



**HAL**  
open science

# Complete mitogenome data from a European specimen of *Ostrinia scapulalis* (Walker, 1859) (Lepidoptera, Pyraloidea, Crambidae, Pyraustinae)

Bernhard Gschloessl, Philippe Audiot, Sabine Nidelet, Gael Kergoat, Rejane  
Streiff

## ► To cite this version:

Bernhard Gschloessl, Philippe Audiot, Sabine Nidelet, Gael Kergoat, Rejane Streiff. Complete mitogenome data from a European specimen of *Ostrinia scapulalis* (Walker, 1859) (Lepidoptera, Pyraloidea, Crambidae, Pyraustinae). *Data in Brief*, 2020, 33, pp.106427. 10.1016/j.dib.2020.106427 . hal-02981600v2

**HAL Id: hal-02981600**

**<https://hal.inrae.fr/hal-02981600v2>**

Submitted on 29 Jun 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0  
International License



Data Article

Complete mitogenome data from a European specimen of *Ostrinia scapularis* (Walker, 1859) (Lepidoptera, Pyraloidea, Crambidae, Pyraustinae)



Bernhard Gschloessl\*, Philippe Audiot, Sabine Nidelet, Gael J Kergoat, Réjane Streiff

CBGP, INRAE, CIRAD, IRD, Montpellier SupAgro, University of Montpellier, Montpellier, France

ARTICLE INFO

Article history:

Received 19 August 2020

Accepted 14 October 2020

Available online 22 October 2020

Keywords:

Annotation

Assembly

Crambidae

Lepidoptera

Mitogenome

Phylogeny

Pyraloidea

ABSTRACT

We present an assembly and annotation of the mitogenome of a European specimen of the Adzuki bean borer, *Ostrinia scapularis* (Walker, 1859). The data were obtained by combining WGS data issue of a *de novo* and a previously published sequence library (Gschloessl et al., 2018). We also provide the phylogenetic positioning of the mitogenome within the *Ostrinia* genus, the *Crambidae* family and with more distant Lepidoptera species.

© 2020 The Authors. Published by Elsevier Inc.

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

\* Corresponding author.

E-mail address: [bernhard.gschloessl@inrae.fr](mailto:bernhard.gschloessl@inrae.fr) (B. Gschloessl).

## Specifications Table

|                                |   |
|--------------------------------|---|
| Subject                        | Insect Science  |
| Specific subject area          | Lepidoptera, Crambidae, Mitogenomics  |
| Type of data                   | <ul style="list-style-type: none"> <li>• FASTQ: DNA sequence reads</li> <li>• FASTA: mitochondrial genome assembly</li> <li>• GFF: mitogenome gene annotations</li> <li>• TABLE: gene annotations, base composition, Lepidoptera species used for phylogenetic analysis</li> <li>• FIGURE: mitogenomic circular map, AT-rich region sequence, phylogenetic tree</li> </ul>  |
| How data were acquired         | Whole-genome shotgun sequencing on Illumina NextSeq 500 v2 platform using the paired-end protocol (2 × 150 base pairs (bp)).  |
| Data format                    | Raw sequence reads and analysed (assembled and annotated) mitogenome.   |
| Parameters for data collection | Genomic DNA was extracted from one <i>Ostrinia scapularis</i> larva using the DNeasy Tissue Kit from Qiagen (Hilden, Germany). A 2 × 150 bp shot-gun paired-end library with an insert size of 300 bp was generated and sequenced by LGC Genomics GmbH (Berlin, Germany) on an Illumina NextSeq 500 v2 platform.  |
| Description of data collection | The <i>Ostrinia scapularis</i> mitogenome was assembled with MITObim v1.9. Genomic regions were annotated with MITOS v2. The circular mitochondrial genome map was generated with OGDRAW. Multiple sequence alignments were generated separately for each protein-coding and rRNA gene with MAFFT v7.471. Phylogenetic trees were built on the concatenated gene alignments with PHYML v3.0 and MrBayes v3.2.7a and visualized with FigTree v1.4.4. |
| Data source location           | Samples were collected in 2015 near Abbeville/France (50° 8' 13.57"N; 1° 50' 23.62"E).  |
| Data accessibility             | Repository name: NCBI BioProject<br>Data identification number: PRJNA637835<br>Direct URL to data:<br><a href="http://www.ncbi.nlm.nih.gov/bioproject/637835">http://www.ncbi.nlm.nih.gov/bioproject/637835</a><br>Repository name: NCBI Genbank<br>Data identification number: MT801073<br>Direct URL to data:<br><a href="http://www.ncbi.nlm.nih.gov/nuccore/MT801073">http://www.ncbi.nlm.nih.gov/nuccore/MT801073</a>                          |

## Value of the Data

- This data represents the mitogenome of a European specimen of *Ostrinia scapularis*.
- The data can benefit to researchers working on Lepidoptera evolution and higher systematics.
- The present data can be used for phylogenetic studies of Lepidoptera mitogenomes especially for getting deeper insights in diversification dynamics of *Ostrinia* species.

