ProteoRE, a Galaxy-based infrastructure for interpreting and exploring mass spectrometry-based proteomics data

Context and motivation
Concurrently with the increased simplicity associated with producing MS-based proteomics data, the bottleneck in many laboratories has shifted to reliable and reproducible interpretation of the data to extract meaningful knowledge. The ProteoRE (Proteomics Research Environment) project is a joint effort between the French proteomics infrastructure (ProFI) and the French bioinformatics Institute (IFB). Its primary aim is to centrally provide the proteomics community with an online research service enabling biologists/clinicians without programming expertise to explore their proteomics data through the Web in a reproducible manner. ProteoRE is built upon the Galaxy framework [1], a well-maintained software platform, providing simple interfaces to tools and online access to computational resources in a collaborative and transparent way.

Use case 1
Explore a human proteome by applying different filters, retrieving tissue expression information, comparing identified proteins to other lists [2], performing enrichment analysis [3] and browsing biological pathways from Reactome [4].

Use case 2
A step-by-step approach for protein biomarkers selection based on their physicochemical properties, tissue expression profiles, and subcellular localization, and which are suitable for mass spectrometry technology.

Future work
- Set up additional use cases and build new components dedicated to pathways analysis and quantitative MS data exploration
- Public release expected by early 2018 (proteore.org registered) Training sessions
- End-user training session addressed to the experimenter community via interactive tours, courses and “bring your own data” session

Conclusion
While Galaxy-based tools offers services for primary proteomics data analysis such as MS data conversion, protein database tools, search algorithms; tools focusing on downstream analysis are still lacking. Specifically addressed to biologists and clinicians, the ProteoRE platform is an attempt to fill this gap in a streamlined manner with the hope of promoting proteomics data in the Life Science community.

References

Acknowledgements
This work is partially funded through the French National Agency for Research (ANR) IFB project, “Infrastructures Nationales en Biologie et Santé”; “Investissements d’Avenir” call.

Contact
yves.vandenbrouck@cea.fr
valentin.loux@inra.fr