The mitochondrial genome of Iberobaenia (Coleoptera: Iberobaeniidae): first rearrangement of protein-coding genes in the beetles
Carmelo Andujar, Paula Arribas, Benjamin Linard, Robin Kundra, Ladislav Bocak, Alfried P Vogler

To cite this version:

HAL Id: hal-01636908
https://hal.archives-ouvertes.fr/hal-01636908
Submitted on 17 Jan 2019

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.
The mitochondrial genome of *Iberobaenia* (Coleoptera: Iberobaeniidae): First rearrangement of protein coding genes in the beetles

Carmelo Andújar,1,2* Paula Arribas,1,2 Benjamin Linard,1 Robin Kundrata,3 Ladislav Bocak,3 Alfried P. Vogler1,2

1 Department of Life Sciences, Natural History Museum, Cromwell Road, London, SW7 5BD, UK
2 Department of Life Sciences, Imperial College London, Silwood Park Campus, Ascot, SL5 7PY, UK
3 Department of Zoology, Faculty of Science, Palacky University, 17. listopadu 50, 771 46 Olomouc, Czech Republic

**Key words:** Coleoptera, gene rearrangement, mitochondrial metagenomics, mitochondrion, Elateroidea.

**Abstract:**

The complete mitochondrial genome of the recently discovered beetle family Iberobaeniidae is described and compared with known coleopteran mitogenomes. The mitochondrial sequence was obtained by shotgun metagenomic sequencing using the Illumina Miseq technology and resulted in an average coverage of 130x and a minimum coverage of 35x. The mitochondrial genome of Iberobaeniidae includes 13 protein-coding genes, 2 rRNAs, 22 tRNAs genes and one putative control region, and showed a unique rearrangement of protein coding genes. This is the first rearrangement affecting the relative position of protein-coding and ribosomal genes reported for the order Coleoptera.

**Running Title:** The mitochondrial genome of *Iberobaenia*
Coleoptera are the most diverse order of insects, with about 179 extant families (Bouchard et al. 2011) and more than 350000 described species. More than 250 complete or almost complete mitochondrial genomes have been sequenced to date, including representatives from ≈100 families. Despite the high diversity and ancient origin of Coleoptera, estimated at around 300 MYA (Hunt et al., 2007; Misof et al., 2014), all known coleopteran mitogenomes share the same gene order and orientation for the protein-coding and rRNA genes, while rearrangements found to date affect only the position of the tRNA genes (Timmermans & Vogler, 2012). The genus *Iberobaenia* Bocak, Kundrata, Andújar & Vogler 2015 has been recently discovered and phylogenetic analyses conducted on mitochondrial and nuclear genes revealed the distinctiveness of the new lineage within the superfamily Elateroidea, resulting in the description of a new beetle family, Iberobaeniidae (Bocak et al. in press).

For sequencing the complete mitochondrial genome of *Iberobaenia* we followed the mitochondrial metagenomic approach (MMG) (Dettai et al., 2012; Zhou et al., 2013; Andújar et al., 2015; Crampton-Platt et al., 2015), which allows to assemble complete mitochondrial genomes from shotgun sequencing of specimen mixtures. The raw DNA extracted from a larval specimen of *Iberobaenia* sp. (Sierra de Cabra, Córdoba, Spain; Voucher specimen deposited in the Natural History Museum of London–BMNH 1042541) was pooled in equimolar concentration with DNA from 20 other Coleoptera from distant lineages mostly in the suborder Adephaga not connected to the current study. The dsDNA concentration was measured using a Qubit 2.0 Fluorometer (Life Technologies Corp., Carlsbad, CA, USA). A TruSeq DNA library was constructed with the pooled DNA and sequenced using the Illumina MiSeq platform (2x300 bp; ~800–950 bp insert size) in 45% of an Illumina flow cell.

