TECHNICAL NOTE

The complete mitochondrial genome and phylogeny of *Geospiza magnirostris* (Passeriformes: Thraupidae)

Wu Liang[1] · Zhao Yunlin[1] · Xu Zhenggang[1,2,3] · Huang Tian[2,3] · Zhou Libo[2,3] · Liu Shiquan[2]

Abstract

Geospiza magnirostris, an island species, was listed as of Least Concern (LC) on the IUCN Red List of Threatened Species; human entrapment, habitat destruction, predators, and competitors from the mainland have reduced its population. In this study, we first determined and described the complete mitochondrial genome of *G. magnirostris*; the circular genome of *G. magnirostris* is 16.798 bp in length, and contains 13 protein-coding genes, two rRNA genes, 22 tRNA genes, and one non-coding control region (D-loop). The overall base composition of the mitochondrial DNA is 29.9% for A, 22.8% for T, 33.2% for C, and 14.1% for G. The percentage of G + C content is 47.3%. A phylogenetic tree was conducted to explore the taxonomic status of *G. magnirostris*, and showed that *G. magnirostris* is clustered in a clade with other Thraupidae species. This information will contribute to future phylogenetic studies of Passeriformes and further conservation strategies for this species.

Keywords *Geospiza magnirostris* · Mitochondrial genome · Phylogeny

Geospiza magnirostris, which belongs to the order Passeriformes and family Thraupidae, is a species of Darwin's finch with a body size from 15 to 16 cm standard length (Rands et al. 2013). As an island species, their habitats are very limited, and portions of them are threatened by human activity and natural conditions, such as human entrapment, predators, competitors from the mainland, and climatic change. The species was listed as of Least Concern (LC) on the IUCN Red List of Threatened Species (IUCN), so we presented its complete mitochondrial genome and explored its taxonomic status with the use of a phylogenetic tree for research and conservation purposes.

The specimen was collected from the Galapagos Islands, Republic of Ecuador. Whole genomic DNA was extracted from the muscle tissue using a DNeasy Plant Mini Kit (Qiagen, Valencia, CA), and genomic DNA data were sequenced using the Illumina MiSeq Platform (Illumina, San Diego, CA). Adapters and low-quality reads were removed using the NGS QC Toolkit (Patel and Jain 2012). A circular map of the mitochondria was generated using OGDRAW (Lohse et al. 2007). The genome was annotated using the MITOS Web Server (Bernt et al. 2013). Each one of the annotated protein-coding genes (PCGs) were compared with published sequences of other vertebrate species using DOGMA (Wyman et al. 2004). Then, maximum likelihood (ML) and neighbor-joining (NJ) trees of 14 Passeriformes species were constructed using MEGA7.0 software with 1000 bootstrap replication (Kumar et al. 2016). The complete mitochondrial genome of *G. magnirostris* was deposited into GenBank under the accession number MG682351.

The complete mitochondrial genome of *G. magnirostris* is 16.798 bp in length (Fig. 1). It consists of 37 mitochondrial genes, including 13 protein-coding genes (PCGs), two rRNA genes, 22 tRNA genes, and one non-coding control region (D-loop). Among these genes, *ND6, tRNA-UUC, tRNA-UUG, tRNA-UGC, tRNA-GUU, tRNA-GCA, tRNA-GUA, tRNA-UGA*, and *tRNA-UGG* are located on the light strand (L-strand), while all of the remaining genes are located on the heavy strand (H-strand). The overall base composition of the mitochondrial DNA is 29.9% for
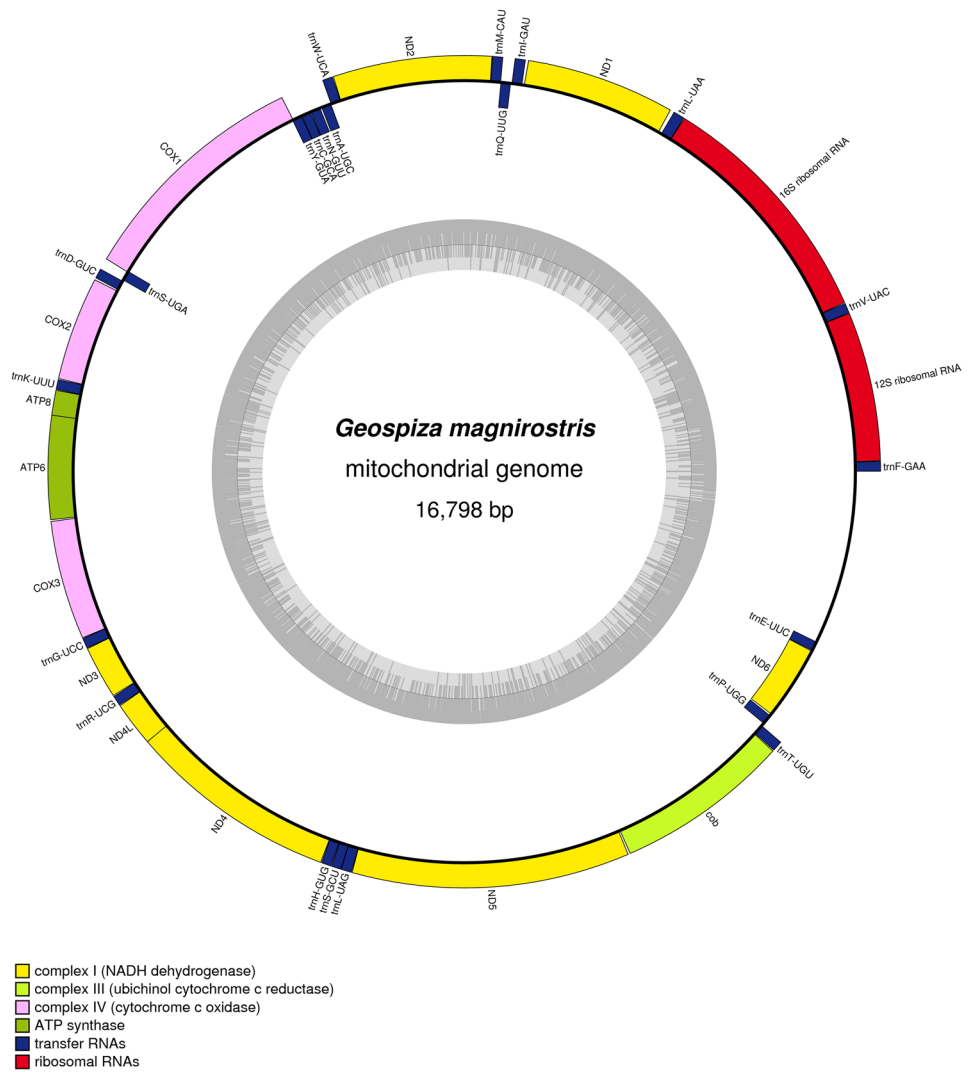
✉ Xu Zhenggang
rssq198677@163.com

