DATA NOTE

The genome sequence of the common green lacewing, *Chrysoperla carnea* (Stephens, 1836) [version 1; peer review: 1 approved, 2 approved with reservations]

Liam M. Crowley ¹,
 University of Oxford and Wytham Woods Genome Acquisition Lab,
 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, Darwin Tree of Life Consortium

¹Department of Zoology, University of Oxford, Oxford, UK

V1 First published: 09 Dec 2021, 6:334
<https://doi.org/10.12688/wellcomeopenres.17455.1>
 Latest published: 09 Dec 2021, 6:334
<https://doi.org/10.12688/wellcomeopenres.17455.1>

Abstract

We present a genome assembly from an individual female *Chrysoperla carnea* (a common green lacewing; Arthropoda; Insecta; Neuroptera; Chrysopidae). The genome sequence is 560 megabases in span. The majority of the assembly (95.70%) is scaffolded into six chromosomal pseudomolecules, with the X sex chromosome assembled. Gene annotation of this assembly by the NCBI Eukaryotic Genome Annotation Pipeline has identified 12,985 protein coding genes.

Keywords

Chrysoperla carnea, common green lacewing, genome sequence, chromosomal, Chrysopidae

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

Open Peer Review

Approval Status

	1	2	3
version 1 09 Dec 2021	 view	 view	 view

1. **John Torcivia** , George Washington University, Washington, USA

2. **Jessica Castellanos-Labarcena** , University of Guelph, Guelph, Canada

3. **Emily Hornett**, University of Liverpool, Liverpool, UK

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: Crowley LM: Investigation, Resources, 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) 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 Crowley LM *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: Crowley LM, University of Oxford and Wytham Woods Genome Acquisition Lab, Darwin Tree of Life Barcoding collective *et al.* **The genome sequence of the common green lacewing, *Chrysoperla carnea* (Stephens, 1836) [version 1; peer review: 1 approved, 2 approved with reservations]** Wellcome Open Research 2021, 6:334 <https://doi.org/10.12688/wellcomeopenres.17455.1>

First published: 09 Dec 2021, 6:334 <https://doi.org/10.12688/wellcomeopenres.17455.1>

Species taxonomy

Eukaryota; Metazoa; Ecdysozoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Neuroptera; Hemerobiformia; Chrysopidae; Chrysopinae; Chrysoperla; *Chrysoperla carnea* (Stephens, 1836) (NCBI:txid189513).

Background

Chrysoperla carnea, a common green lacewing, is a common and widespread lacewing across the Holarctic. It is one of the most common species of lacewing in the UK, found across a wide range of habitats. This species is part of a species complex that contains several cryptic species. These species can be distinguished by differences in the substrate-borne songs produced by adults via abdominal vibrations (Henry *et al.*, 2002). In the UK the *C. carnea* group is currently split into two species, *C. carnea* sensu stricto and *C. lucasina*, both of which appear to be common. A third species, *Chrysoperla pallida*, may also be present. *Chrysoperla carnea* overwinters as an adult in common with all *Chrysoperla* species, but has the unique trait of losing its green pigment and turning yellow-brown during the winter period. The larvae are voracious generalist predators, feeding on aphids and other insects (Rosenheim *et al.*, 1999), including several other pest groups such as spider mites, thrips, whitefly, leafhoppers, psyllids and Lepidoptera. They have been used extensively as biocontrol agents in agricultural and horticultural systems and are commercially produced for this purpose. Adults visit flowers and feed on pollen and nectar. Females have been recorded consuming more pollen than males (Villenave *et al.*, 2005). The eggs are laid on vegetation and are suspended off the surface on characteristic stalks.

