

This project, sponsored by the “la Caixa” Bank Foundation, will store nearly 4,000 genomic data sets.

The EGA, the European archive with the greatest amount of genomic data, and its impact on health

- **Medical genomics, an emerging discipline already being used for diagnosis and treatment of rare diseases and some types of cancer, requires close cooperation between research centres and hospitals.**
- **The European Genome-Phenome Archive (EGA), located in the Centre for Genomic Regulation (CRG) and the European Bioinformatics Institute (EMBL-EBI) is the repository of genomic data of reference in Europe, guaranteeing safety and allowing their use in research.**
- **The project was launched in 2014. Today, the EGA stores nearly 4,000 genomic data sets and over 1,500 studies that occupy over 4.5 petabytes (4,500,000 gigabytes).**
- ***CosmoCaixa Barcelona* hosted today a gathering of scientists, doctors, patients' associations and biomedical companies to discuss the importance of integrating genomic data in medicine in Spain, and their benefits in personalized medicine.**

Barcelona, April 13, 2018 – Àngel Font; corporate director for the area of research and strategy of the "la Caixa" Banking Foundation; Luis Serrano, director of the Centre for Genomic Regulation (CRG) and Arcadi Navarro, secretary of Universities and Research of the Autonomous Government of Catalonia, presented the main themes to be discussed by health experts at the seminar "The Impact of Genomic Data on Health: the Role of the European Genome-Phenome Archive (EGA)" in CosmoCaixa.

Seventeen years after determining the sequence of the human genome, genomic research and the technology that makes it possible have advanced at a remarkable pace. Today, genomic data have great potential to transform the practice of medicine, and are already having an impact in fields such as the diagnosis and treatment of rare diseases, cancer and other pathologies.

The use of genomic data in biomedical research has made it possible to advance in the knowledge of biologic processes and of a number of diseases. But the leap from research to clinical practice requires a broad debate on the challenges and opportunities offered by this new discipline, as well as the importance of safely preserving and sharing this kind of data. "From the Centre for Genomic Regulation (CRG) we work to decode and comprehend the information contained in the genome, and from the European Genome and Phenome Archive (EGA) we work to ensure these data remain safe, and within reach of the scientific community. We have to keep researching but it is also absolutely necessary that scientists share this knowledge with the medical community for the return to society to be real, efficient and agile," states Luis Serrano, director of the CRG.

With this goal, the seminar, "Impact of genomic data on health: the Role of the European Genome and Phenome Archive (EGA)", organized by the "la Caixa" Bank Foundation, and the Centre for Genomic Regulation with the support of Elixir España, will seek to raise awareness and spark debate among the scientific, medical and social communities that makes it possible to drive medical genomics in Spain. Cooperation is a key element for medical genomics to be a reality. Along these lines, Arcadi Navarro, secretary of Universities and Research of the Autonomous Catalan Government and ICREA research professor at the Pompeu Fabra University states that "in an open science world, the EGA is a non-hierarchical network facilitator for cooperation among scientists. It enables a new type of research that will accelerate these cooperative networks, which otherwise would take a long time to come about." The seminar will be participated in by leading speakers from the medical genomics field, representatives from the clinical realm, patients' associations and over 150 participants who have registered for the event.

According to the corporate director for Research and Strategy of the "la Caixa" Bank Foundation, Ángel Font, "preserving an asset as important as health has become one of "la Caixa" Banking Foundation's priorities. For this reason, our strategic plan triples the investment in research up to 90 million euros in 2019. In this case, the project, which is the result of the collaboration and cooperation among different organizations, is another step toward the final, fundamental goal of advancement in cutting-edge biomedical research.

Safe data managed following ethical principles and available for research

Preserving human genomic data and sharing them safely is one of the main challenges for medical genomics to be a reality. The European Genomic and Phenome Archive (EGA) is Europe's seminal public genome base, and one of reference for the rest of the world. Since 2014, and thanks to the support the "la Caixa" Foundation, the Generalitat (Autonomous Government of Catalonia) and the Ministry of Economy, Industry and Competitiveness, the Centre for Genomic Regulation manages the EGA together with the European Bioinformatics Institute (EMBL-EBI) in the United Kingdom. In

Barcelona, the EGA data are stored in the Barcelona Supercomputing Centre - National Supercomputing Centre (BSC-CNS).

EGA is a specialized repository that facilitates the responsible use of data, following the ethical principles on human genomic data, promoting the use of resources by offering them for research, and guaranteeing their long-term availability. The EGA currently stores **nearly 4,000 genomic data sets and over 1,500 studies, coming from about 800,000 people. EGA data** occupy over 4.5 petabytes (4,500,000 gigabytes) and have nearly 12,000 users around the world who consult and use them.

Data that let us improve health

The genomic data stored in the EGA are relevant for the entire scientific and medical community. The repository contains data relevant for the study of cancer and cardiovascular, infectious, inflammatory, neurological and rare diseases. Over half the data refer to cancer and are crucial for advancement in the understanding of this disease.

Keeping them available empowers researchers to do things like identify genetic factors related with diseases, improve diagnoses and search for new treatments. Along these lines, the EGA stores data from international projects and consortia with the most relevance in genomics, such as the International Cancer Genome¹ (ICGC), the International Human Epigenome Consortium² (IHEC), the Wellcome Trust Sanger Institute, and the UK10K project³ on minority diseases in the United Kingdom.

The EGA data, for example, have been used to identify the genetic variants responsible for a number of minority diseases, such as the study published in *Nature Genetics*⁴ in 2017 on an ultrarare disease characterized by babies being born without noses and ocular defects; or genome-wide association studies⁵ (GWAS) such as the one also published by *Nature Genetics* on the genetic susceptibility of the glioblastoma.

Furthermore, CRG researchers developed eDIVA, a resource that allows clinicians to easily manage genomic data. eDIVA has already been used in, for example, a collaboration with a team at the Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS) to diagnose congenital ataxia and other congenital rare diseases. It has also been used in collaboration with researchers at Vall d'Hebron Institut de Recerca (VHIR) to study certain primary immunodeficiencies in babies and children.

More information and details from EGA at <https://ega-archive.org>

1 International Cancer Genome Consortium (<http://icgc.org>)

2 International Epigenome Cancer Consortium (<http://ihcc-epigenomes.org>)

3 UK10K 10,000 Genomes UK Project (<https://www.uk10k.org>)

4 Gordon T C et al. De novo mutations in SMCHD1 cause Bosma arhinia microphthalmia syndrome and abrogate nasal development. *Nature Genetics* (2017). DOI: 10.1038/ng.3765

5 S Melin et al. Genome-wide association with study of glioma subtypes identifies specific differences in genetic susceptibility to glioblastoma and non-glioblastoma tumors. *Nature Genetics* (2017). DOI: 10.1038/ng.3823

FOR MORE INFORMATION AND INTERVIEWS:

Laia Cendrós

Press Office

Centre for Genomic Regulation (CRG)

93 404 60 27 / 669 457 094

laia.cendros@crg.eu

Irene Roch

Communication Department

"la Caixa" Bank Foundation

iroch@fundaciolacaixa.es