CABANA
Capacity building for bioinformatics in Latin America

EMBL-EBI

MOLECULAR DATA MINING TECHNOLOGIES FOR THE STUDY OF GENOMES RELEVANT TO LATIN AMERICAN COUNTRIES

ALAG 2019

Mendoza

Universidad de los Andes

Marco Cristancho, PhD

¿Why is it important to carry out biodiversity (molecular) studies in LATAM?

Marco Cristancho

Universidad de los Andes


BIODIVERSIDAD EN EL MUNDO[45]
5 MÉXICO
2 COLOMBIA
9 ECUADOR
10 VENEZUELA
6 PERÚ
1 BRASIL
4 CHINA
8 INDIA
3 INDONESIA
7 AUSTRALIA
1°
10°

Biodiversity studies importance in Latin America

Mega - Biodiverse countries

EARTH'S MOST BIODIVERSE COUNTRIES								
COUNTRY	Birds	Amphib	Mammals	Reptiles	Fish	Vascular plants	Index	Rank
Brazil	17.6%	13.6%	11.8%	7.9%	13.7%	18.2%	0.83	1
Colombia	18.3%	10.2%	8.1%	5.9%	6.2%	16.6%	0.65	2
Indonesia	16.2%	4.6%	12.2%	7.1%	14.1%	9.5%	0.64	3
Peru	18.1%	7.6%	8.5%	4.7%	4.7%	9.7%	0.53	4
China	12.5%	5.5%	10.0%	4.7%	10.1%	10.4%	0.53	5
Mexico	10.9%	5.0%	9.5%	8.9%	7.9%	8.5%	0.51	6
Australia	7.1%	3.2%	6.4%	10.1%	14.7%	5.1%	0.47	7
India	11.9%	5.2%	7.5%	6.7%	7.4%	6.1%	0.45	8
Ecuador	16.0%	7.2%	6.8%	4.3%	3.3%	6.3%	0.44	9
United States	8.5%	4.0%	8.0%	5.2%	9.3%	6.3%	0.41	10
Venezuela	13.7%	4.8%	6.6%	3.9%	5.2%	6.8%	0.41	11
Bolivia	14.3%	3.2%	6.6%	3.0%	1.2%	5.6%	0.34	12
DR Congo	10.9%	3.2%	7.8%	2.9%	4.5%	3.6%	0.33	13
South Africa	7.6%	1.7%	5.4%	4.4%	6.2%	7.6%	0.33	14
Malaysia	7.1%	3.5%	6.1%	4.7%	5.8%	5.0%	0.32	15
Papua New Guinea	7.2%	4.9%	4.9%	2.7%	8.5%	3.7%	0.32	16
Tanzania	10.6%	2.7%	6.5%	3.5%	5.3%	3.2%	0.32	17
Viet Nam	8.3%	3.0%	5.2%	4.5%	7.3%	3.4%	0.32	18
Thailand	9.3%	1.9%	5.7%	4.2%	6.4%	3.8%	0.31	19
Argentina	10.0%	2.3%	6.8%	4.3%	3.0%	3.0%	0.29	20
Kenya	10.4%	1.5%	6.9%	2.7%	3.2%	2.1%	0.27	21
Cameroon	8.8%	2.9%	6.1%	2.8%	3.1%	2.7%	0.26	22
Panama	8.8%	2.8%	4.5%	2.6%	4.2%	3.2%	0.26	23
Philippines	5.7%	1.5%	3.5%	2.0%	9.9%	2.9%	0.26	24
Costa Rica	8.6%	2.7%	4.1%	2.6%	3.3%	3.9%	0.25	25
Myanmar	10.2%	1.2%	5.4%	3.0%	3.1%	2.3%	0.25	25

MONGABAY

Sistema de Información sobre Biodiversidad de Colombia

frontiers

Microbial diversity exploration of marine hosts at Serrana Bank, a coral atoll of the Seaflower Biosphere Reserve

Astrid Catalina Alvarez-Yela[1*], Jeanneth Mosquera-Rendón[1], Alejandra Noreña-P[1], Marco Cristancho[2], Diana López-Alvarez[1,3*]

[1]Centro de Bioinformática y Biología Computacional - BIOS, Manizales, Colombia

[2]Universidad de los Andes, Bogotá, Colombia

[3]Present address: Facultad de Ciencias Agropecuarias, Universidad Nacional de Colombia, Palmira, Colombia

