Ferran Cardoso Rodriguez

Curriculum Vitae

N4 – London (UK)
+44 07597532973
ferran.cardoso@icr.ac.uk
fercr@tuta.io
github.com/fercrcode
ORCID: 0000-0002-1376-1242

Computational biologist experienced in software development, *omic* and image-based data analysis.
I have published works on characterising the colorectal-cancer phenoscape and on integrating multi-*omic* and clinical data for pan-cancer diagnostics and drug discovery. Driven and inquisitive, I am a firm believer in the FAIR principles and of empowering colleagues, for which I find development and delivery of tools is key.

Current Position

Nov. 2023 – Present

Institute of Cancer Research, Senior Scientific Officer

Digital Pathology, Spatial Multiomics, and Clinical Data.

In my position at the Integrated Pathology Unit, I am exploring novel approaches for integrating these multiple data modalities with the aim of enhancing our understanding of cancer mechanisms and inform downstream integrated diagnostics. Additionally, I seek to develop impactful and open software solutions to empower researchers and colleagues.

Past Experience

Sept. 2019 – Sept. 2023

UCL CI, PhD in Computational Biology

Single-Cell Omic Analysis

Analysed single-cell *omic* profiles of CRC organoids and their TME using established and novel computational approaches in both local and remote environments.
Developed and **Deployed** tools in package repositories, as web-tools or Docker containers.
Visualised results using bespoke and automated interactive report generation.
Disseminated outputs via oral presentations, scientific posters, and scientific research articles.
Collaborated with peers in subjects ranging from drug screening (Ramos Zapatero & Tong *et al.* 23) to CAR T-cell engineering (Michelozzi *et al.* 23).
Community work via public engagement events and BSc. student teaching support.

CRC Stem Cell Polarisation Main Research Project

Characterised **dynamics** regulating stromal and oncogenic **stem cell** polarisation in the context of **colorectal cancer**. Leveraged **remote compute** as well as state-of-the-art **scRNA-seq** analysis workflows including **compositional** and **cell-cell communication** analyses, and **VAE**-based **label transfer** and **integration** with patient cohorts. Outputs in the form of articles, public data share and code repositories, and guided by illustrated **Jupyter** notebooks.
Publication: Cardoso Rodriguez & Qin et al. 2023
Project repository

VRland Method Development

Valley-Ridge score to generate data-driven **Waddington**-like **landscapes** of cellular **plasticity** and **differentiation** from **scRNA-seq** data. In active development to become a cloud-hosted interactive web-app.
Publication: Cardoso Rodriguez & Qin et al. 2023
Project repository

Signalling Knowledge Graphs Method Development

Awarded with the UCL-Yale PhD exchange bursary.
Explored novel methods to characterise inter- and intra-cellular communications using signalling knowledge graphs, including embedding of directed gene networks, and *omic* profile projection.
Assembled and **Embedded** custom signalling KGs leveraging remote GPU compute.
Developed a Python package to compute hierarchy scores on directed graphs.
Disseminated outputs on international conferences.
Collaborated on-site with Smita Krishnaswamy's lab at Yale University.
Blog on conference paper

CyGNAL Analysis Pipeline

Pipeline for the **analysis** and interactive **visualisation** of **mass cytometry** data via PTM signalling and cell-state **classification** via ensemble **ML** methods.
Publication: Sufi & Qin et al. 2021
Project repository

Jan. – Sept. 2019

Imperial College London, Bioinformatics

MSc Computational Projects

Masters degree at Imperial College London uniquely characterised by **three** distinct computational **projects** and reports.

MSc Project 3 JUN. – SEPT. 2019

Genomic Annotation Pipeline

Developed AnnoRE, a **pipeline** for **API-based** querying of databases and downstream **annotation** of **genetic variants** and metabolomics high-throughput data to study **complex trait** diseases.
Project repository

MSc Project 2 APR. – JUN. 2019

scRNAseq data analysis

Analysis of droplet-based **scRNA-seq** datasets to characterise **cardiac** stem populations and **development**.
Web report

MSc Project 1 JAN. – APR. 2019

Flux-Balance app development
Group project developing MetEOr, a Flask-based **web-app** for visualising and performing **Flux Balance Analysis** on whole-organism **metabolic models**. **Back-end** developer responsible for the FBA logic and codebase.
Project repository

Education

2019 – 2023 **PhD Computational Biology**
University College London
PhD programme at Dr. Chris Tape's lab (UCL Cancer Institute).
Viva passed w/o corrections

2022 **UCL-Yale Travel Award**
Collaborative exchange at Prof. Smita Krishnaswamy's lab (YSM/SEAS Yale University)

2018 – 2019 **MSc Bioinformatics and Theoretical Systems Biology**
Imperial College London
Multi-project programme. Merit

2014 – 2018 **BSc Biotechnology**
Universitat de Barcelona
Molecular Biotechnology specialisation.
1st Class (8.7/10)
Erasmus award for thesis on immune infiltrate of craniopharyingiomas (Queen Mary University).
Visiting student at Núria Montserrat's group (i)Pluripotency for organ regeneration (IBEC).

Selected Outputs

Publication **Cardoso Rodriguez & Qin *et al.*, 2023**
An oncogenic phenoscape of colonic stem cell polarization
DOI: Cell

Package **FerranC96/pyKrack**
Computing Krackhardt Hierarchy Score on Directed Graphs
PyPI: pykrack

Pipeline **TAPE-Lab/CyGNAL**
CyTOF Signalling Analysis (CyGNAL) pipeline
DOI: Zenodo | Nat. Protocols

Tool **FerranC96/VRland**
Valley-Ridge Score for Single-Cell Waddington-like Landscapes
DOI: BioRxiv

Skills

Community **Collaboration** both within and outside research group.
Volunteering at public and patient engagement events.
Teaching tutorials for BSc students and ESL teaching experience.
Public speaking at international conferences and events.
FAIR and **FOSS** advocate.

Technical **Languages**: Proficiency in Python and R. LaTeX, Markdown, web technologies.
Reporting: Publication-grade figures, interactive visualisations, web-apps.
Workflows: Remote computing (CPU and GPU-accelerated) in HPC, nextflow pipelines, containerisation.

Analysis **Pipelines**: BCL to FASTQ, custom transcriptome references, sequence alignment.
sc-*omic* data: Dim. reduction, clustering, dif. expression, compositional analysis, cell-cell communications, cellular dynamics, data integration.
ML applications: Knowledge-graph embedding, graph signal processing, classification models, PyTorch.

Dev. **Team Projects**: Version control and collaboration through git, kanban-based tools.
Build and deployment: Multi-language tools, interactive reports, PyPI and conda packaging, nbdev for notebook-centric development, container deployment.