

# Phylogenetic reconstruction of selected species of birds of Colombia with the COI gene

Vanessa Villegas Molina

Genomics and sequence analysis module project

Fundamentals of computational biology

Universidad EAFIT

Medellín

2022

## Introduction

DNA barcode data sets are a tool for taxonomic identification and are especially useful for the conservation of different species [1]. Colombia is the country with the greatest diversity of birds in the world, which represents a tourist attraction and a great interest in knowing each species and its taxonomic identification [2]. Generating in recent years, an increase in the number of identifications for this group, however, it is still believed that there are many gaps in terms of phylogenetic relationships between species, being a challenge for its investigative use by lack of information curation [2]. In this work, we seek to phylogenetically characterize some species of birds from Colombia, in order to better understand their evolutionary relationships. For this, sequences of a specific gene for identification of different species, cytochrome oxidase subunit 1 (COI) [1], were downloaded, multiple alignment was performed and then the reconstruction of the phylogenetic tree was built.

Among the species selected to carry out the phylogenetic reconstruction are: *Automolus ochrolaemus*, *Cantorchilus leucotis*, *Catharus minimus*, *Chaetura pelagica*, *Coereba flaveola*, *Dendrocincla fuliginosa*, *Galbula ruficauda*, *Manacus manacus* (fig. 1) [1]; and as an external group *Pitangus sulphuratus* (fig. 2), a common bird in the Valle de Aburrá [3].

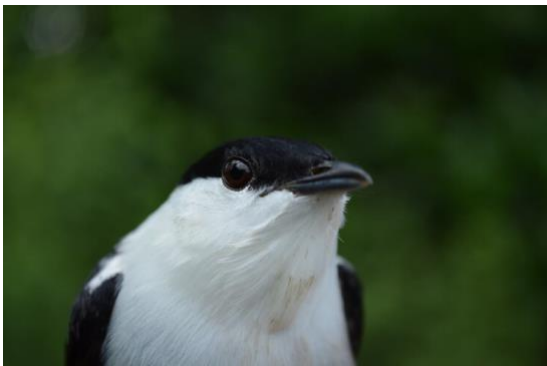


Figure 1. Picture of *Manacus manacus* taken from the paper of Pulgarín-R et al., 2021 [1].



Figure 2. Picture of *Pitangus sulfuratus* taken from the website Birds Colombia [3].

## I Question – Aim

Which is the phylogenetic relationship between these Colombian bird species?

## II Approach

### II. I Programs

NCBI [4], Genious [5]

### II. II Steps

For this project, I chose the article “DNA barcodes of birds from northern Colombia” [1].

[Biodivers Data J.](#) 2021; 9: e64842.

PMCID: PMC8163714

Published online 2021 May 21. doi: [10.3897/BDJ.9.e64842](https://doi.org/10.3897/BDJ.9.e64842)

PMID: [34084068](https://pubmed.ncbi.nlm.nih.gov/34084068/)

## DNA barcodes of birds from northern Colombia

[Paulo Cesar Pulgarín-R.](#)<sup>1</sup> [Martha Olivera-Angel](#),<sup>2</sup> [Luisa Ortiz](#),<sup>2</sup> [Duván Nancíaes](#),<sup>2</sup> [Sara Velásquez-Restrepo](#),<sup>3</sup> and [Juan Fernando Díaz-Nieto](#)<sup>2,3</sup>

And selected 8 species of the 18 presented on the paper. The red asterisks show 7 of the 8 chosen species, for the *Chaetura* sp, another species was selected that is not on the table (*Chaetura pelagica*).

Table 2.

Individuals sampled and barcoded in this study. Individuals with \* represent boreal migrants. Bolded taxa represent inconsistencies between our identification methods (see text).

