




DATA NOTE

The genome sequence of a bird blowfly, *Protocalliphora azurea* (Fallén, 1817) [version 1; peer review: 3 approved]

Steven Falk¹, Darwin Tree of Life Barcoding collective, Wellcome Sanger Institute Tree of Life programme, Wellcome Sanger Institute Scientific Operations: DNA Pipelines collective, Tree of Life Core Informatics collective, Olga Sivell ², Darwin Tree of Life Consortium

¹Independent researcher, Kenilworth, Warwickshire, UK

²Natural History Museum, London, UK

V1 First published: 09 Feb 2023, 8:67
<https://doi.org/10.12688/wellcomeopenres.18872.1>
Latest published: 09 Feb 2023, 8:67
<https://doi.org/10.12688/wellcomeopenres.18872.1>

Abstract

We present a genome assembly from an individual male *Protocalliphora azurea* (a bird blowfly; Arthropoda; Insecta; Diptera; Calliphoridae). The genome sequence is 874 megabases in span. Most of the assembly is scaffolded into seven chromosomal pseudomolecules, including the assembled X and Y sex chromosomes. The mitochondrial genome has also been assembled and is 15.7 kilobases in length. Gene annotation of this assembly on Ensembl identified 11,936 protein coding genes.

Keywords




Protocalliphora azurea, bird blowfly, genome sequence, chromosomal, Diptera





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

Open Peer Review

Approval Status 

	1	2	3
version 1 09 Feb 2023	 view	 view	 view

1. **Xuan Li** , Uppsala University, Uppsala, Sweden
2. **Andrew D Sweet**, Arkansas State University, Jonesboro, USA
3. **Sanjay Kumar Pradhan** , University of Agricultural Sciences, Bangalore, India

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: Falk S: Investigation, Resources; Sivell O: 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>). *The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.*

Copyright: © 2023 Falk S *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: Falk S, Darwin Tree of Life Barcoding collective, Wellcome Sanger Institute Tree of Life programme *et al.* **The genome sequence of a bird blowfly, *Protocalliphora azurea* (Fallén, 1817) [version 1; peer review: 3 approved]** Wellcome Open Research 2023, 8:67 <https://doi.org/10.12688/wellcomeopenres.18872.1>

First published: 09 Feb 2023, 8:67 <https://doi.org/10.12688/wellcomeopenres.18872.1>

Species taxonomy

Eukaryota; Metazoa; Ecdysozoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Oestroidea; Calliphoridae; Chrysomyinae; *Protocalliphora*; *Protocalliphora azurea* (Fallén, 1817) (NCBI: txid226133).

Background

Protocalliphora azurea (Diptera, Calliphoridae) is easily recognisable in the field as a shiny metallic fly with dusting and longitudinal stripes on the thorax. The males are dark blue, while females are bottle-green (Rognes, 1991; Sivell, 2021). This blowfly is common and widely distributed in England and Wales, but uncommon and scattered in Scotland. It has been recorded from March to October (Highland Biological Recording Group, 2021; Sivell, 2021). This species is attracted to flowers, honeydew, rotting fruit, carrion, manure and bird droppings.

Protocalliphora azurea is commonly referred to as a bird blowfly due to its biology. While adult flies feed mainly on flowers, the larvae are obligatory external parasites of birds. They infest variety of bird species (Owen, 1954; Owen & Ash, 1955; Zumpt, 1965). The larvae suck blood of nestlings causing primary myiasis. Usually only a few larvae are involved and the host recovers, although death from blood loss may occur. Subsequently, the birds may also become infested with other fly species, e.g. *Protophormia terraenovae* (species from the same subfamily as *P. azurea*: Chrysomyinae) in secondary myiasis (van Emden, 1954; Zumpt, 1965).