## 1. Data Description

The *O. scapularis* (hereafter named OSCA) mitogenome had a size of 15,305 bp and an A + T content of 80.9%. All 38 genetic regions commonly known for arthropod mitogenomes [2] were identified (Supplementary GFF annotation file): 13 protein-coding genes, 22 tRNA and 2 rRNA (12S rRNA and 16S rRNA) genes, and one A + T-rich control region. The gene order and the composition of start on stop codons of the protein-coding genes corresponded to those observed for other *Ostrinia* spp. [3,4] and Crambidae [5] mitogenomes. Fifteen intergenic spacer regions (see Table 1: positive intergenic nucleotide values) of a total length of 341 bp were identified

**Table 1**Gene features of the *O. scapularis* mitogenome.

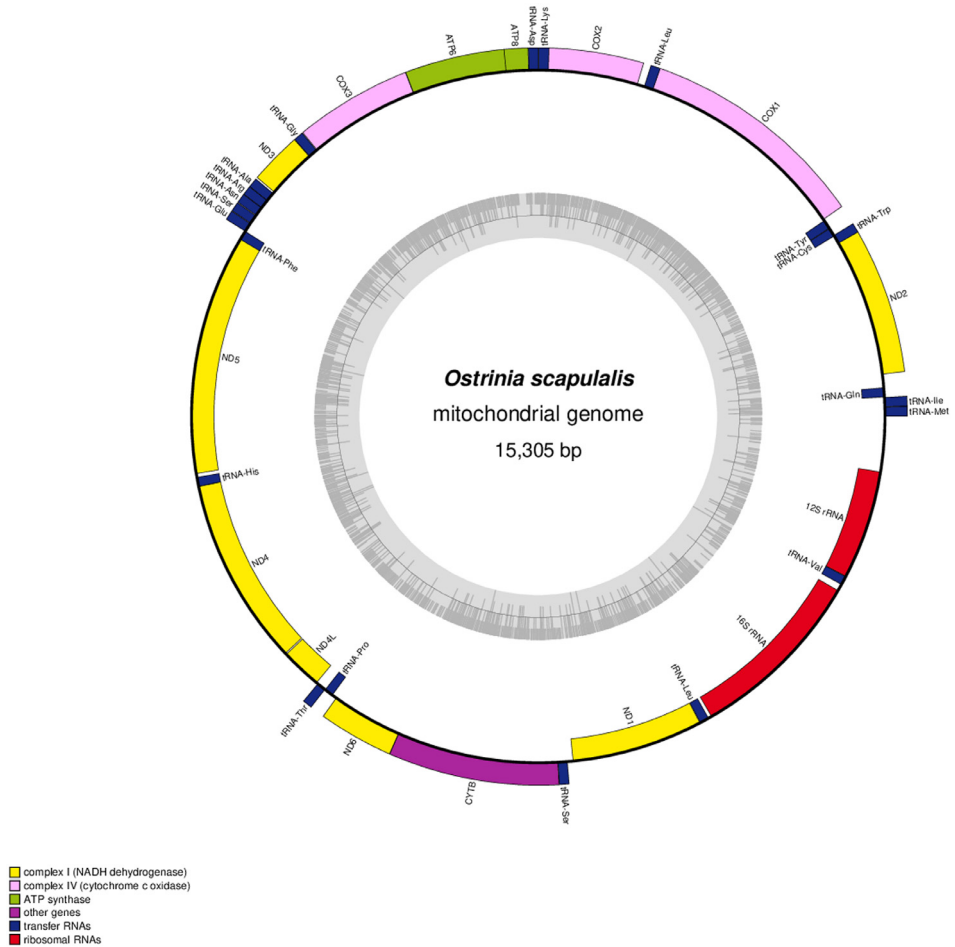
| Gene (anticodon)  | Start  | Stop   | Strand | Size | Intergenic nucleotides | Start/stop codon |
|-------------------|--------|--------|--------|------|------------------------|------------------|
| trnM(cat)         | 1      | 68     | +      | 68   | -1                     |                  |
| trnI(gat)         | 68     | 135    | +      | 68   | 0                      |                  |
| trnQ(ttg)         | 136    | 204    | -      | 69   | 92                     |                  |
| ND2               | 297    | 1289   | +      | 993  | -15                    | ATT/TAA          |
| trnW(tca)         | 1275   | 1341   | +      | 67   | -8                     |                  |
| trnC(gca)         | 1334   | 1398   | -      | 65   | 0                      |                  |
| trnY(gta)         | 1399   | 1465   | -      | 67   | 8                      |                  |
| COX1              | 1474   | 3009   | +      | 1536 | -5                     | CGA/TAA          |
| trnL2(taa)        | 3005   | 3071   | +      | 67   | 48                     |                  |
| COX2              | 3120   | 3788   | +      | 669  | -35                    | ATA/TAA          |
| trnK(ctt)         | 3754   | 3824   | +      | 71   | -1                     |                  |
| trnD(gtc)         | 3824   | 3892   | +      | 69   | 0                      |                  |
| ATP8              | 3893   | 4054   | +      | 162  | -7                     | ATA/TAA          |
| ATP6              | 4048   | 4722   | +      | 675  | -1                     | ATG/TAA          |
| COX3              | 4722   | 5513   | +      | 792  | 2                      | ATG/TAA          |
| trnG(tcc)         | 5516   | 5582   | +      | 67   | 0                      |                  |
| ND3               | 5583   | 5936   | +      | 354  | 10                     | ATT/TAA          |
| trnA(tgc)         | 5947   | 6013   | +      | 67   | -1                     |                  |
| trnR(tcg)         | 6013   | 6077   | +      | 65   | -1                     |                  |
| trnN(gtt)         | 6077   | 6143   | +      | 67   | 2                      |                  |
| trnS1(gct)        | 6146   | 6211   | +      | 66   | 1                      |                  |
| trnE(ttc)         | 6213   | 6279   | +      | 67   | -2                     |                  |
| trnF(gaa)         | 6278   | 6346   | -      | 69   | -17                    |                  |
| ND5               | 6330   | 8057   | -      | 1728 | 24                     | ATT/TAA          |
| trnH(gtg)         | 8082   | 8148   | -      | 67   | -1                     |                  |
| ND4               | 8148   | 9488   | -      | 1341 | 7                      | ATG/TAA          |
| ND4L              | 9496   | 9789   | -      | 294  | 9                      | ATG/TAA          |
| trnI(tgt)         | 9799   | 9865   | +      | 67   | 0                      |                  |
| trnP(tgg)         | 9866   | 9930   | -      | 65   | 32                     |                  |
| ND6               | 9963   | 10,469 | +      | 507  | -1                     | ATT/TAA          |
| CYTB              | 10,469 | 11,617 | +      | 1149 | -1                     | ATG/TAA          |
| trnS2(tga)        | 11,617 | 11,684 | +      | 68   | 38                     |                  |
| ND1               | 11,723 | 12,661 | -      | 939  | 1                      | ATG/TAG          |
| trnL1(tag)        | 12,663 | 12,730 | -      | 68   | 21                     |                  |
| 16S rRNA          | 12,752 | 14,023 | -      | 1272 | 46                     |                  |
| trnV(tac)         | 14,070 | 14,135 | -      | 66   | 0                      |                  |
| 12S rRNA          | 14,136 | 14,914 | -      | 779  | 0                      |                  |
| A + T-rich region | 14,915 | 15,305 | +      | 391  | 0                      |                  |

