The genome sequence of the barbel, Barbus barbus (Linnaeus, 1758)

Pitman, Richard; Hänfling, Bernd; Bista, Iliana

Published in:
Wellcome Open Research

Publication date:
2023

DATA NOTE

The genome sequence of the barbel, *Barbus barbus* (Linnaeus, 1758) [version 1; peer review: 1 approved]

Richard Pitman[1], Bernd Hänfling[2], Iliana Bista[3,4], 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

[1]Environment Agency, Bristol, England, UK
[2]University of Highlands and Islands, Inverness, Scotland, UK
[3]Naturalis Biodiversity Center, Leiden, South Holland, The Netherlands
[4]Wellcome Sanger Institute, Hinxton, England, UK

v1 **First published:** 10 May 2023, **8**:201
https://doi.org/10.12688/wellcomeopenres.19332.1
Latest published: 10 May 2023, **8**:201
https://doi.org/10.12688/wellcomeopenres.19332.1

Abstract

We present a genome assembly from an individual male *Barbus barbus* (the barbel; Chordata; Actinopteri; Cypriniformes; Cyprinidae). The genome sequence is 1,584.9 megabases in span. Most of the assembly is scaffolded into 50 chromosomal pseudomolecules. The mitochondrial genome has also been assembled and is 16.6 kilobases in length.

Keywords

Barbus barbus, barbel, genome sequence, chromosomal, Cypriniformes

This article is included in the Tree of Life gateway.

Open Peer Review

Approval Status ✓

	1
version 1 10 May 2023	✓ view

1. **Guoqing Lu**, University of Nebraska Omaha, Omaha, 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: Pitman R: Investigation, Resources, Writing – Review & Editing; **Hänfling B**: Investigation, Resources, Writing – Review & Editing; **Bista I**: Writing – Original Draft Preparation, 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 (206194, https://doi.org/10.35802/206194) and the Darwin Tree of Life Discretionary Award (218328, https://doi.org/10.35802/218328). IB was supported by a Naturalis Biodiversity Centre Independent Postdoctoral Fellowship (2021–2023).
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

How to cite this article: Pitman R, Hänfling B, Bista I *et al.* **The genome sequence of the barbel, *Barbus barbus* (Linnaeus, 1758) [version 1; peer review: 1 approved]** Wellcome Open Research 2023, **8**:201 https://doi.org/10.12688/wellcomeopenres.19332.1

First published: 10 May 2023, **8**:201 https://doi.org/10.12688/wellcomeopenres.19332.1

Species taxonomy

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Barbinae; *Barbus*; *Barbus barbus* (Linnaeus, 1758) (NCBI:txid40830).

Background

The barbel, *Barbus barbus*, also known as the European barbel, is a freshwater fish species and member of the Cyprinidae family. It is widespread throughout Europe with its distribution ranging from eastern UK and the Pyrenees in the west, to the Black Sea and Anatolia in eastern Europe, and from the Baltic Sea to the Danube valley (Britton & Pegg, 2011). In addition to its native range, the barbel has also established invasive populations in Italy and western UK (De Santis *et al.*, 2019). The preferred habitat of the barbel is fast flowing waters, though different preferences have been recorded for various life stages, with larvae found mostly in the littoral zone and adults in the mid-channel (Britton & Pegg, 2011) in the so called "barbel zone" (Lugowska & Witeska, 2018).

Even though they are classed as "least concern" species in the IUCN red list (Freyhof, 2011), barbels are nevertheless under threat due to various anthropogenic effects, and regional population losses have been observed (Britton & Pegg, 2011; Penaz *et al.*, 2003; Vilizzi *et al.*, 2013). Some of the major threats facing barbel populations include river regulation which could restrict habitat availability (Vilizzi *et al.*, 2013), overfishing, and eutrophication (Penczak & Sierakowska, 2003).

Furthermore, the barbel is considered an important species for recreational angling, especially in countries like the UK and Poland (Lugowska & Witeska, 2018; Penczak & Sierakowska, 2003; Wheeler & Jordan, 1990). As such the barbel represents an important socio-economic resource (Britton & Pegg, 2011), evident also in the presence of several angling societies dedicated to barbel (*e.g.*, Barbel Fishing World, the Barbel Society, Fly Fishing in Poland).

The genus *Barbus* in general is considered very interesting for evolutionary and biogeographic studies (Gandlin *et al.*, 2022). The evolutionary history of the common barbel, in particular, has been linked to the presence of refugia established during several glacial cycles (Kotlik & Berrebi, 2001). Mitochondrial DNA based analysis has suggested the presence of reciprocal monophyly indicating long-term lack of gene flow between main European versus Anatolian populations, along with low levels of divergence within the main European populations (Kotlik & Berrebi, 2001).

