Dylan Lawless

Contact

School of Life Sciences, Global Health Institute,
École Polytechnique Fédérale de Lausanne, Switzerland.
Dylan.Lawless@epfl.ch
LawlessGenomics.com
↻ Online version

Education

2018—current	Postdoctoral scientist EPFL. "Host-pathogen immunology and translational genomics."
2015—2019	PhD University of Leeds and St. James's University Hospital. "Novel genetic discoveries in rare primary immunodeficiencies."
2013—2014	MSc Trinity College Dublin (1st class). "Exploring the therapeutic potential of a peptide derived from a poxviral immune evasion protein."
2009—2013	BSc University College Cork (Honours).

Research

2018—current	École Polytechnique Fédérale de Lausanne, Switzerland, Translational genomics, Fellay lab.
2015—2018	University of Leeds, UK, Genetic discovery in rare PID, Savic lab.
2014—2015	ACM global. Analytical Scientist, clinical trials.
2014	École Polytechnique Fédérale de Lausanne, Switzerland, Intracellular Innate Immunity, Ablasser lab.
2013	Trinity College Dublin, Ireland, Viral Immunology, Bowie lab.
2012	University College Cork, Ireland, Innate immunity in microbiology, Morgan lab.

Expertise

- **Independent research**: State of the art methods in genomic analysis, statistics, study design, team coordination, project management, presentation, scientific writing, deductive reasoning, critical thinking, and multidisciplinary flexibility. Leading a number of international research projects. Project supervision of 5 MSc thesis and PhD, and teaching on numerous courses.
- **Bioinformatics**: Responsible for human bioinformatic experiments and pipelines in PhD lab 2015-2019 and in PostDoc lab 2018-current. ● Variant priorisation and variant effect prediction ● WGS, exome, SNV, CNV, RNAseq, Germline, Somatic mutation, Cancer, Tumor ● Handling genetic data formats FASTQ, BAM, CRAM, SAM, gVCF, ASN.1, json, plink, etc. ● GATK best practices, Popular genomic database and reference panels, Popular variant prediction algorithm/databases, Annotation of variants with custom databases, VEP, dbNSFP, ANNOVAR, etc ● ACMG standards and guidelines for interpretation of variants, Nomenclature standards incl. HGVS, HGNC ● Large-scale NGS initiatives, National and international, Privacy in personal health, Big data analyses, Population genomics, Pedigree analysis, Clinical genetics. ● Pathogen genetics, Viral, Bacterial, Phylogenetics and epidemiology.

- **Statistics**: Association testing, Mixed models, Inference, Variance, Correlation, Regression, Repeated data testing, Sensitivity and specificity analysis, Gene burden testing (extensive use and design), Sequence kernel association testing, GWAS stats, Pathogen variation stats, Dimension reduction, PCA, SVD, Population structure and stats, Genome to genome assoc statistics, Prediction modeling with multimodal data for pre-emptive intervention; Model design, Statistical automation and reproducibility, Data summary, Statistical programming in R.

- **Immunology**: Innate and adaptive; Primary immunodeficiencies and autoinflammatory disease; Functional and computational immunology; Host-pathogen interaction; Clinical interpretation; Immunogenomics and pharmacogenomics.

- **Programming**: R, bash, a wide range of Unix/Linux command-line tools, git version control, server-side editing, web-hosting, html, CSS, Java, Ruby, active interest in learning other popular tools; Go, Julia, SQL etc. High performance computing scheduling, learning scalability with cloud-genomics; AWS, Azure, WDL, Cromwell, Dragon.

- **Data visualisation**: Presentation of data to both scientific and lay audiences: Genomic analysis gallery.

- **Democratized genomics**: I work on additional research problems to provide analysis where it is not available from national services, via LawlessGenomics (non-diagnostic for referral to clinical genetics). I am also building both the front-end and back-end on AWS for my GATK WGS pipeline including interactive evidence report with ∼150 database scores (e.g. gnomAD, ClinVar, ClinGen, UniProt, GO, OMIM, pathogenicity predictors, ACMG reporting standards, etc.)