Genome sequence report

The genome was sequenced from one male *C. carnea* collected from Wytham Woods, Oxfordshire (biological vice-county: Berkshire), UK (latitude 51.772, longitude -1.338). A total of 40-fold coverage in Pacific Biosciences single-molecule long reads and 147-fold coverage in 10X Genomics read clouds were generated. Primary assembly contigs were scaffolded with chromosome conformation Hi-C data. Manual assembly curation corrected 20 missing/misjoins and removed 6 haplotypic duplications, reducing the assembly size by 5.41% and the scaffold number by 2.32% and increasing the scaffold N50 by 0.55%.

The final assembly has a total length of 560 Mb in 337 sequence scaffolds with a scaffold N50 of 94.4 Mb (Table 1). The majority of the assembly sequence (95.70%) was assigned to six chromosomal-level scaffolds, representing five autosomes (numbered by sequence length), and the X sex chromosome (Figure 1–Figure 4; Table 2). There is a very large repeat associated with the X chromosome, which has resulted in the presence of many unlocalised scaffolds in the assembly. The assembly has a BUSCO v5.1.2 (Manni *et al.*, 2021) completeness of 95.9% (single 95.0%, duplicated 0.9%) using the endopterygota_odb10 reference set. While not fully phased, the assembly deposited is of one haplotype. Contigs corresponding to the second haplotype have also been deposited.

Table 1. Genome data for *Chrysoperla carnea*, inChrCarn1.1.

Project accession data	
Assembly identifier	inChrCarn1.1
Species	<i>Chrysoperla carnea</i>
Specimen	inChrCarn1
NCBI taxonomy ID	NCBI:txid189513
BioProject	PRJEB43807
BioSample ID	SAMEA7520372
Isolate information	Female, whole organism
Raw data accessions	
PacificBiosciences SEQUEL II	ERR6406208
10X Genomics Illumina	ERR6054664-ERR6054667
Hi-C Illumina	ERR6054668
Genome assembly	
Assembly accession	GCA_905475395.1
Accession of alternate haplotype	GCA_905475295.1.
Span (Mb)	560
Number of contigs	399
Contig N50 length (Mb)	67.8
Number of scaffolds	337
Scaffold N50 length (Mb)	94.4
Longest scaffold (Mb)	140
BUSCO genome score*	C:95.9%[S:95.0%,D:0.9%],F:1.0%,M:3.1%,n:2124
Genome annotation	
Number of genes	15,736
Number of protein-coding genes	12,985
Average length of gene (bp)	18,136
Average number of exons per transcript	4.07
Average exon size (bp)	337
Average intron size (bp)	4,304
BUSCO annotation score**	C:96.9%[S:96%,D:0.9%],F:0.2%,M:2.8%,n:2124

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

*BUSCO scores based on the endopterygota_odb10 BUSCO set using v5.1.2, run on the inChrCarn1.1 genome assembly using BlobToolKit. A full set of BUSCO scores is available at <https://blobtoolkit.genomehubs.org/view/inChrCarn1.1/dataset/CAJQGA01/busco>.

**BUSCO scores based on the endopterygota_odb10 BUSCO set using v4.1.4, run on the NCBI RefSeq annotation of the inChrCarn1.1 genome assembly (NCBI *Chrysoperla carnea* Annotation Release 100).

Scaffold statistics

- Log10 scaffold count (total 340)
- Scaffold length (total 560M)
- Longest scaffold (140M)
- N50 length (94M)
- N90 length (39M)

BUSCO endopterygota_odb10 (2124)

- Complete (95.9%)
- Fragmented (1.0%)
- Duplicated (0.9%)
- Missing (3.1%)

Figure 1. Genome assembly of *Chrysoperla carnea*, inChrCarn1.1: 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 560,248,701 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 (139,979,878 bp, shown in red). Orange and pale-orange arcs show the N50 and N90 scaffold lengths (94,407,144 and 38,618,709 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 endopterygota_odb10 set is shown in the top right. An interactive version of this figure is available at <https://blobtoolkit.genomehubs.org/view/inChrCarn1.1/dataset/CAJQGA01/snail>.