The availability of whole genome data for the barbel can contribute to better understanding of its phylogenetics and population dynamics, and potentially inform future conservation efforts (Supple & Shapiro, 2018). The genome of *Barbus barbus* was sequenced as part of the Darwin Tree of Life Project, a collaborative effort to sequence all named eukaryotic species in the Atlantic Archipelago of Britain and Ireland. Other assemblies for this species currently are available (*e.g.* GCA_023566175.1). Here we present a chromosomally complete genome sequence for *Barbus barbus*, based on one male specimen from Calverton, Nottingham, UK.

Genome sequence report

The genome was sequenced from one male *Barbus barbus* (Figure 1) collected from Calverton Fish Farm, Nottingham, UK (53.03, –1.05). A total of 47-fold coverage in Pacific Biosciences single-molecule HiFi long reads was generated. Primary assembly contigs were scaffolded with chromosome conformation Hi-C data. Manual assembly curation corrected 68 missing joins or mis-joins and removed eight haplotypic duplications, reducing the scaffold number by 6.19%, and decreasing the scaffold N50 by 43.53%.

The final assembly has a total length of 1,584.9 Mb in 91 sequence scaffolds with a scaffold N50 of 31.5 Mb (Table 1). Most (99.77%) of the assembly sequence was assigned to 50 chromosomal-level scaffolds. Chromosome-scale scaffolds confirmed by the Hi-C data are named in order of size (Figure 2–Figure 5; Table 2). While not fully phased, the assembly deposited is of one haplotype. Contigs corresponding to the second haplotype have also been deposited. The mitochondrial genome was also assembled and can be found as a contig within the multifasta file of the genome submission.

The estimated Quality Value (QV) of the final assembly is 60.2 with *k*-mer completeness of 100%, and the assembly has a BUSCO v5.3.2 (Manni *et al.*, 2021) completeness of 98.7% (single 21%, duplicated 77.6%) using the actinopterygii_odb10 reference set (n = 3,640).

Metadata for specimens, spectral estimates, sequencing runs, contaminants and pre-curation assembly statistics can be found at Tree of Life QC, taxon ID 40830.

Methods

Two juvenile male *Barbus barbus* (fBarBab1 and fBarbab2) were collected from the Environment Agency's National Coarse Fish Rearing Unit in Calverton, Nottingham, UK (latitude 53.03, longitude –1.05) on 25 August 2020. The National

Figure 1. Photograph of the *Barbus barbus* (specimen number SAN0000715, ToLID fBarBab1) specimen used for genome sequencing.

Table 1. Genome data for *Barbus barbus*, fBarBab1.1.

Project accession data		
Assembly identifier	fBarBab1.1	
Species	*Barbus barbus*	
Specimen	fBarBab1	
NCBI taxonomy ID	40830	
BioProject	PRJEB51453	
BioSample ID	SAMEA13335783	
Isolate information	fBarBab1, male: gonad (genome sequencing) fBarBab2, male: spleen (RNA-sequencing); gonad (Hi-C scaffolding)	
Assembly metrics*		***Benchmark***
Consensus quality (QV)	60.2	*≥ 50*
k-mer completeness	100%	*≥ 95%*
BUSCO**	C:98.7%[S:21.0%,D:77.6%], F:0.5%,M:0.8%,n:3,640	*C ≥ 95%*
Percentage of assembly mapped to chromosomes	99.77%	*≥ 95%*
Sex chromosomes	Not identified	*localised homologous pairs*
Organelles	Mitochondrial genome assembled	*complete single alleles*
Raw data accessions		
PacificBiosciences SEQUEL II	ERR9387528, ERR9387529, ERR9387530	
Hi-C Illumina	ERR9248444	
PolyA RNA-Seq Illumina	ERR10123689	
Genome assembly		
Assembly accession	GCA_936440315.1	
Accession of alternate haplotype	GCA_936446605.1	
Span (Mb)	1,584.9	
Number of contigs	199	
Contig N50 length (Mb)	22.8	
Number of scaffolds	91	
Scaffold N50 length (Mb)	31.5	
Longest scaffold (Mb)	47.8	

* Assembly metric benchmarks are adapted from column VGP-2020 of "Table 1: Proposed standards and metrics for defining genome assembly quality" from (Rhie *et al.*, 2021).

** BUSCO scores based on the actinopterygii_odb10 BUSCO set using v5.3.2. C = complete [S = single copy, D = duplicated], F = fragmented, M = missing, n = number of orthologues in comparison. A full set of BUSCO scores is available at https://blobtoolkit.genomehubs.org/view/fBarBab1.1/dataset/CAKZFL01/busco.