- **Cell biology**: Strong background in wet-lab experimentation and design. ● *In vitro* and *in vivo* models, cell lines from human, animal, cancer, primary cell models, PBMC purification and culture, murine models, animal handling, microbiology, virology. ● Cancer immunology ● A wide range of molecular assays; protein expression, ELISA, western blot, histology, immunohistochemistry, immunofluorescence, flow cytometry, FACS, immune biomarkers, cloning, plasmid construction, protein expression, mutagenesis, virology. ● DNA/RNA/cDNA, purification from blood, hair follicle, saliva/buccol, FFPE, tumor, RT-PCR, RT-qPCR , Sanger seq and primer design. ● Capture library design, manual library preparation of genomic DNA, massively parallel somatic sequencing, methylation assays, multiple sequencing platforms, clinical cohort design, gene panel design.

Scientific Outreach

- **Day-to-day**: ● LawlessGenomics.com - A home for topics that are relevant to human precision medicine and genomics. Public invitation for custom genomic analysis on rare disease. ● Genomic analysis gallery - Code and usage of data visualisation for genomic analysis. ● SARSCoV2variants.com - Open-source tracking for emerging SARS-CoV-2 variants that pose a risk based on COVID-19 vaccine genetics. ● Genomic tools - Modular snippets of novel genomic analysis methods. ● Promotional videos for EPFL Global Health Institute to attract international talent - GHI homepage. ● Several links featuring popular work are shown in the highlighted media homepage section.

- **Conferences**: Extended list of conference invitations and participation online.

- **Membership**: European Society of Human Genetics, Swiss Institute of Bioinformatics.

Supervision and Teaching

● 2013-present: extended list of teaching/tutoring roles online. ● 2015-present: supervision of staff and students, 5 MSc and several BSc student projects, MSc and BSc coursework, EPFL (Switzerland) and UoL (UK).

Awards

● **2019** The European Society of Human Genetics Poster. ● **2018** Microsoft Azure Research Award: Data Science and Machine Learning in Predictive Genomics. ● **2017** Wellcome Genome Campus, Cambridge Visitor Grant. ● **2015** University of Leeds 110 Anniversary Postgraduate Research Scholarship. ● **2014** Trinity College Dublin 1st place postgraduate poster prize.

Publications

First-author

- *Blood*. Jun 2020; doi: 10.1182/blood.2020005844. Germline *TET2* loss-of-function causes childhood immunodeficiency and lymphoma. - pdf
- *Journal of Clinical Immunology*. 2019 Aug; doi: 10.1007%2Fs10875-019-00670-z Predicting the occurrence of variants in *RAG1* and *RAG2*. - pdf
- *Frontiers in Immunology*. Jul 2018; doi: 10.3389/fimmu.2018.01527; A case of AOSD caused by a novel splicing mutation in *TNFAIP3* successfully treated with tocilizumab. - pdf
- *Journal of Allergy and Clinical Immunology*. Feb 2018; doi: 10.1016/ j.jaci.2018.02.007; Prevalence and clinical challenges among adult PID patients with recombination-activating gene deficiency. - pdf
- *Journal of Clinical Immunology*. Oct 2017; doi: 10.1007/s10875-017-0427-1; Bialellic Mutations in Tetratricopeptide Repeat Domain 7A (TTC7A) Cause Common Variable Immunodeficiency-Like Phenotype with Enteropathy. - pdf