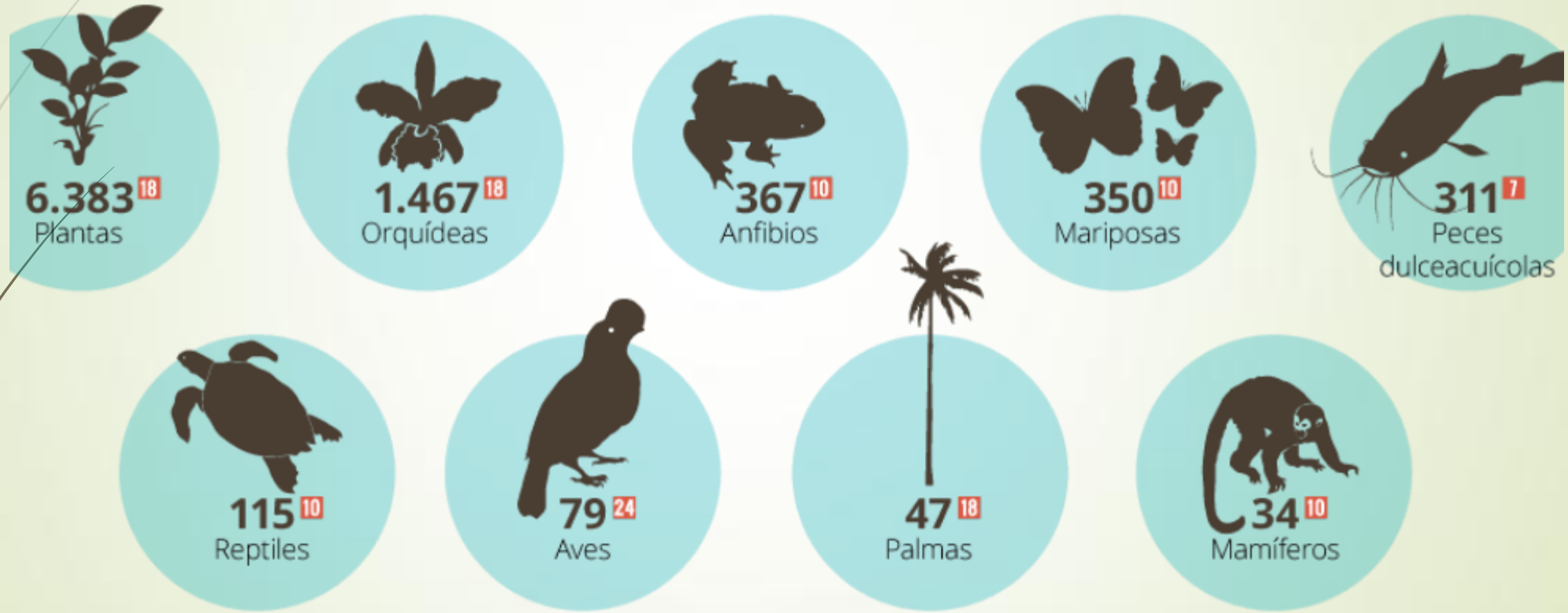
*** Correspondence:**
Astrid Catalina Alvarez-Yela - catalina.alvarez@bios.co
Diana López-Alvarez dianalopez430@gmail.com

Keywords: seaflower, microbial diversity, marine hosts, amplicons, symbiotic interactions

ESPECIES ENDÉMICAS DE LATINOAMÉRICA

NATIONAL GEOGRAPHIC

ESPECIES ENDÉMICAS DE COLOMBIA



SiB Sistema de Información sobre Biodiversidad de Colombia

¿CUÁNTAS ESPECIES ESTÁN AMENAZADAS EN COLOMBIA?

CR EN PELIGRO CRÍTICO

EN EN PELIGRO

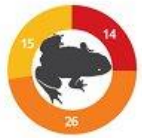
VU VULNERABLES

De acuerdo con la información disponible en la serie de **Libros Rojos de Especies Amenazadas de Colombia** y en la **Resolución 0192 de 2014** se han identificado cerca de **1.200** especies en distintas categorías de amenaza, según los criterios de la unión internacional para la conservación de la naturaleza - UICN.