Protocalliphora azurea is oviparous. Larvae feed approximately for one week, then pupate in the host's nest. Adult flies emerge after 7–12 days (Draber-Mońko, 2004). Duration of development is dependent on temperature and other conditions (Draber-Mońko, 2004; Zumpt, 1965). The pupae are often infested by the parasitoid wasp *Nasonia vitripennis* (Walker, 1836) (Draber-Mońko, 1997). The fly overwinters as adult (Draber-Mońko, 2004). It has been reared from bird nests, also found in houses and in insect hotels (Sivell, 2021; van Emden, 1954).

The high-quality genome assembly described here, generated as part of the Darwin Tree of Life Project, is the first reported for *Protocalliphora azurea*. It will aid research on the taxonomy, biology and ecology of the species.

Genome sequence report

The genome was sequenced from one male *Protocalliphora azurea* specimen (Figure 1) collected from Wytham Woods, Oxfordshire (biological vice-county: Berkshire), UK (latitude 51.769, longitude -1.34). A total of 39-fold coverage in Pacific Biosciences single-molecule HiFi long reads and 78-fold coverage in 10X Genomics read clouds were generated. Primary assembly contigs were scaffolded with chromosome conformation Hi-C data. Manual assembly curation corrected 190 missing joins or mis-joins and removed one haplotypic



Figure 1. Image of the *Protocalliphora azurea* (idProAzur1) specimen used for genome sequencing.

duplications, reducing the scaffold number by 81.3% and increasing the scaffold N50 by 7.15%.

The final assembly has a total length of 874.2 Mb in 23 sequence scaffolds with a scaffold N50 of 164.2 Mb (Table 1). Most (99.96%) of the assembly sequence was assigned to seven chromosomal-level scaffolds, representing five autosomes and the X and Y sex chromosomes. Chromosome-scale scaffolds confirmed by the Hi-C data are named in order of size (Figure 2–Figure 5; Table 2). The assembly has a BUSCO 5.3.2 (Manni *et al.*, 2021) completeness of 98.7% (single 98.4%, duplicated 0.4%) using the OrthoDB v10 Diptera reference set ($n=3,285$). While not fully phased, the assembly deposited is of one haplotype. Contigs corresponding to the second haplotype have also been deposited.

Genome annotation report

The idProAzur1.1 genome assembly was annotated using the Ensembl rapid annotation pipeline (Table 1; https://rapid.ensembl.org/Protocalliphora_azurea_GCA_932274085.1/). The resulting annotation includes 18,968 transcribed mRNAs from 11,936 protein-coding and 1,744 non-coding genes.

Methods

Sample acquisition and nucleic acid extraction

A male *Protocalliphora azurea* (idProAzur1) was caught using a net in Wytham Woods, Oxfordshire (biological vice-county: Berkshire), UK (latitude 51.769, longitude -1.34) on 20 August 2020. The specimen was collected and identified by Steven Falk (independent researcher), and was then snap-frozen on dry ice.

DNA was extracted at the Tree of Life laboratory, Wellcome Sanger Institute (WSI). The idProAzur1 sample was weighed and dissected on dry ice with head tissue set aside for Hi-C sequencing. Thorax was disrupted using a Nippi Powermasher fitted with a BioMasher pestle. High molecular weight (HMW) DNA was extracted using the Qiagen MagAttract HMW DNA extraction kit. Low molecular weight DNA was removed from a 20 ng aliquot of extracted DNA using 0.8X AMPure XP

Table 1. Genome data for *Protocalliphora azurea*, idProAzur1.1.