1 Key Laboratory of Forestry Remote Sensing Based Big Data & Ecological Security for Hunan Province, Central South University of Forestry and Technology, Changsha, China

2 School of Material and Chemical Engineering, Hunan City University, Yiyang, China

3 Engineering Research Center for Internet of Animals, Changsha, China

Fig. 1 Gene map of the *G. magnirostris* mitochondrial genome

- complex I (NADH dehydrogenase)
- complex III (ubichinol cytochrome c reductase)
- complex IV (cytochrome c oxidase)
- ATP synthase
- transfer RNAs
- ribosomal RNAs

Fig. 2 Phylogenetic tree of 14 Passeriformes species based on the ML and NJ trees of 13PCGs.The values at the nodes correspond to ML(0–1) and NJ(1–100) bootstrap support in percentages
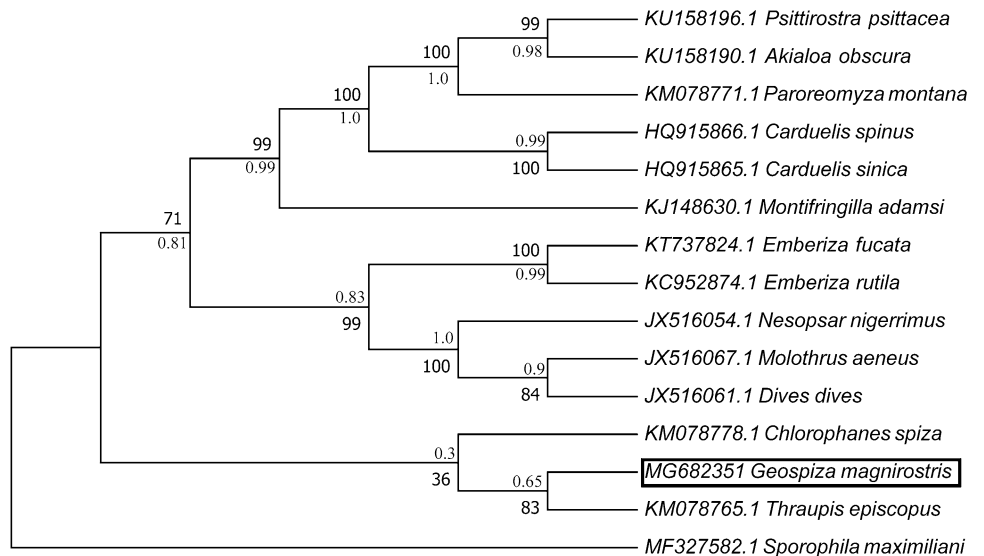
A, 22.8% for T, 33.2% for C, and 14.1% for G. The percentage of G + C content is 47.3%.

To elucidate the phylogenetic position of *G. magnirostris*, the phylogenetic tree was constructed with 14 published completed mitochondrial genomes of Passeriformes (representing five genera) based on the 13 PCGs (Fig. 2). All of the nodes were inferred with strong support by the ML and NJ analysis. As expected, the tree clearly demonstrated that *G. magnirostris*, as well as other Thraupidae species, was clustered in a clade. In all, this genome will provide a reference sequence for the conservation and research of *G. magnirostris*.

Compliance with ethical standards

Conflict of interest The authors declare there is no conflict of interest.

References

Bernt M, Donath A, Juhling F, Externbrink F et al (2013) MITOS: improved de novo metazoan mitochondrial genome annotation. Mol Phylogenet Evol 69(2):313–319

IUCN (2017) The IUCN red list of threatened species. Version 2017-3. http://www.iucnredlist.org. Downloaded on 20 December 2017.

Kumar S, Stecher G, Tamura K (2016) MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol 33(7):1870–1874

Lohse M, Drechsel O, Bock R (2007) OrganellarGenomeDRAW (OGDRAW): a tool for the easy generation of high-quality custom graphical maps of plastid and mitochondrial genomes. Curr Genet 52(5–6):267–274

Patel KR, Jain M (2012) NGS QC toolkit: a toolkit for quality control of next generation sequencing data. PLoS ONE 7(2):e30619

Rands CM, Darling A, Fujita M et al (2013) Insights into the evolution of Darwin's finches from comparative analysis of the *Geospiza magnirostris* genome sequence. BMC Genomics 14(1):95

Wyman SK, Jansen RK, Boore JL (2004) Automatic annotation of organellar genomes with DOGMA. Bioinformatics 20(17):3252