Coarse Fish Rearing Unit at Calverton is funded solely through rod licence duty. The specimens were taken from the nursery pond by Richard Pitman (Environment Agency) using a seine net, and left to recover fully in fresh flowing, clean, borehole water for a week before any sampling commenced. The two specimens were transported alive to the University of Hull where they were identified by Bernd Hänfling (University of Hull) and euthanized in a lethal dose of MS-222. Tissue dissection took place within 30 minutes of euthanasia, and the tissues were immediately shock-frozen

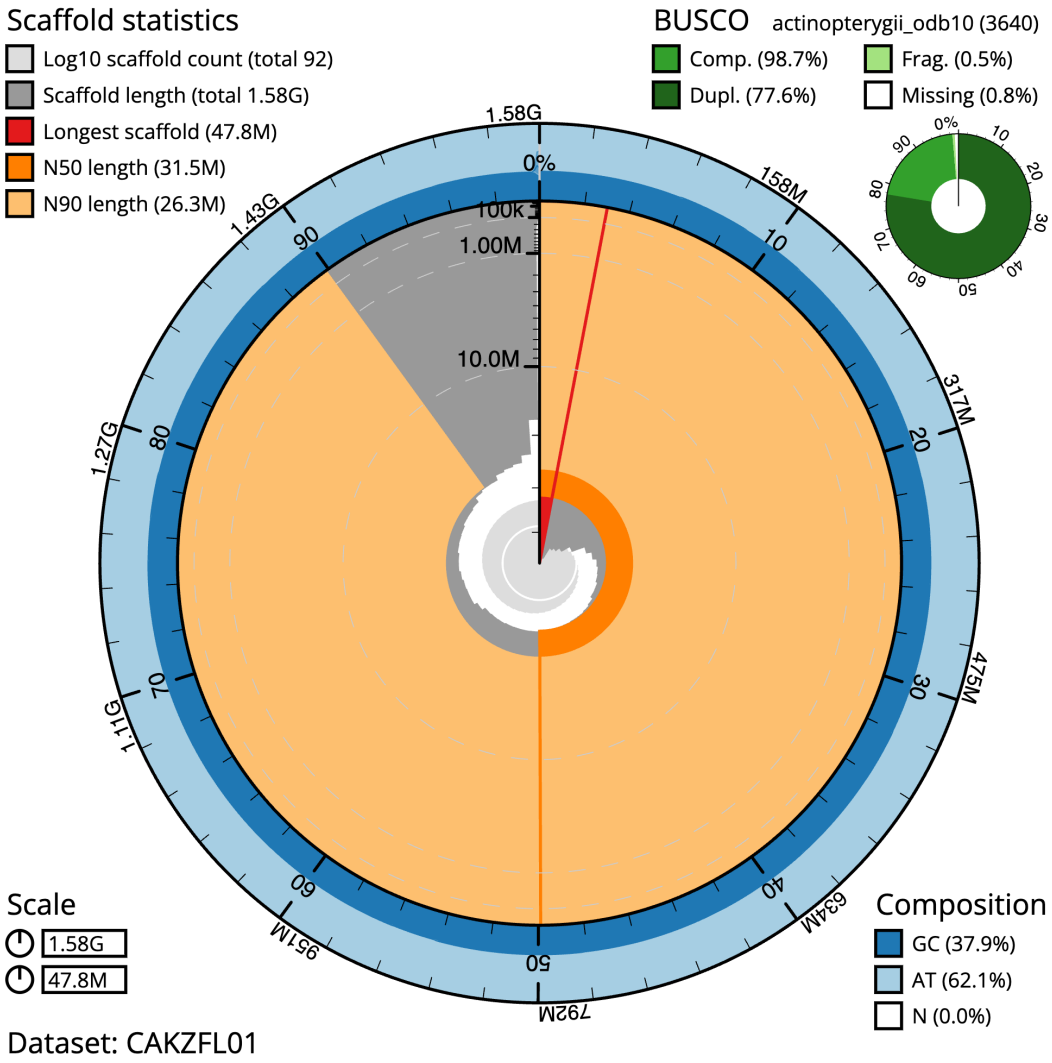

Figure 2. Genome assembly of *Barbus barbus*, fBarBab1.1: metrics. The BlobToolKit Snailplot shows N50 metrics and BUSCO gene completeness. The main plot is divided into 1,000 size-ordered bins around the circumference with each bin representing 0.1% of the 1,584,929,904 bp assembly. The distribution of scaffold lengths is shown in dark grey with the plot radius scaled to the longest scaffold present in the assembly (47,791,430 bp, shown in red). Orange and pale-orange arcs show the N50 and N90 scaffold lengths (31,505,967 and 26,278,021 bp), respectively. The pale grey spiral shows the cumulative scaffold 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 actinopterygii_odb10 set is shown in the top right. An interactive version of this figure is available at https://blobtoolkit.genomehubs.org/view/fBarBab1.1/dataset/CAKZFL01/snail.

in liquid nitrogen. Tissue from fBarBab1 (specimen number SAN0000715) was used for genome sequencing, while tissue from fBarBab2 (specimen number SAN0000701) was used for Hi-C scaffolding and RNA sequencing.

