

## (Ref. 21/2018) Postdoc Position in Machine learning tools to combine sequence and biologically heterogeneous data

The Center for Research in Agricultural Genomics (Crag) is looking for a Postdoc Researcher to be part of the Statistical and Population Genomics Group. For more information: <http://www.cragenomica.es/es/grupos-de-investigacion/statistical-and-population-genomics>

Crag is an independent research institution engaged in leading-edge basic and applied plant and farm animal sciences. Crag is established as a Consortium of the Spanish National Research Council (CSIC), Institute of Agrifood Research and Technology (IRTA), Autonomous University of Barcelona (UAB), and University of Barcelona (UB). The Center is located at the UAB campus, and currently hosts 200 members from across the world.

Research Programs at Crag (from basic science to applied research using plant experimental model systems, crops and farm animals) make extensive use of large sets of genetic and genomic data, including high-throughput sequencing-based methods, high-throughput genotyping, and complete genome sequences and genome re-sequencing (<http://biannualreport2012-2013.cragenomica.es/>).

### Group description:

The numerical genomics group (<http://bioinformatics.cragenomica.es/numgenomics/>) at Centre for Research in Agricultural Genomics (Crag, [www.cragenomica.es](http://www.cragenomica.es)) is primarily interested in the use of high throughput sequencing technology (NGS) for population and statistical genomics. Topics of particular interest is the use of sequence for genomic selection. Crag was recently awarded a Severo Ochoa project for excellence centers in Spain. This project will be carried out in cooperation with international and national groups with a long experience in machine learning that complement our expertise.

### Responsibilities:

We are looking for a Postdoc Researcher with a computational profile and an interest in biological applications.

Big data are characterized not only by their size but also by their heterogeneity and noisiness. These features apply in particular to genomic data: their size has been increasing exponentially with the advent of new sequencing technologies, but also their complexity. We aim at combining several available sources of information, not only the phenotypes and sequence data, but also, e.g., annotation features or metabolic pathways. An important goal is not only to provide efficient predictors but also tools to biologically interpret the results. We aim at developing tools that combine image and sequencing data.

The tasks are included within the framework at the project AGL2016-78709-R.

**Laboratory:** Dr. Miguel Pérez Enciso (<https://www.icrea.cat/Web/ScientificStaff/Miguel-Perez-Enciso-255>)

### Requirements:

Educational requirements: PhD in Biology, Biochemistry or similar.

Qualifications and Experience: Programming ability and experience un python. Experience in machine learning area.

### Valued Skills:

Strong skills in communication, organization, and interpersonal relations.  
Experience in image analysis.

### We offer:

Contract duration: 12 months (extendable if funding available).

Hours/week: Full time

**Submission of applications:**

***Interested candidates, please submit:***

- Letter of motivation describing past and current projects (1.5 pages max.)
- Full CV including a list of publications, experience and techniques.
- Three references (including email address and phone number)

All the candidatures must be registered and send the documents at: <http://www.cragenomica.es/jobs/jobs.php>

**Deadline:** Until suitable candidate found