BIOINFORMATICS UNIT

Fátima Al-Shahrour
Unit Head

Post-Doctoral Fellow
Hector Tejero

OVERVIEW

Bioinformatics is a key discipline for understanding the cancer genome and for the future of cancer therapeutics. Bioinformatics-based approaches have the ability to transform the huge amount of biological data into comprehensive models that provide an in-depth understanding of cancer disease and the complex relationships among genotype and phenotype that are needed to identify cancer driver molecular alterations and new therapeutic targets.

"PanDrugs is a feasible method to identify actionable molecular alterations and to prioritise drugs that facilitate the interpretation of the genomic landscape and clinical decision-making in cancer patients."

The CNIO Bioinformatics Unit (BU) has several goals:

- → To develop new computational methodologies and bioinformatics tools to enable the integration of biological and clinical data.
- → To achieve genome analysis in cancer patients' data to identify new biomarkers and mechanisms of drug response.
- → To provide bioinformatics support with data analysis and interpretation using computational and statistical methods.
- → To maintain the scientific computing facilities at the CNIO and to provide training in bioinformatics tools and methods.

Graduate Students
Guillermo Martín-Serrano (since February), Javier Perales, Kevin Troulé

Technicians
Andrés Cañada (TS)*, Ángel Carro (until January) (TS)*, Coral Fustero (PEJ-L)**, Gonzalo Gómez (TS)*, Osvaldo Graña (TS)*, Elena Piñeiro (TS)*, Miriam Rubio (until April) (TS)*

Titulado Superior (Advanced Degree)
***Plan de Empleo Joven-Licenciado* (Youth Employment Plan-Graduate) (until November)

RESEARCH HIGHLIGHTS

In 2017, the CNIO Bioinformatics Unit published 24 peer-reviewed articles (the full list is available on our web site http://bioinformatics.cnio.es/) as a result of our ongoing research projects and scientific collaborations with CNIO Research Groups and other national and international research institutions.

During this year, we developed several bioinformatics tools for the analysis of next-generation sequencing data in collaboration with the SING group (University of Vigo): **RubioSeq+** for DNA-Seq analysis (Rubio-Camarillo *et al.* 2017), **nextpresso** for RNA-Seq analysis (Graña *et al.* 2017), and **bicycle** for bisulfite-seq analysis (Graña *et al.* 2017). Remarkably, the RUbioSeq+ (http://rubioseq.bioinfo.cnio.es/) tool supports an interactive graphical user interface (GUI) that facilitates its usage for biomedical translational researchers who lack computational or bioinformatics skills.

Also, our Group focuses on gaining a better understanding of the impact of cancer genomics on the making of clinical decisions. To this aim, we have developed two methodologies to guide the selection of therapies, propose sequential treatments and drug repositioning: **PanDrugs** (http://pandrugs.bioinfo.cnio.es) and **SATIE** (http://satie.bioinfo.cnio.es).

All our tools are freely available and have been applied in different genomic studies undertaken in numerous scientific collaborations, such as: the study of the pre-metastatic activity of midkine in cancer (Olmeda D *et al.*, 2017), the identification of a novel oncogenic Braf kinase-inactive mutation (Nieto, P *et al.*, 2017), and also the prediction of drug response in pancreatic cancer patient-derived xenograft mouse models using transcriptomics data (Rajeshkuma NV *et al.*, 2017).

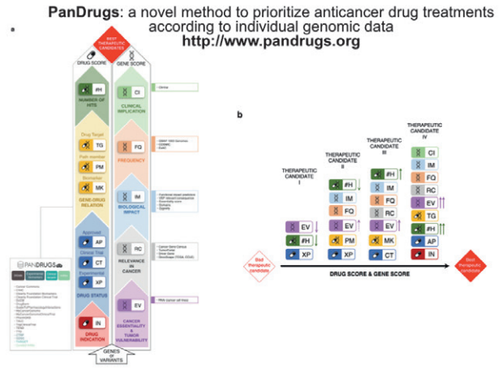

Figure PanDrugs score calculation. (**a**) Gene Score (GScore) uses the biological and clinical impact of mutations, their frequency, relevance and vulnerability in cancer. Drug indication, gene-drug association and number are used to calculate Drug score (DScore). (**b**) 'Best therapeutic candidates' are based on the accumulated and weighted scoring of GScore and DScore.

Many more bioinformatics analyses were performed together with other CNIO groups, such as: the development of a targeted sequencing assay for Pheochromocytoma and Paraganglioma diagnostics (Currás-Freixes *et al.*, 2017) and the study of NUP98-HOXA9 fusion in leukaemia (Rio-Machin A *et al.*, 2017), among others. ■

SELECTED PUBLICATIONS*

- Gómez-López G, Dopazo J, Cigudosa JC, Valencia A, Al-Shahrour F (2017). Precision medicine needs pioneering clinical bioinformaticians. *Brief Bioinform*. PMID: 29077790.
- Graña O, López-Fernández H, Fdez-Riverola F, González Pisano D, Glez-Peña D (2017). bicycle: a bioinformatics pipeline to analyze bisulfite sequencing data. *Bioinformatics*. PMID: 29211825.
- Perales-Patón J, Piñeiro-Yañez E, Tejero H, López-Casas PP, Hidalgo M, Gómez-López G, Al-Shahrour F (2017). Pancreas cancer precision treatment using avatar mice from a Bioinformatics Perspective. *Public Health Genomics* 20, 81-91.
- Rubio-Camarillo M, López-Fernández H, Gómez-López G, Carro Á, Fernández JM, Torre CF, Fdez-Riverola F, Glez-Peña D (2017). RUbioSeq+: A multiplatform application that executes parallelized pipelines to analyse next-generation sequencing data. *Comput Methods Programs Biomed* 138, 73-81.
- Graña O, Rubio-Camarillo M, Fdez-Riverola F, Pisano DG, Glez-Peña D (2017). Nextpresso: next generation sequencing expression analysis pipeline. *Curr Bioinform*. DOI 10.2174/1574893612666170810153850.
- Andrés-León E, Gómez-López G, Pisano DG (2017). Prediction of miRNA-mRNA Interactions Using miRGate. *Methods Mol Biol* 1580, 225-237.

*please see BU's web site for a list of all publications.