The gene names, the coordinates, the orientation, the size, the gene overlaps/intergenic spacers (column 6) and the start/stop codons of the *O. scapularis* mitogenome are listed.

**Table 2**Base composition and skewness indices for each specific gene region of the *O. scapularis* mitogenome.

| Type                 | Size [bp] | A%   | G%   | T%   | C%   | A + T% | AT skew | GC skew |
|----------------------|-----------|------|------|------|------|--------|---------|---------|
| Entire mitogenome    | 15,305    | 41.7 | 7.7  | 39.2 | 11.4 | 80.9   | 0.031   | -0.197  |
| Protein-coding genes | 11,139    | 34.5 | 10.8 | 44.9 | 9.9  | 79.4   | -0.131  | 0.045   |
| tRNA                 | 1480      | 41.7 | 10.8 | 39.9 | 7.6  | 81.6   | 0.022   | 0.172   |
| rRNA                 | 2051      | 41.7 | 9.9  | 43.4 | 4.9  | 85.1   | -0.021  | 0.338   |
| A + T-rich region    | 391       | 42.7 | 2.0  | 49.6 | 5.6  | 92.3   | -0.075  | -0.467  |

which ranged from 1 to 92 bp with the longest being located between the trnQ and ND2 genes. Furthermore, region overlaps (see Table 1: negative intergenic nucleotide values) of 1 to 35 bp could be observed for fifteen gene pairs. More annotation details can be found in Table 1 and are also represented as mitogenome map in Fig. 1. Base compositions, AT and GC skews to each genomic region type are shown in Table 2. The tRNA secondary structures showed the typical



**Fig. 1.** Map of the *O. scapularis* mitogenome generated with OGDRAW. Annotated genes are highlighted.

clover-leaf except for trnS1 in which the dihydrouridine (DHU) arm was replaced by an unstable loop (Supplementary Fig. 1). The characteristics of the A+T-rich region are shown in Fig. 2. Finally, we built a phylogeny based on 63 lepidopteran mitogenomes (including the generated OSCA mitogenome, Table 3), with a focus on Pyraloidea (Fig. 3 and Supplementary Fig. 2).