Co-author

- *eLife*. Dec 2021; doi: 10.7554/elife.72559. Biallelic mutations in calcium release activated channel regulator 2A (CRACR2A) cause a primary immunodeficiency disorder - pdf
- *Scientific reports*. Feb 2021; doi: 10.1038/s41598-021-84070-7 The influence of human genetic variation on Epstein-Barr virus sequence diversity. - pdf
- *Arthritis & Rheumatology*. Sep 2020; doi: 10.1002/art.41531 A novel RELA truncating mutation in familial Behçet's Disease-like mucocutaneous ulcerative condition. - pdf
- *Journal of clinical immunology*. Dec 2019; doi: 10.1007/s10875-019-00735-z Expanding Clinical Phenotype and Novel Insights into the Pathogenesis of ICOS Deficiency. - pdf
- *Blood*. Dec 2018; doi: 10.1182/blood-2018-07-866939 A novel *RAG1* mutation reveals a critical in vivo role for HMGB1/2 during V(D)J recombination. - pdf
- *Science Translational Medicine*. 2016 Mar 30;8(332):332ra45. doi: 10.1126/scitranslmed.aaf1471; Familial autoinflammation with neutrophilic dermatosis reveals a regulatory mechanism of pyrin activation. - pdf

Pre-prints

- Genome-wide association study of pediatric sepsis. link
- Rare variants in antiviral response genes permit severe lower respiratory tract infection in children. link
- Viral genetic determinants of persistent humanorthopneumovirus infection. link
- Genome-wide association study of susceptibility to respiratory syncytial virus infection during infancy.

- Incomplete recovery of Zebrafish retina following cryoinjury.
- Prevalence of *CFTR* variants in PID patients with bronchiectasis - an important modifying co-factor.
- A mitochondrial mutational signature of temperature in ectothermic and endothermic vertebrates. biorxiv

Supplemental - Interests and experience

Tools and databases

I enjoy documenting and curating datasets from a large number of the most popular genomic resources for variant annotation and interpretation. For example, a well curated systems make it simple to reference the tools that I use most often:

General genomics: Ensembl, RefSeq, HGNC, The Human Protein Atlas, GTEx, GHR Genes, UniProt, BioCarta, GNF/Atlas, Gene Ontology, KEGG, BioMart, BioCarta, GNF/Atlas, Gene Ontology, BioMart, KGGSeq, HaploReg, **Annotation resources**: dbNSFP, ANNOVAR, UCSC hgVai, Ensembl VEP, SnpEff, SnpSift **Genotype analysis**: PLINK, KING, SHAPEIT2 : IMPUTE2, GCTA, LocusZoom, LDLink, FUMA, imputation; Sanger, MACH, TOPMed. **Population/conservation data**: Exac, gnomAD, PHAST, Mitomap, fitCons, phastCons, kaviar3, dbSNP, dbVar, EVS, NHLBI ESP6500, DGV, Bravo, FinnGen, 1000 Genomes Project (1KG), UKBB. **Drug-gene**: CPIC Genes-Drugs, AACT, EMA Approved Drugs, FDA Approved Drugs, FDA Pharmacogenomics, PharmGKB, DGI. **Cancer**: TP53, ICGC somatic, CancerHotspots, OncoTree, cBioPortal, GDC, Cancer Gene Census, COSMIC, FusionGDB, CIViC. **Association**: ClinVar, ClinGen, GWAS Catalog, gene2phenotype, GenCC, Mastermind, GE PanelApp, CKB, Mondo, CGD, HPO, DailyMed, DECIPHER, PMKB, Consensus, HGMD. **Viral genetics**: asn2fsa, clc novo assemble, ClustalO, GenBank, IQ-Tree, MAFFT, NextClade CLI, Tbl2asn, VIGOR, **Protein structure**: RCSB PDB, alphafold, RoseTTA, FoldX, DynaMut. **Prediction**: GERP, DANN SNVs, Domino, CADD, dbscSNV, Essential Genes, GDI, GHIS, HIPred, LoFTool, P(HI) Score, P(rec) Score, RVIS, ALoFT, BayesDel, DEOGEN2, Eigen, FATHMM, FATHMM, HIPred, KEGG, LIST-S2, LRT, LoFTool, M-CAP, MPC, MVP, MetaLR,SVM,RNN, MutPred, MutationAssessor, MutationTaster, PROVEAN, PrimateAI, REVEL, SIFT, SiPhy, bStatistic, Polyphen-2, EVE.

