



**Date:** 23 November 2012

**Headline:** An international competition reaffirms the potential of bioinformatics in the diagnosis of disease

**Byline:** n/a

(Nanowerk News) The biosciences are generating enormous amounts of data at unprecedented speeds. Making sense of these data and extracting useful and reliable information from databases is an increasingly difficult and complex task. Backed by the scientific community, IBM Research y Philip Morris International (PMI) R&D launched a project called IMPROVER (Industrial Methodology for PROcess VERification in Research) in May with the aim of challenging the world's best computational researchers to demonstrate the power of their methods to exploit genomic information to extract predictive and clinical indicators that are reliable and verifiable.

The first challenge posed by IMPROVER, "Diagnosis Signature", consisted of classifying a group of patients for four diseases by using genomic data obtained from clinical samples in so-called "blind tests". The team led by IRB Barcelona and Anaxomics Biotech achieved fourth place worldwide, in the competition that saw the participation of 54 groups, mainly from Europe and the US. The project was announced in Nature Biotechnology and the results from the competition will be published in high impact journals.

According to David Rossell, head of the Biostatistics and Bioinformatics Unit at IRB Barcelona who designed the probability prediction algorithms, "this type of international challenge is very effective as a proof of concept, to demonstrate that it is possible to make predictions and to make them well, when the world's most advanced and effective techniques are used."

Patrick Aloy, Group Leader at the join IRB Barcelona-BSC program in computational biology, who specializes in molecular networks in disease added: "This competition fosters the creation of a scientific body of evidence for the biomedical community, including both researchers and clinicians, that adds value to and instills confidence in biocomputational techniques and their industrial applications.

**Using biological information in statistical algorithms**

The challenge consisted of using computational methods to evaluate and verify samples from patients with psoriasis, multiple sclerosis, chronic pulmonary obstruction and lung cancer. The IRB Barcelona team's method was two-fold: first, Anaxomics Biotech provided published information on proteins involved in each of the diseases, and Dr. Aloy expanded this biological documentation with unpublished data from other molecules that are known to be involved. Second, Dr. Rossell

integrated the biological information into his probability prediction tools. “It’s a question of separating the wheat from the chaff: we need to identify which subgroups of the thousands of genes we are dealing with are important in order to use them in our analysis. We need to combine our knowledge taken from databases with data we get from experiments to get the best results,” explains Rossell.

The first IMPROVER challenge has highlighted the predictive power of biocomputation given that the majority of participants were able to diagnose patients with at least 90% certainty, and in some cases 100%.

The organizers plan to continue the IMPROVER project for the next four years. They have already announced the second challenge, which will be launched in 2013.

-- ENDS --