



**Press Release  
CEA – Biofacet agreement**

**Genomics and Big Data: software to help you see more clearly**

**Évry, 10 October 2016** – The CEA’s Centre national de Génotypage (French national genome research platform) and the Company Biofacet have announced the start of the second phase of a high speed data processing and sequencing software development (NGS: Next Generation Sequencing). Following the validation of a pilot study carried out in 2015 on exomes<sup>1</sup>, the aim of this second phase is to produce a software platform for storing, organising and querying sequence variants from NGS data on complete genomes.

While sequencing programmes are developing in a sustained manner, it must be said that the mass of data collected remains difficult to exploit for the scientific community. The explosion in data production, coupled with the specificity of the field, prevent conventional database technologies from operating efficiently. Paradoxically, the mass of knowledge accumulated on myriads of national or international projects generates an intrinsic obstacle to its exploitation. No system exists that is able to exploit in a detailed manner and cross-reference the sequence variant data<sup>2</sup>. Pinpointing the variants of a genome and large scale sequence matching so as to detect common variant profiles constitutes a powerful research tool, and also a diagnostic aid for treating patients.

Given the severe limitations of the current systems, the CEA and Biofacet have developed the specifications of a database management system capable of very large scale storage and querying of national and international data banks of variants taken from the sequencing of exomes or complete genomes (WGS Whole Genome Sequencing). The resulting technology, contained in the Biofacet™ software, makes it possible to aggregate and query studies on thousands of samples. By optimised coupling of digital and phenotypic data, the technology developed allows more specifically:

- The “deep” querying of data, i.e. the possibility of establishing queries for all the values produced by the SNP-callers and for every genome position (3 billion nucleotides for the human genome);
- Queries mixing genotypes and phenotypes, thus allowing the targeting of potential variants of sequences linked to diseases;
- The incremental addition of samples.

For Jean-François Deleuze, director of CNG: *“The development of precision medicine cannot become a reality without tools capable of rigorously analysing variant data on a very large scale. By its massive coverage, Whole Genome sequencing introduces a technological breakthrough, not covered by the conventional “Big Data” tools. Producing and analysing these data routinely at the CNG for many years, we know all about the difficulty of managing*

---

<sup>1</sup> Exomes contain all the functional genes of the organism (1.5% of the genome)

<sup>2</sup> Variants are mutations that can play a part in the development of a pathology



these data. We are very pleased to collaborate with a French start-up in this high added value field”.

Jean-Jacques Codani, Chairman of Biofacet SAS: *“Although the Biofacet™ software has already obtained CLIA certification for clinical diagnostics on the other side of the Atlantic, the challenge posed by CNG is something else entirely. We took up this challenge because at the CNG we found two essential components for the deployment of this technology: firstly know-how in the production of NGS data and unquestionable scientific skills, and an excellent environment and technical expertise for high performance calculation”.*

As the pilot application was successfully deployed at the CEA, the partners envisage gradually increasing momentum involving the processing of thousands of WGS genomes in production at the CNG. In doing this, they will validate a software component that can meet the analytical challenges posed by the advent of genomic medicine, and more generally by the study of genetic variations which are significant for the bio-industry.

### **About CEA**

The French Atomic Energy and Alternative Energy Commission (CEA, Commissariat à l'énergie atomique et aux énergies alternatives) is a public research agency involved in four major fields: defence and security, nuclear and renewable energy, technological research for industry and fundamental research.

Relying on acknowledged expert capabilities, the CEA takes part in setting up collaborative projects with a large number of academic and industrial partners. With 16,000 researchers and employees, it is a major player in European research and enjoys a growing presence on the world scene.

Find out more: [www.cea.fr](http://www.cea.fr)

Press contact: Nicolas Tilly – +33 (0)1 64 50 17 16 – [nicolas.tilly@cea.fr](mailto:nicolas.tilly@cea.fr)

### **About Biofacet**

Biofacet SAS is a software publisher specialised in Bioinformatics.

Biofacet SAS develops and markets the Biofacet™ software, new generation technology that its founders have been developing since 1998, in the company GQ Life Sciences Inc. With an algorithmic process designed to solve the problems posed by the analysis of large scale genomic data, Biofacet guides its customers using its extensive expertise and experience of the international market.

Find out more: <http://www.biofacet.com/>