

Miguel Ángel González Arias

PERSONAL INFORMATION

Place and date of birth: Mexico city, Mexico.

Mail: gamamiguelangel@gmail.com

Socials: X (Twitter), [@miangoar](#). Bluesky, [@miangoar](#). GitHub, [@miangoar](#). Personal website, [miangoaren.github.io](#).

EDUCATION

Master degree in Biochemical Sciences (graduated with honors)

Cuernavaca, Morelos

Institute of Biotechnology, National Autonomous University of Mexico (UNAM)

Aug. 2020 – Mar. 2024

- Thesis: [Comparative evaluation of protein language models for betalactamase classification and prediction of their catalytic activity](#)
- Advisors: [Alejandro Angel Garciarrubio Granados](#) and [Lorenzo Patrick Segovia Forcella](#)
- Second advisors: [José Arcadio Farías Rico](#) and [Francisco Xavier Soberón Mainero](#)

Bachelors degree in Biology

Coyoacan, Mexico City

Faculty of Sciences, UNAM

Aug. 2015 – Jan. 2021

- Thesis: [Genome reconstruction from metagenomes of the Gulf of Mexico](#)
- Advisor: [Lorenzo P. Segovia Forcella](#)

Certified proficiency in English (equivalent to B2)

Coyoacan, Mexico City

National School of Languages, Linguistics and Translation, UNAM

Nov. 2019

PUBLICATIONS/PREPRINTS

[\(2025\) Large scale analysis of the \$\beta\$ -lactamase sequence space with protein language models](#)

Final draft begin revised by co-authors.

SKILLS AND EXPERIENCE

Biology: molecular biology, biochemistry, protein evolution, protein structure and function, phylogenomics, microbial ecology, metagenomics, next generation sequencing, comparative genomics, systems biology.

Informatics: python, R (basics), bash (Linux-GNU), computer hardware, data warling, data visualization, machine learning, deep learning, high performance computing (Oracle Grid Engine), cloud computing (Microsoft Azure and Google Colab), office suite tools, ChatGPT.

Languages: Spanish (native), English (B2).

Soft skills: organization, self-taught, story telling, curiosity, adaptability, friendly.

TEACHING EXPERIENCE

- **(Bachelor's thesis advising) Díaz López, Emiliano (2024). Evaluation of virtual phlogenomic methodologies for the identification and classification of Archaea**
Gustavo A. Madero, Mexico city, National School of Biological Sciences (IPN), May. 2024
- **(Lecture) Introduction to microbial genomics and taxogenomics**
Coyoacan, Mexico city, Institute of Marine Sciences and Limnology, UNAM, Sep. 2023
- **(Lecture) Introduction to protein science with machine learning**
Cuernavaca, Morelos, Center for Genomic Sciences, UNAM, Feb. 2023
- **(Lecture) Introduction to microbial ecology and metagenomics**
Coyoacan, Mexico city, Faculty of Sciences, UNAM, Jun. 2021
- **(Lecture) Genome reconstruction from metagenomes**
International, Virtual event in colaboration with WinterGenomics, Jun. 2020

SCIENCE COMMUNICATION EXPERIENCE

Resource: [Awesome AI-based Protein Science](#)

Social media

Genome workshop

American Educational Center

Cuernavaca, Morelos

Nov. 2019

Festival of sciences and humanities

UNAM science Museum (Universum)

Coyoacan, Mexico city

Oct. 2019

Resource: [Personal blogpost on protein science and other topics](#)

Social media

Resource: [Annual biology student tips series](#)

Social media

COURSES AND SCHOLARSHIPS

Awarded scholarship: 1st Latin american workshop on genome mining

Mathematics Research Center

Guanajuato, Mexico

2022

Awarded scholarship: Representation Learning in Biology ISMB/ECCB 2021

Virtual event

International

2021

Awarded scholarship: Introduction to cloud computing with Microsoft Azure (AZ-900)

Virtual event

Mexico

2021

Awarded scholarship: Scholarship for master's studies by the UNAM

Institute of Biotechnology, UNAM

Cuernavaca, Morelos

2021

Course: Tools for the analysis of metagenomic data

Institute of Biotechnology, UNAM

Cuernavaca, Morelos

2020

Awarded scholarship: 4th International Symposium on Bioinformatics

National Institute of Public Health

Cuernavaca, Morelos

2019

Course: Whole metagenome shotgun analysis and comparative genomics

Gulf of Mexico Research Consortium

Cuernavaca, Morelos

2018

CODE EXAMPLES

- [Phylogenetic analysis and ancestral sequence reconstruction](#)
- [Protein structure analysis of predicted models from ESMAtlas](#)
- [Batched protein structure prediction with ESMFold and Google drive](#)
- [Comparison of contact maps and attention patterns derived from protein language models](#)
- [Download thousands of predicted structures from the AlphaFold Database](#)
- [Amino acid conservation parsing into PDB files](#)
- [Make ChimeraX visualization of the "protein folding" process by ESMFold](#)

ACADEMIC REFERENCES

- **Lorenzo Patrick Segovia Forcella**
Institute of Biotechnology, UNAM
Contact: lorenzo.segovia@ibt.unam.mx
- **Alejandro Angel Garcarrubio Granados**
Institute of Biotechnology, UNAM
Contact: alejandro.garciarrubio@ibt.unam.mx
- **José Arcadio Farías Rico**
Center for Genomic Sciences, UNAM
Contact: jafarias@ccg.unam.mx
- **José Luis Puente García**
Institute of Biotechnology, UNAM
Contact: jose.puente@ibt.unam.mx
- **Claudia Martínez Anaya**
Institute of Biotechnology, UNAM
Contact: claudia.martinez@ibt.unam.mx
- **Luis David Alcaraz Peraza**
Faculty of Sciences, UNAM
Contact: lalcaraz@ciencias.unam.mx