Dr Emeline Favreau
emeline.a.favreau@gmail.com
https://emelinefavreau.github.io/

Professional Summary
Ten-year experience as **Genomic Researcher** with a strong background in **Rare Disease** genomics, large-scale genomic data analysis and evolutionary biology. Proven success in **cross-disciplinary collaborations** with clinicians, statisticians, and lab scientists. Skilled in **short-read** and **long-read whole genome** and transcriptome analysis, **short and structural variant calling**, and multi-omics integration. Adept in using cluster computing, containers, and **Nextflow** workflow languages for reproducible, high-throughput analyses in human genomics. Manages projects with flexibility to re-prioritise when needed.

Current Position
Bioinformatician in Rare Diseases — Wallace Lab, University of Cambridge
May 2023 – Present
The sole port of call for leading and supporting genomic analyses across 6 interdisciplinary research groups focused on the rare disease group of Inborn Errors of Immunity (IEI). Consultant for >10 projects from analysis design, pipeline development, data analysis, manuscript preparation. Key projects which benefited from expert project management:

- **Rare disease variant discovery**: Remapped >13,000 NIHR Bioresource genomes to GRCh38; performed small and structural variant calling for > 6,000 genomes, with Dr. Ernest Turro (Mount Sinai, USA); investigated mode of inheritance and *de novo* variants in proband-parents trios; produced reports for and participated in monthly MDTs with clinicians and researchers (INTREPID project).
- **Cancer Immune-related adverse effects analysis**: Variant annotation, ancestry analysis, and PRS on a small cohort of patients presenting a variable set of cancers, for an Australian clinical team associated with Prof. Matthew Cook's lab (Cambridge, UK).
- **Immunology statistical analysis**: Statistical analysis of clinical data (B cell data, immunoglobulin, genetic variants, Human Phenotype Ontology terms) from >2,000 IEI patients, with Prof. Chris Wallace (Cambridge, UK) and Prof. Siobhan Burns (UCL, UK).
- **Benchmark and pipeline design**: Benchmarking of tools for the remapping project; updated and deployed R-based batch normalisation workflows on HPC cluster for PhD students and postdocs in Prof. Matthew Cook's lab.
- **Data infrastructure**: Designed and implemented a secure archiving solution for genomic and clinical data (> 2 PB of data including flow cytometry, WGS, RNAseq); supporting R&D, ensuring data harmonization and reproducibility for all 6 research groups; extensive documentation.

Previous Experience

Postdoctoral Researcher (NERC Fellow) — Sumner Lab, UCL

Oct 2019 – Apr 2023

- Conducted large-scale comparative genomics, transcriptomics and population genetics on social insects (bees and wasps).
- Developed bioinformatic pipelines for RNA-seq, genome assembly (long and short reads), annotation, and machine learning for phenotype prediction.
- Defined the lab's data management strategy for long-term genomic and transcriptomic data preservation.

Technical Skills

Languages & Tools:

- Bash, R, SQL, Python.
- HPC scheduling, notions of AWS.
- Nextflow sarek, rnaseq, pgsc_calc.
- Docker, Singularity, conda, git.

Genomic Analysis:

- Whole genome variant calling (SNPs, SVs).
- WGS Annotation focusing on rare disease genetics with ClinVar, VEP, gnomAD
- Hypothesis-driven analyses: ancestry analysis, gene enrichment analysis, GWAS, PRS.

Transcriptomics & Multi-Omics:

- RNA-seq QC, differential expression (DESeq2), co-expression networks (WGCNA).
- Multi-omics integration across transcriptomics and genomics.
- Long-read experience, sequencing and assembling *de novo* ant genome on MinION.

Reproducibility & Documentation:

- Git/GitHub, RMarkdown, workflow management, scientific writing and publication.
- Data stewardship and FAIR principles for patient/clinical data; expert in dealing with >10-year old data sets with multiple identifiers.

Selected Publications

1. April 2025: Poster at **Genomics of Rare Disease (Hinxton, UK)**. *Feasibility, costs and benefits of remapping genome sequences from GRCh37 to GRCh38 in an Inborn Errors of Immunity cohort*.
2. Favreau, E., Geist, K.S., Wyatt, C.D.R., Toth, A.L., Sumner, S., and Rehan, S.M. (2022) Co-expression gene networks and machine-learning algorithms unveil a core genetic toolkit for reproductive division of labour in rudimentary insect societies. Genome biology and evolution, evac174. DOI: https://doi.org/10.1093/gbe/evac174
3. Favreau, E., Lebas, C., Stolle, E., Priyam, A., Pracana, R., Aron, S., and Wurm, Y. (2022) No supergene despite social polymorphism in the big-headed ant *Pheidole pallidula*. bioRxiv. DOI: https://doi.org/10.1101/2022.12.06.519286

Education

PhD in Evolutionary Genomics (NERC) — Dr. Yannick Wurm & Dr. Max Reuter, Queen Mary University of London
2015 – 2019
Research on ant genome evolution and sociality. Field work in 3 continents. Sequenced a *de novo* genome for an ant species with long-reads (MinION) and short-reads (Illumina). Population genetics and GWAS to investigate the basis of one-queen / multiple-queen system.

MRes Biosystematics, Imperial College London & Natural History Museum
2013 – 2014
Evolutionary biology, Insect taxonomy, Phylogenetics. Learnt the fundamentals of gold-standard documentation practice for long-term storage.

BSc Environmental Sciences (First Class Hons), Oxford Brookes University
2008 – 2011