Figure 2. Genome assembly of *Chrysoperla carnea*, inChrCarn1.1: 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/inChrCarn1.1/dataset/CAJQGA01/blob>.

Genome annotation report

The inChrCarn1.1 genome has been annotated using the NCBI RefSeq annotation pipeline (Table 1; NCBI *Chrysoperla carnea* Annotation Release 100). The resulting annotation includes 17,649 transcribed mRNAs from 12,985 protein-coding and 2,751 non-coding genes. There is an average of 1.31 transcripts per gene and 4.07 exons per transcript. The annotated

genome has a BUSCO v4.1.4 completeness of 96.9% using the endopterygota_odb10 reference set.

Methods

Sample acquisition, DNA extraction and sequencing

A single female *C. carnea* was collected from Wytham Woods, Oxfordshire (biological vice-county: Berkshire), UK (latitude

Figure 3. Genome assembly of *Chrysoperla carnea*, inChrCarn1.1: 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/inChrCarn1.1/dataset/CAJQGA01/cumulative>.

51.772, longitude -1.338) by Liam Crowley, University of Oxford, using a pooter. The sample was identified by the same individual, and preserved on dry ice.

DNA was extracted from the whole organism of inChrCarn1 at the Wellcome Sanger Institute Scientific Operations core

from the whole organism using the Qiagen MagAttract HMW DNA kit, according to the manufacturer's instructions. Pacific Biosciences HiFi circular consensus and 10X Genomics read cloud DNA sequencing libraries were constructed according to the manufacturers' instructions. Sequencing was performed by the Scientific Operations core at the WSI on Pacific

Figure 4. Genome assembly of *Chrysoperla carnea*, inChrCarn1.1: Hi-C contact map. Hi-C contact map of the inChrCarn1.1 assembly, visualised in HiGlass. Chromosomes are shown in order of size from left to right and top to bottom.

Table 2. Chromosomal pseudomolecules in the genome assembly of *Chrysoperla carnea*, inChrCarn1.1.

INSDC accession	Chromosome	Size (Mb)	GC%
FR997754.1	1	139.98	29.1
FR997755.1	2	100.74	29
FR997756.1	3	94.41	28.9
FR997757.1	4	77.48	28.8
FR997758.1	5	76.43	29.2
FR997759.1	X	38.62	31.1
FR997760.1	MT	0.02	21.1
-	Unplaced	32.58	27.8

Biosciences SEQUEL II and Illumina HiSeq X instruments. Hi-C data were generated using the Arima v2 Hi-C kit and sequenced on a HiSeq X instrument.

Genome assembly

Assembly was carried out with Hifiasm (Cheng *et al.*, 2021); 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 longranger align, calling variants with freebayes (Garrison & Marth, 2012). The assembly was then scaffolded with Hi-C data (Rao *et al.*, 2014) using SALSA2 (Ghurye *et al.*, 2019). The assembly was checked for contamination and corrected using the gEVAL system (Chow *et al.*, 2016) as described previously (Howe *et al.*, 2021). Manual curation (Howe *et al.*, 2021) was performed using gEVAL, HiGlass (Kerpedjiev *et al.*, 2018) and Pretext. The mitochondrial genome was assembled using MitoHiFi (Uliano-Silva *et al.*, 2021), 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 *C. carnea* assembly was annotated by the NCBI Eukaryotic Genome Annotation Pipeline, an automated pipeline

Table 3. Software tools used.

Software tool	Version	Source
Hifiasm	0.12	Cheng et al., 2021
purge_dups	1.2.3	Guan et al., 2020
SALSA2	2.2	Ghurye et al., 2019
longranger align	2.2.2	https://support.10xgenomics.com/genome-exome/software/pipelines/latest/advanced/other-pipelines
freebayes	1.3.1-17-gaa2ace8	Garrison & Marth, 2012
MitoHiFi	1.0	Uliano-Silva et al., 2021
gEVAL	N/A	Chow et al., 2016
HiGlass	1.11.6	Kerpedjiev et al., 2018
PretextView	0.1.x	https://github.com/wtsi-hpag/PretextView
BlobToolKit	2.6.2	Challis et al., 2020

that annotates genes, transcripts and proteins on draft and finished genome assemblies. The annotation was generated from transcripts and proteins retrieved from NCBI Entrez by alignment to the genome assembly, [as described here \(Pruitt et al., 2014\)](#).

