



Santiago López Begines, PhD

Neuroscientist & Data Scientist | Biomedical Data Pipelines |
Neurotech · MedTech · Pharma | R · Python ·
Electrophysiology

+34 645692225

santiago.lopez.begines@gmail.com

linkedin.com/in/santibegines

slopezbegines.github.io

Seville, Spain

SUMMARY

Neuroscientist and Data Scientist with 15+ years in biomedical research and 4–5 years specializing in machine learning, multi-omics analysis, and production data pipelines.

First author in Science Advances (2025); co-authored 8+ publications in EMBO Journal, eLife, and Cell Death & Disease.

Unique hybrid profile: deep experimental background in whole-cell patch-clamp electrophysiology, proteomics, and transgenic mouse models — combined with production-level data science in Python and R.

Key results:

- 90% reduction in electrophysiology processing time via automated pipeline (IBiS, Seville)
- 70% reduction in omics data cleaning time via automated R scripts (LCSB, University of Luxembourg)
- ML-based biomarker discovery in neurodegeneration (Batten disease / CLN4)

Proficient in Python (scikit-learn, XGBoost, LightGBM, TensorFlow) and R (tidyverse, tidymodels, Seurat, Shiny) for automated pipelines, statistical modeling, and data visualization on large-scale proteomics, transcriptomics, and electrophysiological datasets.

SKILLS

PROGRAMMING & DATA SCIENCE

- Languages: Python (NumPy, Pandas, scikit-learn, TensorFlow, XGBoost, LightGBM) | R (tidyverse, tidymodels, Seurat, Shiny, ggplot2, caret) | SQL
- ML & Statistics: Predictive modeling, classification, time-series analysis, statistical testing, cross-validation frameworks, NLP/sentiment analysis
- Infrastructure: Git/GitHub, REST APIs, Jupyter, Linux/Ubuntu
- Visualization: R Shiny dashboards, ggplot2, matplotlib, seaborn, scientific documentation

BIOINFORMATICS & OMICS

- Multi-omics integration: proteomics, single-cell transcriptomics (scRNA-seq), metabolomics
- Orbitrap/MaxQuant / label-free quantification / differential protein expression
- Biomarker discovery pipelines (DEP, limma, feature selection)
- Automated analysis pipelines — reproducible workflows in R/Python

NEUROSCIENCE & EXPERIMENTAL

- Whole-cell patch-clamp electrophysiology (Axopatch 200B, pCLAMP, HEKA) — paired recordings in mouse brain slices
- Transgenic mouse models | surgical procedures (cranial window, AAV injections)

- Protein biochemistry: co-IP, affinity pull-down, SNARE complex, Western blot
 - Confocal and electron microscopy · Molecular biology: PCR/RT-PCR, cloning, sequencing
-

WORK EXPERIENCE

Senior Data Scientist — Biomedical Pipelines | University of Luxembourg (LCSB) | 2023 – 2025

Built and maintained proteomics analysis pipelines in R/Python for label-free quantification, differential protein expression, and biomarker discovery in neurodegenerative disease (Orbitrap/MaxQuant output).

- 70% reduction in data cleaning time via automated R scripts for omics data validation and transformation
- ML models (classification, regression) for disease progression prediction — statistical validation via cross-validation frameworks
- SOP documentation for reproducible analytical workflows
- EU-funded project collaboration across international multidisciplinary teams

Senior Neuroscientist & Data Scientist | Instituto de Biomedicina de Sevilla (IBiS) | 2017 – 2023

Led data analysis for proteomics and transcriptomics studies on neurodegeneration mechanisms; designed protein interaction experiments (co-IP, pull-down, SNARE complex).

- 90% reduction in processing time via automated electrophysiology analysis pipeline (Python/R + pCLAMP integration)
- Whole-cell patch-clamp paired recordings in mouse brain slices — large-scale electrophysiological datasets
- First-author publication in *Science Advances* (2025); contributed to 8+ peer-reviewed publications
- Supervised PhD candidates; trained teams in data analysis methods

PhD Researcher — Neuroscience & Quantitative Data Analysis | University of Barcelona | 2009 – 2017

Doctoral thesis on phototransduction retinal complex — protein–protein interactions (co-IP, affinity pull-down) combined with quantitative electrophysiology and imaging analysis.

- Designed experimental workflows integrating wet-lab biochemistry with quantitative data analysis
 - Transgenic mouse colony management and surgical protocols
 - PhD in Biomedicine (2016); MSc in Neurosciences (2011)
-

PUBLICATIONS AND CONFERENCES

- **Co-authored 8+ peer-reviewed publications** in high-impact biomedical journals (ORCID: orcid.org/0000-0001-8809-8919)
 - Presented research findings at international conferences (10+ oral and poster presentations)
-

EDUCATION

- **MSc in Big Data and Data Science** — UNED, Spain (2025). Specialized in ML pipelines, statistical modeling, and large-scale data analysis applied to biomedical research.
- **PhD in Biomedicine** — University of Barcelona, Spain (2016)
- **MSc in Neurosciences** — University of Barcelona, Spain (2011)
- **BSc in Biochemistry** — University of Seville, Spain (2009)