Code	Field ID	BOLD ID	BOLD hit <sup>1</sup> (%)	NCBI ID	NCBI hit <sup>1</sup> (%)	Consensus sp BOLD ID	Seq length (bp)	BIN
* LCA35	<i>Automolus ochrolaemus</i>	<i>A. ochrolaemus</i>	100	<i>A. ochrolaemus</i>	99.10	<i>A. ochrolaemus</i>	671	<a href="#">BOLD:ADM4531</a>
* LCA9	<i>Cantorchilus leucotis</i>	<i>C. leucotis</i>	100	<i>C. leucotis</i>	95.55	<i>C. leucotis</i>	690	<a href="#">BOLD:ABX4224</a>
* LCA12	<i>Catharus minimus</i>	<i>C. minimus</i>	100	<i>C. minimus</i>	100	<i>C. minimus</i> *	657	<a href="#">BOLD:AAA9441</a>
LCA30	<i>Catharus minimus</i>	<i>C. minimus</i>	100	<i>C. minimus</i> *	100	<i>C. minimus</i> *	660	<a href="#">BOLD:AAA9441</a>
LCA3	<i>Catharus ustulatus</i>	<i>C. ustulatus</i> *	100	<i>C. ustulatus</i> *	100	<i>C. ustulatus</i> *	702	<a href="#">BOLD:AAA9440</a>
LCA26	<i>Chaetura</i> sp	<i>C. brachyura</i>	100	<i>C. brachyura</i>	100	<i>C. brachyura</i>	644	<a href="#">BOLD:AAK0488</a>
LCA27	<i>Chaetura</i> sp	<i>C. brachyura</i>	100	<i>C. brachyura</i>	100	<i>C. brachyura</i>	642	<a href="#">BOLD:AAK0488</a>
LCA28	<i>Chaetura</i> sp	<i>C. brachyura</i>	100	<i>C. brachyura</i>	100	<i>C. brachyura</i>	652	<a href="#">BOLD:AAK0488</a>
* LCA24	<i>Coereba flaveola</i>	<i>C. flaveola</i>	100	<i>C. flaveola</i>	100	<i>C. flaveola</i>	651	<a href="#">BOLD:AAA4006</a>
* LCA33	<i>Dendrocincla fuliginosa</i>	<i>D. fuliginosa</i>	99.85	<i>D. fuliginosa</i>	99.15	<i>D. fuliginosa</i>	673	<a href="#">BOLD:ABZ6107</a>
LCA20	<i>Elaenia flavogaster</i>	<i>E. flavogaster</i>	99.85	<i>E. flavogaster</i>	98.93	<i>E. flavogaster</i>	681	<a href="#">BOLD:AAB3859</a>
LCA6	<i>Elaenia flavogaster</i>	<i>E. flavogaster</i>	100	<i>E. flavogaster</i>	99.39	<i>E. flavogaster</i>	696	<a href="#">BOLD:AAB3859</a>
* LCA18	<i>Galbula ruficauda</i>	<i>G. ruficauda</i>	100	<i>G. ruficauda</i>	97.55	<i>G. ruficauda</i>	675	<a href="#">BOLD:ABX4491</a>
* LCA31	<i>Manacus manacus</i>	<i>M. aurantiacus</i>	100	<i>M. manacus</i>	100	<i>M. aurantiacus</i>	667	<a href="#">BOLD:AAB9291</a>
LCA36	<i>Manacus manacus</i>	<i>M. aurantiacus</i>	100	<i>M. manacus</i>	100	<i>M. aurantiacus</i>	663	<a href="#">BOLD:AAB9291</a>
LCA38	<i>Manacus manacus</i>	<i>M. aurantiacus</i>	100	<i>M. manacus</i>	100	<i>M. aurantiacus</i>	667	<a href="#">BOLD:AAB9291</a>
LCA4	<i>Momotus subrufescens</i>	<i>M. momota</i>	100	<i>M. momota</i>	96.92	<i>M. momota</i>	681	<a href="#">BOLD:ABX4186</a>

Then, I looked up each of the 8 species plus one external species, also from Colombia, on the NCBI database [4] and on nucleotide. The following image is the search of the external species, *Pitangus sulphuratus*.

National Library of Medicine  
National Center for Biotechnology Information

Nucleotide

Nucleotide

pitangus sulphuratus

Create alert Advanced

Species
Animals (101)
Protists (2)
Viruses (4)
Customize ...

Molecule types
genomic DNA/RNA (107)
Customize ...

Source databases
INSDC (GenBank) (107)
Customize ...