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.

Data availability

European Nucleotide Archive: *Chrysoperla carnea* (common green lacewing). Accession number [PRJEB43807](#); <https://identifiers.org/ena.embl/PRJEB43807>.

The genome sequence is released openly for reuse. The *C. carnea* 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.4789929>.

Members of the Darwin Tree of Life Barcoding collective are listed here: <https://doi.org/10.5281/zenodo.4893704>.

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>.

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](#)

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](#)

Cheng H, Concepcion GT, Feng X, et al.: **Haplotype-Resolved de Novo**

Assembly Using Phased Assembly Graphs with Hifiasm. *Nat Methods*. 2021; **18**(2): 170–75.

[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](#)

Garrison E, Marth G: **Haplotype-Based Variant Detection from Short-Read Sequencing.** arXiv: 1207.3907. 2012.

[Reference Source](#)

Ghurye J, Rhie A, Walenz BP, *et al.*: **Integrating Hi-C Links with Assembly Graphs for Chromosome-Scale Assembly.** *PLoS Comput Biol*. 2019; **15**(8): e1007273.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Guan D, McCarthy SA, Wood J, *et al.*: **Identifying and Removing Haplotypic Duplication in Primary Genome Assemblies.** *Bioinformatics*. 2020; **36**(9): 2896–98.

[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Henry CS, Brooks SJ, Duelli P, *et al.*: **Discovering the True *Chrysoperla Carnea* (Insecta: Neuroptera: Chrysopidae) Using Song Analysis, Morphology, and Ecology.** *Ann Entomol Soc Am*. 2002; **95**(2): 172–91.

[Publisher Full Text](#)

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, 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): 4647–54.

[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](#)

Rosenheim JA, Limburg DD, Colfer RG: **Impact of Generalist Predators on a Biological Control agent, *Chrysoperla Carnea*: Direct Observations.**

Ecological Applications: A Publication of the Ecological Society of America. 1999; **9**(2): 409–17.

[Publisher Full Text](#)

Uliano-Silva M, Nunes JGF, Krasheninnikova K, *et al.*: **marcelauliano/MitoHiFi: mitohifi_v2.0.** 2021.

[Publisher Full Text](#)

Villénave J, Thierry D, Al Mamun A, *et al.*: **The Pollens Consumed by Common Green Lacewings *Chrysoperla* spp. (Neuroptera: Chrysopidae) in Cabbage Crop Environment in Western France.** *Eur J Entomol*. 2005; **102**(3): 547–552.

[Publisher Full Text](#)

Open Peer Review

Current Peer Review Status:

Version 1

Reviewer Report 19 April 2023

<https://doi.org/10.21956/wellcomeopenres.19301.r54901>

© 2023 Hornett E. 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.

Emily Hornett

University of Liverpool, Liverpool, UK

This article reports the sequencing, assembly and annotation of the genome of a green lacewing, *Chrysoperla carnea*.

This article is very concise and clearly written. There are a few comments:

The Background section would benefit from polishing. The first three lines contain the word "common" three times, and it appears again twice more in this paragraph. A pedantic point, but easily resolved e.g. change "in common with" to "similarly to".

The Background section would also benefit from two additions: 1) mention of whether lacewing genomes have been sequenced before (they have...just - e.g. Wang et al 2022 - but still underexplored) and 2) why this lacewing is being sequenced and what can we learn from this genome?

The main text states a male is sequenced, while in the Methods it says female - presumably it was a female sequenced given the lack of a Y chromosome.