## 2. Experimental Design, Materials and Methods

### 2.1. DNA sampling and sequencing

*Ostrinia scapularis* larval samples were collected in 2015 from a stem of mugwort in North of France (50° 8'13.57"N; 1°50'23.62"E). Genomic DNA was extracted from one larva using the DNeasy Tissue Kit from Qiagen (Hilden, Germany) according to the manufacturer's protocol. Subsequently, a 2 × 150 bp shot-gun paired-end library with an insert size of 300 bp, was generated using the DNA extract and sequenced by LGC Genomics GmbH (Berlin, Germany) on

**Table 3**

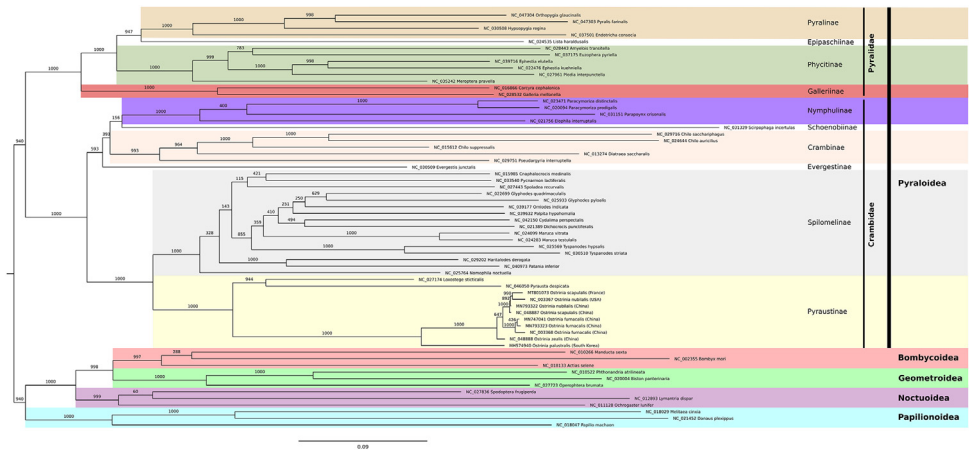
Lepidoptera mitogenomes which were used to build the phylogenetic trees. For *Ostrinia* spp. the countries are also given in which the samples were taken.

