




DATA NOTE

The genome sequence of the common toad, *Bufo bufo* (Linnaeus, 1758) [version 1; peer review: 2 approved]

Jeffrey W. Streicher ¹, Wellcome Sanger Institute Tree of Life programme, Wellcome Sanger Institute Scientific Operations: DNA Pipelines collective, Tree of Life Core Informatics collective, Darwin Tree of Life Consortium

¹Department of Life Sciences, Natural History Museum, London, UK

V1 First published: 20 Oct 2021, 6:281
<https://doi.org/10.12688/wellcomeopenres.17298.1>
Latest published: 20 Oct 2021, 6:281
<https://doi.org/10.12688/wellcomeopenres.17298.1>

Abstract

We present a genome assembly from an individual male *Bufo bufo* (the common toad; Chordata; Amphibia; Anura; Bufonidae). The genome sequence is 5.04 gigabases in span. The majority of the assembly (99.1%) is scaffolded into 11 chromosomal pseudomolecules. Gene annotation of this assembly by the NCBI Eukaryotic Genome Annotation Pipeline has identified 21,517 protein coding genes.

Keywords

Bufo bufo, common toad, genome sequence, chromosomal



This article is included in the [Tree of Life gateway](#).

Open Peer Review

Reviewer Status  

Invited Reviewers

1

2

version 1

20 Oct 2021



report



report

1. **Zexi Cai** , Aarhus University, Tjele, Denmark
2. **H. Bradley Shaffer**, University of California, Los Angeles, Los Angeles, USA

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

Corresponding author: Darwin Tree of Life Consortium (mark.blaxter@sanger.ac.uk)

Author roles: Streicher JW: Investigation, Resources, Writing – Original Draft Preparation, Writing – Review & Editing;

Competing interests: No competing interests were disclosed.

Grant information: This work was supported by the Wellcome Trust through core funding to the Wellcome Sanger Institute (206194) and the Darwin Tree of Life Discretionary Award (218328).

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2021 Streicher JW *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Streicher JW, Wellcome Sanger Institute Tree of Life programme, Wellcome Sanger Institute Scientific Operations: DNA Pipelines collective *et al.* **The genome sequence of the common toad, *Bufo bufo* (Linnaeus, 1758) [version 1; peer review: 2 approved]** Wellcome Open Research 2021, 6:281 <https://doi.org/10.12688/wellcomeopenres.17298.1>

First published: 20 Oct 2021, 6:281 <https://doi.org/10.12688/wellcomeopenres.17298.1>

Species taxonomy

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Neobatrachia; Hyloidea; Bufonidae; Bufo; *Bufo bufo* Linnaeus 1758 (NCBI:txid8384).

Introduction

The common toad, *Bufo bufo* (Anura: Bufonidae) is widely distributed throughout Europe. It has a biphasic life cycle that includes aquatic, benthic larvae and terrestrial adults. Bufonids like *B. bufo* are notable amongst anurans in that they (1) lack maxillary teeth, (2) have Bidder's organs, and (3) have paired paratoid glands that contain alkaloid toxins. *Bufo bufo* has been used extensively in comparative vertebrate research including as a model system in sensory biology (Ewert, 1974).

Based on populations from mainland Europe, the nuclear genome size of *B. bufo* was previously estimated to be between 5.82 and 7.75 picograms (= 5.69 and 7.58 gigabases; (Gregory, 2021)). This is slightly larger than our 5.04 gigabase assembly. The eleven pseudomolecules in our assembly match the expected number of chromosomes in *B. bufo* ($2N = 22$; six macro- and five micro-chromosomes; (Birstein & Mazin, 1982; Makino & Others, 1951)).

This is the third nuclear genome sequence to be reported from a bufonid anuran (Edwards *et al.*, 2018; Lu *et al.*, 2021). The *B. bufo* reference genome reported here has been used to study pseudogenization of the tooth enamel gene amelogenin in bufonids (Shaheen *et al.*, 2021). The genome of a common toad from the UK is particularly timely as a tool for understanding the dynamics of population declines observed over the last two decades (Carrier & Beebee, 2003; Petrovan & Schmidt, 2016).

Genome sequence report

The genome was sequenced from one male *B. bufo* collected from the Natural History Museum (NHM) Wildlife Garden, London, UK (Figure 1A, B). A total of 64-fold coverage in Pacific Biosciences single-molecule long reads (N50 28 kb) and 56-fold coverage in 10X Genomics read clouds (from molecules with an estimated N50 of 29 kb) were generated. Primary assembly contigs were scaffolded with chromosome conformation Hi-C data. Manual assembly curation corrected 3498 missing/misjoins and removed 513 haplotypic duplications, reducing the assembly length by 2.4% and the scaffold number by 49.5%, and increasing the scaffold N50 by 38.9%.

