DOI 10.18699/BGRS/SB-2020-148

Mitogenomics and phylogenetics of vulnerable and endangered birds of genera *Aratinga* and *Psittacus*

Sofia Ochkalova
Applied Genomics Lab, SCAMT,
ITMO University,
Saint-Petersburg, Russia
ochkalova@scamt-itmo.ru

Aleksey Komissarov
Applied Genomics Lab, SCAMT,
ITMO University,
Saint-Petersburg, Russia
komissarov@scamt-itmo.ru

Taras Oleksyk
Biology Department, University of
Puerto Rico, Mayaguez, Puerto Rico
Department of Biological Sciences,
Oakland University, Rochester, USA
oleksyk@oakland.edu

***Abstract* — Among parrots, there are more threatened and endangered species than in any other avian order. Exploration of these outstanding birds exhibiting advanced cognitive abilities is a relevant scientific problem. In this study, we de novo assembled 19 mitochondrial genomes from Illumina raw reads, integrated them with information obtained from related genomes, and did a phylogenetic reconstruction of species from *Aratinga* and *Psittacus* genera. Additionally, we created a Python pipeline for assembly mtDNA bearing tandemly repeated regions from whole genome sequencing data. Then we annotated genes and finally performed phylogenetic analysis to provide insights into the evolution and demographic history of the species.**

Keywords* — *Aratinga, Psittacus, mitochondrial DNA, phylogenics, parrots

Motivation and aim

Motivation

Since long ago, parrots (*Psittaciformes*) have aroused human attention. Attractiveness of plumages, ability to mimic speech and intelligence make them popular pets. Unfortunately, the consequence is a radical drop in their populations due to the extensive illegal harvest of wild birds in addition to habitat loss. Our work is supposed to contribute to a major project on genome database of endangered parrots, which will help to develop a method to recognize species prohibited for trade using PCR tests.

Aim

We aimed to generate good-quality mitogenomes of the vulnerable tropical parrots from *Aratinga* and *Psittacus* genera fort phylogenetic analysis.

Materials and methods

Samples for DNA sequencing were obtained from ten and nine different individuals of unidentified species from *Aratinga* and *Psittacus* genera respectively. For all samples sequencing was performed using the Illumina HiSeq2000 platform in paired-end mode.

The initial QC of NGS data was performed using FastQC [1]. Reads were pre-processed for downstream analysis using in-house script tool for trimming Illumina adapters.

Mitogenome assembly utilized low-coverage whole genome data. We attempted to reconstruct mitochondrial genome using de novo assembler NOVOPlasty 3.8.3 for organelle genomes [2], but did not succeed. The output contig was only 15450 bp length and did not include 2 protein-coding genes. Assuming that the breakdown of the assembly was due to the inability of NOVOPlasty to master repeats in control region, we developed a pipeline to assemble mitogenomes considering this feature. Annotation was conducted with MITOS Annotator [3].

Phylogenetic tree was reconstructed using different combinations of mtDNA regions of our samples and other species from *Psittaciformes* (*Amazona*, *Ara*, *Psittacara*, *Eupsittula*, *Arapornis*). Extracted genes were aligned using MAFFT [4], processed by Gblocks to remove hypervariable regions and concatenated. Obtained alignments were used to reconstruct a maximum likelihood tree with RAxML v8.2 [5].

Results

The circular assembled mtDNA size is 17,011 bp for *Aratinga* and about 16,850 bp for *Psittacus*. Similar to the existing data, we identified 13 predicted protein-coding genes (PCGs), two ribosomal RNA (rRNA) genes, and 22 tRNAs and one long tandem repeat 2 kb long and located before control region.

Phylogenetic reconstruction revealed that the sampled individuals of *Aratinga* genus refer to *Guaruba guarouba* monotypic group. For samples of *Psittacus* genus we discovered two different subspecies of *Psittacus erithacus*.

Acknowledgment

Russian Foundation for Basic Research Grants 17-00-00144 as part of 17-00-00148.

References

[1] S. Andrews, F. Krueger, A. Seconds-Pichon, F. Biggins, and S. Wingett, “FastQC. A quality control tool for high throughput sequence data. Babraham Bioinformatics,” *Babraham Institute*, 2015. .

[2] N. Dierckxsens, P. Mardulyn, and G. Smits, “NOVOPlasty: De novo assembly of organelle genomes from whole genome data,” *Nucleic Acids Res.*, 2017.

[3] M. Bernt *et al.*, “MITOS: Improved de novo metazoan mitochondrial genome annotation,” *Mol. Phylogenet. Evol.*, 2013.

[4] K. Katoh, K. I. Kuma, H. Toh, and T. Miyata, “MAFFT version 5: Improvement in accuracy of multiple sequence alignment,” *Nucleic Acids Res.*, 2005.

[5] A. Stamatakis, “RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies,” *Bioinformatics*, 2014.