

Guest lecture

Sitting on a “gold mine” of proteomics data: Now that the data is there for everyone, what can we do with it?

29 September 2017

9.30-10.30 AM in the BMB seminar room



Juan Antonio Vizcaino, PhD

**Proteomics Team Leader - Vizcaino Team: Proteomics
EMBL-EBI**

Abstract: First of all, I will describe the work that we have done in the last years to create a robust infrastructure to enable data sharing of mass spectrometry (MS) proteomics data in the public domain, including the development of the world-leading PRIDE database and related tools, open data standards and the establishment of the worldwide ProteomeXchange Consortium of proteomics resources, aiming to standardise data submission and dissemination practises in the field. Thanks, among other efforts, to the success of PRIDE and ProteomeXchange, the proteomics community is now widely embracing open data policies, an opposite scenario to the situation just a few years ago.

Then, I will explain how this plethora of data is being increasingly reused by the research community, e.g. in proteogenomics approaches, to build spectral libraries, for tool benchmarking or in innovative meta-analysis studies, among other applications. In this context of data reuse, we aim that PRIDE increasingly becomes a 'Hub' for proteomics data for all life scientists, by disseminating quality-filtered proteomics data into highly popular EMBL-EBI resources such as UniProt, Ensembl and the Expression Atlas.

We also aim to facilitate data reuse by third parties by building open, reproducible and scalable proteomics data analysis and quality control pipelines. As a proof of concept, we are using the OpenMS framework to build the first version of these pipelines, which are being deployed in the EMBL-EBI "Embassy Cloud", with the idea that in the future they can be made available in other cloud infrastructures, and that can be freely reused by any interested researcher in the community. These pipelines are connected to PRIDE, bringing the analysis tools closer to the data.

Host: Assistant Professor Veit Schwämmle, Department of Biochemistry and Molecular Biology, SDU.