DNA was extracted at the Tree of Life laboratory, Wellcome Sanger Institute (WSI). The fBarBab1 sample was weighed and dissected on dry ice. Gonad tissue was cryogenically disrupted to a fine powder using a Covaris cryoPREP Automated Dry Pulveriser, receiving multiple impacts. High molecular weight (HMW) DNA was extracted using the Qiagen MagAttract HMW DNA extraction kit. HMW DNA was sheared into an average fragment size of 12–20 kb in a Megaruptor 3 system with speed setting 30. Sheared DNA was purified by solid-phase reversible immobilisation using AMPure PB beads with a 1.8X ratio of beads to sample to remove the shorter fragments and concentrate the DNA sample. The concentration of the sheared and purified DNA was assessed using a Nanodrop spectrophotometer and Qubit Fluorometer and Qubit dsDNA High Sensitivity Assay kit. Fragment size distribution was evaluated by running the sample on the FemtoPulse system.

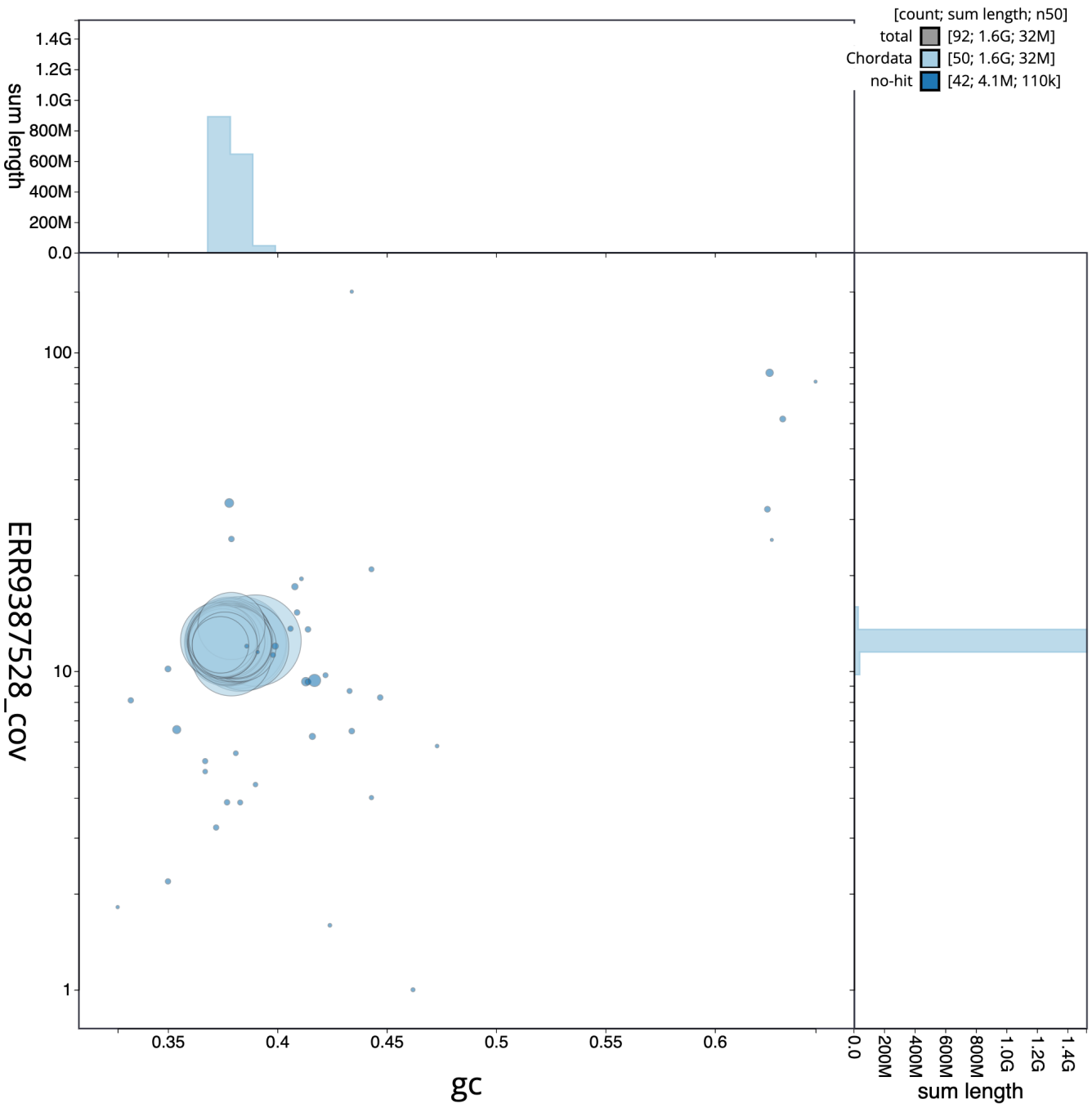

Figure 3. Genome assembly of *Barbus barbus*, fBarBab1.1: GC coverage. BlobToolKit GC-coverage plot. Scaffolds are coloured by phylum. Circles are sized in proportion to scaffold length. Histograms show the distribution of scaffold length sum along each axis. An interactive version of this figure is available at https://blobtoolkit.genomehubs.org/view/fBarBab1.1/dataset/CAKZFL01/blob.

RNA was extracted from spleen tissue of fBarBab2 in the Tree of Life Laboratory at the Wellcome Sanger Institute (WSI) using TRIzol, according to the manufacturer's instructions. RNA was then eluted in 50 μl RNAse-free water and its concentration assessed using a Nanodrop spectrophotometer and Qubit Fluorometer using the Qubit RNA Broad-Range (BR) Assay kit. Analysis of the integrity of the RNA was done using Agilent RNA 6000 Pico Kit and Eukaryotic Total RNA assay.

Sequencing

Pacific Biosciences HiFi circular consensus DNA sequencing libraries were constructed according to the manufacturers' instructions. Poly(A) RNA-Seq libraries were constructed using the NEB Ultra II RNA Library Prep kit. DNA and RNA sequencing were performed by the Scientific Operations core at the WSI on Pacific Biosciences SEQUEL II (HiFi) and Illumina NovaSeq 6000 (RNA-Seq) instruments. Hi-C data were also

