

# Julio Gamboa

Preferred name: Julen

## Data Science Skills

- **Programming languages:** R/Tidyverse, Python, GNU Bash scripting, Julia, Perl.
- **Relational databases:** SQL/MySQL.
- **Version control:** Git, GitHub, GitLab.
- **Reporting and dashboards:** R Shiny, Flexdashboard, Quarto, RMarkdown, Tableau.
- **Statistical and Machine learning approaches:** R (mlr, xgboost, stan), Python (scikit-learn, xgboost, keras, tensorflow, huggingface).

[j.a.r.gamboa@gmail.com](mailto:j.a.r.gamboa@gmail.com)

(+44) 7584 122253

(+1) 979 6766610

Llangollen, UK\*

**Github/Portfolio:**

[evoclock \(JGamboa\)](#) ·

[GitHub](#)

**Personal website:**

<https://evoclock.github.io/>

[jgamboa.github.io/](https://jgamboa.github.io/)

*\*Willing to relocate*

---

## Essential skills

My top essential skills are: **Project management, emotional intelligence, empathy, public speaking, creativity and problem solving.**

I have demonstrated these through teaching. I have handled complex situations affecting academic performance due to personal circumstances of 30+ students per semester. My approach has produced consistent positive student feedback.

I have led my own research projects at Texas A&M University with minimal supervision. I have successfully managed to work around RAM/GPU constraints when performing clustering and classification of an extremely large dataset by formulating a three stage clustering pipeline that integrates approaches from different disciplines (CD-HIT, BIRCH, and HDBSCAN), saving both time and resources. I have also published and presented my research at national and international conferences.

---

## Work Experience

**Texas A&M University, USA** — Ph.D researcher Aramayo Lab | January 2020 - Present

- I have been assessing how the informational content (annotated protein domains) of animal proteomes (the library of proteins of an organism) evolves through a novel use of bioinformatics and machine learning methods.
- I have analysed human genomic variation data (>10TB of data) to classify these mutations using a combination of characterisation, clustering, molecular modelling, and annotation analyses to answer to what extent differences in where different populations have evolved happen to have promoted significant adaptive changes in the function and architecture of different families of proteins.

**Texas A&M University, USA** — Ph.D researcher Blackmon Lab | August 2018 - 2019

- I was the second author on an amphibian karyotype database developed and deployed for the Blackmon Lab. Available there: <https://evobir.shinyapps.io/AmphibianDB/>
- I performed extensive mathematical modelling to explore the effect that different population genetic phenomena can have on populations with skewed sex ratios of mating individuals, for example populations of red deer. The results are available here: <https://evoclock.shinyapps.io/osrbias/>

**Texas A&M University, USA** — Selected teaching experience | August 2018 - 2023

- **Computational Genomics (BIOL350)/Genomics(BIOL650)**
- **Digital Biology (BIOL647).**
- **Critical Writing in Biology (BIOL401)**
- **Communicating Biological Research to the Public (BIOL402)**

---

## EDUCATION

**Texas A&M University, College Station, Texas, USA | 2018 - Present**

Ph.D in Biology (in progress)

Main focus: Comparative genomics analysis of clinal evolution of circadian proteins.

**University College London, London, UK | 2013-2017**

MSci Cell and Developmental Biology - 2:1

Thesis: Altered circadian control of cell cycle progression in the context of wound healing *in vitro*.

---

**References are available upon request**