| Super-family  | Family       | Subfamily       | Species                                   | Genbank no. |
|---------------|--------------|-----------------|---|-------------|
| Bombycoidea   | Bombycidae   | Bombycinae      | <i>Bombyx mori</i>                        | NC_002355   |
| Bombycoidea   | Saturniidae  | Saturniinae     | <i>Actias selene</i>                      | NC_018133   |
| Bombycoidea   | Sphingidae   | Sphinginae      | <i>Manduca sexta</i>                      | NC_010266   |
| Geometroidea  | Geometridae  | Ennominae       | <i>Biston panterinaria</i>                | NC_020004   |
| Geometroidea  | Geometridae  | Ennominae       | <i>Phthonandria atrilineata</i>           | NC_010522   |
| Geometroidea  | Geometridae  | Larentiinae     | <i>Operophtera brumata</i>                | NC_027723   |
| Noctuoidea    | Erebidae     | Lymantriinae    | <i>Lymantria dispar</i>                   | NC_012893   |
| Noctuoidea    | Noctuidae    | Noctuidae       | <i>Spodoptera frugiperda</i>              | NC_027836   |
| Noctuoidea    | Notodontidae | Thaumetopoeinae | <i>Ochrogaster lunifer</i>                | NC_011128   |
| Papilionoidea | Nymphalidae  | Danainae        | <i>Danaus plexippus</i>                   | NC_021452   |
| Papilionoidea | Nymphalidae  | Nymphalinae     | <i>Melitaea cinxia</i>                    | NC_018029   |
| Papilionoidea | Papilionidae | Papilioninae    | <i>Papilio machaon</i>                    | NC_018047   |
| Pyraloidea    | Crambidae    | Crambinae       | <i>Chilo auricilius</i>                   | NC_024644   |
| Pyraloidea    | Crambidae    | Crambinae       | <i>Chilo sacchariphagus</i>               | NC_029716   |
| Pyraloidea    | Crambidae    | Crambinae       | <i>Chilo suppressalis</i>                 | NC_015612   |
| Pyraloidea    | Crambidae    | Crambinae       | <i>Diatraea saccharalis</i>               | NC_013274   |
| Pyraloidea    | Crambidae    | Crambinae       | <i>Pseudargyria interruptella</i>         | NC_029751   |
| Pyraloidea    | Crambidae    | Evergestinae    | <i>Evergestis junctalis</i>               | NC_030509   |
| Pyraloidea    | Crambidae    | Nymphulinae     | <i>Elophila interruptalis</i>             | NC_021756   |
| Pyraloidea    | Crambidae    | Nymphulinae     | <i>Paracymoriza distinctalis</i>          | NC_023471   |
| Pyraloidea    | Crambidae    | Nymphulinae     | <i>Paracymoriza prodigalis</i>            | NC_020094   |
| Pyraloidea    | Crambidae    | Nymphulinae     | <i>Parapopynx crisonalis</i>              | NC_031151   |
| Pyraloidea    | Crambidae    | Pyraustinae     | <i>Pyrauste sticticalis</i>               | NC_027174   |
| Pyraloidea    | Crambidae    | Pyraustinae     | <i>Ostrinia furnacalis</i> (China)        | NC_003368   |
| Pyraloidea    | Crambidae    | Pyraustinae     | <i>Ostrinia furnacalis</i> (China)        | MN793323    |
| Pyraloidea    | Crambidae    | Pyraustinae     | <i>Ostrinia furnacalis</i> (China)        | MN747041    |
| Pyraloidea    | Crambidae    | Pyraustinae     | <i>Ostrinia nubilalis</i> (USA)           | NC_003367   |
| Pyraloidea    | Crambidae    | Pyraustinae     | <i>Ostrinia nubilalis</i> (China)         | MN793322    |
| Pyraloidea    | Crambidae    | Pyraustinae     | <i>Ostrinia palustralis</i> (South Korea) | MH574940    |
| Pyraloidea    | Crambidae    | Pyraustinae     | <i>Ostrinia scapularis</i> (France)       | MT801073    |
| Pyraloidea    | Crambidae    | Pyraustinae     | <i>Ostrinia scapularis</i> (China)        | NC_048887   |
| Pyraloidea    | Crambidae    | Pyraustinae     | <i>Ostrinia zealis</i> (China)            | NC_048888   |
| Pyraloidea    | Crambidae    | Pyraustinae     | <i>Pyrausta despicata</i>                 | NC_046050   |
| Pyraloidea    | Crambidae    | Schoenobiinae   | <i>Scirpophaga incertulas</i>             | NC_031329   |
| Pyraloidea    | Crambidae    | Spilomelinae    | <i>Cnaphalocrocis medinalis</i>           | NC_015985   |
| Pyraloidea    | Crambidae    | Spilomelinae    | <i>Cydalima perspectalis</i>              | NC_042150   |
| Pyraloidea    | Crambidae    | Spilomelinae    | <i>Dichocrocis punctiferalis</i>          | NC_021389   |
| Pyraloidea    | Crambidae    | Spilomelinae    | <i>Glyphodes pyloalis</i>                 | NC_025933   |
| Pyraloidea    | Crambidae    | Spilomelinae    | <i>Glyphodes quadrimaculalis</i>          | NC_022699   |
| Pyraloidea    | Crambidae    | Spilomelinae    | <i>Haritalodes derogata</i>               | NC_029202   |
| Pyraloidea    | Crambidae    | Spilomelinae    | <i>Maruca testulalis</i>                  | NC_024283   |
| Pyraloidea    | Crambidae    | Spilomelinae    | <i>Maruca vitrata</i>                     | NC_024099   |
| Pyraloidea    | Crambidae    | Spilomelinae    | <i>Nomophila noctuella</i>                | NC_025764   |
| Pyraloidea    | Crambidae    | Spilomelinae    | <i>Omiodes indicata</i>                   | NC_039177   |
| Pyraloidea    | Crambidae    | Spilomelinae    | <i>Palpita hypohomalia</i>                | NC_039632   |
| Pyraloidea    | Crambidae    | Spilomelinae    | <i>Patania inferior</i>                   | NC_040973   |
| Pyraloidea    | Crambidae    | Spilomelinae    | <i>Pycnarmon lactiferalis</i>             | NC_033540   |
| Pyraloidea    | Crambidae    | Spilomelinae    | <i>Spoladea recurvalis</i>                | NC_027443   |
| Pyraloidea    | Crambidae    | Spilomelinae    | <i>Tyspanodes hypsalis</i>                | NC_025569   |
| Pyraloidea    | Crambidae    | Spilomelinae    | <i>Tyspanodes striata</i>                 | NC_030510   |
| Pyraloidea    | Pyrallidae   | Epipaschiinae   | <i>Lista haraldusalis</i>                 | NC_024535   |
| Pyraloidea    | Pyrallidae   | Galleriinae     | <i>Corcyra cephalonica</i>                | NC_016866   |
| Pyraloidea    | Pyrallidae   | Galleriinae     | <i>Galleria mellonella</i>                | NC_028532   |
| Pyraloidea    | Pyrallidae   | Phycitinae      | <i>Amyelois transitella</i>               | NC_028443   |
| Pyraloidea    | Pyrallidae   | Phycitinae      | <i>Ephestia elutella</i>                  | NC_039716   |
| Pyraloidea    | Pyrallidae   | Phycitinae      | <i>Ephestia kuehniella</i>                | NC_022476   |
| Pyraloidea    | Pyrallidae   | Phycitinae      | <i>Euzophera pyriella</i>                 | NC_037175   |
| Pyraloidea    | Pyrallidae   | Phycitinae      | <i>Meroptera pravella</i>                 | NC_035242   |
| Pyraloidea    | Pyrallidae   | Phycitinae      | <i>Plodia interpunctella</i>              | NC_027961   |
| Pyraloidea    | Pyrallidae   | Pyralinae       | <i>Endotricha consocia</i>                | NC_037501   |
| Pyraloidea    | Pyrallidae   | Pyralinae       | <i>Hyposopygia regina</i>                 | NC_030508   |
| Pyraloidea    | Pyrallidae   | Pyralinae       | <i>Orthopygia glaucinalis</i>             | NC_047304   |
| Pyraloidea    | Pyrallidae   | Pyralinae       | <i>Pyralis farinalis</i>                  | NC_047303   |