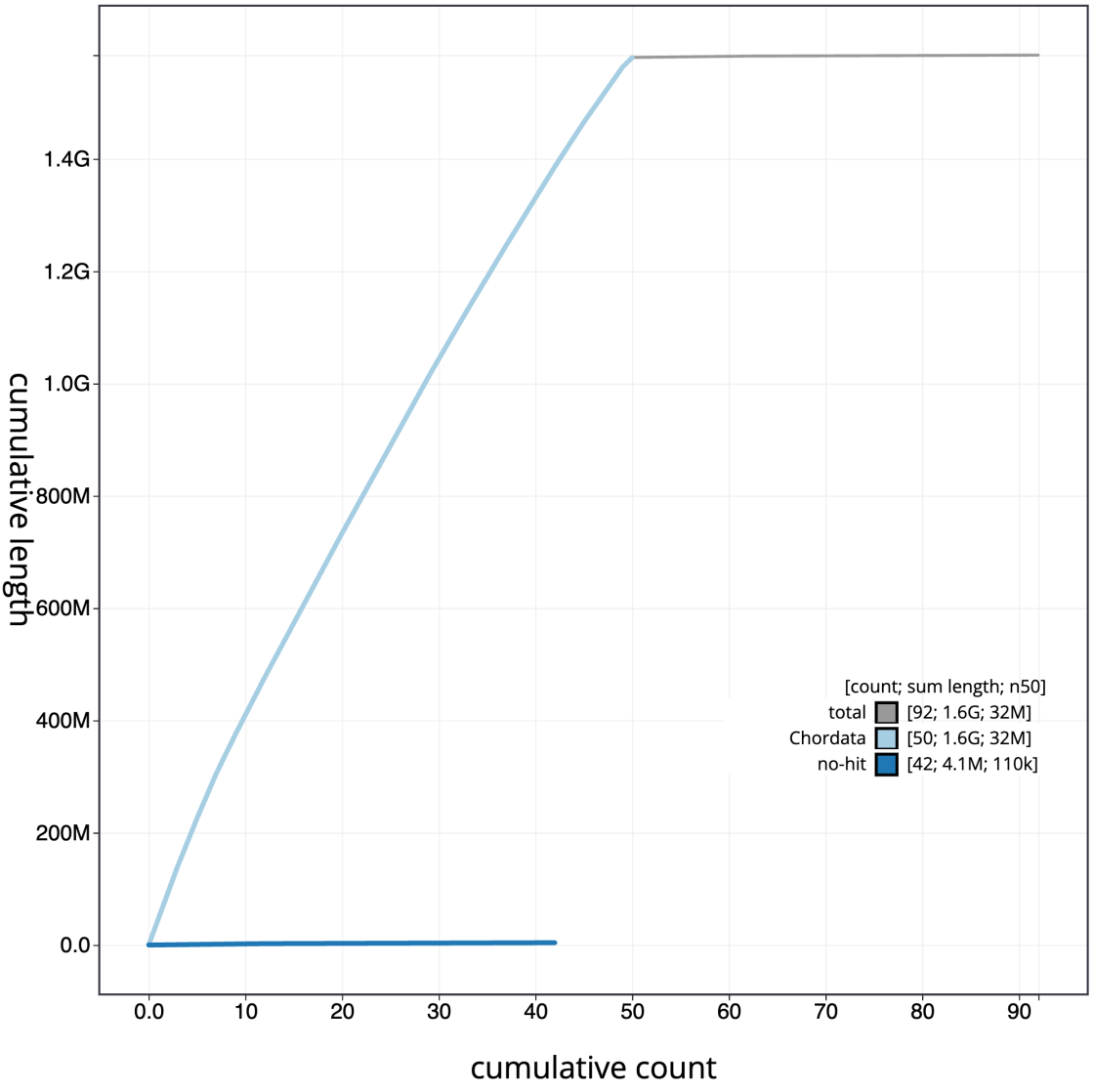

Figure 4. Genome assembly of *Barbus barbus*, fBarBab1.1: cumulative sequence. BlobToolKit cumulative sequence plot. The grey line shows cumulative length for all scaffolds. Coloured lines show cumulative lengths of scaffolds assigned to each phylum using the buscogenes taxrule. An interactive version of this figure is available at https://blobtoolkit.genomehubs.org/view/fBarBab1.1/dataset/CAKZFL01/cumulative.

generated from gonad tissue of fBarBab2 using the Arima v2 kit and sequenced on the Illumina NovaSeq 6000 instrument.

Genome assembly, curation and evaluation

Assembly was carried out with Hifiasm (Cheng *et al.*, 2021) and haplotypic duplication was identified and removed with purge_dups (Guan *et al.*, 2020). The assembly was then scaffolded with Hi-C data (Rao *et al.*, 2014) using YaHS (Zhou *et al.*, 2023). The assembly was checked for contamination as described previously (Howe *et al.*, 2021). Manual curation was performed using HiGlass (Kerpedjiev *et al.*, 2018) and Pretext (Harry, 2022). The mitochondrial genome was assembled using MitoHiFi (Uliano-Silva *et al.*, 2022), which runs MitoFinder (Allio *et al.*, 2020) or MITOS (Bernt *et al.