



### Business challenge

Centre de Regulació Genòmica (CRG) wanted to help worldwide research institutions share genome sequencing data. How can such large amounts of data be transmitted quickly and cost-effectively?

### Transformation

With a high-speed IBM® Aspera® file transfer platform, CRG can share files at up to 3 Gbit/s, giving researchers rapid access to the data they need for their cutting-edge genetics projects.



David Camargo  
IT Director  
Centre de Regulació Genòmica

### Business benefits:

**Up to 500**

TB of discrete sequencing data easily packaged for sharing

**Up to 3**

Gbit/s transfer speeds achieved

**Supports**

leading-edge genetics research

## Centre de Regulació Genòmica supports cutting-edge research with lightning-fast file transfer

Created in 2000, Centre de Regulació Genòmica (CRG, in English: the Centre for Genomic Regulation) is a renowned international biomedical research institute. Employing over 600 people, the organization aims to advance knowledge for the benefit of society, public health and economic prosperity.

*“Working with Tmediat to deploy IBM Aspera was a key enabler of the European Genome-phenome Archive.”*

David Camargo  
IT Director  
Centre de Regulació Genòmica

Share this



## Bringing genetic data to a global audience

Around the world, researchers are working to identify how variation in the genetic makeup of individuals affects human health. As genetic sequencing technologies become faster, cheaper and more accurate, the volume of data available for analysis is growing rapidly. How can research institutions share this data quickly, conveniently and cost-effectively?

David Camargo, IT Director at Centre de Regulació Genòmica, takes up the story: “Although the cost of genetic sequencing is falling, collecting data for research into the detection, prevention and treatment of diseases still requires significant investment of time and resources.

“To help solve this challenge, CRG partnered with the European Bioinformatics Institute [EMBL-EBI] to improve the services of the European Genome-phenome Archive [EGA] to offer easier access for a global research audience. The goal was to enhance the EGA’s service for storing and sharing genomic and phenotypic data from biomedical research projects. This would enable research groups to leverage existing sequencing data that is relevant to their areas of study—helping to avoid duplication of effort, facilitating advanced research, and lowering the barriers to entry for smaller institutions.”

Ángel Carreño Torres, Systems and Operations Manager at Centre de Regulació Genòmica and the European Genome-phenome Archive, continues: “The EGA contains a range of different data formats—from raw DNA nucleotide sequences to more refined formats containing information about the way that sequences relate to a human reference genome.

“One of the immediate challenges that we faced was the large and fast-growing volume and variety of the EGA archive, which currently contains over two petabytes of data. To perform their work effectively, research institutions require huge amounts of data very quickly, which may be distributed across multiple small files, a small number of very large files, or a mixture of the two. In fact, one of our most requested studies consists of 500 terabytes of this genetic data. From an early stage, it was clear that traditional file transfer protocols would be unable to deliver the required performance.”

## Selecting a proven solution

For a number of years, the EMBL-EBI had relied on IBM Aspera High-Speed File Transfer solutions to meet its internal file transfer needs. When CRG joined the project, the organization decided to perform a proof-of-concept exercise to determine whether the IBM solution was still the right platform to achieve the project’s goals.

“Our colleagues at the EMBL-EBI had already used IBM Aspera widely, and we wanted to discover if we could extend the benefits that the EMBL-EBI had seen into our work on the EGA,” recalls Ángel Carreño Torres.

During the proof-of-concept phase, CRG engaged expert consultants from IBM Business Partner TmediaT to perform a deep test of the IBM Aspera solution. During this phase, CRG and TmediaT discovered, reported, and reproduced a number of technical issues, enabling the organization to fix them ahead of time and reduce the risk of delays to the project.

“Working with the TmediaT team was a very positive experience,” comments David Camargo. “They helped us to confirm that IBM Aspera would deliver the targeted levels of performance for the EGA. In addition, TmediaT ran a technical workshop to provide the details we needed about the Aspera API and development processes. Any investment in a major software component is a big decision, and our collaboration with TmediaT gave us the confidence that we had made the right choice.”

## Lightning-fast file transfer

Following a successful testing phase, a joint team from the EGA and CRG deployed IBM Aspera High-Speed File Transfer on a high-performance, hardened computing environment.