112 [24]
Aves

68 [25]
Aves
Bosques húmedos de los andes y la costa pacífica

40 [26]
Mamíferos



55 [27]
Anfibios

43 [28]
Reptiles

7 [29]
Corales

14 [29]
Moluscos marinos

7 [29]
Decápodos marinos

9 [30]
Arácnidos

12 [31]
Mariposas

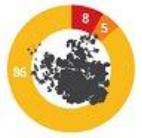
5 [30]
Escarabajos

19 [29]
Himenópteros

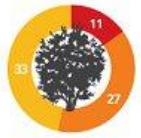
26 [31]
Decápodos dulceacuícolas

28 [32]
Peces marinos
Incluye rayas y tiburones

53 [32]
Peces dulceacuícolas

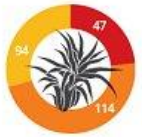


99 [33] [34]
Musgos y afines



71 [35]
Fanerógamas
Árboles, arbustos y lianas

88 [37]
Frailejones, zamias y palmas



255 [38]
Bromelias, labiadas y pasifloras

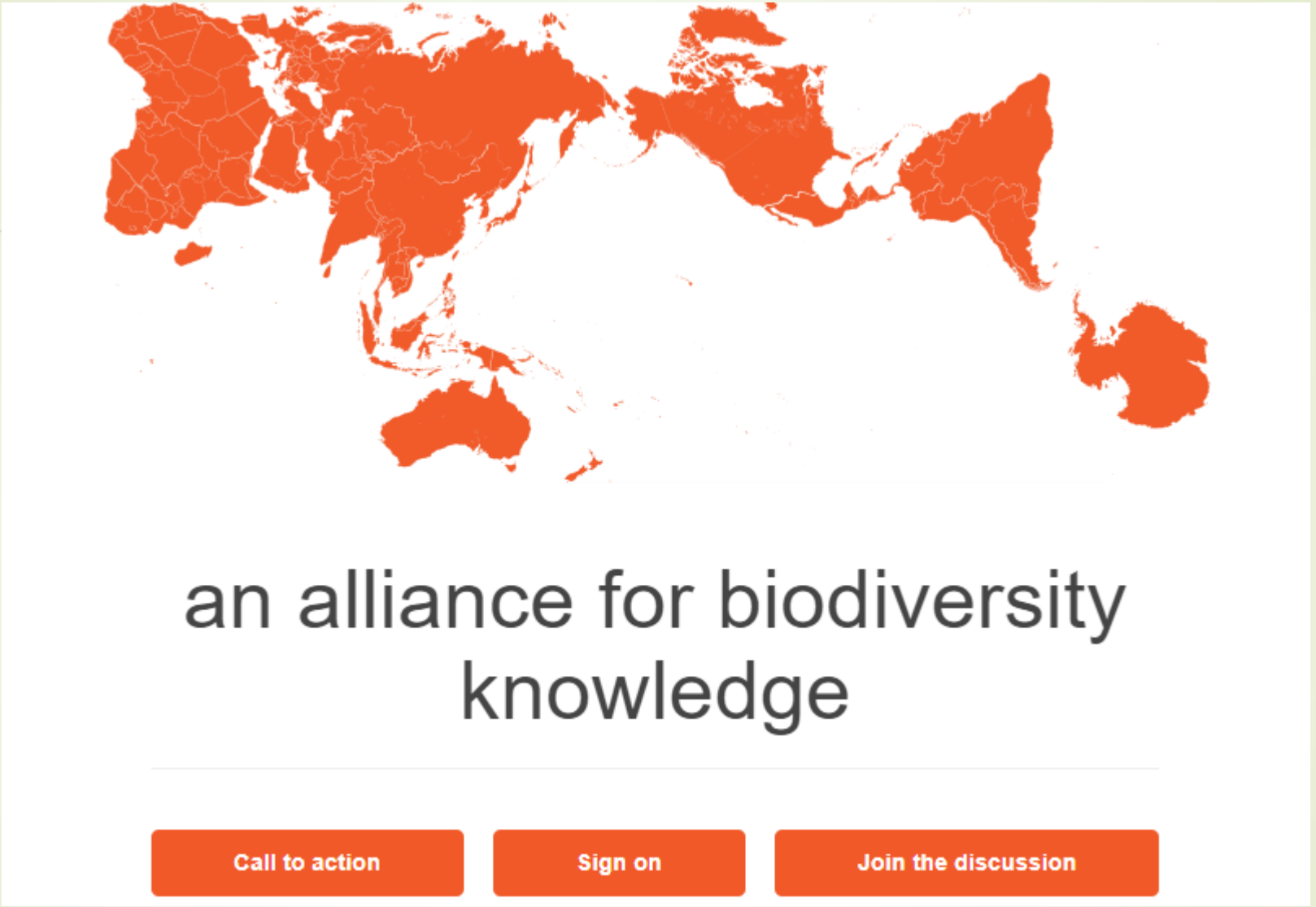
34 [39] [40]
Especies maderables

44 [40]
Magnolias y afines

207 [41]
Orquídeas

SiB Sistema de Información sobre Biodiversidad de Colombia

Biodiversity Informatics



https://www.biodiversityinformatics.org/en/#

PLOS COMPUTATIONAL BIOLOGY A Peer-Reviewed, Open Access Journal

Building cyberinfrastructure in emerging developing countries to stimulate innovation and socioeconomic growth

Jose De Vega[1]*, Robert P Davey[1], Jorge Duitama[2], Dairo Escobar[3], Marco A Cristancho-Ardila[4], Graham J Etherington[1], Alice Minotto[1], Nelson E Arenas-Suarez[5], Juan D Pineda-Cardenas[6], Javier Correa-Alvarez[7], Anyela V Camargo Rodriguez[8], Juan P Mallarino-Robayo[4], Emiliano Barreto-Hernandez[9], Monica Muñoz-Torres[10], Narcis Fernandez-Fuentes[11], Federica Di Palma[1], and The Colombian Cyberinfrastructure Consortium for Biodiversity[§]

[1] Earlham Institute, Norwich Research Park, Norwich, UK.

[2] Systems and Computing Engineering Department, Universidad de Los Andes, Bogotá, Colombia.

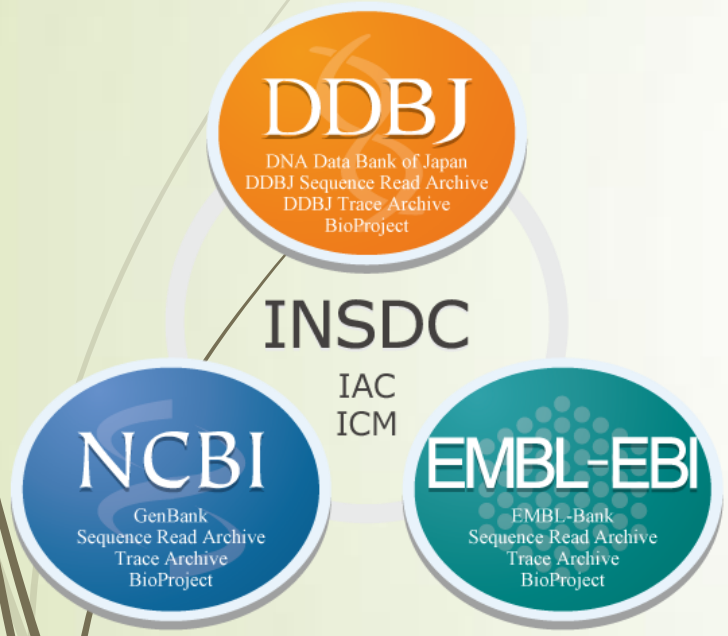
[3] SiB Colombia - Instituto de Investigación de Recursos Biológicos Alexander von Humboldt, Bogotá, Colombia.

[4] Faculty of Sciences, Universidad de Los Andes, Bogotá, Colombia.

[5] Faculty of Sciences, Universidad Antonio Nariño, Bogotá, Colombia.

[6] Apolo Supercomputing Centre, EAFIT University, Medellín, Colombia.