The final assembly has a total length of 5.04 Gb in 1307 sequence scaffolds with a scaffold N50 of 636 Mb (Table 1). The majority, 99.1%, of the assembly sequence was assigned to 11 chromosomal-level scaffolds (numbered by sequence length) (Figure 2–Figure 5; Table 2). The assembly has a BUSCO (Simão *et al.*, 2015) v5.1.2 completeness of 90.1% using the tetrapoda_odb10 reference set. However, a BUSCO (v4.0.2) score of 95.3% using the same reference set was obtained for the annotated gene set of the aBufBuf1.1 assembly (see section *Genome annotation*), indicating that the assembly has a high level of completeness and that some genes were missed during BUSCO analysis of the whole genome assembly. While not fully phased, the assembly deposited is of one haplotype. Contigs corresponding to the second haplotype have also been deposited.

Genome annotation

The *B. bufo* assembly was annotated by the NCBI Eukaryotic Genome Annotation Pipeline, an automated pipeline that

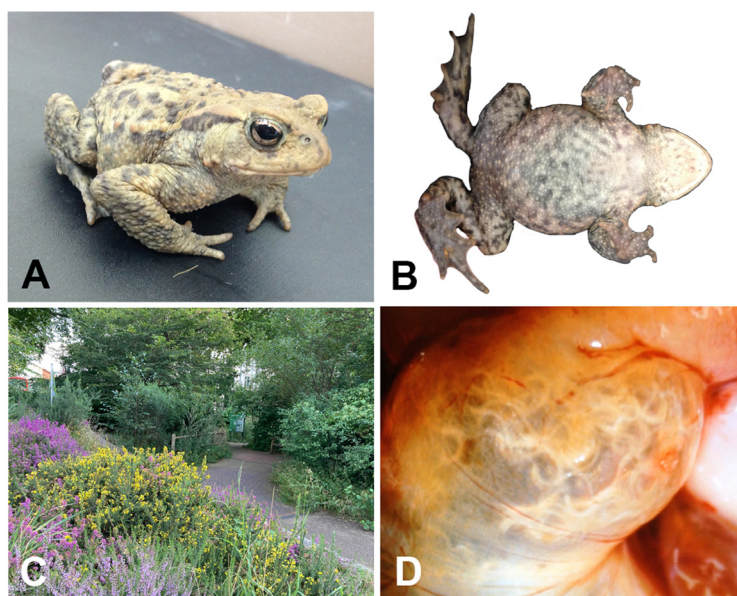


Figure 1. (A) Male voucher specimen of *Bufo bufo* (NHMUK 2013.484; Field ID, JWS 758; Snout-Vent Length 55.5 mm) from which the genome was sequenced. (B) Ventral surface of NHMUK 2013.484. (C) The individual was collected from the Natural History Museum Wildlife Garden, London, England. (D) Large numbers of an unidentified nematode parasite were present in the stomach of NHMUK 2013.484.

Table 1. Genome data for *Bufo bufo*, aBufBuf1.1.

Project accession data	
Assembly identifier	aBufBuf1.1
Species	<i>Bufo bufo</i>
Specimen	aBufBuf1
NCBI taxonomy ID	NCBI:txid8384
BioProject	PRJEB42238
BioSample ID	SAMEA7521636
Isolate information	Male, heart tissue; NHMUK 2013.484
Raw data accessions	
PacificBiosciences SEQUEL II	ERR7012639, ERR7015063-ERR7015065
10X Genomics Illumina	ERR6002753-ERR6002766, ERR6003048, ERR6003049
Hi-C Illumina	ERR6002767-ERR6002770
BioNano	ERZ3003198
Genome assembly	
Assembly accession	GCA_905171765.1
Accession of alternate haplotype	GCA_905171715.1
Span (Mb)	5,045
Number of contigs	5,502
Contig N50 length (Mb)	3.96
Number of scaffolds	1,307
Scaffold N50 length (Mb)	636
Longest scaffold (Mb)	843
BUSCO genome score*	C:90.1%[S:88.5%,D:1.6%],F:3.2%,M:6.7%,n:5310
Genome annotation	
Number of genes	30,286
Number of protein-coding genes	21,517
Average length of gene (bp)	57,667
Average number of exons per gene	12
Average exon size (bp)	241
Average intron size (bp)	8,995
BUSCO annotation score**	C:95.3%[S:93.2%,D:2.1%],F:0.7%,M:4.0%,n:5310

C= complete [S= single copy, D=duplicated], F=fragmented, M=missing, n=number of orthologues in comparison.

*BUSCO scores based on the terapoda_odb10 BUSCO set using v5.1.2, run on the aBufBuf1.1 genome assembly using BlobToolKit. A full set of BUSCO scores is available at <https://blobtoolkit.genomehubs.org/view/aBufBuf1.1/dataset/CAJIMN01/busco>.