12S rRNA-14,915-

TAATACTTAAAATTTTAACTTAGATTTTTTTTTTTTTTTTTTATATTA  
 AATATTTAATATAAATTATTAATTTTAAATATTTAATTTCTTTTCTTT  
 TTCATAATATTAATATTAATAATTAATATCATTTATACAGCGATTTATAA  
 TCATTGAAATAAATAATTAATTATTAAGTTTAATAATTAATTAATTAATGA  
 TTTATTAATATATTTATAAATAAATTTCTATTATATATATATATATA  
 TTAAATATAATAATTAATTAATATTTAATATATATATATATATATA  
 AACCATTCTAATTTTTTTTCTTTTAATAATAAATTAATAAATAAGCTAA  
 ATAAAGCTTATATAAACCATTCCTAATTTTTTTTCTTTTAATAATAAA-  
 15,305-trnM

**Fig. 2.** Characteristics of the A+T-rich region of the *O. scapularis* mitogenome. The AT-rich region consists of a TTAGA (red) motif which is followed by a 18 bp poly-T stretch (brown). Subsequently, it contains five ATTTA motifs (blue) of which the last one is followed by a (AT)<sub>10</sub> microsatellite stretch (purple) and terminates with a polyT/polyA region (green).



**Fig. 3.** Phylogenetic tree representing 63 Lepidoptera mitogenomes which was built on concatenated alignments of the nuclear sequences of the 13 protein-coding and both rRNA genes applying a Maximum Likelihood approach with PHYML. For *Ostrinia* spp. the countries are also given in which the samples were taken.

an Illumina NextSeq 500 v2 platform. In total, 372,260,856 genomic reads were generated of which 306,203,188 remained after quality trimming with Trimmomatic v0.36 [6] (specific parameters were: PE -pred33, ILLUMINACLIP:\$ADAPTERF:2:26:10 HEADCROP:5 LEADING:26 TRAILING:26 SLIDINGWINDOW:5:28 MINLEN:30, with \$ADAPTERF containing all Illumina adapters and their reverse complement, respectively).

2.2. Mitogenome assembly and feature analyses

The OSCA mitogenome was assembled in two steps as follows. First, the scaffolds of a previously published *O. scapularis* genome draft [1], OSCA v1.2 (BioProject PRJNA390510), were searched for putative mitochondrion sequences by using *blastn* v2.2.28 [7] against the mitochondrial genome sequence of the closely related Asian corn borer *Ostrinia furnacalis* (Guenée) (NCBI

accession number NC\_003368). Genome scaffolds with high quality mitochondrion hits were selected and assembled by LGC with CAP3 v2013 [8] into a mitogenome sequence of 14,864 bp. In a second step, this draft mitogenome was refined using MITObim v1.9 [9] (using the default k-mer size of 31 bp and the `-pair` parameter) and the Trimmomatic-cleaned paired-end Illumina reads. Gene identification as well as tRNA structure prediction were carried out on the MITOS web server v2 [10], using as reference set 'RefSeq 81 Metazoa' and the genetic code '5' for invertebrates. Nucleic base contents were determined with the `wordcount` program of the EMBOSS toolkit v6.6.0.0 [11]. AT/GC skewness indices describing the genomic compositional asymmetry of a mitogenome or specific sequence region were calculated as:  $AT\ skew = [A - T] / [A + T]$  and  $GC\ skew = [G - C] / [G + C]$ , where each base letter represents the respective base count. The circular map of the OSCA mitogenome was generated with the web tool OGDRAW v1.3.1 [12].

