

Saul Pierotti

Computational biologist — PhD fellow at EMBL-EBI
Born in 1995 — Italian citizen

Email: saulpierotti@gmail.com
Website: saulpierotti.github.io
GitHub: [saulpierotti](https://github.com/saulpierotti)
ORCID: [0000-0002-5942-0793](https://orcid.org/0000-0002-5942-0793)
LinkedIn: [saulpierotti](https://www.linkedin.com/in/saulpierotti)

RESEARCH EXPERIENCE

European Bioinformatics Institute (EMBL-EBI) University of Cambridge

Cambridge, UK
Oct 2021 — Ongoing

Predoctoral fellow; Supervisor: Prof. Ewan Birney

QTL mapping and exploration of GxE, GxG, and indirect (social) genetic effects in medaka fish.

- Developed and made publicly available 3 Nextflow computational pipelines for genotype imputation, linear mixed model GWAS, and causal variant detection and prioritization (1 published and 1 preprinted)
- Collaborated closely with experimental research groups and delivered a complex GWAS project involving more than 2200 samples that led to the identification of 16 QTLs in medaka fish, 4 of which experimentally validated (manuscript in writing)
- Discovered pervasive gene by environment and genetic interactions in medaka fish, one of which experimentally validated
- Delivered training for the Nextflow workflow language and gene-by-environment interactions to an internal and external scientific audience
- Collaborated internationally with 3 research groups from Germany and Japan
- Led the writing of multiple manuscripts and directly handled interactions with scientific journals from submission to publication
- Presented talks and posters at numerous internal and external events

Science for Life Laboratory (SciLifeLab)

Stockholm, Sweden
Jan 2021 — Jun 2021

Master thesis; Supervisor: Prof. Arne Elofsson

Prediction of protein contact maps from deep mutational scanning data using machine-learning

- Developed a protein contact map prediction method using gradient-boosted trees, neural networks, and other machine-learning approaches
- Achieved equivalent performance to a previously published method but without using protein structural features (which are hard and expensive to obtain)

Max Planck Institute for Evolutionary Biology

Plön, Germany
Jun 2019 — Sep 2019

Intern; Supervisors: Prof. Arne Traulsen, Dr. Jenna Gallie

Experimental evolution of *Escherichia coli* and *Pseudomonas fluorescens* in long-term batch cultures

- Cultured bacteria and detected colony shape and number changes following evolution on different culture media

University of Perugia

Perugia, Italy
Oct 2018 — Oct 2019

Bachelor thesis; Supervisor: Dr. Manlio di Cristina

Characterization of vacuolar proteins in *Toxoplasma gondii*

- Performed PCR, CRISPR, gel electrophoresis, immunofluorescence staining, and other molecular biology techniques
- Obtained and assessed the effect of gene knockouts in *Toxoplasma gondii*
- Handled infectious *Toxoplasma gondii* samples and infected human cell lines

EDUCATION

European Bioinformatics Institute (EMBL-EBI) University of Cambridge

Cambridge, UK
Expected Oct 2025

PhD candidate at the University of Cambridge as part of the the EMBL International PhD Program (EIPP) — Applicant acceptance rate 2.7%
<https://www.embl.org/about/info/embl-international-phd-programme/>

University of Bologna

International Master Degree in Bioinformatics (MSc)
Final mark: 110/110 cum laude

Bologna, Italy
Jul 2021

University of Perugia

Bachelor's degree in Biotechnology (BSc)
Final mark: 110/110 cum laude and honorable mention

Perugia, Italy
Oct 2019

SKILLS HIGHLIGHTS

- R, Python, and bash programming (6 years)
- HPC (slurm, lsf) and cloud environments (aws, UKB RAP) (5 years)
- Nextflow workflow language (3 years, contributor to nf-core modules and author of numerous pipelines)
- Containerization technologies and virtual environments (Conda, Docker, Singularity) (5 years)
- Whole-genome sequencing data analysis and GATK best practices (3 years)
- Statistical methods for genome-wide association studies (linear mixed models)
- Supervised and unsupervised machine-learning methods (neural networks, ensemble classifiers, support vector machines, gradient-boosted trees, and others)
- Wet-lab techniques in molecular and cellular biology (PCR, CRISPR, immunofluorescence, Western blot, cell cultures, and others)