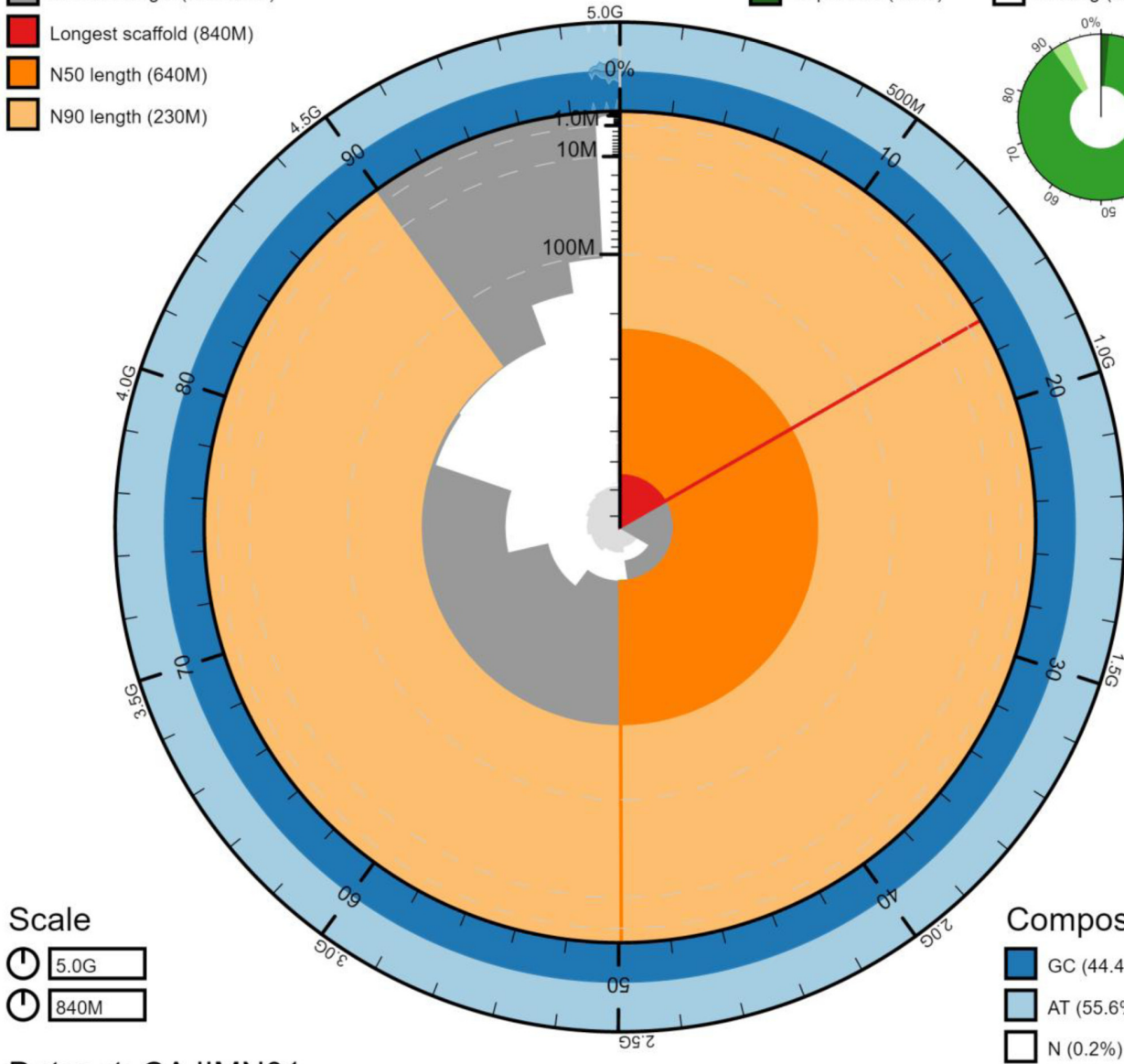
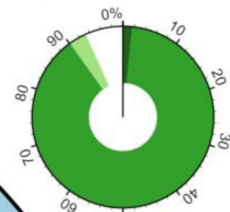
**BUSCO scores based on the terapoda_odb10 BUSCO set using v4.0.2, run on the NCBI RefSeq annotation of the aBufBuf1.1 genome assembly (NCBI *Bufo bufo* Annotation Release 100).

Scaffold statistics

- Log10 scaffold count (total 1.3k)
- Scaffold length (total 5.0G)
- Longest scaffold (840M)
- N50 length (640M)
- N90 length (230M)

BUSCO tetrapoda_odb10 (5310)

- Complete (90.1%)
- Fragmented (3.2%)
- Duplicated (1.6%)
- Missing (6.7%)



Scale

- 5.0G
- 840M

Composition

- GC (44.4%)
- AT (55.6%)
- N (0.2%)

Dataset: CAJIMN01

Figure 2. Genome assembly of *Bufo bufo*, aBufBuf1.1: metrics. The BlobToolKit Snailplot shows N50 metrics and BUSCO gene completeness. The main plot is divided into 1,000 size-ordered bins around the circumference with each bin representing 0.1% of the 5,044,762,059 bp assembly. The distribution of chromosome lengths is shown in dark grey with the plot radius scaled to the longest chromosome present in the assembly (843,366,180 bp, shown in red). Orange and pale-orange arcs show the N50 and N90 chromosome lengths (635,713,434 and 230,778,867 bp), respectively. The pale grey spiral shows the cumulative chromosome count on a log scale with white scale lines showing successive orders of magnitude. The blue and pale-blue area around the outside of the plot shows the distribution of GC, AT and N percentages in the same bins as the inner plot. A summary of complete, fragmented, duplicated and missing BUSCO genes in the tetrapoda_odb10 set is shown in the top right. An interactive version of this figure is available at <https://blobtoolkit.genomehubs.org/view/Bufo%20bufo/dataset/CAJIMN01/snail>.