### 2.3. Phylogenetic analyses

In order to perform comparative analyses of *O. scapularis* with other Lepidoptera the nucleotide sequences of the 13 PCGs and 2 rRNA genes of 63 mitogenomes - comprising 59 Lepidoptera species - (Table 3) were downloaded from the NCBI 'nucleotide' database (<https://www.ncbi.nlm.nih.gov/nucleotide>, May 2020). Our taxon sampling was focused on the superfamily Pyraloidea, encompassing 34 species belonging to the Crambidae (including *O. scapularis*) and 13 species belonging to the Pyralidae. Regarding *Ostrinia* spp., we included in total nine mitogenomes belonging to five *Ostrinia* species. As outgroups we used 12 representative species of the lepidopteran superfamilies Bombycoidea, Geometroidea, Noctuoidea and Papilionoidea; based on the comprehensive study of Wahlberg et al. [13], we used the three Papilionoidea representatives to root our trees. The PCG and rRNA nucleotide sequences of all species were grouped into 15 gene-specific sequence sets. Each gene set was in separate globally aligned with the program MAFFT v7.471 [14] by applying the Needleman-Wunsch algorithm (`mafft-ginsi` for PCGs and `mafft-qinsi` for rRNA sequences) with a maximum of 1000 refinement iterations (`-maxiterate 1000`), leaving the other parameters set to their default value. Subsequently, for each species all gene-specific alignments were concatenated following the same gene order using in-house Python and Perl scripts.

Phylogenetic trees were generated with the PHYML web server v3.0 [15] and a local installation of MrBayes v3.2.7a [17]. PHYML was run using the Smart Model Selection [16] option with the Akaike Information Criterion and the tree searching method BIONJ being set and by applying 1000 bootstraps. Concerning the MrBayes analysis, a separate partition was set for each gene. For each partition the evolutionary model was set to the GTR substitution model (`nst=6`) with gamma-distributed rate variation (`rates=invgamma`) across sites and a proportion of invariable sites (`GTR +I +Γ`). The standard nucleotide substitution model (`nucmodel=4by4`) was chosen and priors settings were left as by default. Trees were built using 100,000 samples (`ngen=1,000,000`, `samplefreq=10`). Further parameters were: `printfreq=100`, `diagnfreq=1000`, `nchains=4`, `savebrlens=yes`, `starttree=random`, `startparams=reset`, `outgroup pamac` (corresponding to *Papilio machaon*), `sump burnin=25,000`, `sumt burnin=25,000`. The phylogenetic tree was visualized using the program FigTree v1.4.4 [18].

### CRedit Author Statement

BG conceived the study. RS led the ANR research program on European *Ostrinia* genomics. RS, PA, SN collected the samples in the field. SN, PA prepared the samples and extracted the DNA. BG was responsible for the bioinformatics analyses, i.e. improvement of the mitogenome assembly, annotation, characterization, comparative analyses. BG and GJK did the phylogenetic analyses. All authors wrote and approved the manuscript.



## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships which have, or could be perceived to have, influenced the work reported in this article.

## Acknowledgments

This work was supported by a grant from the French National Research Agency (ANR, project Adapt-Ome, ANR-13-BSV7-0012-01). Most of the bioinformatics analyses were performed on the CBGP HPC computational platform (Montferrier-sur-Lez, France, [https://www6.montpellier.inrae.fr/cbpg\\_eng/Platforms/Computing-platform](https://www6.montpellier.inrae.fr/cbpg_eng/Platforms/Computing-platform)). We are grateful to the bioinformatics platform ABiMS (Roscoff, France, <http://abims.sb-roscoff.fr>) for providing computational resources.

## Supplementary Materials

Supplementary material associated with this article can be found, in the online version at doi:[10.1016/j.dib.2020.106427](https://doi.org/10.1016/j.dib.2020.106427).

