Alan O'Callaghan

Curriculum Vitae

alan.ocallaghan@outlook.com | alanocallaghan.github.io | alanocallaghan | alanocallaghan | Alan O'Callaghan

Skills

- Programming: R, Java, Ruby, Python, C++.
- Statistics: Bayesian modelling, GLM, survival analysis, high dimensional statistics.
- Computing: Linux/Unix, LaTeX, git, make, snakemake, unit testing, continuous integration.
- Web: HTML/CSS, JS, htmlwidgets (R).

Education

Ph.D., biomedical data science

MRC Human Genetics Unit, University of Edinburgh

2018 – 2022

- Statistical methods for, and applied statistical analysis of, single cell RNAseq with complex experimental designs.
- Implementation of scalable Bayesian inference (adaptive Metropolis within Gibbs, divide and conquer MCMC).
- Assisting other students and contributing to research and computing environment in the institute.

M.Sc., Bioinformatics and computational genomics

Queen's University Belfast — *Distinction*

2014 – 2015

- Dissertation: Java-based eye-tracking and image analysis of histopathology whole slide images.
- Lectures: programming, image analysis, statistical learning techniques.
- Programming projects: Implementation of alignment, clustering, linear regression and cross-validation algorithms in R; cell detection, Ki67 quantification and image segmentation in Java.

B.Sc. (hons), biopharmaceutical chemistry

National University of Ireland, Galway — *2.1*

2008 – 2013

- Research Project: Structure elucidation and structure-activity relationship of bioactive marine compound.
- Lectures: (bio)chemistry, toxicology, pharmacology, molecular modelling, drug discovery and design.

Experience

Research software engineer

Centre for Genomics and Experimental Medicine, University of Edinburgh — *April 2023 – present*

- Software development associated with QuPath, an open source digital pathology image analysis tool.
- Implementing and maintaining machine learning pipelines using Deep Java Library.
- Building and packaging native library dependencies for the QuPath project.
- Developing and delivering training materials for QuPath users.
- Facilitating research by providing technical support and guidance, and leading research when time permits.
- Developing and maintaining documentation and continuous integration pipelines for associated projects.

Postdoctoral research associate

MRC Biostatistics Unit, University of Cambridge — *2022 – present*

- Computational optimisation of Bayesian models for the identification of eQTLs from RNAseq data.
- Computational analysis of RNAseq data using gene signatures.

Research assistant
UNIVERSITY OF EDINBURGH *April 2021 – September 2021*

- Developing a Carpentries-style course in high-dimensional statistics.
- Certification as a trained Carpentries instructor.

Bioinformatician
FIOS GENOMICS *October 2015 – August 2018*

- Communicating regularly with clients to plan projects and explain results.
- Exploratory and statistical analysis for academics and commercial clients.
- Statistical analysis & visualisations for client publications & commercial research projects.
- Developer, author & maintainer of proprietary R packages.
- Minor system administration tasks.

Software

R packages

Package	Platform	Role	Average monthly downloads
heatmaply	CRAN	Contributor	8392
scater	Bioconductor	Maintainer	6927
densvis	Bioconductor	Maintainer	1909
contrast	CRAN	Maintainer	1215
iheatmapr	CRAN	Maintainer	622
bayefdr	CRAN	Maintainer	343
snifter	Bioconductor	Maintainer	319
BASiCS	Bioconductor	Contributor	252
BASiCStan	Bioconductor	Maintainer	147

Support forums

StackOverflow reputation: 3,038
CrossValidated reputation: 1,226

Publications

- **O'Callaghan, A.**, Vallejos, C. A., "Scalable Inference of Transcriptional Variability with BASiCS". in: *Journal of Theoretical Biology* 611 (Aug. 2025), p. 112157. ISSN: 00225193. DOI: 10.1016/j.jtbi.2025.112157. (Visited on 06/06/2025)
- Goldsborough, T., **O'Callaghan, A.**, Inglis, F., Leplat, L., Filby, A., Bilen, H., Bankhead, P., *A Novel Channel Invariant Architecture for the Segmentation of Cells and Nuclei in Multiplexed Images Using InstanSeg*. Sept. 2024. DOI: 10.1101/2024.09.04.611150. (Visited on 02/06/2025)
- Aquilina, M., Wu, N. J., Kwan, K., Bušić, F., Dodd, J., Nicolás-Sáenz, L., **O'Callaghan, A.**, Bankhead, P., Dunn, K. E., *GelGenie: An AI-powered Framework for Gel Electrophoresis Image Analysis*. Sept. 2024. DOI: 10.1101/2024.09.06.611479. (Visited on 02/06/2025)
- Goldsborough, T., Philps, B., **O'Callaghan, A.**, Inglis, F., Leplat, L., Filby, A., Bilen, H., Bankhead, P., *InstanSeg: An Embedding-Based Instance Segmentation Algorithm Optimized for Accurate, Efficient and Portable Cell Segmentation*. Aug. 2024. DOI: 10.48550/arXiv.2408.15954. arXiv: 2408.15954 [cs]. (Visited on 02/06/2025)
- Kaczmarzyk, J. R., **O'Callaghan, A.**, Inglis, F., Gat, S., Kurc, T., Gupta, R., Bremer, E., Bankhead, P., Saltz, J. H., "Open and Reusable Deep Learning for Pathology with WSInfer and QuPath". In: *npj Precision Oncology* 8.1 (Jan. 2024), p. 9. ISSN: 2397-768X. DOI: 10.1038/s41698-024-00499-9. (Visited on 02/06/2025)
- Harris, B. T., Rajasekaran, V., Blackmur, J. P., **O'Callaghan, A.**, Donnelly, K., Timofeeva, M., Vaughan-Shaw, P. G., Din, F. V. N., Dunlop, M. G., Farrington, S. M., *Transcriptional Dynamics of Colorectal Cancer Risk Associated Variation at 11q23.1 Are Correlated with Tuft Cell Abundance and Marker Expression* in Silico. Preprint. Bioinformatics, Mar. 2022. DOI: 10.1101/2022.03.29.485182
- **O'Callaghan, A.**, Eling, N., Marioni, J. C., Vallejos, C. A., "BASiCS Workflow: A Step-by-Step Analysis of Expression Variability Using Single Cell RNA Sequencing Data". In: *F1000Research* 11 (Jan. 2022), p. 59. ISSN: 2046-1402. DOI: 10.12688/f1000research.74416.1

