# Julio Gamboa

I am a computational biologist by training with expertise in a wide range of tools and languages such as R, Python, and GNU Bash, as well as statistical and machine learning methods such as XGBoost and HDBSCAN. What truly motivates me is thinking deeply about the underlying processes that generate any dataset I work with. I believe in the power of interdisciplinary knowledge, as it allows me to approach challenges from different perspectives and find innovative solutions.

i.a.r.gamboa@gmail.com (+44) 7939 386019 (+1) 979 6766610 Llangollen, United Kingdom https://github.com/evoclock

#### **EXPERIENCE**

#### **Texas A&M University, USA** — Ph.D student Aramayo Lab

January 2020 - Present

- Analysis of informational content in eukaryotic proteomes from an evolutionary perspective.
  Manuscript in progress.
- Analysis of human variation contained in the gnomAD v3.1.2 database (>2TB of compressed data) focused on clustering of variants using a combination of clustering algorithms
   (CD-HIT/BIRCH/HDBSCAN). My aim is to assess whether latitudinal changes (clinal variation) over the evolutionary history of *H. sapiens* has led to differential patterns of protein function.
- Tools: R/Tidyverse, GNU Bash, AWK, SED, regex, Perl, org-mode, emacs, Python, Git, GitHub, SQL.

#### **Texas A&M University, USA** — Ph.D student Blackmon Lab

August 2018 - 2019

- Second author on amphibian karyotype database developed and deployed for the Blackmon Lab.. https://evobir.shinyapps.io/AmphibianDB/
- Forward time population genetic modelling of biased sex ratios in populations experiencing different selection coefficients, allelic dominance relationships, population sizes, and sex determination mechanisms. https://evoclock.shinyapps.jo/osrbjas/
- **Tools**: R/Tidyverse, Shiny.

#### **Texas A&M University, USA** — Selected teaching and mentoring experience

August 2018 - 2023

- Computational Genomics (BIOL350)/Genomics(BIOL650) 3 semesters Introductory course for applied genome assembly and comparative genomics.
- **Digital Biology (BIOL647)** 2 semesters Fast-paced graduate level course that takes the principles learned in BIOL650 applied to bioinformatic analyses using the command line.

- **Critical Writing in Biology (BIOL401)** 1 semester This course aims to teach the student how to prepare and write a manuscript for publication in scientific journals.
- Communicating Biological Research to the Public (BIOL402) 1 semester This course guides students to take high level scientific writing and transform it into content for general audiences.

## Volunteering

### Texas A&M University, College Station, Texas, USA

Open Source for Open Science Workshop 2020 - Volunteer - Generated materials for Carl Hjelmen's presentation.

(https://eeb.tamu.edu/open-source-open-science/open-source-open-science-workshop-2020/)

Open Source for Open Science Workshop 2019 - Volunteer

(https://eeb.tamu.edu/open-source-open-science/open-source-open-science-workshop-2019/)

#### **EDUCATION**

## Texas A&M University, College Station, Texas, USA

Ph.D in Biology (in progress)

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

2018 - Present

#### University College London, London, UK

MSci Cell and Developmental Biology - 2:1

Thesis: Altered circadian control of cell cycle progression in the context of wound healing *in vitro* 2013-2017

#### Skills

- R/Tidyverse, Shiny, Bash, SSH, AWK, SED, Perl, Python, Julia, SQL, LaTeX, Git/GitHub, SLURM
- Descriptive and inferential statistics/Bayesian statistics and mathematical modelling.
- Machine learning models in R and Python (scikit-learn, mlr, and xgboost)

## Publications and posters

R.D. Perkins, J.R. Gamboa, M.M. Jonika, J. Lo, A. Shum, R.H. Adams, H. Blackmon. The Amphibian Karyotype Database. Chromosome Research. https://doi.org/10.1007/s10577-019-09613-1

Poster presentation at Evolution Conference 2019 in Providence, RI, USA.

Poster presentation at STEGG 2019 in College Station, Texas, USA.