



centre nacional d'anàlisi genòmica  
centro nacional de análisis genómico



## “Bioinformatician – Quality Control & Primary Analysis”

### Centro Nacional de Análisis Genómico (CNAG-CRG)

#### The Institute

The Centro Nacional de Análisis Genómico (CNAG-CRG) is one of the largest Genome Sequencing Centers in Europe. CNAG-CRG researchers participate in major International Genomic Initiatives such as the International Cancer Genome Consortium (ICGC), the International Human Epigenome Consortium (IHEC), the International Rare Diseases Research Consortium (IRDiRC) and the European Infrastructure for life-science information (ELIXIR), as well as in several EU-funded projects. The CNAG-CRG has 6 Illumina sequencers (including a NovaSeq 6000 and two HiSeq 4000) and several Oxford Nanopore Technologies sequencers that produce ultra-long reads (a GridION and 3 MinION). The HPC cluster has 3472 computing cores and 7.6 Petabytes of data storage and is connected to the Barcelona Supercomputing Center.

The CNAG-CRG is integrated with the Centre for Genomic Regulation (CRG), an international biomedical research institute of excellence, based in Barcelona, Spain, with more than 400 scientists from 44 countries. The CRG shares principles of an interdisciplinary, motivated and creative scientific team that is supported by high-end and innovative technologies and a flexible and efficient administration.

CRG has been conferred with a badge of 'HR Excellence in Research' by the European Commission, in recognition to its progress in implementing the European Charter for Researchers and the Code of Conduct for Recruitment of Researchers, that among others consists of transparent, merit-based Recruitment procedures and attractive work-life balance working conditions.

For further information: [www.cnag.crg.eu](http://www.cnag.crg.eu) and [www.crg.eu](http://www.crg.eu)

#### The role

We are seeking a bioinformatician, computational biologist or software developer with knowledge of genomics to work within the Production Bioinformatics Team. The successful candidate will be responsible for the development of integrated tools and pipelines for automated processing and quality control of large amounts of Next Generation Sequencing data from a wide range of applications (Whole Genome, Exome and Panel sequencing, RNA-Seq, Bisulphite Sequencing, Single Cell RNA-Seq, Genotyping-By-Sequencing, HiC, etc.), projects (inherited diseases, cancer, infectious diseases, evolution, population genomics, structural genomics, epigenetics, etc.) and organisms. The position also requires in-depth custom analysis of datasets and projects, typically for troubleshooting difficult cases or when new developments are required.

The main tasks of the role include:

- Management, processing, quality control and analysis of large amounts of Next Generation Sequencing data
- Development and/or testing of new software and solutions for the quickly changing field of



centre nacional d'anàlisi genòmica  
centro nacional de análisis genómico



#### DNA sequencing

- Ongoing development of the data processing and quality control pipeline on an HPC cluster
- Assure automated continuous operation and integration with the in-house developed Laboratory Information Management System (LIMS)
- Contribute to the development of Standard Operating Procedures and the CNAG-CRG ISO certifications and accreditations (currently ISO 9001 and ISO 17025)
- Communication and troubleshooting with the Sequencing Unit, bioinformatics groups and CNAG-CRG collaborators.

#### About the team / lab / department

The Production Bioinformatics team is in charge of the day to day production and processing of polished, high quality sequence data, and acts as a liaison between the sequencing laboratory, other bioinformatics teams and the CNAG-CRG collaborators. The team is managed by Matthew Ingham and is part of the CNAG-CRG Bioinformatics Analysis Unit led by Dr. Sergi Beltran.

Responsibilities include working with lab technicians, sequencer manufacturers and system administrators to maintain a stable informatics environment for the generation and processing of sequencing data by the Illumina and Oxford Nanopore instruments. The team also ensures final data quality based on the review of basecalling confidence, error rates, alignment results, and coverage metadata. This is mainly done through semi-automated workflows developed and maintained by the team and running in the CNAG-CRG's High Performance Computing (HPC) cluster. To remain on the cutting edge of next generation sequencing, the CNAG-CRG is constantly testing, upgrading and improving its instruments and protocols. Therefore, the team also conducts customised analysis and investigation, not only for specific issues and projects, but also to improve the workflows and adapt to new developments in the field. In addition, the team is responsible for developing an in-house Lab Information Management System (LIMS), which is used throughout the CNAG-CRG to track projects, samples and run information along with analysis results. The operations of the Production Bioinformatics Team are included within the CNAG-CRG ISO 9001 certification and the ISO 17025 accreditation.

These various aspects of the Production Bioinformatics team maintain the steady production of high quality data at the CNAG-CRG, and we're looking for an experienced candidate ready to contribute in this dynamic environment.

#### Whom would we like to hire?

- You hold a Master's degree in Bioinformatics or related discipline or Bachelor's degree with 2+ years of practical experience
- You have good knowledge of genomics
- You are very able in Unix and Windows operating systems and networks
- You have proven experience with Unix shell scripts, Python, and SQL
- You have excellent teamwork, organizational and communicative skills
- You are fluent in English



centre nacional d'anàlisi genòmica  
centro nacional de análisis genómico



### Desirable but not required

- You have experience in genomic data analysis
- You have experience in managing large sets of data
- You have knowledge of Perl and R, IPython and/or Matplotlib
- You have experience with batch processing on a cluster
- You have good software development practices

### The Offer

- **Contract duration:** 1 year initial contract with possibility to extend
- **Estimated annual gross salary:** Salary is commensurate with qualifications and consistent with our pay scales.
- **Target start date:** as soon as possible

We provide a highly stimulating environment with state-of-the-art infrastructures, and unique professional career development opportunities.

We offer and promote a diverse and inclusive environment and welcomes applicants regardless of age, disability, gender, nationality, race, religion or sexual orientation.

We are committed to reconcile a work and family life for our employees and are offering the opportunity to benefit from annual leave.

### Application Procedure

All applications must include:

1. A motivation letter
2. A complete CV including contact details
3. Contact details of two referees.

All applications must be addressed to Human Resources and be submitted online on the CRG Career site - <http://www.crg.eu/en/content/careers/job-opportunities>

**Deadline:** Please submit your application by the 1st of June,2019



HR EXCELLENCE IN RESEARCH

baldiri reixac, 4  
pcb - tower i, 2nd floor  
08028 barcelona

t +34 93 4020542  
f +34 93 4037279  
[www.cnag.crg.eu](http://www.cnag.crg.eu)