“DNA sequencing research can produce large amounts of data extremely quickly. What 15 years ago required almost USD3 billion of investment and 13 years of work, it’s now possible to sequence a human genome in a few hours certainly for less than USD1,000” adds Ángel Carreño Torres.

Today, the EGA and CRG have the secure, high-performance platform they need to curate and share valuable genetics data with institutions around the world.

“When a research team has gathered anonymized sequencing data from individuals whose consent agreements authorize data release for research use, they can upload it to the EGA via our online portal,” explains Ángel Carreño Torres. “Other research institutions can then apply to use the data, and these requests are reviewed by the organization that monitored the study in which the data was collected.

“Once a request is approved, we re-encrypt the data with a per-request private key, and load it into a container. The requester can then download the contents of the container using the IBM Aspera client, decrypt it in a secure vault, and perform their analyses.”

## Supporting cutting-edge research

Thanks to its IBM Aspera solution, the EGA and CRG are achieving their goal of making research data available to research groups around the world at speed and scale.

“Our transfer rates depend on the number of concurrent users accessing the system; typically, we deliver speeds of one gigabit per second, with peaks of up to three gigabits per second when demand for the system is higher than usual,” says Ángel Carreño Torres. “Although we have seen other file transfer solutions that can deliver these excellent levels of performance, none of them offered the ease of usability that we obtain from IBM Aspera. All what our requesters need to do is just install Aspera client on their system and start downloading their granted data.”

Ángel Carreño Torres continues:  
“It’s difficult to overstate the levels of performance that IBM Aspera can deliver.

“On one occasion, we noticed that a requester’s performance suddenly dropped, and we got in touch with them to find out if there was anything that we could do to help. It turned out that the file transfer was so fast and was taking up so much bandwidth that it had overwhelmed their network—so we had to continue the transfer at a slower rate! We’ve even heard that some of our requesters were able to secure investment to upgrade their networking hardware as a result of their positive experiences with IBM Aspera.”

“In the coming years, the institutions that rely on our data will continue to enhance their computing, network and storage infrastructure. We believe that our solution will enable them to maximize the value of their investment in these technologies by benefiting from the full file-transfer speeds that IBM Aspera can deliver.”

David Camargo concludes: “Working with TmediaT to implement IBM Aspera has been a key enabler of the EGA project—enabling us to achieve our goal of supporting leading-edge genetics research.”

*“It’s difficult to overstate the levels of performance that IBM Aspera can deliver.”*

Ángel Carreño Torres  
Systems and Operations Manager  
Centre de Regulació Genòmica and the  
European Genome-phenome Archive

## Solution components

- IBM® Aspera® High-Speed File Transfer

### Take the next step

To learn more about solutions from TmediaT, please visit: [tmediat.es](http://tmediat.es)

To learn more about the European Genome-phenome Archive, please visit: <https://ega-archive.org>

To learn more about IBM Aspera solutions, please contact your IBM representative or IBM Business Partner, or visit the following website: [ibm.com/software/aspera](http://ibm.com/software/aspera)

### Connect with us



© Copyright IBM Corporation 2016, IBM Corporation, 1 New Orchard Road, Armonk, NY 10504 U.S.A. Produced in the United States of America, April 2016.

IBM, the IBM logo, [ibm.com](http://ibm.com), and Aspera are trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the Web at “Copyright and trademark information” at [ibm.com/legal/us/en/copytrade.shtml](http://ibm.com/legal/us/en/copytrade.shtml).

This document is current as of the initial date of publication and may be changed by IBM at any time. Not all offerings are available in every country in which IBM operates. The performance data and client examples cited are presented for illustrative purposes only. Actual performance results may vary depending on specific configurations and operating conditions. THE INFORMATION IN THIS DOCUMENT IS PROVIDED “AS IS” WITHOUT ANY WARRANTY, EXPRESS OR IMPLIED, INCLUDING WITHOUT ANY WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND ANY WARRANTY OR CONDITION OF NON-INFRINGEMENT. IBM products are warranted according to the terms and conditions of the agreements under which they are provided.



ZZC03364-USEN-02

