Multi-omic functional analysis to predict metabolic enzyme, kinases and transcription factor activity changes.

Aurelien Dugourd[1], Julio Saez-Rodriguez[1,2]

[1]RWTH Aachen University, Faculty of Medicine, Joint Research Centre for Computational Biomedicine (JRC-COMBINE), Aachen, Germany

[2]European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Hinxton, UK

In the past decade, quantitative omic dataset (transcriptomic, proteomic, phosphoproteomic and metabolomic) have been generated with increasing frequency in various experimental contexts. Thus it is critical to develop consistent methodologies to analysis and integrate such datasets in order to tap into their full potential. To address this challenge, we have developed an approach to recover and integrate functional information about the main actors of signaling and metabolism (kinases, metabolic enzymes and transcription factors) through analytical enrichment strategies and integration of inferred activity changes into multi-omic prior knowledge network with causal reasoning.

In particular, we have developed a new algorithm named MetaboMatch to perform functional analysis of metabolic enzyme by integrating metabolomic data with reaction networks. This method builds up on previous efforts made in the context of target sets enrichment analysis (such as kinase and transcription factor target enrichment analysis), using metabolic enzyme target sets specifically tailored to take into account the particular behavior of metabolic reaction networks (such as interdependency of metabolite concentrations and depletion/accumulation mechanisms). Thus MetaboMatch allows to easily infer metabolic enzyme activity changes from untargeted metabolomic data in a flexible range of experimental set-up, taking full advantage of the well characterised structure of metabolic reaction networks and preserving consistency between their expected behavior and the algorithm's statistical assumptions.

We applied this integration pipeline in the context of a multi omic dataset generated from kidney cancer and metastatic cell models to systematically derive mechanistic signature discriminating different disease states and pinpointing key molecular drivers of metastasis.