Matthew Parker

The University of Sheffield
Firth Court
Western Bank
Sheffield
S10 2TN

Email: mparker2@sheffield.ac.uk
Website: mparkerbio.com
LinkedIn: mparker02
Twitter: @1m1a2t
GitHub: mparker2

Education;

PhD Bioinformatics and Molecular Biology; October 2014 - 2018 (expected)
Department of Molecular Biology, The University of Sheffield
Analysis of high throughput sequencing datasets to characterise the biological significance of unusual DNA structures called G Quadruplexes. Gained a broad range of skills including experience of both bioinformatics and wet lab biology.

BSci (Hons) 1st, Biochemistry; September 2011 - June 2014
Department of Molecular Biology, The University of Sheffield.
Graduated second in year with an average score of 74.

A-Levels; September 2009 – June 2011
The King's School, Grantham
2 A* incl. Maths, 3 A incl. Chemistry, Biology, 1 A (AS level) Further Maths.

GCSE; September 2007 – June 2009
The King's School, Grantham
8 A* incl. Maths, English, 5 A

Skills;

- **Programming**: Uses the scripting languages Python and R every day for the manipulation and visualisation of complex datasets. Fluent in the use of Scientific Python packages including numpy, scipy, pandas, matplotlib, scikit-learn and keras. Example projects can be found on GitHub.
- **Computing**: Experienced in working with Unix systems, including High Performance Computing clusters (running Sun Grid Engine). Has collaborated on tools and pipelines for reproducible distributed bioinformatic analyses.
- **Research**: Very capable researcher, able to critically appraise data and scientific literature to ask and answer new research questions, understand new techniques, and quickly prototype and build new tools.
- **Communication**: Experienced in explaining technical concepts to a non-technical audience, for example, presented satellite land cover prediction model to audience of conservation specialists whilst at Natural England. Regularly presented work to both laboratory group and department during PhD. Has more than 50 hours experience teaching both Python and R for biology students at Undergraduate level.
- **Teamwork**: During the Science To Data Science training course, worked as part of a small team to rapidly deliver customer behaviour models to a London start-up. As part of the departmental PhD society, helped to organise and run a weekend retreat to the Peak District for approximately 80 PhD students.

Experience;

Science To Data Science; August-September 2017
Took part in 5 week intensive re-training programme for PhD students interested in entering the field of Data Science. As part of the course, worked with a London start-up company using their data to predict customer behaviour. Course Fees were funded by the Sheffield University Postgraduate Researcher Experience Programme.

Internship at Natural England; January - April 2017
As part of the Earth Observation team, helped build a model to predict land cover and habitat classification from ESA Satellite Data. This involved preprocessing and unsupervised segmentation of satellite imagery using scikit-image, followed by segment class prediction using scikit-learn and R package randomForest. Software developed during the project is to be used by environmental agencies in various countries and British territories. Code and examples can be found on GitHub and mparkerbio.com. Funded by the BBSRC Professional Internships for PhD Students scheme.

HackMed Sheffield; April 2017
Part of a team placing first at medical themed Hackathon. Built a pipeline to identify antibiotic resistance genes in bacterial samples using Oxford Nanopore long read technology. View the project on DevPost.

Placement at Sheffield Institute of Translational Neuroscience; June - July 2013
Short project to characterise the development of neuronal precursor cells (using *in situ*hybridisation, immunostaining) in a mutant with neurodevelopmental defects related to schizophrenia. Funding provided by the Wellcome Trust Biomedical Vacation Scholarship scheme.

Example Projects;

- Contributed to an open source package for de-duplicating Unique Molecular Identifiers commonly used in single cell sequencing experiments. These contributions introduced indexing methods which reduce the number of pairwise calculations between sequences required, thereby resulting in a significant increase in speed and memory performance.
- Built a convolutional neural network for the automation of Stomatal Index counting from light microscopy images. This is a common and labour intensive task for scientists in the field of plant leaf development. Code is available on GitHub.
- Used the Hungarian Algorithm to help University teaching staff optimally assign undergraduate students to preferred third year research projects.

Papers;

Simon Lesbirel, Nicolas Viphakone, **Matthew Parker**, ... & Stuart A. Wilson (2018). The m6A-methylase complex recruits TREX and regulates mRNA export. *Scientific Reports*, 8(1): 13827

Conference Posters;

Matthew Parker, Colette Baxter, Manoj Valluru, Karim Sorefan (2017). Two tetrad G Quadruplexes modulate expression and splicing of EXTENSIN genes in Arabidopsis thaliana. *G4thering Prague*

References;

Dr Karim Sorefan

Lecturer
Dept. of Molecular Biology
University of Sheffield
k.sorefan@sheffield.ac.uk

Dr Ian Sudbery

Lecturer
Dept. of Molecular Biology
University of Sheffield
i.sudbery@sheffield.ac.uk

Richard Alexander

Lead Data Scientist
Natural England
richard.alexander@
naturalengland.org.uk