## References

- [1] B. Gschloessl, F. Dorkeld, P. Audiot, A. Bretaudeau, C. Kerdelhué, R. Streiff, De novo genome and transcriptome resources of the Adzuki bean borer *Ostrinia scapularis* (Lepidoptera: crambidae), *Data Brief* 17 (2018) 781–787, doi:[10.1016/j.dib.2018.01.073](https://doi.org/10.1016/j.dib.2018.01.073).
- [2] J.L. Boore, Animal mitochondrial genomes, *Nucleic Acids Res.* 27 (1999) 1767–1780, doi:[10.1093/nar/27.8.1767](https://doi.org/10.1093/nar/27.8.1767).
- [3] C. Li, L. Li, Y. Ren, Z. Lu, Y. Song, L. Liu, S. Lv, Y. Yu, X. Men, Characterization of the complete mitochondrial genome of Asia Corn Borer, *Ostrinia furnacalis* (Lepidoptera: Crambidae), *Mitochondrial DNA Part B* 5 (2020) 936–937, doi:[10.1080/23802359.2020.1718025](https://doi.org/10.1080/23802359.2020.1718025).
- [4] N. Zhou, Y. Dong, P. Qiao, Z. Yang, Complete mitogenomic structure and phylogenetic implications of the genus *Ostrinia* (Lepidoptera: crambidae), *Insects* 11 (2020), doi:[10.3390/insects11040232](https://doi.org/10.3390/insects11040232).
- [5] M. Yang, L. Song, J. Mao, Y. Shi, C. Wu, Y. Zhang, L. Huang, W. Peng, X. Liu, Complete mitochondrial genome of the soybean leafroller, *Omiodes indicata* (Lepidoptera: pyraloidea: crambidae), and phylogenetic analysis for pyraloidea, *Int. J. Biol. Macromol.* 115 (2018) 53–60, doi:[10.1016/j.ijbiomac.2018.03.041](https://doi.org/10.1016/j.ijbiomac.2018.03.041).
- [6] A.M. Bolger, M. Lohse, B. Usadel, Trimmomatic: a flexible trimmer for illumina sequence data, *Bioinformatics* 30 (2014) 2114–2120, doi:[10.1093/bioinformatics/btu170](https://doi.org/10.1093/bioinformatics/btu170).
- [7] S.F. Altschul, W. Gish, W. Miller, E.W. Myers, D.J. Lipman, Basic local alignment search tool, *J. Mol. Biol.* 215 (1990) 403–410, doi:[10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2).
- [8] X. Huang, A. Madan, CAP3: a DNA sequence assembly program, *Genome Res.* 9 (1999) 868–877, doi:[10.1101/gr.9.9.868](https://doi.org/10.1101/gr.9.9.868).
- [9] C. Hahn, L. Bachmann, B. Chevreux, Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads—a baiting and iterative mapping approach, *Nucleic Acids Res.* 41 (2013) e129, doi:[10.1093/nar/gkt371](https://doi.org/10.1093/nar/gkt371).
- [10] M. Bernt, A. Donath, F. Jühling, F. Externbrink, C. Florentz, G. Fritzsch, J. Pütz, M. Middendorf, P.F. Stadler, MITOS: improved de novo metazoan mitochondrial genome annotation, *Mol. Phylogenet. Evol.* 69 (2013) 313–319, doi:[10.1016/j.ympev.2012.08.023](https://doi.org/10.1016/j.ympev.2012.08.023).
- [11] P. Rice, I. Longden, A. Bleasby, EMBOS: the European molecular biology open software suite, *Trends Genet.* 16 (2000) 276–277, doi:[10.1016/s0168-9525\(00\)02024-2](https://doi.org/10.1016/s0168-9525(00)02024-2).
- [12] S. Greiner, P. Lehwerk, R. Bock, OrganellarGenomeDRAW (OGDRAW) version 1.3.1: expanded toolkit for the graphical visualization of organellar genomes, *Nucleic Acids Res.* 47 (2019) W59–W64, doi:[10.1093/nar/gkz238](https://doi.org/10.1093/nar/gkz238).
- [13] N. Wahlberg, C.W. Wheat, C. Peña, Timing and patterns in the taxonomic diversification of Lepidoptera (butterflies and moths), *PLoS One* 8 (2013) e80875, doi:[10.1371/journal.pone.0080875](https://doi.org/10.1371/journal.pone.0080875).
- [14] K. Katoh, D.M. Standley, MAFFT multiple sequence alignment software version 7: improvements in performance and usability, *Mol. Biol. Evol.* 30 (2013) 772–780, doi:[10.1093/molbev/mst010](https://doi.org/10.1093/molbev/mst010).
- [15] S. Guindon, F. Lethiec, P. Duroux, O. Gascuel, PHYML Online—a web server for fast maximum likelihood-based phylogenetic inference, *Nucleic Acids Res.* 33 (2005) W557–W559, doi:[10.1093/nar/gki352](https://doi.org/10.1093/nar/gki352).
- [16] V. Lefort, J.-E. Longueville, O. Gascuel, SMS: smart model selection in PhyML, *Mol. Biol. Evol.* 34 (2017) 2422–2424, doi:[10.1093/molbev/msx149](https://doi.org/10.1093/molbev/msx149).
- [17] F. Ronquist, M. Teslenko, P. van der Mark, D.L. Ayres, A. Darling, S. Höhna, B. Larget, L. Liu, M.A. Suchard, J.P. Huelsenbeck, MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space, *Syst. Biol.* 61 (2012) 539–542, doi:[10.1093/sysbio/sys029](https://doi.org/10.1093/sysbio/sys029).
- [18] A. Rambaut, A.J. Drummond, FigTree Version 1.4. 0, 2012.