Data visualisation

Communication of actionable and interpretable results is critical for success. I spend a significant proportion of my time on both concepts and execution. A selection of published work can be seen on my Genomic analysis gallery.

The data usage tracking networks demo gathers every instance of data usage for long-term tracking, comparing individual and cross-project usage. It allows you to see what is heavily reused, unused, and potentially identifying new secondary uses.

While bioinformatic evidence is often sufficient, a simplified demonstration can be more interpretable and persuasive. For example, a dominant variant was visualised by confirmatory Sanger sequencing, western blotting and demonstration of alternative splicing. The same analysis also used a range of bioinformatic evidence databases, population genetics, and pathogenic variant prediction.

Democratised genomics

Genetic evidence builder is a static demo version of a free clinical genome analysis platform that does not retain data or track users. Identifying candidate causal variants uses a large number of databases which are inaccessible to non-bioinformaticians. I make the results and evidence available and interpretable to non-experts. The live version is being built with modular components hosted on AWS with my GATK WGS pipeline to handle large-scale data requests and report reputable evidence sources, confidence, and according to standardised guidelines.

Robust statistical genomic analysis requires willingness to present data and evidence; the attached link is an example of real-time sharing for our collaborators that is currently in progress for my latest pre-print. This contains preliminary results and links to code and data repository and manuscript.

"Statistics on the table, please" - Pearson, 1910. The complexity of genomic data means that it is difficult to *succinctly* present the chain of events that produced evidence. Raw datasets are extremely large, reference and annotation datasets are even larger, processing involves heavy computation, and statistical analysis might be complex or involve private data. Most lay users are interested but often favour answers that are *"Yes or No"* instead of *confidence intervals* and *Bayesian priors*. I spend a great amount of time learning how to present results where every step is accounted for, all data is accessible for replication, and simultaneously provide a satisfactory summary. Therefore, I learn about database management, web development for mobile and desktop, responsive design, live data access, and interactive data viz/statistics to fulfil this need.

National-scale genomics

The majority of my work has been for national-scale analysis. I have worked on Genomics England, specifically in NIHR BioResource Rare Disease with two early papers. This included national-scale analysis of all registered immunodeficiency patients in the UK (presented at ESID 2016), CFTR registry database, and WGS analysis of multiple disorders.

For the last three years I have been working on several muti-ethnic population-based studies, including the SNF Swiss Pediatric Sepsis Study (SPSS), The infant susceptibility to pulmonary infections and asthma following RSV exposure in infancy birth cohort (INSPIRE), and Environmental Influences on Child Health Outcomes (ECHO), via Vanderbilt School of Medicine.

I also lead several pathogen sequencing and epidemiology projects, mainly for RSV and SARS-CoV-2, variant association analysis, pathogenicity, pre-emptive mutation screening, rapid tracking. For example, my primer analysis tool [demo output] aligns all (~15,000) public RSV sequences collected from 1956-2021 and checks that sequencing primers capture all known isolates.

Emerging genomic tech

For the last four years I have been developing statistically robust novel methods in protein-pathway network analysis for discovery of genetic determinants of rare diseases. The methods detect damaging variants that are easily characterised, as well as non-biased detection for variants-of-unknown-significance which would normally be uninterpretable according to current analysis standards. This work has included contributions from students under my supervision, including three MSc thesis projects and one PhD candidate project. Datasets include ~400 patients with rare disease, WGS, exome, GWAS, clinical data, human, bacterial, and viral genetics.

Simultaneously, I have been working on host-pathogen interactions with newly designed analysis methods and using natural quasi random infection by RSV, a major burden on global health, and other similar pathogens. This involves long-term surveillance of ~2000 infants in USA and Europe. We use modern methods for simultaneous analysis of exome, GWAS, RT-qPCR, serology, inflammatory immune biomakers, clinical data, viral seq data. The results been very successful and are being expanded with additional cohorts.