*, 2013) and uses these annotations to select the final mitochondrial contig and to ensure the general quality of the sequence. To evaluate the assembly, MerquryFK was used to estimate consensus quality (QV) scores and *k*-mer completeness (Rhie *et al.*, 2020). The genome was analysed within the BlobToolKit environment (Challis *et al.*, 2020) and BUSCO scores (Manni *et al.*, 2021; Simão *et al.*, 2015) were calculated. Table 3 contains a list of software tool versions and sources.

Ethics and compliance issues

The materials that have contributed to this genome note have been supplied by a Darwin Tree of Life Partner. The submission of materials by a Darwin Tree of Life Partner is subject to the Darwin Tree of Life Project Sampling Code of Practice. By agreeing with and signing up to the Sampling Code of

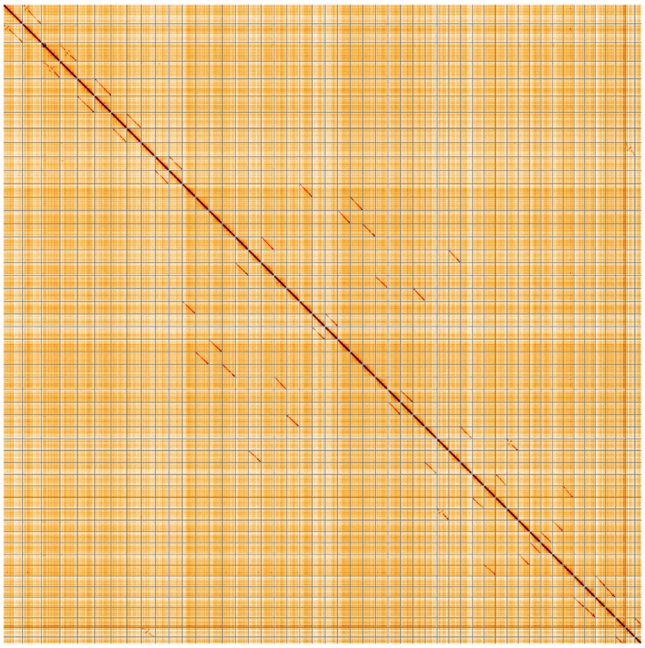

Figure 5. Genome assembly of *Barbus barbus*, fBarBab1.1: Hi-C contact map. Hi-C contact map of the fBarBab1.1 assembly, visualised using HiGlass. Chromosomes are shown in order of size from left to right and top to bottom. An interactive version of this figure may be viewed at https://genome-note-higlass.tol.sanger.ac.uk/l/?d=B1D0NuYgSkeSiqt33YY0pw.