PUBLICATIONS

S. Pierotti, B. Welz, M. Osuna-López, T. Fitzgerald, J. Wittbrodt, E. Birney; Genotype imputation in F2 crosses of inbred lines, *Bioinformatics Advances*, 2024, <https://doi.org/10.1093/bioadv/vbae107>

PREPRINTS

* these authors contributed equally to this work

S. Pierotti, T. Fitzgerald, E. Birney; FlexLMM: a Nextflow linear mixed model framework for GWAS, *preprint on arXiv*, 2024, <https://doi.org/10.48550/arXiv.2410.01533>, *under review at Bioinformatics*

S. Pierotti*, I. Brettell*, T. Fitzgerald, C. Herder, N. Aadepe, C. Pylatiuk, J. Wittbrodt, E. Birney, F. Loosli; Measurement and classification of bold-shy behaviours in medaka fish, *preprint on bioRxiv*, 2024, <https://doi.org/10.1101/2024.10.18.618696>, *submitted to Bioinformatics*

PLANNED PUBLICATIONS

Co-first author paper on the detection of 16 QTLs for heart rate in medaka, their CRISPR validation, and the characterization of GxE and GxG effects

Collaboration with Bettina Welz (Wittbrodt group, Heidelberg University, Germany)

First author application note on a genetic variant prioritization pipeline

Code available here: <https://github.com/birneylab/varexplore>

First author paper on a GWAS for behavioural phenotypes in medaka

Collaboration with the group of Felix Loosli (KIT, Karlsruhe, Germany)

CONFERENCE PRESENTATIONS

The Biology of Genomes, Cold Spring Harbor Laboratory (Cold Spring Harbor, USA), May 7th to 11th, 2024 — poster

<https://meetings.cshl.edu/abstracts.aspx?meet=GENOME&year=24>

NRBP Medaka International Workshop, National Institute for Basic Biology (Okazaki, Japan), April 26th to 27th, 2024 — talk

<https://sites.google.com/nibb.ac.jp/nbrp-medaka-workshop/home-program>

TEACHING EXPERIENCE

EMBL-EBI course “Exploring Gene and Environmental Exposure Interactions to Understand Human Health and Disease”, February 2024

<https://www.ebi.ac.uk/training/events/exploring-gene-and-environmental-exposure-interactions-understand-human-health-and-disease/>

EMBL-EBI course “Using fish models to explore the impact of Gene by Environment (GxE)”, October 2022

<https://www.ebi.ac.uk/training/events/medaka-workshop>

Nextflow practical for the EBI predoc course (online), November 2022

https://saulpierotti.github.io/nextflow_workshop_ebi_predoc_course_2022

LANGUAGES

English: fluent, CEFR level C2. IELTS score 8.0. CAE with grade A awarded by Cambridge ESOL.

Italian: native speaker. **Others:** currently studying Turkish, German and Dutch at a beginner level.

REFEREES

Ewan Birney — *current PhD supervisor and director of EMBL-EBI (Cambridge, UK)*

birney@ebi.ac.uk

Jochen Wittbrodt — *close experimental collaborator and group leader at Heidelberg University (Heidelberg, Germany)*

jochen.wittbrodt@cos.uni-heidelberg.de

Virginie Uhlmann — *member of thesis advisory committee and director of the University of Zurich BioVisionCenter (Zurich, Switzerland)*

virginie.uhlmann@uzh.ch

Arne Elofsson — *Master’s thesis supervisor and group leader at SciLifeLab (Stockholm, Sweden)*

arne@bioinfo.se