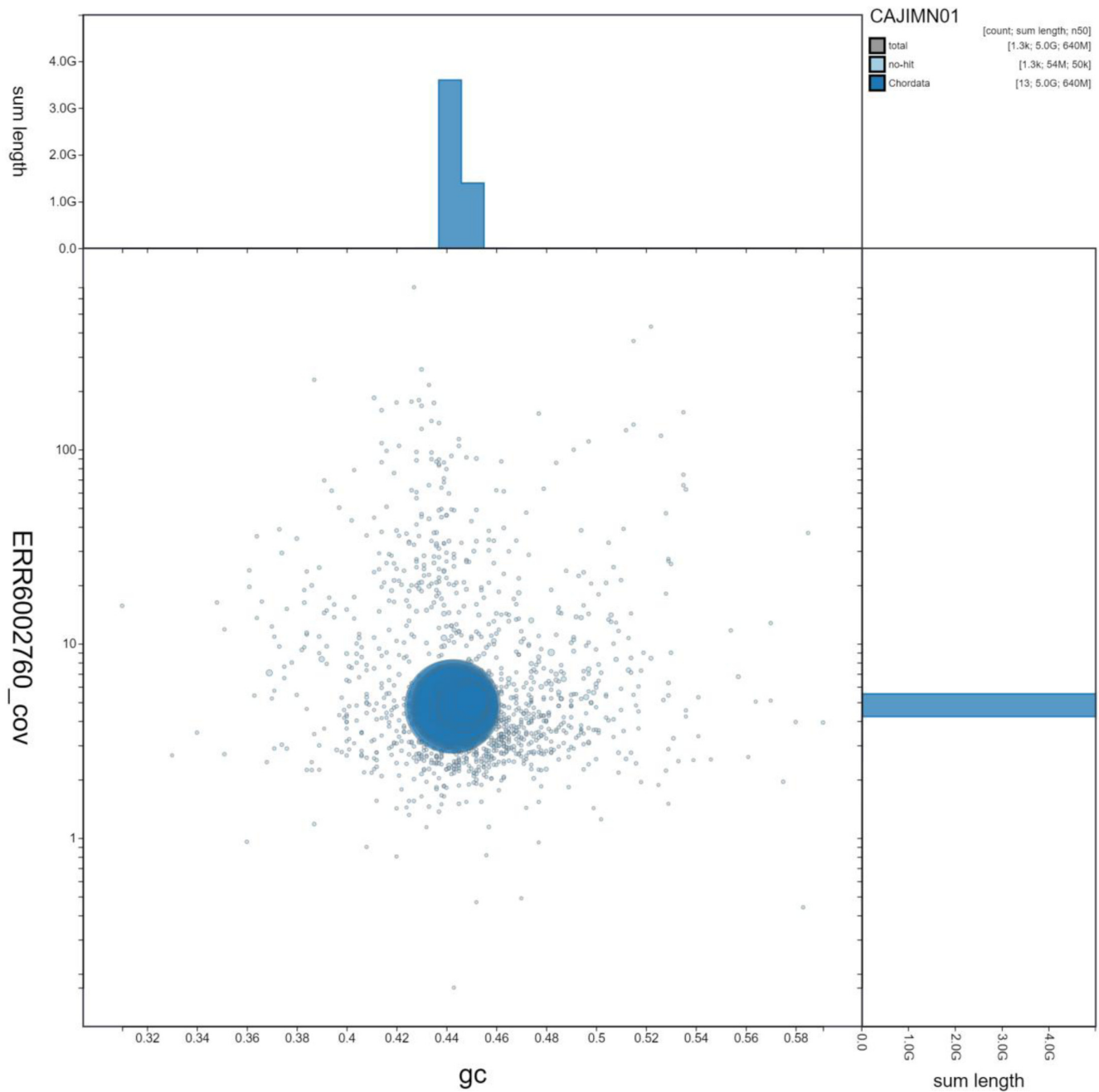


Figure 3. Genome assembly of *Bufo bufo*, aBufBuf1.1: GC-coverage. BlobToolKit GC-coverage plot. Scaffolds are coloured by phylum. Circles are sized in proportion to scaffold length. Histograms show the distribution of scaffold length sum along each axis. An interactive version of this figure is available at <https://blobtoolkit.genomehubs.org/view/Bufo%20bufo/dataset/CAJIMN01/blob>.

annotates genes, transcripts and proteins on draft and finished genome assemblies. The annotation (NCBI *Bufo bufo* Annotation Release 100; Table 1) was generated from transcripts and proteins retrieved from NCBI Entrez by alignment to the genome assembly, as described (Pruitt *et al.*, 2014).