Project accession data		
Assembly identifier	idProAzur1.1	
Species	<i>Protocalliphora azurea</i>	
Specimen	idProAzur1	
NCBI taxonomy ID	226133	
BioProject	PRJEB48588	
BioSample ID	SAMEA7746778	
Isolate information	male, whole organism	
Assembly metrics		Benchmark
BUSCO*	C:98.7%[S:98.4%,D:0.4%], F:0.5%,M:0.8%,n:3,285	$C \geq 95\%$
Percentage of assembly mapped to chromosomes	99.96%	$\geq 95\%$
Sex chromosomes	X and Y chromosomes	<i>localised homologous pairs</i>
Organelles	Mitochondrial genome assembled	<i>complete single alleles</i>
Raw data accessions		
PacificBiosciences SEQUEL II	ERR7254656, ERR7254657	
10X Genomics Illumina	ERR7253241–ERR7253244	
Hi-C Illumina	ERR7253240	
PolyA RNA-Seq Illumina	ERR10123665	
Genome assembly		
Assembly accession	GCA_932274085.1	
Accession of alternate haplotype	GCA_932273785.1	
Span (Mb)	874.2	
Number of contigs	411	
Contig N50 length (Mb)	8.4	
Number of scaffolds	23	
Scaffold N50 length (Mb)	164.2	
Longest scaffold (Mb)	212.8	
Genome annotation		
Number of protein-coding genes	11,936	
Non-coding genes	1,744	
Gene transcripts	18,968	

* BUSCO scores based on the diptera_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/idProAzur1.1/dataset/CAKNZY01/busco>.

purification kit prior to 10X Chromium sequencing; a minimum of 50 ng DNA was submitted for 10X sequencing. 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

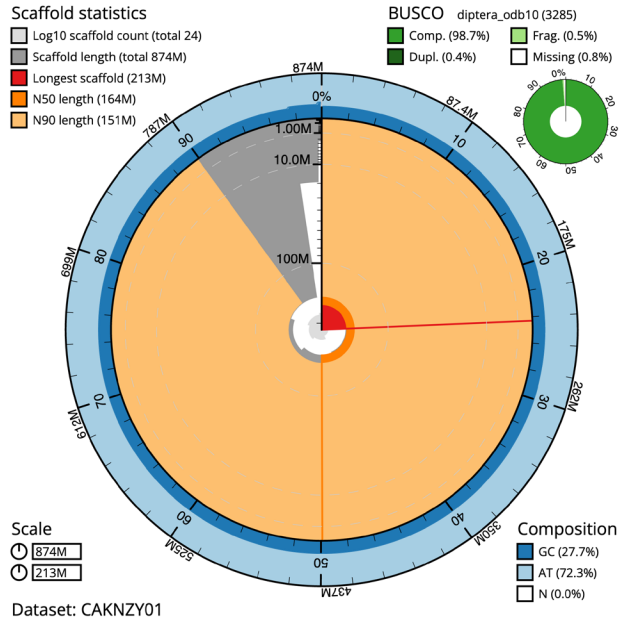


Figure 2. Genome assembly of *Protocalliphora azurea*, idProAzur1.1: metrics. The BlobToolKit Snailplot shows N50 metrics and BUSCO gene completeness. The main plot is divided into 1,000 size-ordered bins around the circumference with each bin representing 0.1% of the 874,252,225 bp assembly. The distribution of scaffold lengths is shown in dark grey with the plot radius scaled to the longest scaffold present in the assembly (212,808,213 bp, shown in red). Orange and pale-orange arcs show the N50 and N90 scaffold lengths (164,199,054 and 151,389,329 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 diptera_odb10 set is shown in the top right. An interactive version of this figure is available at <https://blobtoolkit.genomehubs.org/view/idProAzur1.1/dataset/CAKNZY01/snail>.