Molecular Databases



BOLD SYSTEMS

PATRIC
Pathosystems Resource Integration Center

KEGG
Kyoto Encyclopedia of Genes and Genomes

JGI

(GIGA)^n DB

ENA
European Nucleotide Archive
EMBL-EBI

Sequence

Environmental

Contig_Set

BioSample

Coding

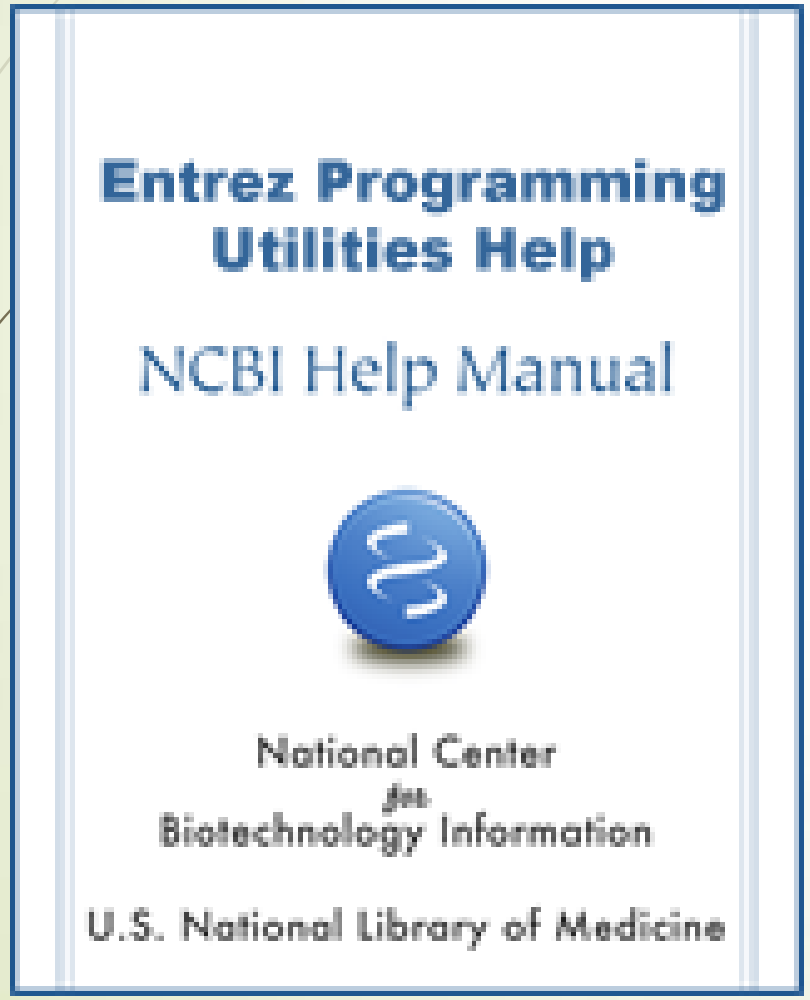
Sample

Marco Cristancho

Universidad de los Andes

Data Mining Tools

E-Direct



EBI APIs

Ensembl REST APIs

rest

Latin America Molecular Diversity in Public Databases

Noreña – P *et al. BMC Genomics* 2018, **19**(Suppl 8):859
https://doi.org/10.1186/s12864-018-5194-8

BMC Genomics

RESEARCH Open Access

Colombia, an unknown genetic diversity in the era of Big Data

Alejandra Noreña – P[1], Andrea González Muñoz[1*], Jeanneth Mosquera-Rendón[1], Kelly Botero[1] and Marco A. Cristancho[1,2]

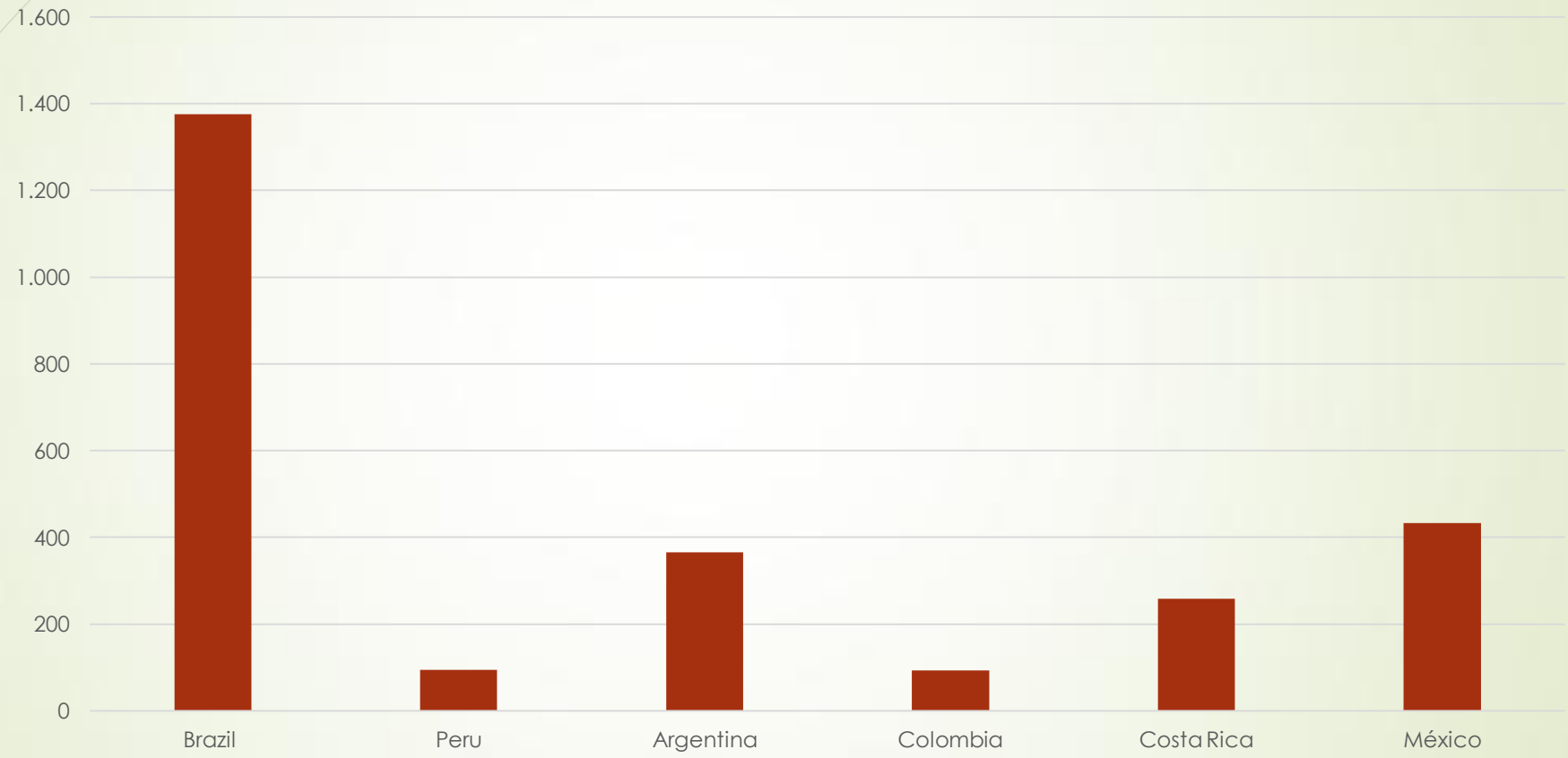
From Selected articles from the IV Colombian Congress on Bioinformatics and Computational Biology & VIII International Conference on Bioinformatics SolBio 2017
Santiago de Cali, Colombia. 13-15 September 2017