Methods

Sample acquisition

A single male *B. bufo* was collected from a stable, isolated population in the NHM Wildlife Garden, London, UK (latitude 51.49586, longitude -0.178622, elevation 17 m) by Jeffrey

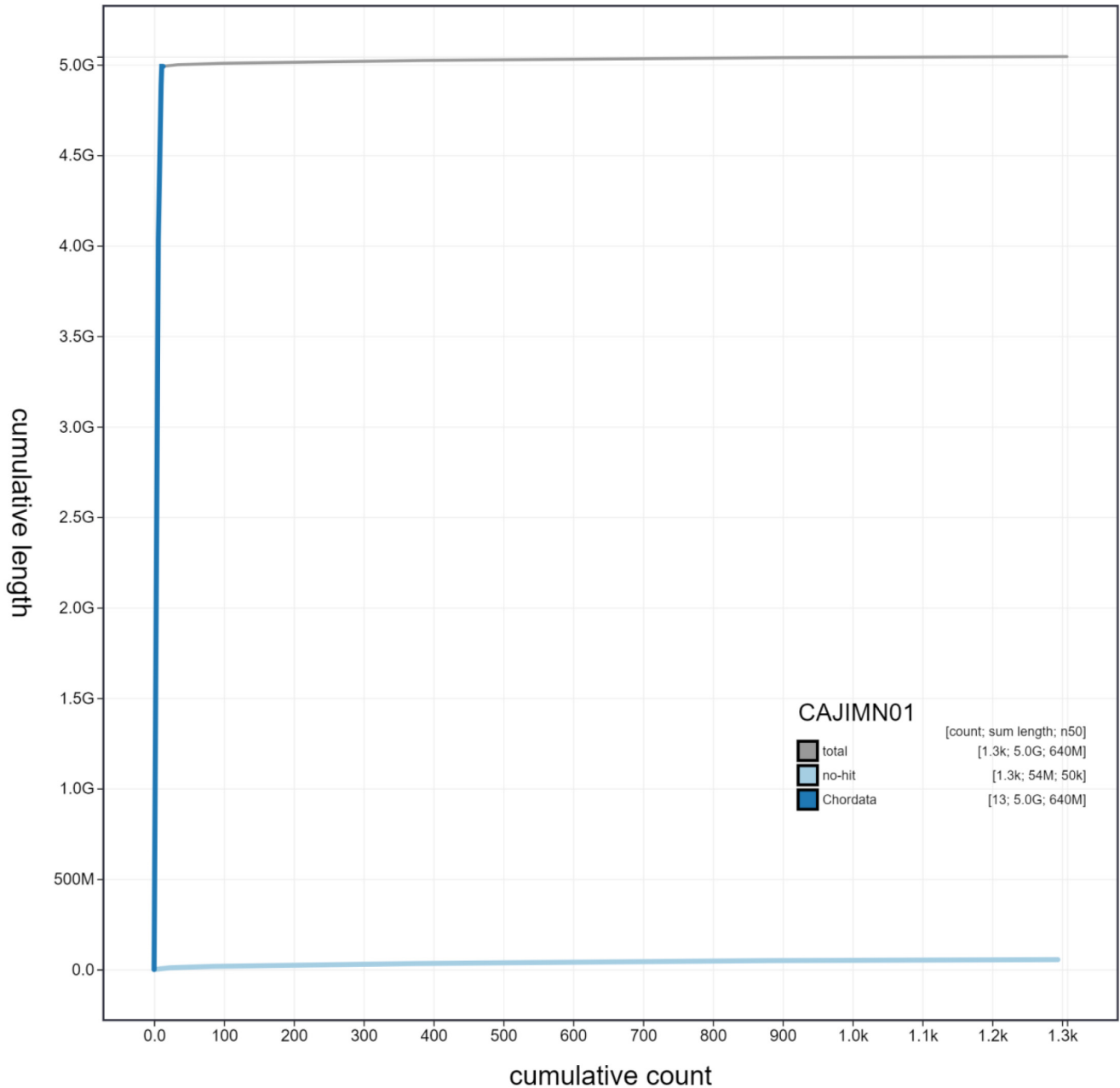


Figure 4. Genome assembly of *Bufo bufo*, aBufBuf1.1: cumulative sequence. BlobToolKit cumulative sequence plot. The grey line shows cumulative length for all scaffolds. Coloured lines show cumulative lengths of scaffolds assigned to each phylum using the buscogenes taxrule. An interactive version of this figure is available at <https://blobtoolkit.genomehubs.org/view/Bufo%20bufo/dataset/CAJIMN01/cumulative>.

W. Streicher on 1 July 2015 (Figure 1C). The specimen of *B. bufo* (NHMUK 2013.484, Field ID: JWS 758) was 55.5 mm snout–vent length (determined using Miyamoto digital callipers to the nearest 0.1 mm) and contained many nematode parasites in its stomach (Figure 1D). The specimen was collected with permission from the NHM Wildlife Garden management team and is part of a long-term monitoring project run by the

Department of Life Sciences and the Angela Marmont Centre for UK Biodiversity. It was humanely euthanised using a saturated solution of tricaine mesylate (MS-222). Multiple tissues including heart, thigh muscle, liver, eyes, kidney, testes, Bidder’s organ, and intestines were sampled into an ammonium sulfate-based RNA + DNA preservation buffer. After ~24 hours of storage at 4°C, the tissues were transferred