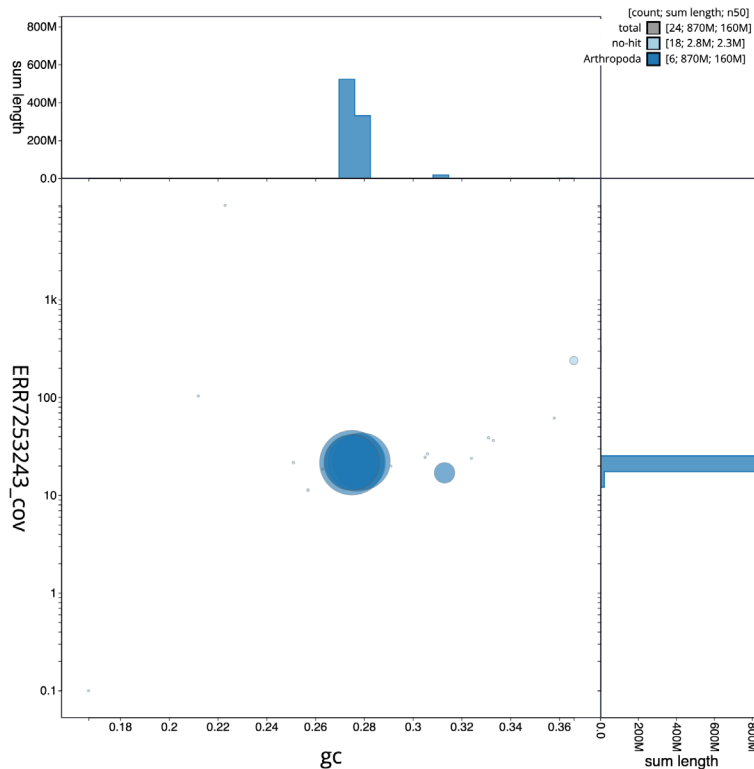


Figure 3. Genome assembly of *Protocalliphora azurea*, idProAzur1.1: GC coverage. BlobToolKit GC-coverage plot. Scaffolds are coloured by phylum. Circles are sized in proportion to scaffold length. Histograms show the distribution of scaffold length sum along each axis. An interactive version of this figure is available at <https://blobtoolkit.genomehubs.org/view/idProAzur1.1/dataset/CAKNZY01/blob>.

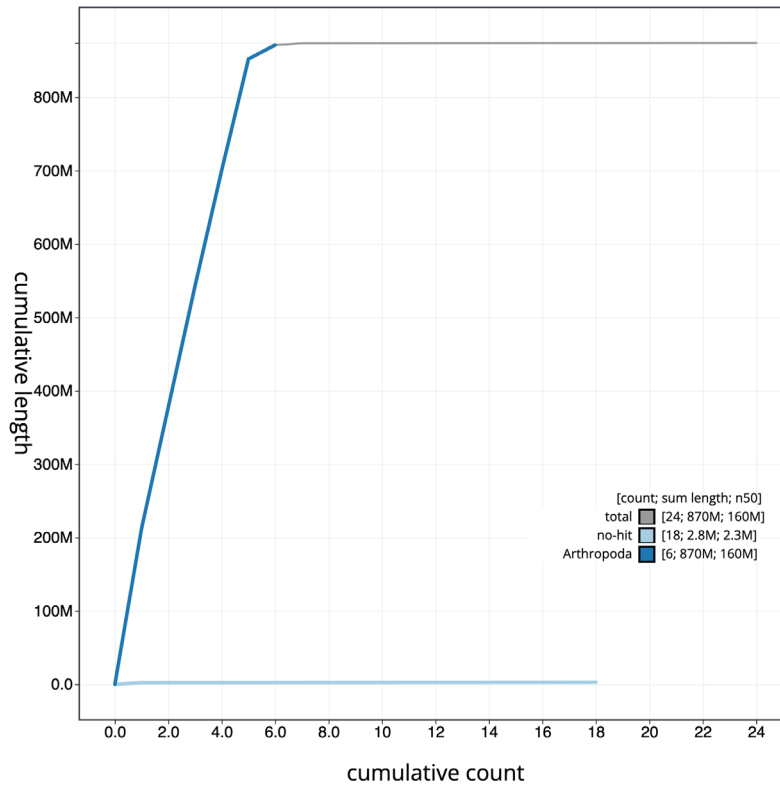


Figure 4. Genome assembly of *Protocalliphora azurea*, idProAzur1.1: cumulative sequence. BlobToolKit cumulative sequence plot. The grey line shows cumulative length for all scaffolds. Coloured lines show cumulative lengths of scaffolds assigned to each phylum using the buscogenes taxrule. An interactive version of this figure is available at <https://blobtoolkit.genomehubs.org/view/idProAzur1.1/dataset/CAKNZY01/cumulative>.