Abstract

Background: Latin America harbors some of the most biodiverse countries in the world, including Colombia. Despite the increasing use of cutting-edge technologies in genomics and bioinformatics in several biological science fields around the world, the region has fallen behind in the inclusion of these approaches in biodiversity studies. In this study, we used data mining methods to search in four main public databases of genetic sequences such as: NCBI Nucleotide and BioProject, Pathosystems Resource Integration Center, and Barcode of Life Data Systems databases. We aimed to determine how much of the Colombian biodiversity is contained in genetic data stored in these public databases and how much of this information has been generated by national institutions. Additionally, we compared this data for Colombia with other countries of high biodiversity in Latin America, such as Brazil, Argentina, Costa Rica, Mexico, and Peru.

ENA
European Nucleotide Archive
EMBL-EBI

Environmental



Marco Cristancho
Universidad de los Andes

ENA
European Nucleotide Archive
EMBL-EBI

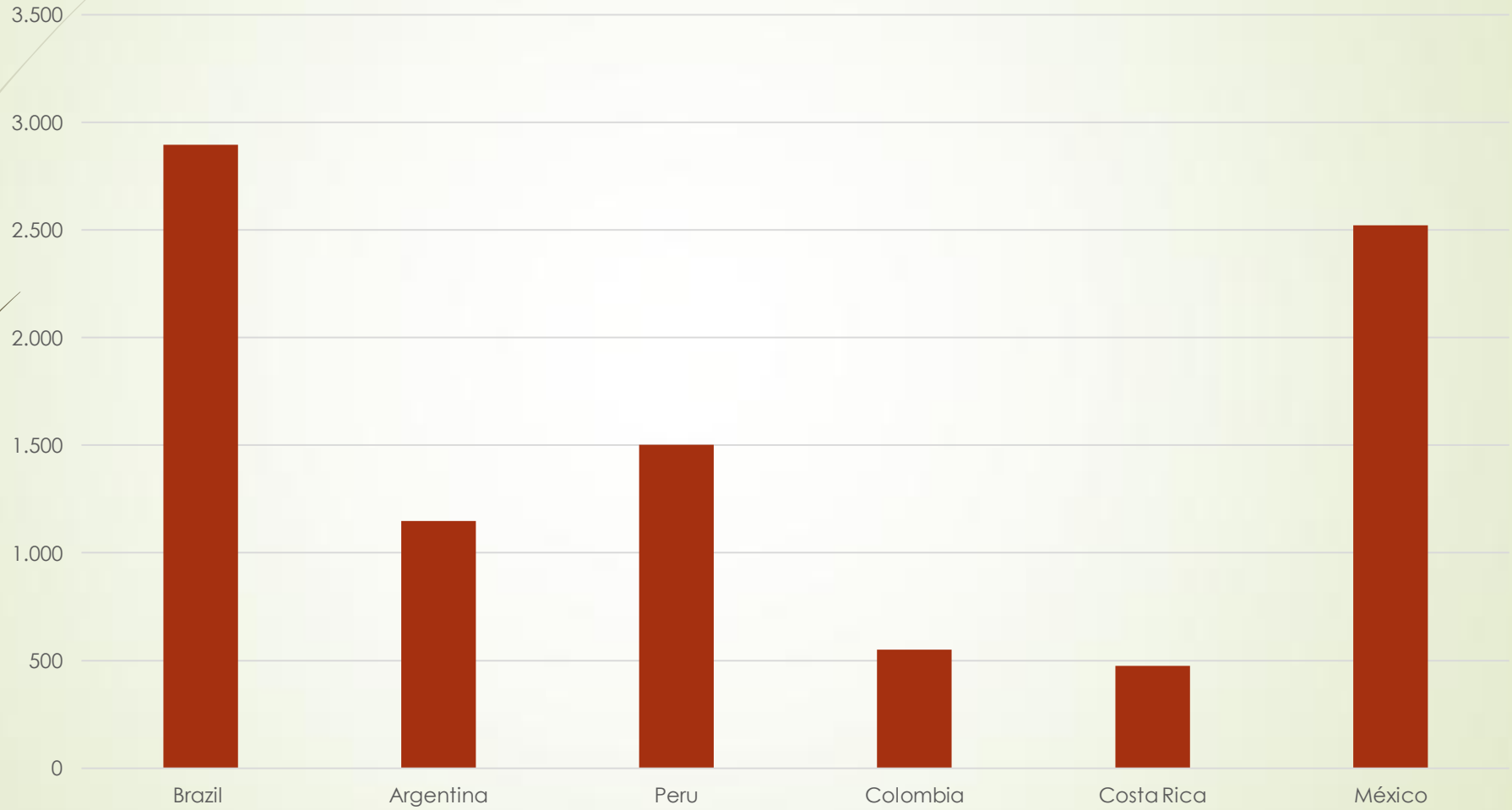
Environmental

Marco Cristancho

Universidad de los Andes

ENA
European Nucleotide Archive
EMBL-EBI

Sample

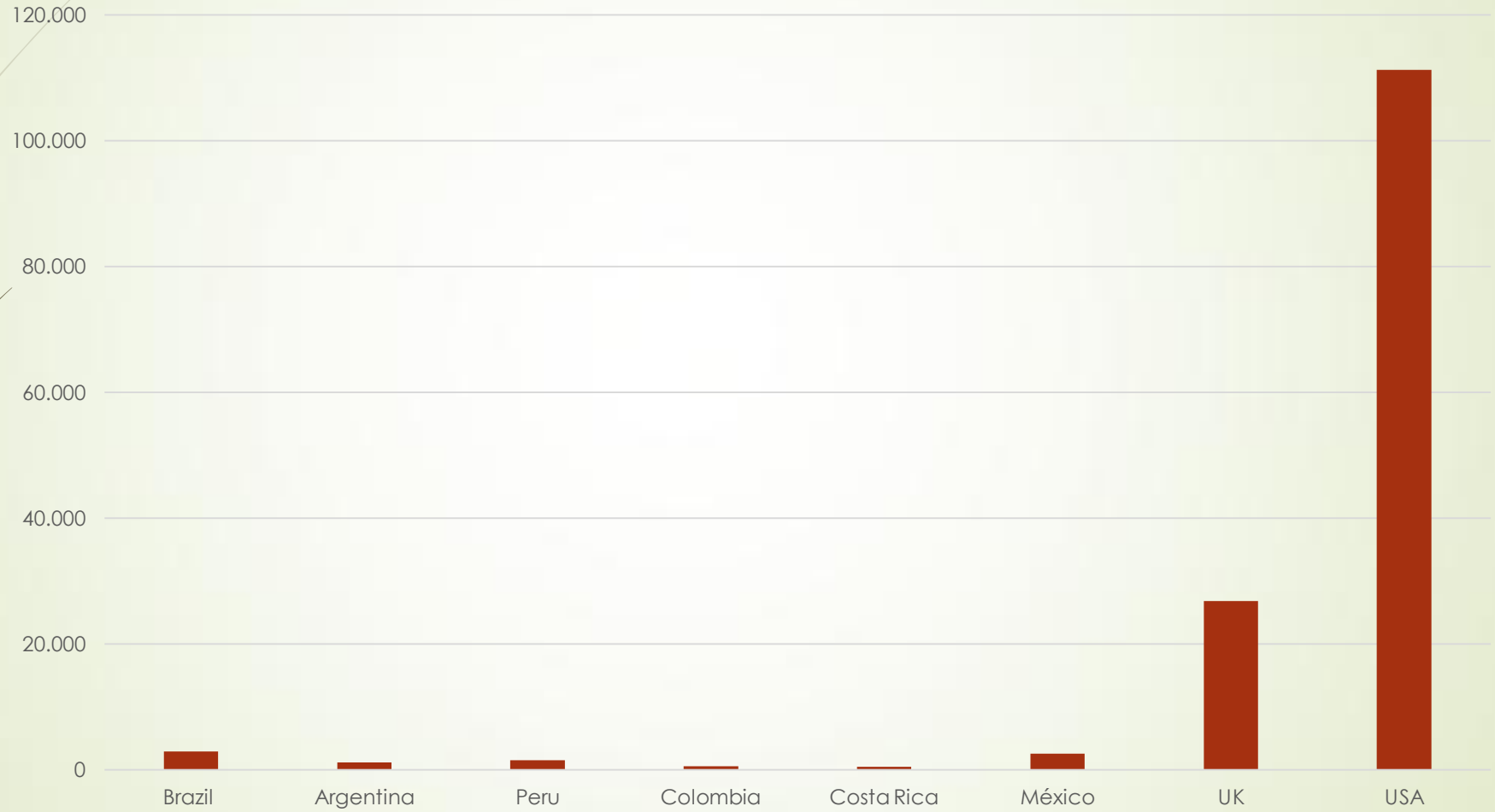


Marco Cristancho

Universidad de los Andes

ENA
European Nucleotide Archive
EMBL-EBI

Sample

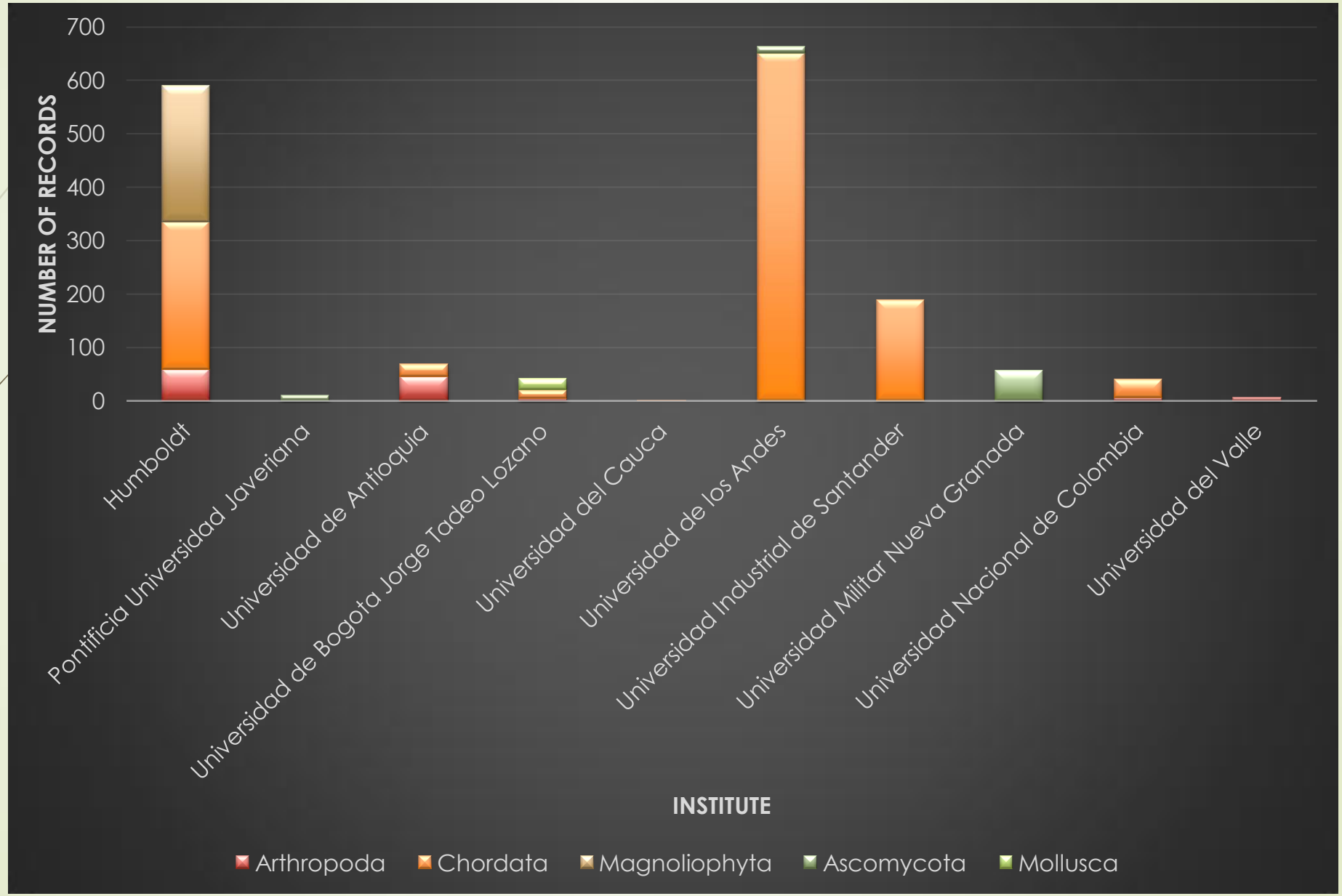


Marco Cristancho

Universidad de los Andes

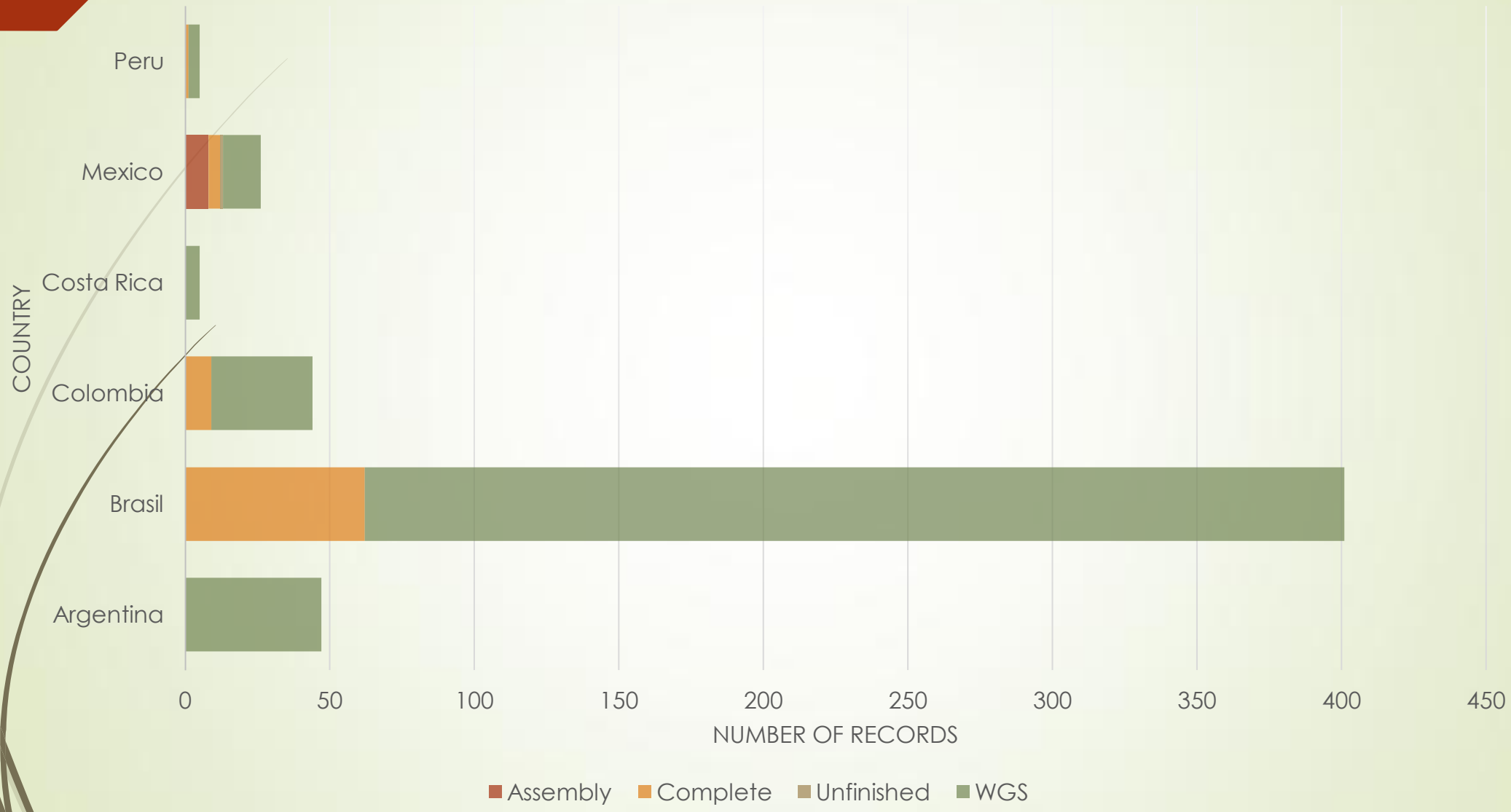
BOLD SYSTEMS

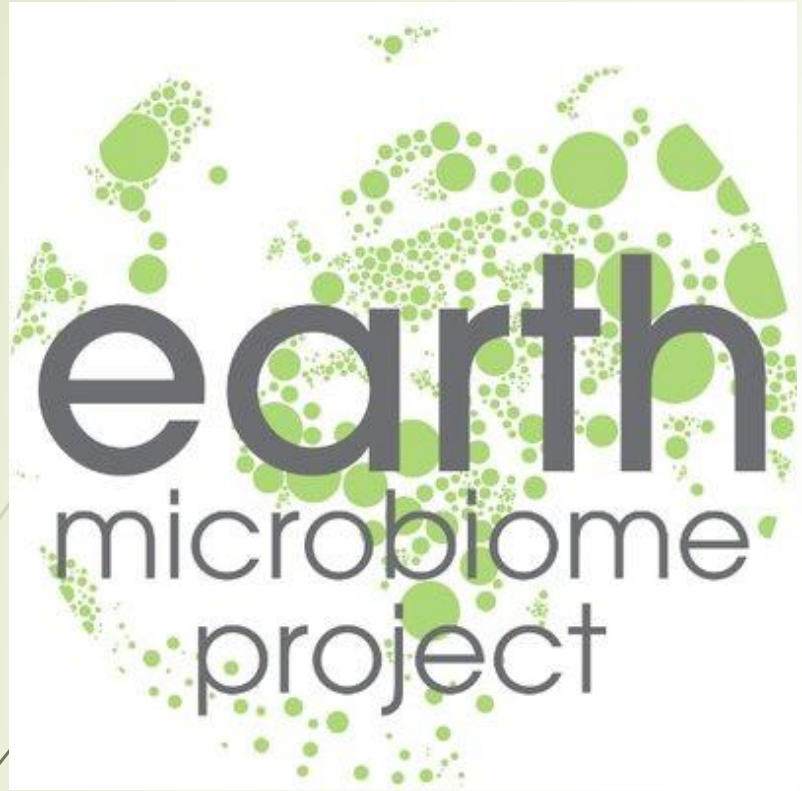
2.095 species for Colombia (April 2019)
1.673 (published paper)



Universidad de los Andes

Genome Statistics





VERTEBRATE
GENOMES
PROJECT

A PROJECT OF THE G10K CONSORTIUM

CABANA

Capacity building for Bioinformatics in Latin America

GCRF Global Challenges Research Fund

RESEARCH COUNCILS UK

EMBL-EBI

What is CABANA?

CABANA
Capacity building for bioinformatics in Latin America

- A capacity-strengthening project focused on bioinformatics
- Orchestrated by an international consortium of ten organisations – nine in Latin America, one in the UK
