

Candidate Information

Position:	Research Fellow (Computational Microbial Genomics)
School/Department:	School of Biological Sciences
Reference:	26/113431
Closing Date:	Monday 6 July 2026
Salary:	£41,519 per annum
Anticipated Interview Date:	Friday 17 July 2026
Duration:	Fixed term available for 12 months or until 31 July 2027, whichever is soonest

JOB PURPOSE:

The Creevey Research Group at Queen's University Belfast is seeking an experienced Research Fellow in computational genomics to join a funded interdisciplinary project. The project, conducted in collaboration with the University of Lincoln and University of Liverpool, focuses on developing pangenome-based computational approaches for the detection of artificially introduced genetic elements in bacteria.

The post holder will take a leading role in large-scale bacterial genome data acquisition, quality control, annotation, and network-based comparative genomic analysis. These activities directly underpin the development and evaluation of machine learning models for genomic anomaly detection. The role is primarily bioinformatic and genomic in nature; experience in AI and machine learning is valued but not required.

The successful candidate will be required to obtain UK security clearance prior to or shortly after commencement of employment. Any offer of employment will be conditional upon receipt of satisfactory clearance.

MAJOR DUTIES:

1. Genomic Data Management and Analysis

- Identify, curate, and process large-scale bacterial genome datasets from public repositories and curated sources.
- Implement quality control procedures and preprocessing pipelines for diverse genomic data types.
- Develop and maintain comprehensive data catalogues with appropriate metadata standards.

2. Genome Annotation and Comparative Genomics

- Design and apply annotation workflows to characterise gene content, function, and evolutionary relationships across bacterial taxa.
- Perform pangenome construction and gene family clustering to produce outputs suitable for downstream modelling.

3. Network-Based Genomic Analysis

- Construct and analyse genomic network representations.
- Apply network analysis methods to identify structural and functional patterns across large genome collections.
- Develop scalable approaches for network construction applicable to pan-species prokaryotic datasets.

4. Computational Pipeline Development

- Design, implement, and maintain robust, scalable bioinformatics workflows for execution on HPC infrastructure.
- Develop data processing approaches suitable for integration with AI and machine learning frameworks.
- Document pipelines and tools to professional standards, ensuring reproducibility and transferability.

5. Collaboration and AI Integration

- Work closely with AI and machine learning specialists to deliver processed datasets compatible with model training pipelines.
- Contribute to the iterative refinement of data representations based on model performance feedback.
- Liaise with collaborators at partner institutions to coordinate research activities across work packages.

6. Research Output and Dissemination

- Prepare high-quality research publications for peer-reviewed journals.
- Present research findings at national and international conferences and project review meetings.
- Develop documentation and user guides for computational tools and workflows.

7. Training, Mentorship, and Professional Development

- Supervise and support postgraduate students and junior team members in computational genomics methods.
- Engage with the broader research community through networking, collaboration, and professional development activities.
- Contribute to grant applications and funding proposals as required.

ESSENTIAL CRITERIA:

1. Have or about to obtain* a PhD in computational biology or bioinformatics or a closely related discipline, with a strong emphasis on computational analysis (*must be obtained within 3 months of commencement of employment).
2. Extensive experience in handling and analysing large-scale genomic datasets, including data acquisition from public repositories, quality filtering, and assembly assessment.
3. Demonstrated experience in genome annotation workflows, including tools for gene prediction, functional annotation, and ortholog or gene family assignment.
4. Experience in pangenome analysis, including the construction of pangenomes and interpretation.
5. Experience developing and applying network-based approaches to comparative genomic data.
6. Advanced programming skills in Python, R, or other languages appropriate for large-scale biological data analysis, with experience working in Linux/Unix environments.
7. Demonstrated experience designing and implementing reproducible bioinformatics pipelines.
8. Experience working with high-performance computing environments and job schedulers (e.g., Slurm or equivalent).
9. Experience with version control systems (e.g., Git) and reproducible research practices.
10. Ability to communicate complex technical information clearly, both in writing and verbally, including to interdisciplinary audiences.
11. Ability to work independently and as part of a multidisciplinary, multi-institutional team.
12. Strong scientific writing skills, with a track record of, or demonstrable capacity for, peer-reviewed publication commensurate with career stage.
13. Willingness to travel for collaborative activities with partner institutions.
14. Willingness to obtain UK security clearance as a condition of employment and to work with sensitive data under appropriate data security protocols.

DESIRABLE CRITERIA:

1. Experience with machine learning or deep learning frameworks, or direct collaboration with machine learning researchers in a biological context.
2. Experience managing or contributing to very large-scale genome datasets, such as those from AllTheBacteria, NCBI RefSeq, or BV-BRC.
3. Knowledge of microbial genomics and associated databases relevant to bacterial diversity and evolution.
4. Experience with cloud computing platforms or containerisation technologies (e.g., Docker, Singularity).
5. Experience developing and maintaining open-source bioinformatics software or analysis packages.
6. Understanding of FAIR data principles and modern data management practices.
7. Experience in collaborative, multi-institutional research projects.
8. Awareness of biosecurity and responsible research practices in the context of genomic data.

Additional information:

Informal enquiries can be directed to: Dr. James Gillespie - j.gillespie@qub.ac.uk.