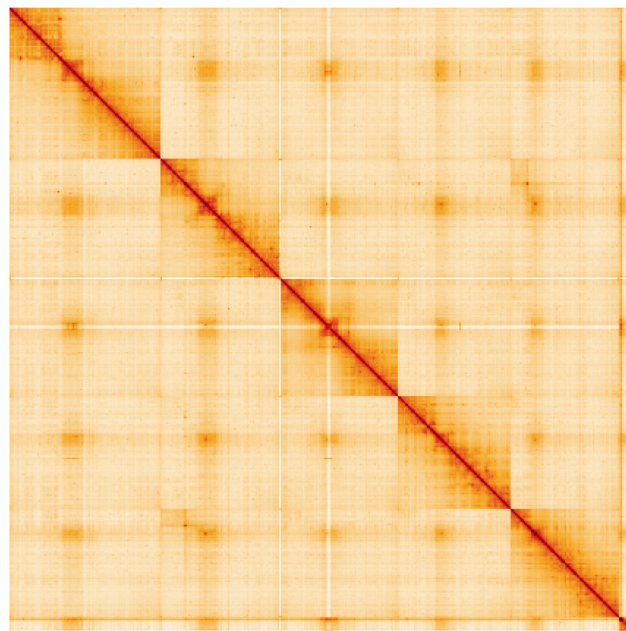


Figure 5. Genome assembly of *Protocalliphora azurea*, idProAzur1.1: Hi-C contact map. Hi-C contact map of the idProAzur1.1 assembly, visualised using HiGlass. Chromosomes are shown in order of size from left to right and top to bottom. An interactive version of this figure may be viewed at https://genome-note-higlass.tol.sanger.ac.uk/?d=ftusVv38SaqwPwj6_Ij3OA.

Table 2. Chromosomal pseudomolecules in the genome assembly of *Protocalliphora azurea*, idProAzur1.

INSDC accession	Chromosome	Size (Mb)	GC%
OW026519.1	1	212.81	27.5
OW026520.1	2	166.18	27.7
OW026521.1	3	164.2	27.9
OW026522.1	4	157.91	27.5
OW026523.1	5	151.39	27.5
OW026524.1	X	18.98	31.3
OW026525.1	Y	2.32	36.6
OW026526.1	MT	0.02	22.3
-	unplaced	0.45	28.4

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.

RNA was extracted from abdomen tissue of idProAzur1 in the Tree of Life Laboratory at the 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 and 10X Genomics read cloud 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 (10X and RNA-Seq) instruments. Hi-C data were also generated from head tissue of idProAzur1 using the Arima v2 kit and sequenced on the Illumina NovaSeq 6000 instrument.

Genome assembly

Assembly was carried out with Hifiasm (Cheng *et al.*, 2021) and haplotypic duplication was identified and removed with purge_dups (Guan *et al.*, 2020). One round of polishing was performed by aligning 10X Genomics read data to the assembly with Long Ranger ALIGN, calling variants with freebayes (Garrison & Marth, 2012). The assembly was then scaffolded with Hi-C data (Rao *et al.*, 2014) using YaHS (Zhou *et al.*, 2022). 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 performed annotation using MitoFinder (Allio *et al.*, 2020). The genome was analysed and BUSCO scores generated within the BlobToolKit environment (Challis *et al.*, 2020). Table 3 contains a list of all software tool versions used, where appropriate.

Genome annotation

The Ensembl gene annotation system (Aken *et al.*, 2016) was used to generate annotation for the *Protocalliphora azurea* assembly GCA_932274085.1. Annotation was created primarily through alignment of transcriptomic data to the genome, with gap filling via protein to-genome alignments of a select set of proteins from UniProt (UniProt Consortium, 2019).

Ethics/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 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. 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.

Table 3. Software tools and versions used.