- Open to new partners – from within and beyond our partner countries
- Funded by the Global Challenges Research Fund – part of the UK AID budget
- 1 Oct 2017 – 31 Dec 2021

CABANA consortium

CABANA
Capacity building for bioinformatics in Latin America

EMBL-EBI

Universidad de los Andes

Cinvestav

VALE

UNIVERSIDAD DE COSTA RICA

UNICAMP

USMP
UNIVERSIDAD DE SAN MARTIN DE PORRES

UNIVERSIDAD DE BUENOS AIRES

CIP
INTERNATIONAL POTATO CENTER
A CGIAR RESEARCH CENTER

INTA
Instituto Nacional de Tecnología Agropecuaria

CABANA

Capacity building for bioinformatics in Latin America

CABANA
Capacity building for bioinformatics in Latin America

UNDARK

VARIABLES / *News & Features*

Changing Latin America's Culture of Insular Science

The region's scientists lament that their research is too often disconnected from the larger world. In the age of Zika, that needs to change.

03.21.2017 / BY Aleszu Bajak

1 COMMENT

CABANA
Capacity building for bioinformatics in Latin America

CABANA Workshop: Analysis of Crop Genomics Data

Date: Monday 2 - Friday 6 December 2019

Venue: Universidad de los Andes - UniAndes - Cra 1 Nº 18A - 12, 111711, Bogota, CUN, Colombia

Application opens: Thursday 25 July 2019

Application deadline: Sunday 29 September 2019

Participation: Open application with selection

Contact: Marco Cristancho, Derrin Crawte

Registration fee: $112000 COP

Registration closed

Find other similar courses »

CABANA Workshop: Introduction to Metagenomics

Date: Monday 2 - Friday 6 December 2019

Venue: Faculty of Natural Sciences - University of Buenos Aires (FCEN-UBA) - Ciudad Universitaria. Nuñez - Ciudad Autónoma de Buenos Aires - Pabellón II, 1428, Buenos Aires, Buenos Aires, Argentina

Application opens: Tuesday 03 September 2019

Application deadline: Sunday 29 September 2019

Participation: Open application with selection

Contact: Derrin Crawte, Maria Bernardi

Registration fee: ARS $1300

Registration closed

Find other similar courses »

Marco Cristancho

ma.cristancho29@uniandes.edu.co

CABANA

Capacity building for bioinformatics in Latin America

https://www.cabana.online/

Thanks for listening!