



# Sam Robson

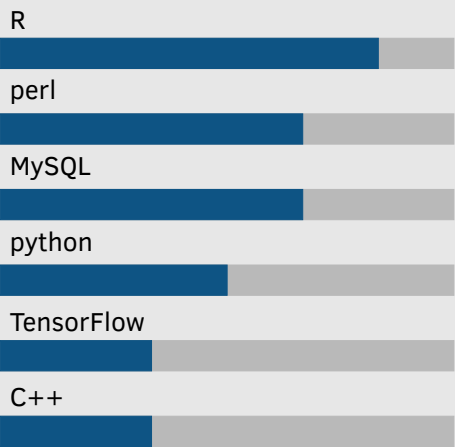
Bioinformatics Lead at  
Centre for Enzyme  
Innovation

- Portsmouth, United Kingdom
- [linkedin.com/in/samrobson/](https://www.linkedin.com/in/samrobson/)
- [irunfasterthanmycode.github.io](https://github.com/samrobson/irunfasterthanmycode)

## About me

Computational biologist with a strong mathematics and statistics background. Extensive experience of maintaining, processing, and analysing Big Data from next generation DNA sequencing. Expertise in a wide variety of data mining, data visualisation, deep learning, and machine learning methods to identify statistically significant trends in high-dimensional data. Bioinformatics Lead at the Centre for Enzyme Innovation and Faculty Bioinformatics Lead at the University of Portsmouth.

## Skills



RStudio\*5 Git\*4 LaTeX\*4

(\*\*)[The skill scale is from 0 (Fundamental Awareness) to 6 (Expert).]

## Education

- 2004-2008 PhD in *Mathematical Biology and Biophysical Chemistry*  
University of Warwick
- 2003-2004 MSc in *Mathematical Biology and Biophysical Chemistry* Class: 1<sup>st</sup>  
University of Warwick
- 1861-1863 BSc in *Mathematics (Hons)* Class: 2:1  
University of Warwick

## Experience

- Since 2017 Centre for Enzyme Innovation, University of Portsmouth Bioinformatics Lead  
Responsible for a small team of bioinformatics researchers to use machine learning algorithms to identify potential novel plastic-degrading enzymes from microbiological genome data
- 2014-2017 Sam Robson Consulting (Self-Employed) Statistics Consultant  
Identified key factors influencing doctor burnout through the use of multivariate mixed-effects regression analysis in the largest study of doctor burnout yet conducted
- 2010-2017 The Gurdon Institute, University of Cambridge Bioinformatician  
Developed and maintained pipeline and database for in-depth processing, mining and analysis of high-dimensional genome wide DNA sequencing data
- 2008-2010 Wellcome Trust Sanger Institute Statistical/Mathematical Biologist  
Responsible for maintaining, processing and normalizing genome-scale Big Data, including sample QC, concordance analysis with previously published SNP data, data optimization and case-control association testing

## Analysis Skills

Big Data wrangling, maintenance and analysis of extremely large data sets, pipeline development for high-throughput DNA sequencing data, normalization of complex data sets, data visualisation, machine learning, regression analysis (linear and generalised linear models), classification models (*unsupervised*: K-means, hierarchical clustering, mixture models; *supervised*: random forest, K-nearest neighbour, SVM), PCA dimensional reduction

## Leadership Skills

Bioinformatics Lead and Board Member at the Centre for Enzyme Innovation in charge of a small group of bioinformatics researchers, supervisor for a number of PhD students, Faculty Bioinformatics Lead working with and advising researchers throughout the University on a number of distinct projects, explaining complex technical outputs to non-experts and management

## Communication Skills

Excellent communication skills at the interface between Life Sciences, able to effectively explain complex analysis concepts to non-specialists, able to maintain extremely high standards when working across a large number of disparate projects, ability to work effectively across disciplines, experience collaborating with industry, excellent ability to identify and solve problems, ability to work on own initiative or as a keen team player, highly motivated

## Awards

- 2019 Award of £6 million from Research England E3 Fund
- 2019 Award of £5,000 Google Cloud Platform research credits
- 2017 Awarded CStat and CSci Membership of the Royal Statistical Society



# Sam Robson

Bioinformatics Lead at  
Centre for Enzyme  
Innovation



Portsmouth, United Kingdom



[linkedin.com/in/samrobson/](https://www.linkedin.com/in/samrobson/)