The protocols followed are fairly standard, and the techniques sound. However, the Methods are very brief and there are few parameters described, which makes replication tricky. Given this article type, which is technical in nature, a bit more detail would be useful for readers.

For instance, although it is stated that the assembly was checked for contamination, no outcome is noted. Was there contamination? And were any cobionts (non-target organisms) present in the dataset?

In the DNA extraction section of the Methods, the wording needs clearing up, e.g. "whole organism" is repeated.

All in all, a useful addition to the genome sequencing project.

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

Partly

Are the protocols appropriate and is the work technically sound?

Yes

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

Partly

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

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Insect genomics, symbiosis, evolution, microbiomes

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

Reviewer Report 06 March 2023

<https://doi.org/10.21956/wellcomeopenres.19301.r54904>

© 2023 Castellanos-Labarcena J. 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.

Jessica Castellanos-Labarcena

University of Guelph, Guelph, Canada

This study presents a genome assembly and annotation for the species *Chrysoperla carnea*. The authors use three sequencing technologies to generate the data and a combination of various software tools to provide an almost complete genome scaffolded into six chromosomal pseudomolecules. The methodology is adequately described for each step, and informative visualizations and tables are included.

Below you can find my recommendations for the improvement of the document.

1. The abstract and methods sections state that the individual used for the study is female, but the 'Genome sequence report' section states male. This needs to be corrected before indexing.
2. In the 'Background' section, I suggest including a citation for the statement regarding using green lacewing as a biocontrol agent. I also suggest mentioning if any other genome has already been reported for this species complex.

3. In the 'Genome Sequence Report' section could be relevant to mention why the endopterygota_odb10 reference set was used and the effect that has on the results.
4. I suggest fixing the axis labels in Figure 2 (a raw sequence identifier is used in the y-axis; the 'sum length' is inverted) and improving the figure caption. The colours used are very similar and difficult to distinguish when the bubbles overlap. A legend referencing the bubble size to the scaffold length could be beneficial. Also, some of the bars on the histogram are barely visible.
5. In the 'Methods' section, I recommend providing more details about the morphological identification of the specimens. Were pictures taken?
6. In the 'Methods' section, the assembly identifier is used in reference to the species name "DNA was extracted from the whole organism of inChrCarn1..." this should be corrected. The phrase "the whole organism" is repeated twice in this sentence.
7. Figure 4 caption could include more details, for example, of how this figure is generated (visualization of the Hi-C data mapped to sequence data).
8. In the 'Genome Assembly' section, details about some of the parameters used in the software settings should be included when suitable.

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?

Partly

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

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: macrogenetics, bioinformatics, 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, however I have significant reservations, as outlined above.

Reviewer Report 16 September 2022

<https://doi.org/10.21956/wellcomeopenres.19301.r52091>

© 2022 Torcivia J. 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.

John Torcivia

George Washington University, Washington, DC, USA

The authors present a genome assembly (along with downstream annotation) for a species of lacewing. There is some background provided on the species and usefulness of full assemblies of the genome of any species is helpful to report.

Protocols used appear to be fairly standard and generally accepted methodologies to perform this assembly and downstream annotation of the genome. While these techniques do not guarantee a perfect assembly, they offer a strong starting point for future research to refine as needed.

Particular techniques are provided along with supporting papers. One potential improvement to the publication would be additional technical information (which parameters were chosen for each tool for example) to help with potential replication. Many parameters are related to computation cost but some could have an effect on the actual assembly itself. That said, it is unlikely that those selected deviate in a significant way from the standard parameters used for these tools and I don't have concern with the general replicability of this dataset. I like the inclusion of a table of the software tools used along with their versions and think this is a helpful addition to these types of articles.

The sequencing data is uploaded in FASTQ format which is a typical and well understood format. This is deposited into a public repository and available for download and additional utilization. Additional downstream annotation is provided in an NIH database and is represented in generally accepted and understood formats.

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, cancer genetics, biostatistics, NGS analysis

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.