Table 2. Chromosomal pseudomolecules in the genome assembly of *Barbus barbus*, fBarBab1.

INSDC accession	Chromosomes	Size (Mb)	GC%
OW387168.1	1	47.79	38.3
OW387169.1	2	46.57	38.3
OW387170.1	3	46.01	39
OW387171.1	4	42.98	38.5
OW387172.1	5	42.38	37.8
OW387173.1	6	41.29	37.8
OW387174.1	7	38.97	37.7
OW387175.1	8	35.83	37.6
OW387176.1	9	34.6	37.9
OW387177.1	10	33.7	37.7
OW387178.1	11	33.4	37.8
OW387179.1	12	33.25	37.8
OW387180.1	13	32.86	37.6
OW387181.1	14	32.81	38
OW387182.1	15	32.26	37.8
OW387183.1	16	31.98	37.8
OW387184.1	17	31.82	38
OW387185.1	18	31.79	37.9
OW387186.1	19	31.68	37.8
OW387187.1	20	31.64	37.8
OW387188.1	21	31.58	37.7
OW387189.1	22	31.51	38
OW387190.1	23	31.24	37.9
OW387191.1	24	31.2	38
OW387192.1	25	31.19	37.3
OW387193.1	26	31.13	37.6
OW387194.1	27	30.98	37.7
OW387195.1	28	30.92	37.7
OW387196.1	29	30.83	37.8
OW387197.1	30	30.09	37.8
OW387198.1	31	29.41	37.9
OW387199.1	32	29.4	37.9
OW387200.1	33	29.29	38
OW387201.1	34	28.99	38
OW387202.1	35	28.66	38.1
OW387152.1	36	28.63	37.9
OW387153.1	37	28.32	38.3
OW387154.1	38	28.24	37.6
OW387155.1	39	28.15	38.1
OW387156.1	40	27.81	37.8
OW387157.1	41	27.76	37.8
OW387158.1	42	27.52	38.1
OW387159.1	43	26.61	37.5
OW387160.1	44	26.28	37.5
OW387161.1	45	26.26	37.6
OW387162.1	46	25.3	37.5
OW387163.1	47	24.93	37.5
OW387164.1	48	24.54	37.9
OW387165.1	49	23.17	37.6
OW387166.1	50	17.27	37.4
OW387167.1	MT	0.02	43.6
-	unplaced	4.12	42.2

Table 3. Software tools: versions and cources.

Software tool	Version	Source
BlobToolKit	4.0.7	https://github.com/blobtoolkit/blobtoolkit
BUSCO	5.3.2	https://gitlab.com/ezlab/busco
Hifiasm	0.16.1-r375	https://github.com/chhylp123/hifiasm
HiGlass	1.11.6	https://github.com/higlass/higlass
Merqury	MerquryFK	https://github.com/thegenemyers/MERQURY.FK
MitoHiFi	2	https://github.com/marcelauliano/MitoHiFi
PretextView	0.2	https://github.com/wtsi-hpag/PretextView
purge_dups	1.2.3	https://github.com/dfguan/purge_dups
YaHS	yahs-1.1.91eebc2	https://github.com/c-zhou/yahs

Practice, the Darwin Tree of Life Partner agrees they will meet the legal and ethical requirements and standards set out within this document in respect of all samples acquired for, and supplied to, the Darwin Tree of Life Project. All efforts are undertaken to minimise the suffering of animals used for sequencing. Each transfer of samples is further undertaken according to a Research Collaboration Agreement or Material Transfer Agreement entered into by the Darwin Tree of Life Partner, Genome Research Limited (operating as the Wellcome Sanger Institute), and in some circumstances other Darwin Tree of Life collaborators.

Data availability

European Nucleotide Archive: *Barbus barbus* (barbel). Accession number PRJEB51453; https://identifiers.org/ena.embl/PRJEB51453. (Wellcome Sanger Institute, 2022)

The genome sequence is released openly for reuse. The *Barbus barbus* genome sequencing initiative is part of the Darwin Tree of Life (DToL) project. 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.

Author information

Members of the Darwin Tree of Life Barcoding collective are listed here: https://doi.org/10.5281/zenodo.4893703.

Members of the Wellcome Sanger Institute Tree of Life programme are listed here: https://doi.org/10.5281/zenodo.4783585.

Members of Wellcome Sanger Institute Scientific Operations: DNA Pipelines collective are listed here: https://doi.org/10.5281/zenodo.4790455.

Members of the Tree of Life Core Informatics collective are listed here: https://doi.org/10.5281/zenodo.5013541.

Members of the Darwin Tree of Life Consortium are listed here: https://doi.org/10.5281/zenodo.4783558.