Software tool	Version	Source
BlobToolKit	3.2.6	Challis <i>et al.</i> , 2020
freebayes	1.3.1-17-gaa2ace8	Garrison & Marth, 2012
Hifiasm	0.15.3	Cheng <i>et al.</i> , 2021
HiGlass	1.11.6	Kerpedjiev <i>et al.</i> , 2018
Long Ranger ALIGN	2.2.2	https://support.10xgenomics.com/genome-exome/software/pipelines/latest/advanced/other-pipelines
MitoHiFi	2	Uliano-Silva <i>et al.</i> , 2022
PretextView	0.2	Harry, 2022
purge_dups	1.2.3	Guan <i>et al.</i> , 2020
YaHS	1	Zhou <i>et al.</i> , 2022

Data availability

European Nucleotide Archive: *Protocalliphora azurea* (bird blowfly). Accession number [PRJEB48588](https://identifiers.org/ena.embl/PRJEB48588); <https://identifiers.org/ena.embl/PRJEB48588>. (Wellcome Sanger Institute, 2022). The genome sequence is released openly for reuse. The *Protocalliphora azurea* 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. Raw data and assembly accession identifiers are reported in [Table 1](#).

Author information

Members of the University of Oxford and Wytham Woods Genome Acquisition Lab are listed here: <https://doi.org/10.5281/zenodo.4789928>.

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

- Aken BL, Ayling S, Barrell D, *et al.*: **The Ensembl gene annotation system.** *Database (Oxford)*. 2016; **2016**: baw093.
[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](#)
- 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](#)
- Draber-Mońko A: ***Protocalliphora azurea* (Fall.) (Diptera, Calliphoridae) and other insects found in nests of sparrows, *Passer domesticus* (L.) and *Passer montanus* (L.) in the vicinity of Warsaw.** *International Studies on Sparrows*. 1997; **22-23**: 3–10.
- Draber-Mońko A: **Calliphoridae.** *Plujki (Insecta: Diptera)*. 2004.
[Reference Source](#)
- Garrison E, Marth G: **Haplotype-based variant detection from short-read sequencing.** 2012.
[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](#)
- Highland Biological Recording Group: **HBRG Insects Dataset.** Occurrence dataset. 2021; (Accessed: 25 December 2022).
[Reference Source](#)
- Howe K, Chow W, Collins J, *et al.*: **Significantly improving the quality of genome assemblies through curation.** *GigaScience*. 2021; **10**(1): g1aa153.
[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](#)
- Manni M, Berkeley MR, Seppy 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](#)
- Owen DF: **Protocalliphora in birds' nests.** *British Birds*. 1954; **47**: 236–243.
- Owen DF, Ash JS: **Additional records of Protocalliphora (Diptera) in birds' nests.** *British Birds*. 1955; **48**(5): 225–229.
[Reference Source](#)
- 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](#)
- Rognes K: **Blowflies (Diptera, Calliphoridae) of Fennoscandia and Denmark.** *Fauna Entomologica Scandinavica*. Leiden: Brill/Scandinavian Science Press Ltd, 1991.
[Reference Source](#)
- Sivell O: **Blow flies (Diptera: Calliphoridae, Polleniidae, Rhiniidae).** *Handbooks for the Identification of British Insect*. St Albans: Royal Entomological Society. 2021; **10**(16).
[Reference Source](#)
- Uliano-Silva M, Ferreira JGRN, Krasheninnikova F, *et al.*: **MitoHiFi: a python pipeline for mitochondrial genome assembly from PacBio High Fidelity reads.** *bioRxiv*. [Preprint]. 2022.
[Publisher Full Text](#)
- UniProt Consortium: **UniProt: a worldwide hub of protein knowledge.** *Nucleic Acids Res*. 2019; **47**(D1): D506–D515.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- van Emden FI: **Diptera Cyclorrhapha, Calypttrata (I) Section (a). Tachinidae and Calliphoridae.** In: *Handbooks for the identification of British insects*. Entomological Society of London, 1954.
[Reference Source](#)
- Wellcome Sanger Institute: **The genome sequence of a bird blowfly, *Protocalliphora azurea* (Scopoli, 1763).** *European Nucleotide Archive*. [Dataset]. 2022.
<https://identifiers.org/ena.embl/PRJEB48588>
- Zhou C, McCarthy SA, Durbin R: **YaHS: yet another Hi-C scaffolding tool.** *bioRxiv*. [Preprint]. 2022.
[Publisher Full Text](#)
- Zumpt F: **Myiasis in Man and Animals in the Old World. A Textbook for Physicians, Veterinarians and Zoologists.** London: Butterworths, 1965.
[Reference Source](#)