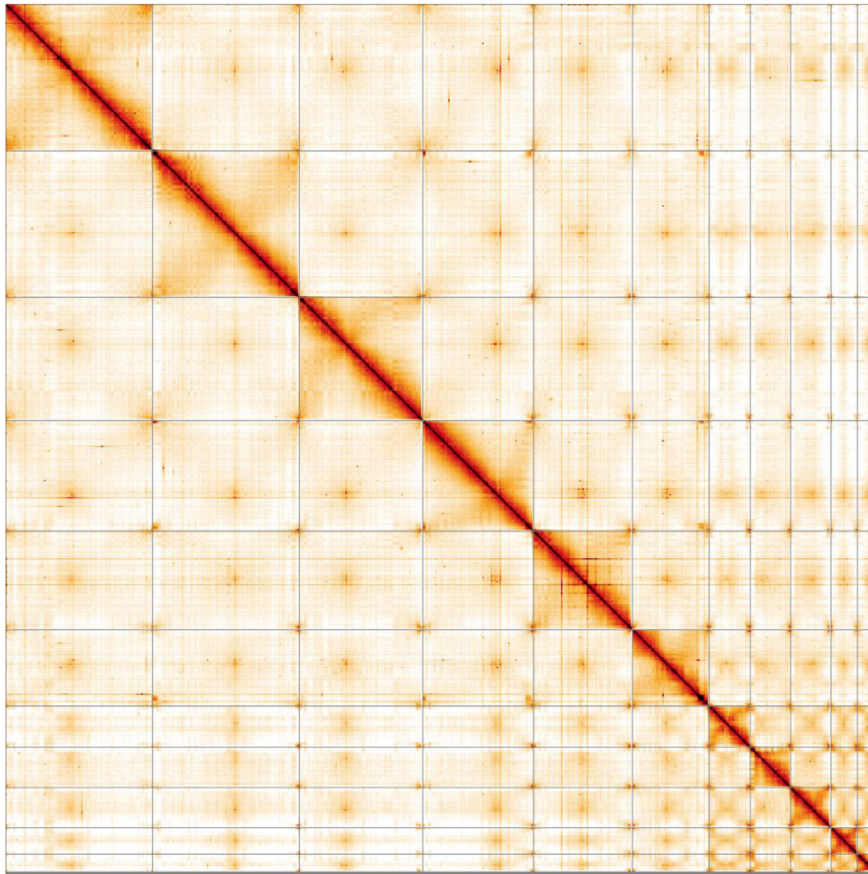


Figure 5. Genome assembly of *Bufo bufo*, aBufBuf1.1: Hi-C contact map. Hi-C contact map of the aBufBuf1.1 assembly, visualised in HiGlass. Chromosomes are presented in size order from left to right and top to bottom.

Table 2. Chromosomal pseudomolecules in the genome assembly of *Bufo bufo*, aBufBuf1.1.

INSDC accession	Chromosome	Size (Mb)	GC%
LR991667.1	1	843.37	44.5
LR991668.1	2	842.56	44.3
LR991669.1	3	707.96	44.4
LR991670.1	4	635.71	44.4
LR991671.1	5	567.30	44.4
LR991672.1	6	439.63	44.8
LR991673.1	7	236.60	44.7
LR991674.1	8	231.67	44.8
LR991675.1	9	230.78	44.8
LR991676.1	10	151.57	44.8
LR991677.1	11	103.21	45
LR991678.1	MT	0.02	42.7
-	Unlocalised	54.40	45.1

to -80°C until they were sent for genome sequencing. Sample tissue has been accessioned by the NHM Molecular Collections Facility (NHMUK 2013.484).

DNA extraction and sequencing

DNA was extracted from heart tissue using the Bionano Prep Animal Tissue DNA Isolation kit according to the manufacturer's instructions. Pacific Biosciences CLR long read and 10X Genomics read cloud sequencing libraries were constructed according to the manufacturers' instructions. Hi-C data were generated from heart tissue using the Arima v2 Hi-C kit. Extraction and sequencing was performed by the Scientific Operations DNA Pipelines at the Wellcome Sanger Institute on Pacific Biosciences SEQUEL I and Illumina HiSeq X instruments. DNA was labeled for Bionano Genomics optical mapping following the Bionano Prep Direct Label and Stain (DLS) Protocol and run on one Saphyr instrument chip flowcell.

Genome assembly

Assembly was carried out following the Vertebrate Genome Project pipeline v1.6 (([Rhie et al., 2021](#))) with Falcon-unzip ([Chin et al., 2016](#)), haplotypic duplication was identified and removed with `purge_dups` ([Guan et al., 2020](#)) and a first round of scaffolding carried out with 10X Genomics read clouds using `scaff10x`. Hybrid scaffolding was performed using the BioNano DLE-1 data and `BioNano Solve`. Scaffolding with Hi-C data (([Rao et al., 2014](#))) was carried out with `HiLine`, then 3D-DNA ([Dudchenko et al., 2017](#)). The Hi-C scaffolded

assembly was polished with arrow using the PacBio data, then polished with the 10X Genomics Illumina data by aligning to the assembly with `longranger align`, calling variants with `freebayes` ([Garrison & Marth, 2012](#)) and applying homozygous non-reference edits using `bcftools consensus`. Two rounds of the Illumina polishing were applied. The mitochondrial genome was assembled at The Rockefeller University using the `mitoVGP` pipeline ([Formenti et al., 2021](#)). The assembly was checked for contamination and corrected using the `gEVAL` system ([Chow et al., 2016](#); ([Howe et al., 2021](#))). Manual curation was performed using evidence from Bionano (using the Bionano Access viewer), using `HiGlass` ([Kerpedjiev et al., 2018](#)) and `Pretext`, as described previously ([Howe et al., 2021](#)). [Figure 2–Figure 4](#) and BUSCO values were generated using `BlobToolKit` ([Challis et al., 2020](#)). [Table 3](#) contains a list of software tools and versions, where applicable.