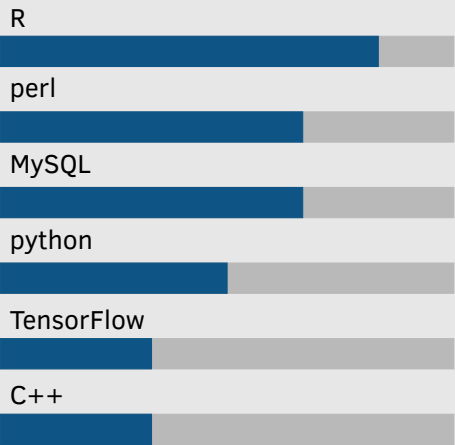


[irunfasterthanmycode.github.io](https://github.com/samrobson/irunfasterthanmycode)

## About me

Computational biologist with a strong mathematics and statistics background. Extensive experience of maintaining, processing, and analysing Big Data from next generation DNA sequencing. Expertise in a wide variety of data mining, data visualisation, deep learning, and machine learning methods to identify statistically significant trends in high-dimensional data. Bioinformatics Lead at the Centre for Enzyme Innovation and Faculty Bioinformatics Lead at the University of Portsmouth.

## Skills



RStudio★5 Git★4 LaTeX★4

(\*)[The skill scale is from 0 (Fundamental Awareness) to 6 (Expert).]

## Publications

2019 *Interaction of Sox2 with RNA binding proteins in mouse embryonic stem cells*, Experimental Cell Research. 381, 1, 129-138

2019 *METTL1 Promotes let-7 MicroRNA Processing via m7G Methylation*, Molecular Cell. 74, 6, 1278-1290

2019 *Dystrophic mdx mouse myoblasts exhibit elevated ATP/UTP-evoked metabotropic purinergic responses and alterations in calcium signalling*, Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease

2018 *SRPK1 maintains acute myeloid leukemia through effects on isoform usage of epigenetic regulators including BRD4*, Nature Communications. 9, 1, 5378

2018 *Phosphorylation of histone H4T80 triggers DNA damage checkpoint recovery*, Molecular Cell. 72, 4, 625-635

2018 *Inhibition of the acetyltransferase NAT10 normalizes progeric and aging cells by rebalancing the Transportin-1 nuclear import pathway*, Science Signaling. 11, 537, eaar5401

2017 *Promoter-bound METTL3 maintains myeloid leukaemia by m6A-dependent translation control*, Nature. 552, 126-131

2017 *A lncRNA fine tunes the dynamics of a cell state transition involving Lin28, let-7 and de novo DNA methylation*, eLife. 6, e23468

2017 *RNA binding by the histone methyltransferases Set1 and Set2*, Molecular and Cellular Biology. 37, 14, e00165-17

2016 *A chemical probe for the ATAD2 bromodomain*, Angewandte Chemie International Edition. 55, 38, 11382-11386

2016 *Discovery of I-BRD9, a selective cell active chemical probe for bromodomain containing protein 9 inhibition*, Journal of Medicinal Chemistry. 59, 4, 1425-1439

2015 *Generation of a selective small molecule inhibitor of the CBP/p300 bromodomain for leukemia therapy*, Cancer Research. 75, 23, 5106-5119

2014 *The breast cancer oncogene EMSY Represses transcription of anti-metastatic microRNA miR-31*, Molecular Cell. 53, 5, 806-818

2014 *Recurrent mutations, including NPM1c, activate a BRD4-dependent core transcriptional program in acute myeloid leukemia*, Leukaemia. 28, 2, 311-320

2014 *Glutamine methylation in histone H2A is an RNA-polymerase-I-dedicated modification*, Nature. 505, 7484, 564-568

2014 *BET protein inhibition shows efficacy against JAK2V617F-driven neoplasms*, Leukaemia. 28, 1, 88-97

2013 *The non-coding snRNA 7SK controls transcriptional termination, poisoning, and bidirectionality in embryonic stem cells*, Genome Biology. 14, 9, R98

2012 *Human RNA Methyltransferase BCDIN3D Regulates MicroRNA Processing*, Cell, 151 (2), 278-288

2012 *Three distinct patterns of histone H3Y41 phosphorylation mark active genes*, Cell Reports. 2, 3, p. 470-477

2011 *Deciphering c-MYC-regulated genes in two distinct tissues*, BMC Genomics. 12, 1, 476

2011 *Inhibition of BET recruitment to chromatin as an effective treatment for MLL-fusion leukaemia*, Nature, 478 (7370), 529-533

2010 *Origins and functional impact of copy number variation in the human genome*, Nature, 464 (7289), 704-712

2010 *Genome-wide association study of copy number variation in 16,000 cases of eight common diseases and 3,000 shared controls*, Nature, 464 (7289), 713-720

2010 *Nucleosome-interacting proteins regulated by DNA and histone methylation*, Cell, 143 (3), 470-484

2006 *c-Myc and downstream targets in the pathogenesis and treatment of cancer*, Recent Patents on Anti-Cancer Drug Discovery. 1, 3, 305-326