Over the last decade, bioinformatic and statistical methods have developed to expand our understanding of health beyond SNV interpretation. My work in this area has included analysis of VDJ recombination (1, 2, 3) and TCR repertoire (2). I also increasingly encounter minimal residual disease, not necessarily in cancer alone, but for potent dominant loss-of-function (RelA) and gain-of-function variants that can be pre-emptively detected (MEFV, A20). I am interested in consolidating WGS, exome, and somatic pipelines, while merging RNAseq (and other omic data) to provide a single interpretation process and provide practical real-world applications like pre-emptive intervention, somatic risk, ctDNA profiling, PRS, or tissue-specific disease.

I work on keeping up with new technologies. My interests include long read seq, methyl-seq epigenetic markers (I have designed protocols for cheap genome-wide methylation analysis during work on complete germline *TET2* LoF), RNA abundance and differential expression, design of universal library preparations, conversion of germline capture for somatic analysis in clinical cases (and test cases for long-term somatic risk [1, 2]).

I work on machine learning projects such as clinical prediction in adult and pediatric sepsis (EPFL / ETHZ). This merges genetics (exome, genome, protein pathway analysis) and clinical features for patients in ICU to predict survival and organ failure. There is a clinical need for these tools; however, the only feasible implementation of real-time prediction using genetic/biomarker ∼outcome will be through simple, integrated services that are commercial or nationalised.

License, accreditation interest

My interest in practical translational medicine requires learning about the logistics of scalability and legal requirements for diagnostic genetics (e.g. Swiss federal laws under Sec 810.1). I have prepared pipelines that use both commercial and academic database licences and am capable of producing tools that adhere to legal requirements. ISO 15189 is a commonly sought standard accreditation for genetic analysis labs which is carried out by recognised accreditation services. Here in French-speaking Switzerland, the most recently accredited NGS service is Geneva health 2030 genome center for clinical grade sequencing. Other additional ISO accreditation standards concern Genomic information representation, including 23092-4 (for reference software) or 23092 (for transport and storage of genomic information). For example, I teach our students in pharmacogenomic product design to define what might be best for commercial genomics in Switzerland/EU, and refer to BlueprintGenomics as an example with international licensing and accreditation. I follow GA4GH for information about many legal and ethic topics, and have worked on some genomic privacy projects for Swiss Personalized Health Network and Personalized Health and Related Technologies of the ETH Board. I am capable of adhering to quality control management requirements and have interest how to provide resources to the major markets including EU, North-South America, Asia, and emerging markets in Africa.

Clinical genetics

Since 2013 I have been doing primary research partnered with accredited clinical genetic labs and direct interaction with patients. This includes meeting patients, collecting blood or tissue samples, custom gene panel design, functional wet-lab assays (e.g. source), and tailored genomic analysis. I have worked with a number of hospitals world-wide (UK, Ireland, Austria, Germany, Australia, USA) and most frequently with: [1] National health service Leeds NGS labs, [2] University of Leeds NGS research facility, [3] National health service Newcastle NGS labs, and [4] CHUV Unité de médecine de précision.

Scientific writing

I work on producing frequent, good quality first author papers in leading journals - Google scholar. I also have interest in publishing technical and lay scientific writing as listed in my scientific outreach section (CV). My research involves patients with rare disease and as such, I sometimes produce reports for clinical genetic labs, e.g. St. James's University Hospital UK. For project funding, I frequently do grant writing and reviewing in UK and for Swiss national science foundation (SNSF) (e.g. 2018-2022 Postdoc mobility, Sinergia, Ambizione, National Research Programmes, etc.) (mysnf.ch). I enjoy writing greatly and love to learn about technical document writing, publishing, language formats, typographic style, and software; i.e. LaTeX, Markdown, html, css, version control, InDesign, etc.

Lausanne, Jan 2022