Ethical/compliance issues

The materials that have contributed to this genome note were 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.

Table 3. Software tools used.

Software tool	Version	Source
Falcon-unzip	falcon-kit 1.4.2	(Chin et al., 2016)
purge_dups	1.0.0	(Guan et al., 2020)
scaff10x	4.2	https://github.com/wtsi-hpag/Scaff10X
Bionano Solve	3.3_10252018	https://bionanogenomics.com/downloads/bionano-solve/
3D-DNA	180922	(Dudchenko et al., 2017)
Arrow	gcpp 1.9.0-SL-release-8.0.0+1-37-gd7b188d	https://github.com/PacificBiosciences/GenomicConsensus
Longranger	2.2.2	https://support.10xgenomics.com/genome-exome/software/pipelines/latest/advanced/other-pipelines
freebayes	v1.3.1-17-gaa2ace8	(Garrison & Marth, 2012)
bcftools-consensus	1.10.2	http://samtools.github.io/bcftools/bcftools.html
gEVAL	N/A	(Chow et al., 2016)
HiGlass	1.11.6	(Kerpedjiev et al., 2018)
PretextView	0.1	https://github.com/wtsi-hpag/PretextView
BlobToolKit	2.6.1	(Challis et al., 2020)

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: *Bufo bufo* (common toad). Accession number [PRJEB42238](https://identifiers.org/ena.embl/PRJEB42238); <https://identifiers.org/ena.embl/PRJEB42238>.

The genome sequence is released openly for reuse. The *B. bufo* genome sequencing initiative is part of the [Darwin Tree of Life](#) (DToL) project and the [Vertebrate Genomes Project](#). All raw sequence data and the assembly have been deposited in INSDC databases. Raw data and assembly accession identifiers are reported in [Table 1](#).

Author information

Members of the Wellcome Sanger Institute Tree of Life programme collective are listed here: <https://doi.org/10.5281/zenodo.5377053>.

Members of Wellcome Sanger Institute Scientific Operations: DNA Pipelines collective are listed here: <https://doi.org/10.5281/zenodo.4790456>.

Members of the Tree of Life Core Informatics collective are listed here: <https://doi.org/10.5281/zenodo.5013542>.

Members of the Darwin Tree of Life Consortium are listed here: <https://doi.org/10.5281/zenodo.4783559>.

Acknowledgements

JWS thanks Caroline Ware, Nicky Reilly, Naomi Lake, and the NHM Wildlife Garden Team for permitting specimen collection. JWS also thanks Donney Nicholson for assistance with tissue shipment as well as David Gower, Simon Loader (Department of Life Sciences), and John Tweddle (Angela Marmont Centre) for logistical assistance.