Sequence Type
Nucleotide (107)

Summary 20 per page Sort by Default order Send to:

TAXONOMY
Was this helpful?

[Pitangus sulphuratus](#)

*Pitangus sulphuratus* is a species of bird in the family *Tyrannidae*.

Taxonomy ID: [371930](#)

I proceeded to search for a bird species identification gene and chose cytochrome oxidase subunit 1 (COI),

- ☐ [Pitangus sulphuratus isolate h57548-Pisu\\_ \(40\) cytochrome c oxidase subunit I \(COI\) gene, partial cds; mitochondrial](#)
1. [542 bp linear DNA](#)
- Accession: EU232973.1 GI: 159156673
- [PubMed](#) [Taxonomy](#)
- [GenBank](#) [FASTA](#) [Graphics](#) [PopSet](#)

to download the sequence for each of the 9 species in GenBank format.

GenBank

**Pitangus sulphuratus isolate h57548-Pisu\_(40) cytochrome c oxidase (COI) gene, partial cds; mitochondrial**

GenBank: EU232973.1

FASTA Graphics PopSet

Go to: ☐

LOCUS EU232973 542 bp DNA linear VRT 26-JUL-2016

DEFINITION Pitangus sulphuratus isolate h57548-Pisu\_(40) cytochrome c oxidase subunit I (COI) gene, partial cds; mitochondrial.

ACCESSION EU232973

VERSION EU232973.1

KEYWORDS

SOURCE mitochondrion Pitangus sulphuratus

ORGANISM Pitangus sulphuratus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archelosauria; Archosauria; Dinosauria; Saurischia; Theropoda; Coelurosauria; Aves; Neognathae; Passeriformes; Tyrannidae; Pitangus.

REFERENCE 1 (bases 1 to 542)

AUTHORS Chavez,A.V., Clozato,C.L., Lacerda,D.R., Sari,E.H. and Santos,F.R.

Send to:

Choose Destination ☒ File ☐ Clipboard ☐ Analysis Tool

Download 1 item.

Format

Show GI ☐

Create File

PubMed

Taxonomy

PopSet

Next, I opened the sequences with the program Geneious [5].

sequence-9.gb	3 KB	GenBank File	hoy, 3:10 p. m.
sequence.gb	4 KB	GenBank File	hoy, 3:06 p. m.
sequence-8.gb	3 KB	GenBank File	hoy, 3:06 p. m.
sequence-7.gb	3 KB	GenBank File	hoy, 3:06 p. m.
sequence-6.gb	4 KB	GenBank File	hoy, 3:06 p. m.
sequence-5.gb	3 KB	GenBank File	hoy, 3:06 p. m.
sequence-4.gb	3 KB	GenBank File	hoy, 3:06 p. m.
sequence-3.gb	3 KB	GenBank File	hoy, 3:06 p. m.
sequence-2.gb	3 KB	GenBank File	hoy, 3:06 p. m.

And proceeded to make a MUSCLE alignment (multiple alignment) with the following settings.

Geneious Prime Trial - 5 days 8 hours left

Alignment

Geneious Alignment MUSCLE Alignment Clustal Omega Realign Region

Translation Align Consensus Align

Muscle 3.8.425 by Robert C. Edgar

Maximum number of iterations: 3,000

☒ Optimize profile-dependent parameters ?

☐ Use custom MUSCLE executable:  Change...