The Illumina output was quality assessed with FastQC 0.10.1 (Babraham Institute 2013) and processed with Trimmomatic 0.30 (Lohse et al. 2012) for Illumina adapter removal. Reads we subsequently filtered using Blast 2.2.27 (Altschul et al. 1990) against a reference database including 245 nearly complete coleopteran mitochondrial genomes retrieved from GenBank. Surviving mitochondrial reads were subsequently assembled using Celera Assembler 7.0 (Myers 2000), Newbler 2.7 (Margulies et al. 2006) and IDBA-UD (Peng et al. 2012). Celera analyses were done using the RunCa command with the bogart method for unitigger computations, and Unitig Repeat/Unique Toggling was conducted from unitigs longer than 500 bp; other parameters were set to default options. Newbler analyses were run using -mi 95 -ml 150 and IDBA using --maxk 300 --mink 50.
Mitochondrial contigs longer than 500 bp obtained with the three assemblers were pooled and super-contiged using Geneious 7.1.9 (http://www.geneious.com), saving the 50% majority rule consensus from the resulting contigs. The mitogenome of *Iberobaenia* was identified using the *cox1* barcode sequence obtained from the same specimen with PCR-Sanger sequencing (Genbank accession no. XXX), and was annotated using gene predictions with MITOS (Bernt et al., 2013) and manually refined in Geneious. Additionally, mitochondrial Illumina reads were mapped against the mitogenome of *Iberobaenia* in Geneious (using a 99% similarity threshold).

The widely overlapping contigs obtained from the three independent assemblers, after super-contiging and circularization, resulted in a complete mitochondrial sequence of 16931bp (Accession no. KT825140) including 13 protein-coding, 2 rRNAs and 22tRNAs genes. The overall base composition was 46.9% for A, 14.1% for C, 7.5% for G, 31.5% for T. All protein-coding genes were encoded on the H-strand with exception of *nad1*, *nad4*, *nad4l* and *nad5*. All tRNA genes were encoded on the H-strand with the exception of tRNA-Cys, tRNA-Phe, tRNA-His, tRNA-Val, tRNA-Gln, tRNA-Pro and tRNA-Leu2. Our predictive annotation resulted in the canonical ATN start codon (isoleucine) for all genes (ATA: *nad2*, *cox1*, *cob*, *atp6*, *atp8* and *nad3*; ATG: *cox2* and *cox3*; ATT: *nad5*, *nad4* and *nad4l*; ATC: *nad6*) but *nad1*, where a TTG start codon was identified as previously reported in other Coleoptera (Sheffield et al., 2008). With regard to the 3’ end, *nad2* was predicted with an incomplete stop codon "T", three genes showed the stop codon TAG (*nad5*, *nad1* and *cob*) and all other genes were inferred with the stop codon TAA. The longest gene was *nad5* (1647bp) and the shortest *atp8* (150bp). A maximum-likelihood phylogenetic tree was obtained in RaxML (Stamatakis, 2006) including some of the closest relatives to Iberobaniidae according to Bocak et al. (in review) (Fig. 1). The best tree was selected over 100 alternative runs and node support was calculated with 10000 bootstrap replicates.

The structure of the mitogenome of *Iberobaenia* showed a unique gene rearrangement with regard to the constant pattern found in all other mitochondrial genomes for the order Coleoptera (Timmermans & Vogler, 2012). The rearrangement affected the protein-coding, ribosomal and tRNAs genes. The gene order for the mitogenome of *Iberobaenia* sp., including tRNAs, is: I, M, *nad2*, W, C, Y, *cox1*, L1, *cox2*, K, D, *atp8*, *atp6*, *cox3*, *nad3*, N, R, A, E, F, G, S, *nad5*, H, *nad4l*, *nad1*, control region, Q, *nad4*, T, P, *nad6*, *cob*, S2, L2, *rrnL*, V, *rrnS*. The high sequencing read coverage (average:130x; max: 213x; min: 35x) provides high reliability to the obtained sequence. This is the first mitochondrial gene rearrangement reported in the order Coleoptera excluding those affecting only to tRNAs (Timmermans &
Vogler, 2012). This finding adds to the unique features of the enigmatic beetle family Iberobaeniidae recently described from the Iberian Peninsula.

**Acknowledgements**

We thank Alex Aitken, Kevin Hopkins and Peter Foster (Natural History Museum) for their technical assistance. This work was funded by the Leverhulme Trust (grant F/00696/P), the NHM Biodiversity Initiative and the Natural Environment Research Council (grant NE/M021955). PA was supported by a postdoctoral grant from The Royal Society (Newton International Program, UK).

**Declaration of interest**

The authors report no conflict of interest. The authors alone are responsible for the content and writing of the paper.

**References:**


Figure captions:

Figure 1. Maximum-likelihood phylogenetic tree including some of the closest relatives to Iberobaniidae according to Bocak et al. (in review). Numbers on nodes represent bootstrap support. Right: Schematic representation of *Iberobaenia* sp. male specimen.