

CATEGORY: AI and Bioinformatics

PP05.01: Enhanced insight generation through automated transformation of historical experiments into Quantitative Knowledge Base

Wednesday, 20th September 2023. 11:45 - 13:00 KST

Anna Quagliari¹, Aaron Triantafyllidis¹, Bradley Green¹, Mark R. Condina^{1,2}, Paula Burton Ngov¹, Giuseppe Infusini¹ and Andrew I. Webb^{1,3,4}

¹Mass Dynamics, Melbourne, Victoria 3000, Australia

²Clinical & Health Sciences, University of South Australia, Adelaide, SA 5095, Australia.

³The Walter and Eliza Hall Institute of Medical Research, Parkville, Victoria 3052, Australia

⁴Department of Medical Biology, University of Melbourne, Melbourne, Victoria 3010, Australia

Following recent technological advances in proteomics, there has been a dramatic increase in the volume of proteomics datasets. As this wealth of data increases, it is plausible to suggest that new analyses will share facets with previous studies, such as design, cell lines, or tissue types, whether these studies are public or reside with lab ownership. However, the responsibility of querying these past experiments and harnessing potential insights for interpretation relies heavily on the researcher's personal knowledge and predispositions, thus potentially overlooking significant results that could otherwise be of relevance.

The Mass Dynamics platform is already engineered to build a quantitative knowledge-base from a user's or a lab's historical data as they analyze their experiments. It also provides the capacity to interrogate and explore the analysis of proteins across the user's past experiments and those publicly available on the platform. This combination of features facilitates extensive exploration and effective utilization of data.

Here, we present a novel set of features within the application, enabling users to merge multiple experiments into a unified analysis. This capability enhances the results interpretation of single experiments by combining multiple datasets within the application and their re-analysis using various statistical methods. Simultaneously, it provides users with the means to conveniently validate the merged dataset's quality through interactive visual tools and undertake multi-experiment evaluations. The platform also supports result interpretation in the context of the researcher's historical and public experiments, and annotation using external knowledge bases such as Reactome, STRING, and Gene Ontologies. Crucially, this set of tools creates an unprecedented capability that achieves comprehensive data exploration and insight generation.

To demonstrate the platform's capabilities, we have employed datasets from public repositories to illustrate not only the platform's ability to seamlessly merge and analyze experiments but also its facilitation of insightful data interpretation.