Open Peer Review

Current Peer Review Status: ✓ ✓ ✓

Version 1

Reviewer Report 13 September 2023

<https://doi.org/10.21956/wellcomeopenres.20926.r65700>

© 2023 Pradhan S. 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.



Sanjay Kumar Pradhan 

University of Agricultural Sciences, Bangalore, India

The manuscript is well structured, and all the methodologies are clearly mentioned.

1. Why was only the male fly sampled? Authors should explain why a single fly was sampled, rather than having some replications.
2. What taxonomic and ecological studies details are available currently in the domain, what are the loopholes the author finds regarding available information about *Protocalliphora azurea*, and how will the current findings fill the gap?"

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: Gene Editing, Insect metagenomics

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 11 September 2023

<https://doi.org/10.21956/wellcomeopenres.20926.r65687>

© 2023 Sweet A. 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.



Andrew D Sweet

Arkansas State University, Jonesboro, Arkansas, USA

The authors sequenced, assembled, and annotated the genome and mitogenome from an individual male specimen of a species of bird blowfly (*Protocalliphora azurae*) using a combination of PacBio HiFi, 10X, RNA-Seq, and Hi-C data.

This genome report clearly lays out the approaches and assembly statistics for a species of bird blowfly. The chromosome-level assembly seems of high quality. I have a few comments that I think can help improve and clarify the report.

1. Can you provide more details about the manual assembly curation? How exactly did you correct for missing joins and haplotypic duplications?
2. Is it possible to use a clearer image of the organism for Figure 1? I understand the use of having images of the vials, but it is difficult to see features of the adult fly.
3. It is unclear you utilized RNA-Seq until the Methods section. I recommend including this information at the beginning of the Genome sequence report section.
4. The mitochondrial genome did not have annotations associated with the GenBank accession provided, so I was unable to assess the mitogenome annotation. Please ensure the annotation details are available with the GenBank record.
5. GC-content seems to be on the low side, although not the lowest for insects. Was the nucleotide content unsurprising for this taxon?

Minor comments:

Background:

Paragraph 2, Line 2, "They infest a variety..."

Paragraph 3, Line 4: "overwinters as an adult"

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

Yes

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?

Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Evolutionary genomics, mitogenomics, host-parasite coevolution, phylogenomics, population genomics, ornithology, entomology

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 15 June 2023

<https://doi.org/10.21956/wellcomeopenres.20926.r60239>

© 2023 Li X. 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.



Xuan Li 

Uppsala University, Uppsala, Sweden

In the article, the authors present a chromosomal-level assembly of male genome of the blowfly, *Protocalliphora azurea*. Collected data (i.e. Pacbio Hifi, 10X reads and Hi-C reads) are sufficient and the output is of good quality. The applied methodology, data statistics, and accessibility are well documented. Such resources will be handy for future studies related to *P. azurea*.

1. However, in the intro, besides description of life traits, it is not clear what kinds of research have been conducted and how this new genome could facilitate studies of *P. azurea* in the future.
2. The XY assemblies (18.98 Mb and 2.32 Mb respectively) are significantly smaller than those of autosomes (ranging from 151.39 Mb to 212.81 Mb). Knowing the difficulties when it comes to sex chromosome assembly, the authors should address how well their assembly represents the sequences of sex chromosomes. This is important especially considering it will serve as a reference for genomic studies to come.

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: Evolutionary genetics, sex chromosome evolution, insect sex determination

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.
