

Roberto Sánchez Navarro

Bioinformatician - Data Analyst



robertosanchezn

robertosanchezn

robertosan97@gmail.com

+34 644 902 167

About me

Bioinformatician looking for real-world challenges. I'm curious about the data, so I often take the extra step until it makes sense. Then, I like presenting it in a way that tells a story.

I'm continuously under training, learning new technologies every week, and when I pick one up, I enjoy passing that knowledge along to my peers.

Soft skills

Attentive to detail
Focused on quality
Good communicator
Highly motivated

Areas of specialization

Data Science • Data Analysis
• Bioinformatics • Big Data

Interests

Running • Cycling • Hiking

EXPERIENCE

2022–2023

PhD Candidate

AALBORG UNIVERSITY • Denmark

Development of a python application for the detection and analysis of exopolysaccharide gene clusters in bacterial metagenomes. This entailed:

- Retrieving all exopolysaccharide clusters from the published literature and extracting all existing genetic knowledge about them.
- Development of a Snakemake pipeline that uses DIAMOND, hmmer, biopython, and visualization tools, among others, to mine public genome databases to collect sequence data to train models.
- Exploratory analysis of these sequences using the phylogenetic python suite ete3 to redefine the classification of these clusters.
- Using all this knowledge to develop de novo HMMs to be used in the new application.



2021–2022

Research Assistant

AALBORG UNIVERSITY • Denmark

Genome mining of metagenome-assembled genomes from wastewater treatment plants looking for biosynthetic gene clusters.



EDUCATION

2023–2024

Big Data Analysis

MSc. • University of Murcia



2019–2021

Biotechnology

MSc. • Aalborg University



2015–2019

Biotechnology

BSc. • University of Murcia



2017–2018

Exchange Year

ISEP • University of Tennessee



PROGRAMMING

R

4 years



python

3 years



bash

3 years



LaTeX

4 years



SQL

1 year



TECHNICAL SKILLS

Databases
Bioinformatics
Data Wrangling
Visualization
Reporting
Big Data
Office Tools
DB Clients
Machine Learning
Development
Scraping
Other

MySQL • MongoDB • HBase • Neo4j
NGS • (meta)genomics • genome mining
pandas • dplyr • tidy • purrr
matplotlib • plotly • ggplot2 • ggiraph • Shiny
R Markdown • jupyter • Google colab • Quarto
Hadoop • slurm • Spark • Snakemake
Overleaf • Excel • PowerPoint • Slidy
sqlalchemy • pymongo • dbplyr • happybase
caret • Keras • scikit-learn • TensorFlow
conda/mamba • Docker • Git
Scrapy • BeautifulSoup • rvest
regex • API integration (Requests & http)

LANGUAGES

English	C1 (Advanced) + full professional proficiency
Spanish	Mother tongue
French	B1 Certificate
Danish	Beginner

PUBLICATIONS

Sánchez-Navarro et. al, 2022
Long-Read Metagenome-Assembled Genomes Improve Identification of Novel Complete Biosynthetic Gene Clusters in a Complex Microbial Activated Sludge Ecosystem. *mSystems*. 10.1128/msystems.00632-22 [↗](#)

SHINY APPS

gtdb-viz	↗	Interactive visualization of the phylogenomic database GTDB
in-dashboard	↗	Simple BI dashboard mock-up for an university project.