References

- Birstein VJ, Mazin AL: **Chromosomal Polymorphism of *Bufo Bufo*: Karyotype and C-Banding Pattern of *B. B. Verrucosissima***. *Genetica*. 1982. [PubMed Full Text](#)
- Carrier JA, Beebe TJ: **Recent, Substantial, and Unexplained Declines of the Common Toad *Bufo Bufo* in Lowland England**. *Biological Conservation*. 2003; **111**(3): 395–99. [PubMed Abstract](#) | [Publisher Full Text](#)
- Challis R, Richards E, Rajan J, et al.: **BlobToolKit - Interactive Quality Assessment of Genome Assemblies**. *G3 (Bethesda)*. 2020; **10**(4): 1361–74. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Chin CS, Peluso P, Sedlazeck FJ, et al.: **Phased Diploid Genome Assembly with Single-Molecule Real-Time Sequencing**. *Nat Methods*. 2016; **13**(12): 1050–54. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Chow W, Brugger K, Caccamo M, et al.: **gEVAL — a Web-Based Browser for Evaluating Genome Assemblies**. *Bioinformatics*. 2016; **32**(16): 2508–10. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Dudchenko O, Batra SS, Omer AD, et al.: **De Novo Assembly of the *Aedes Aegypti* Genome Using Hi-C Yields Chromosome-Length Scaffolds**. *Science*. 2017; **356**(6333): 92–95. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Edwards RJ, Tuipulotu DE, Amos TG, et al.: **Draft Genome Assembly of the Invasive Cane Toad, *Rhinella Marina***. *Gigascience*. 2018; **7**(9): giy095. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Ewert JP: **The Neural Basis of Visually Guided Behavior**. *Sci Am*. 1974; **230**(3): 34–42. [PubMed Abstract](#) | [Publisher Full Text](#)
- Formenti G, Rhie A, Balacco J, et al.: **Complete Vertebrate Mitogenomes Reveal Widespread Repeats and Gene Duplications**. *Genome Biol*. 2021; **22**(1): 120. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Garrison E, Marth G: **Haplotype-Based Variant Detection from Short-Read Sequencing**. 2012; arXiv: 1207.3907. [Reference Source](#)
- Gregory: **Animal Genome Size Database**. 2021. [Reference Source](#)
- Guan D, McCarthy SA, Wood J, et al.: **Identifying and Removing Haplotype Duplication in Primary Genome Assemblies**. *Bioinformatics*. 2020; **36**(9): 2896–98. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free 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](#)
- Lu B, Jiang J, Wu H, et al.: **A Large Genome with Chromosome-Scale Assembly Sheds Light on the Evolutionary Success of a True Toad (*Bufo Gargarizans*)**. *Mol Ecol Resour*. 2021; **21**(4): 1256–73. [PubMed Abstract](#) | [Publisher Full Text](#)
- Makino S, and Others: **An Atlas of the Chromosome Numbers in Animals**. *An Atlas of the Chromosome Numbers in Animals.*, no. 2nd ed.(1st American ed.). 1951. [Reference Source](#)
- Petrovan SO, Schmidt BR: **Volunteer Conservation Action Data Reveals Large-Scale and Long-Term Negative Population Trends of a Widespread Amphibian, the Common Toad (*Bufo Bufo*)**. *PLoS One*. 2016; **11**(10): e0161943. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Pruitt KD, Brown GR, Hiatt SM, et al.: **RefSeq: An Update on Mammalian Reference Sequences**. *Nucleic Acids Res*. 2014; **42**(Database issue): D756–63. [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): 1665–80. [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): 737–46. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Shaheen J, Mudd AB, Diekwisch TGH, et al.: **Pseudogenized Amelogenin Reveals Early Tooth Loss in True Toads (Anura: Bufonidae)**. *Integr Comp Biol*. 2021; icab039. [PubMed Abstract](#) | [Publisher Full Text](#)
- Simão FA, Waterhouse RM, Ioannidis P, et al.: **BUSCO: Assessing Genome Assembly and Annotation Completeness with Single-Copy Orthologs**. *Bioinformatics*. 2015; **31**(19): 3210–12. [PubMed Abstract](#) | [Publisher Full Text](#)

Open Peer Review

Current Peer Review Status:  

Version 1

Reviewer Report 04 January 2022

<https://doi.org/10.21956/wellcomeopenres.19125.r46560>

© 2022 Shaffer H. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



H. Bradley Shaffer

Department of Ecology and Evolutionary Biology, UCLA La Kretz Center for California Conservation Science, Institute of the Environment and Sustainability, University of California, Los Angeles, Los Angeles, CA, USA

This is a straightforward genome release report of an important European anuran. The methods and presentation of the data are obviously formulaic, but in this case, that's fine. It could be indexed as is, although I have a few small comments that might increase its visibility or clarity, particularly for the amphibian community that will be its primary user group.

The Introduction is very thin and not really very informative or compelling. I would include more on the family Bufonidae, for example. It is the third most speciose family out of 54 currently recognized on Amphibiaweb, there's been a lot of recent controversy over genus-level taxonomy - that kind of thing. I'd like to see a bit more on conservation, given how hard toads have been hit by chytrid. What about the portion of the family phylogeny that is now covered by the three species with genomes, and important targets for additional lineages?

Methods:

- Usually one records snout-urostyle length in an anuran, rather than snout-vent. Or at least that's the measurement I prefer. They are almost identical.
- I could live without Figure 4, but it's fine if that is what you want. I don't think it's terribly necessary.

Small notes:

- Benthic larvae - I've never heard that term for anuran larva. Unless they are very specialized, it applies to pretty much all tadpoles. I would drop this.
- Last sentence of the Genome Annotation section should be corrected to say "...as described by Pruitt *et al.* (2014):

"The annotation (NCBI Bufo bufo Annotation Release 100; Table 1) was generated from

transcripts and proteins retrieved from NCBI Entrez by alignment to the genome assembly, as described by (Pruitt et al., 2014)."

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?

Yes

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

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Amphibian and reptile evolution and ecology; conservation genomics and genetics; conservation 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.

Reviewer Report 01 November 2021

<https://doi.org/10.21956/wellcomeopenres.19125.r46626>

© 2021 Cai Z. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Zexi Cai 

Center for Quantitative Genetics and Genomics, Aarhus University, Tjele, Denmark

The authors presented a high quality genome assembly of the common toad. The data generation and analysis pipelines applied in this report are sound and the information for the resource reuse are clearly stated in the report.

However, in the "Genome annotation" part, the authors stated the method which should go to method part instead of a summary of the annotated genomic feature. Also, in the method part, the detail of running BUSCO should be stated clearly.

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

Yes

Are the protocols appropriate and is the work technically sound?

Yes

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

Yes

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

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Bioinformatics, quantitative genetics and genomics.

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.
