(1):giaa153.

IUCN. *Podarcis raffonei*. The IUCN Red List of Threatened Species; 2009;e.T61552A12514822.

Kerpedjiev P, Abdennur N, Lekschas F, McCallum C, Dinkla K, Strobelt H, Luber JM, Ouellette SB, Azhir A, Kumar N, et al. HiGlass: web-based visual exploration and analysis of genome interaction maps. *Genome Biol*. 2018;19(1):125.

Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods*. 2015;12(4):357–360.

Krzywinski MI, Schein JE, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA. Circos: an information aesthetic for comparative genomics. *Genome Res*. 2009;19(9):1639–1645. doi:10.1101/gr.092759.109

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

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R; 1000 Genome Project Data Processing

Subgroup. The sequence alignment/map format and SAMtools. *Bioinformatics*. 2009;25(16):2078–2079.

Lo Cascio P, Grita F, Guarino L, Speciale C. A little is better than none: new insights into the natural history of the Aeolian wall lizard *Podarcis raffonei* from La Canna stack (Squamata Sauria). *Naturalista sicil*. 2014;38(2):355–366.

Manni M, Berkeley MR, Seppey M, Simão FA, Zdobnov EM. 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.

Mapleson D, Venturini L, Kaithakottil G, Swarbreck D. Efficient and accurate detection of splice junctions from RNA-Seq with Portcullis. *GigaScience*. 2018;7(12):giy131. doi:10.1093/gigascience/giy131

Miller JR, Delcher AL, Koren S, Venter E, Walenz BP, Brownley A, Johnson J, Li K, Mobarry C, Sutton G. Aggressive assembly of pyrosequencing reads with mates. *Bioinformatics*. 2008;24(24):2818–2824.

Ou S, Su W, Liao Y, Chougule K, Agda JRA, Hellinga AJ, Lugo CSB, Elliott TA, Ware D, Peterson T, et al. Benchmarking transposable element annotation methods for creation of a streamlined, comprehensive pipeline. *Genome Biol*. 2019;20(1):275.

Pasquesi GIM, Adams RH, Card DC, Schield DR, Corbin AB, Perry BW, Reyes-Velasco J, Ruggiero RP, Vandewege MW, Shortt JA, et al. Squamate reptiles challenge paradigms of genomic repeat element evolution set by birds and mammals. *Nat Commun*. 2018;9(1):2774.

Pinto BJ, Keating SE, Nielsen SV, Scantlebury DP, Daza JD, Gamble T. Chromosome-level genome assembly reveals dynamic sex chromosomes in neotropical leaf-litter geckos (Sphaerodactylidae: Sphaerodactylus). *J Hered*. 2022;113(3):272–287.

Podnar M, Pinsker W, Mayer W. Complete mitochondrial genomes of three lizard species and the systematic position of the Lacertidae (Squamata). *J Zool Syst Evol Res*. 2009;47(1):35–41.

Ranallo-Benavidez TR, Kamil SJ, Michael CS. GenomeScope 2.0 and Smudgeplot for reference-free profiling of polyploid genomes. *Nat Commun*. 2020;11(1):1432.

Rhie A, McCarthy SA, Fedrigo O, Damas J, Formenti G, Koren S, Uliano-Silva M, Chow W, Fungtammasan A, Kim J, et al. Towards complete and error-free genome assemblies of all vertebrate species. *Nature*. 2021;592(7856):737–746.

Rhie A, Walenz BP, Koren S, Phillippy AM. Merqury: reference-free quality, completeness, and phasing assessment for genome assemblies. *Genome Biol*. 2020;21(1):245.

Roscito JG, Sameith K, Pippel M, Francoijs K-J, Winkler S, Dahl A, Papoutsoglou G, Myers G, Hiller M. The Genome of the tegu lizard *Salvator merianae*: combining Illumina, PacBio, and optical mapping data to generate a highly contiguous assembly. *GigaScience*. 2018;7(12):giy141.

Salvi D, Pinho C, Mendes J, James Harris D. Fossil-calibrated time tree of Podarcis wall lizards provides limited support for biogeographic calibration models. *Mol Phylogenet Evol*. 2021;161:107169.

Smit AFA, Hubley R, Green P. RepeatMasker. n.d. http://repeatmasker.org. [accessed 2022 March 1].

Yan H, Bombarely A, Li S. DeepTE: a computational method for de novo classification of transposons with convolutional neural network. *Bioinformatics*. 2020;36(15):4269–4275.

Yang W, Feiner N, Pinho C, While GM, Kaliontzopoulou A, James Harris D, Salvi D, Uller T. Extensive introgression and mosaic genomes of Mediterranean endemic lizards. *Nat Commun*. 2021;12(1):2762.