JACK GISBY

 github.com/jackgisby   jackgisby.com   j.gisby20@imperial.ac.uk

 London, United Kingdom   (+44)7748 559964

RESEARCH INTERESTS

Data scientist and computational biologist interested in data-driven research and the -Omics technologies. My PhD work involves integrating genomic, transcriptomic and proteomic datasets to identify molecular pathways that play a driving role in inflammatory disease. I am also interested in developing data processing and mining methods to facilitate the extraction of biological insights from -Omics datasets.

EDUCATION

Imperial College London   *July 2020 - August 2023*
PhD Candidate - Department of Immunology and Inflammation
· Thesis: Using Multi-Omics to Understand Inflammatory Disease
· Supervised by Dr James Peters

University of Birmingham   *September 2016 - June 2020*
Biochemistry with Professional Placement (MSci) - First Class
· Dissertation: A Prior Knowledge-Based Computational Workflow for *de novo* Structural Elucidation of Small Molecules in Mass Spectrometry Metabolomics

PUBLICATIONS

Longitudinal proteomic profiling of dialysis patients with COVID-19 reveals markers of severity and predictors of death
First author   *eLife* 2021 - doi:10.7554/eLife.64827
· Lead the analysis of a high-dimensional Olink proteomics dataset with a complex cohort consisting of repeated measurements at inconsistent time points.
· Applied joint models and linear mixed models to identify predictors of death and key proteins that changed over time following COVID-19 symptom onset. These included KRT19, a marker of epithelial injury, and ACE2, the cell receptor for SARS-CoV-2.
· Utilised supervised learning algorithms to identify biomarkers of severe disease.

Mendelian randomisation identifies alternative splicing of the FAS death receptor as a mediator of severe COVID-19
Second author (equal contribution)   *medRxiv* 2021 - doi:10.1101/2021.04.01.21254789
· Developed a parallelised pipeline for two sample Mendelian Randomisation and subsequently investigated COVID-19 susceptibility and severity using Olink & SOMA pQTLs.
· Used fine mapping and colocalisation tools in tandem with public data repositories to investigate causal variants of disease.

The widespread nature of Pack-TYPE transposons reveals their importance for plant genome evolution
First author   *bioRxiv* 2021 - doi:10.1101/2021.06.18.448592
· Designed an algorithm for the specific annotation of repetitive elements that capture chromosomal DNA; implemented this as an R package available as part of the Bioconductor project.

- Mined open source genetic data to demonstrate the abundance of these elements for multiple superfamilies and genomes. Found that recent insertions of these elements have impacted the evolution of genes.

Plasma Lectin Pathway Complement Proteins in Patients With COVID-19 and Renal Disease
Co-author   *Frontiers in Immunology* 2021 - doi:10.3389/fimmu.2021.671052

CONFERENCES AND PRESENTATIONS

- **Presentation - Biomarkers of the Future 2021** - Longitudinal proteomic profiling of dialysis patients with COVID-19 reveals markers of severity and predictors of death
- **Poster - HUPO Reconnect 2021** - Longitudinal proteomic profiling of dialysis patients with COVID-19 reveals markers of severity and predictors of death
- **Poster - UK-CIC Immunology 2021** - Longitudinal proteomic profiling of dialysis patients with COVID-19 reveals markers of severity and predictors of death
- **Presentation - Longitudinal Studies 2021, Wellcome Genome Campus** - Longitudinal proteomic profiling of dialysis patients with COVID-19 reveals markers of severity and predictors of death

ADDITIONAL RESEARCH EXPERIENCE

A Prior Knowledge-Based Computational Workflow for *de novo* Structural Elucidation of Small Molecules in Mass Spectrometry Metabolomics
Dissertation project, Dr Ralf Weber   *October 2019 - June 2020*
- Developed and optimised a Python package, *MetaboBlend*, to propose candidate molecules using tandem mass spectra.
- Processed large, high-dimensional mass-spectrometry datasets for the validation of *MetaboBlend*.
- Used Python, SQLite and Bash to implement the *MetaboBlend* workflow on a Linux high-performance computing cluster for the generation of large structure databases.
- Created a "metabolite-likeness" model based on a deep learning implementation of the one-class support vector data description algorithm, adapted to model metabolites with a transformer architecture.

Development of an Immunoturbidimetric Assay for Serum Amyloid A for an Automated Clinical Analyser
Industrial Placement, The Binding Site   *August 2018 - August 2019*
- Developed assays for inflammatory biomarkers, such as Serum Amyloid A, for use in clinical laboratories.
- Interacted with, and regularly delivered presentations to, researchers from diverse fields and non-scientific staff to ensure the timely completion of my development project.
- Took responsibility for planning and carrying out experiments and statistical analyses in a research environment.

TEACHING EXPERIENCE

- **Lecture:** RNA-seq in Molecular Epidemiology - Developed a two-part lecture explaining the basics of RNA sequencing and its application to investigate genetic variants. Set homework and final exam questions for the module.
- **Supervision:** Co-supervisor of a medical undergraduate - On-going project using Olink proteomics to investigate Lupus.
- **Volunteer:** Code Club - Developed my teaching abilities and provided help with coding skills to 9-13 year-olds as part of the Code Club voluntary initiative.

SPECIFIC COMPUTATIONAL SKILLS

Operating Systems	Windows, Ubuntu
Programming Languages	Python, R, SQLite, Bash, working knowledge of C++ & Julia
Other Tools	Git, Linux HPC, Conda, Docker, Nextflow
General Software	MS Office, LaTeX
Data types	Olink & SOMA proteomics, RNA-seq, MS metabolomics, whole genome sequences, GWAS
Bioinformatics/Statistics	Package development in R and Python, Linear and mixed models, Mendelian randomisation & colocalisation, Joint models, Supervised learning (scikit-learn, tensorflow, pytorch), WGCNA

ADDITIONAL COURSES

· "From large datasets to biological insight" by Wellcome Connecting Science & EMBL-EBI

· "Introduction to Assessment and Feedback" by Imperial College London

· "Introduction to Teaching and Learning" by Imperial College London

· "Profiling and optimisation in Python" by Imperial College London

· "Deep Learning in the Life Sciences" by MITx

· "Software Development with C++" by University of Birmingham Research Computing

· "The Biostars Handbook: Bioinformatics Data Analysis" by Istvan Albert

· "Introduction to Computer Science and Programming Using Python" by MITx

· "Introduction to Computational Thinking and Data Science" by MITx

· "Data Analysis for Life Sciences and Genomics Data Analysis" by HarvardX

REFERENCES

Dr James Peters (PhD Supervisor)
Clinical Reader in Rheumatology
Faculty of Medicine
Imperial College London
Hammersmith
W12 0NN
j.peters@imperial.ac.uk

Dr Marco Catoni (Undergraduate Project)
Lecturer in Plant Biology
School of Biosciences
University of Birmingham
Edgbaston
B15 2TT
m.catoni@bham.ac.uk