References

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

Bernt M, Donath A, Jühling F, *et al.*: **MITOS: Improved *de novo* metazoan mitochondrial genome annotation.** *Mol Phylogenet Evol.* 2013; **69**(2): 313–319.
PubMed Abstract | Publisher Full Text

Britton JR, Pegg J: **Ecology of European Barbel *Barbus Barbus*: Implications for River, Fishery, and Conservation Management.** *Reviews Fish Sci.* 2011; **19**(4): 321–330.
Publisher 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

De Santis V, Gutmann Roberts C, Britton JR: **Influences of angler subsidies on the trophic ecology of European barbel *Barbus barbus*.** *Fish Res.* 2019; **214**: 35–44.
Publisher Full Text

Freyhof J: ***Barbus barbus* (errata version published in 2016).** The IUCN Red List of Threatened Species 2011: e.T2561A97789324. 2011.

Gandlin AA, Japoshvili B, Epitashvili G, *et al.*: **Phylogeography of the Kura Barbel *Barbus cyri* De Filippi as Inferred from mtDNA Data.** *Inland Water Biol.* 2022; **15**(1): 11–22.
Publisher Full Text

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

Harry E: **PretextView (Paired REad TEXTure Viewer): A desktop application for viewing pretext contact maps.** 2022; (Accessed: 19 October 2022).
Reference Source

Howe K, Chow W, Collins J, *et al.*: **Significantly improving the quality of genome assemblies through curation.** *GigaScience.* Oxford University Press, 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

Kotlík P, Berrebi P: **Phylogeography of the barbel (*Barbus barbus*) assessed by mitochondrial DNA variation.** *Mol Ecol.* 2001; **10**(9): 2177–2185.
PubMed Abstract | Publisher Full Text

Lugowska K, Witeska M: **The effect of temperature on early development of barbel *Barbus barbus* (L.).** *Aquac Res.* 2018; **49**(7): 2495–2502.
Publisher 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

Peňáz M, Pivnička K, Baruš V, *et al.*: **Temporal changes in the abundance of barbel, *Barbus barbus,* in the Jihlava River, Czech Republic.** *Folia Zoologica.* 2003; **52**(4): 441–448.
Reference Source

Penczak T, Sierakowska K: **Anglers' records as a tool for assessing changes in fish populations.** *J Appl Ichthyol.* 2003; **19**(4): 250–254.
Publisher 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

Simão FA, Waterhouse RM, Ioannidis P, *et al.*: **BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs.** *Bioinformatics.* 2015; **31**(19): 3210–3212.
PubMed Abstract | Publisher Full Text

Supple MA, Shapiro B: **Conservation of biodiversity in the genomics era.** *Genome Biol.* 2018; **19**(1): 131.
PubMed Abstract | Publisher Full Text | Free Full Text

Uliano-Silva M, Ferreira JGRN, Krasheninnikova K, *et al.*: **MitoHiFi: a python pipeline for mitochondrial genome assembly from PacBio High Fidelity reads.** *bioRxiv.* [Preprint]. 2022.
Publisher Full Text

Vilizzi L, Copp GH, Britton JR: **Age and growth of European barbel *Barbus barbus* (Cyprinidae) in the small, mesotrophic River Lee and relative to other populations in England.** *Knowl Manag Aquat Ecosyst.* 2013; (409): 09.
Publisher Full Text

Wellcome Sanger Institute: **The genome sequence of the barbel, *Barbus barbus* (Linnaeus, 1758).**European Nucleotide Archive. [dataset], accession number PRJEB51453, 2022.

Wheeler A, Jordan DR: **The status of the barbel, *Barbus barbus* (L.) (Teleostei, Cyprinidae), in the United Kingdom.** *J Fish Biol.* 1990; **37**(3): 393–399.
Publisher Full Text

Open Peer Review

Current Peer Review Status: ✔

Version 1

Reviewer Report 06 June 2023

https://doi.org/10.21956/wellcomeopenres.21417.r57156

✔ **Guoqing Lu**
Department of Biology, University of Nebraska Omaha, Omaha, Nebraska, USA

The paper titled "The genome sequence of the barbel, *Barbus barbus* (Linnaeus, 1758)" is well-written with a comprehensive and detailed description of the genome sequencing and assembly process. The authors employed a combination of Illumina and PacBio sequencing technologies, leading to the generation of a high-quality genome assembly. The assessment shows the draft genome is of high quality. The authors have made the datasets generated during this study freely available for download from the European Bioinformatics Institute (EBI).

It is worth noting that although gene annotation is currently lacking, the authors acknowledge ongoing efforts in this aspect of the project. While RNA-Seq is briefly mentioned, its specific role and contributions are not highlighted in the paper.

Overall, this study is a well-executed and noteworthy contribution to the field.

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: Genomics, bioinformatics, fish 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.