Maximum number of trees to build: 1

Sequence Order: ☒ Group sequences by similarity ☐ Preserve sequence order

Optimization: ☐ Diagonal ☒ Anchor

Keep tree from: ☐ Iteration 1 ☐ Iteration 2

Distance measure: Iteration 1: kmer4\_6 Subsequent: pctid\_kimura

Clustering method: Iteration 1 & 2: UPGMB Subsequent: UPGMB

Tree Rooting method: Iteration 1 & 2: pseudo Subsequent: pseudo

Sequence weighting scheme: Iteration 1 & 2: CLUSTALW Subsequent: CLUSTALW

Terminal gaps: Half penalty Objective score: spm

Anchor spacing: 32 Gap open score: -1

Diagonals: Min length: 24 Margin: 5

Minimum column anchor scores: Min best: 90 Min smoothed: 90

Hydrophobicity: Multiplier: 1.2 Window size: 5

Maximum Memory (MB): 500

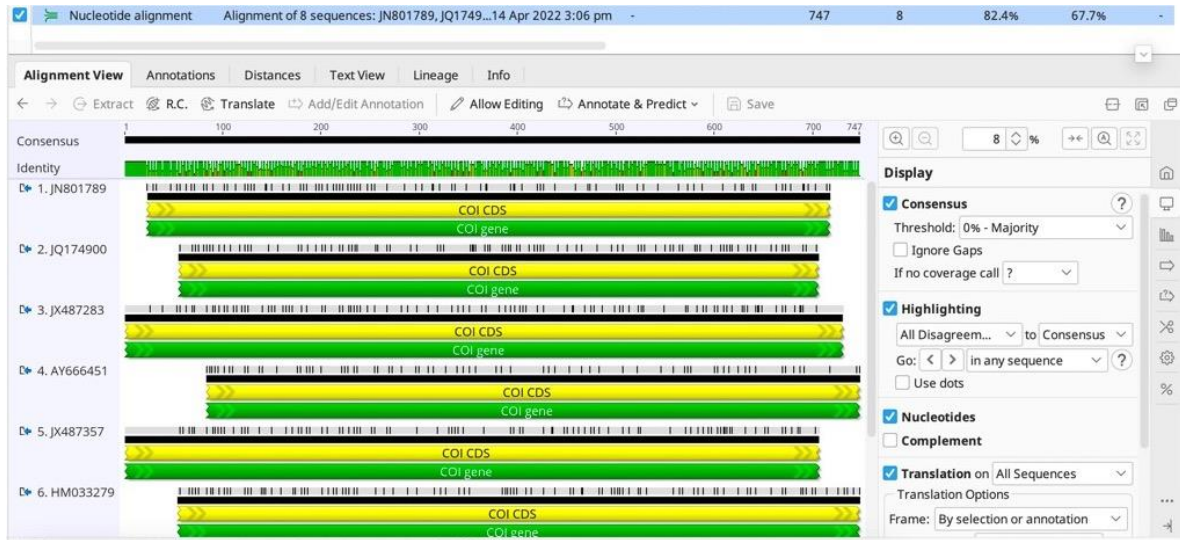
☐ Quit and return best alignment found if not finished after 1 hours

Extra Options:

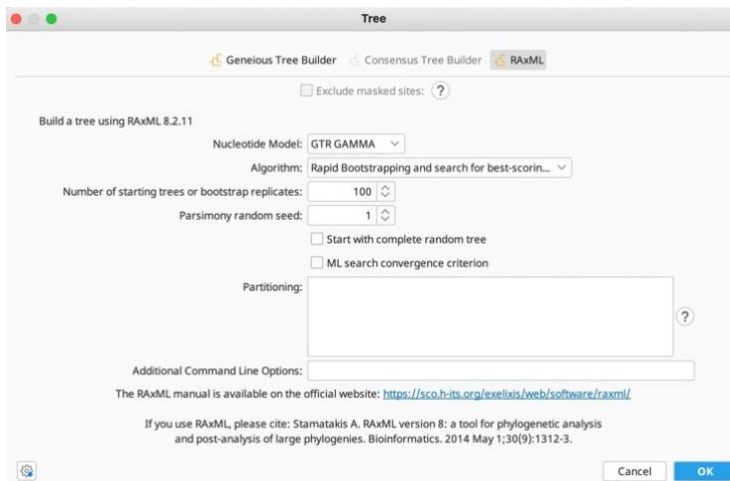
Fewer Options ^

Cancel OK

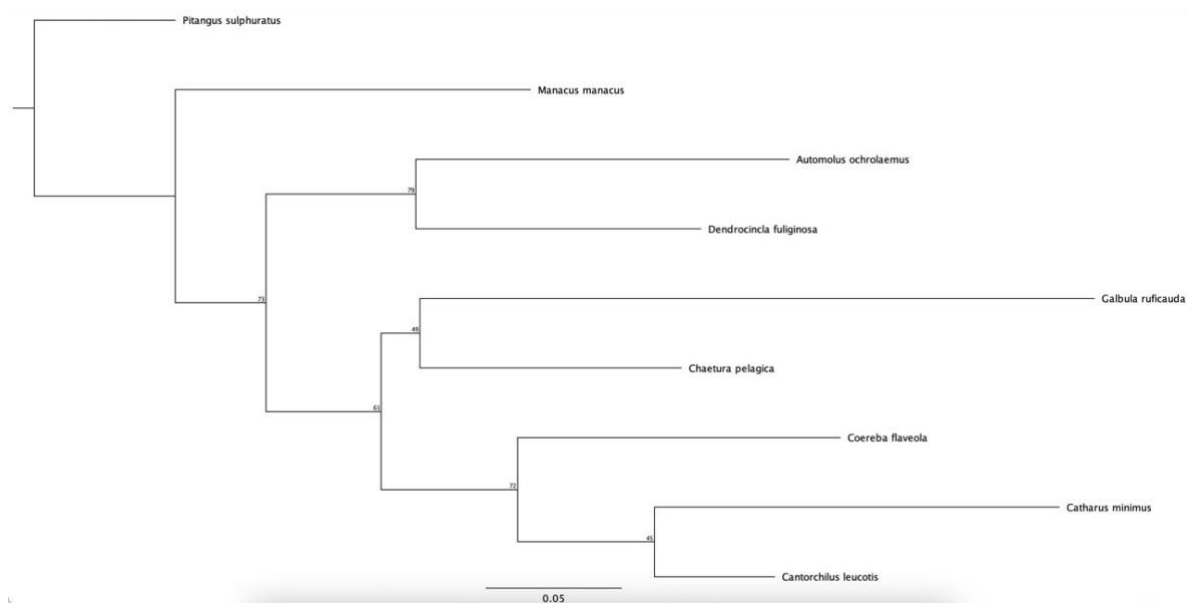
The resulting alignment.



For the next step, I built a tree using the plugin RAXML 8.2.11 with the following settings.



And the tree that resulted is shown in the following image.



### III Results & Discussion

The tree that was obtained, shows the possible phylogenetic relationships of the 9 species that were selected. As it can be seen, the external group is not as closely related as the other selected species are among each other. *Manacus Manacus* would come in second as the least related, but overall, there are close relationships shown.

### Conclusion

The bootstrap values obtained in the generated phylogeny are not high support values (we consider >75 as high values) so we could not say that this is the true evolutionary history of these species. However, this may be a viable approach to knowing the true phylogeny of these specific birds and possibly filling in information gaps about their phylogeny. In addition, it can be seen with this work, how the use of DNA barcode data sets, tools such as NCBI database [4], Geneious [5] and the identification gene of different species (COI), helps to reconstruct the phylogeny of different species, among many other things.

### References

- [1] Pulgarín-R, P. C., Olivera-Angel, M., Ortiz, L., Nanclares, D., Velásquez-Restrepo, S., & Díaz-Nieto, J. F. (2021). DNA barcodes of birds from northern Colombia. *Biodiversity data journal*, 9, e64842. <https://doi.org/10.3897/BDJ.9.e64842>
- [2] Acevedo-Charry, O., Colón-Piñeiro, Z., Ocampo, D., Pinzon, M. & Ayerbe-Quiñones, F. (2020). Avifauna colombiana. En: Moreno, L. A. & Andrade, G. I. (Eds.). *Biodiversidad 2019. Estado y tendencias de la biodiversidad continental de Colombia*. Bogotá, D. C., Colombia: Instituto de Investigación de Recursos Biológicos Alexander von Humboldt. 92p.

[3] Bichofué/Great Kiskadee/Pitangus sulphuratus. (2022). Retrieved 19 April 2022, from <https://birdscolumbia.com/2016/08/03/bichofue-pitangus-sulphuratus/>

[4] National Center for Biotechnology Information (NCBI)[Internet]. Bethesda (MD): National Library of Medicine (US), National Center for Biotechnology Information; [1988] – [cited 2022 Apr 16]. Available from: <https://www.ncbi.nlm.nih.gov/>

[5] Geneious Prime 2022.1.1. (<https://www.geneious.com>)