- Reijns, M. A. M., Thompson, L., Acosta, J. C., Black, H. A., Sanchez-Luque, F. J., Diamond, A., Parry, D. A., Daniels, A., O'Shea, M., Uggenti, C., Sanchez, M. C., **O'Callaghan, A.**, McNab, M. L. L., Adamowicz, M., Friman, E. T., Hurd, T., Jarman, E. J., Chee, F. L. M., Rainger, J. K., Walker, M., Drake, C., Longman, D., Mordstein, C., Warlow, S. J., McKay, S., Slater, L., Ansari, M., Tomlinson, I. P. M., Moore, D., Wilkinson, N., Shepherd, J., Templeton, K., Johannessen, I., Tait-Burkard, C., Haas, J. G., Gilbert, N., Adams, I. R., Jackson, A. P., "A Sensitive and Affordable Multiplex RT-qPCR Assay for SARS-CoV-2 Detection". In: *PLOS Biology* 18.12 (Dec. 2020). Ed. by Bill Sugden, e3001030. ISSN: 1545-7885. DOI: 10.1371/journal.pbio.3001030
- Galili, T., **O'Callaghan, A.**, Sidi, J., Sievert, C., "Heatmaply: An R Package for Creating Interactive Cluster Heatmaps for Online Publishing". In: *Bioinformatics* 34.9 (May 2018). Ed. by Jonathan Wren, pp. 1600–1602. ISSN: 1367-4803, 1460-2059. DOI: 10.1093/bioinformatics/btx657. (Visited on 12/03/2022)
- Morgan, R., Keen, J., Halligan, D., **O'Callaghan, A.**, Andrew, R., Livingstone, D., Abernethie, A., Maltese, G., Walker, B., Hadoke, P., "Species-Specific Regulation of Angiogenesis by Glucocorticoids Reveals Contrasting Effects on Inflammatory and Angiogenic Pathways". In: *PLOS ONE* 13.2 (Feb. 2018). Ed. by Christina L Addison, e0192746. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0192746

Teaching

Author/instructor/helper

QuPath workshops — *Various 2025 – ongoing*

- Writing, reviewing and maintaining documentation and workshop materials for QuPath.
- Organizing and delivering QuPath tutorials, talks and workshops.

instructor/helper

Carpentries workshops — *Online, 2021 – ongoing*

- Data carpentries R for social sciences (presenter)
- Python Intro for Libraries (helper)
- Statistics with R (helper)
- Data analysis and visualisation with Python for Genomics (helper)

IGMM statistical seminar series

Lectures — *Online, 2020*

- Exploratory data analysis
- Experimental design, hypothesis testing, statistical power

Applied analysis workshop for single cell RNAseq

One day workshop — *Institute of Genetics and Molecular Medicine, 2020*

- Exploratory analysis
- Normalisation
- Feature selection
- Differential expression analysis
- Clustering

Teaching materials

Author/maintainer — *2021 – ongoing*

- High dimensional statistics with R
- Orchestrating single cell analysis, multisample chapter

Conferences and meetings

From Images to Knowledge

Workshop — *Milan, 2024*

QuPath for Python Programmers

European Mathematical Genetics Meeting
TALK *Cambridge, 2022*
Optimising eQTL discovery with BaseQTL using a screening approach

Institute of Genetics and Molecular Medicine internal meetings
POSTER *Edinburgh, 2019 – 2020*
Scalable Bayesian analysis of single cell RNAseq
Bayesian analysis of multi-donor scRNAseq data (2nd prize)

University of Edinburgh Centre for Statistics conference
POSTER *Edinburgh, June 2019*
Scalable Bayesian analysis of single cell RNAseq

Quantitative Genomics
PRESENTATION *Francis Crick Institute, June 2019*
Scalable Bayesian analysis of single cell RNAseq

Genomic Medicine
POSTER *Edinburgh, May 2019*
Scalable Bayesian analysis of single cell RNAseq

Edinbr (R user group)
PRESENTATION *Edinburgh, March 2019*
Building interactive modules using htmlwidgets

Edinburgh Bioinformatics
PRESENTATION *Edinburgh, November 2018*
Interactive data visualisation using R and plotly.

Extracurricular

Society committees
EDINBURGH/GALWAY *Various*

- **IGMM Postgraduate Society (Edinburgh)**: Organising student events, advocating for improvements in student working conditions and well-being.
- **IGMM social committee (Edinburgh)**: Organising social events for students/postdocs/staff. Editor of noteworthy student Christmas movie, 2018.
- **Lotus society (Galway)**: Organising yoga classes and events for students/postdocs/staff.

Honors & Awards

- Best savoury food (mattar paneer), Institute Christmas Party, 2018
- School of Chemistry Medal in Molecular Modelling and Drug Design, NUI Galway, 2011
- Entrance Scholarship, NUI Galway, 2008