DOI: 10.1002/eji.202370300

European Journal of Immunology

**P 153**

## Pipeline for proteomic in silico cell-type deconvolution

*TUMURBAATAR T. [1,2], PELIN H. [3], BARROS DE ANDRADE E SOUSA L. [3], CHINTALAGIRI N. [1], MÜLLER-REIF J. [4,5], GEYER P. [4,5], GENZEL-BOROVICZÉNY O. [1], MANN M. [4,6], PIRAUD M. [3], KLEIN C. [1], NUSSBAUM C. [1], PANGRATZ-FUEHRER S. [1], KIM-HELLMUTH S. [1,2]*

*[1] Department of Pediatrics, Dr. von Hauner Children's Hospital, University Hospital LMU Munich, Munich, Germany; [2] Institute of Translational Genomics, Department of Computational Health, Helmholtz Munich, Munich, Germany, Munich, Germany; [3] Helmholtz AI, Helmholtz Zentrum München, Neuherberg, Germany; [4] Proteomics and Signal transduction, Max-Planck-Institute of Biochemistry, Martinsried, Germany; [5] OmicEra Diagnostics GmbH, Planegg, Germany; [6] Proteomics Program, Novo Nordisk Foundation Centre for Protein Research, University of Copenhagen Faculty of Health and Medical Sciences, Copenhagen, Denmark*

**Objective:** Analysing the human blood proteome grants us an in-depth look into the health status of an individual and opens many avenues for the development of novel diagnostic and therapeutic methods. However, affinity-based techniques such as ELISA and immunoblotting, which are commonly used for single protein quantifications, are increasingly being replaced by mass spectrometry (MS)-based methods that can provide comprehensive results in even small sample volumes. Despite the attractiveness of MS-based workflows, there are prominent issues that still need to be addressed, most notably, the abundance of missingness in MS datasets as well as the lack of cell type resolution in bulk tissues MS. Therefore, our goal is to develop a pipeline that is able to impute these missing values and perform cell type deconvolution from MS blood proteomics data.

**Methods and results:** Based on pre-existing MS proteomics datasets in addition to comparative studies on the effectiveness of various mathematical imputation and deconvolution methods, we developed an R-based pipeline that is capable of imputing missing protein expression values and furthermore, with the help of a signature matrix, calculate the proportions of specific cell types within a given sample. The pipeline is modular, allowing for customization and sample-specific optimization. Various normalization and data preparation workflows are included as well. Accuracy was assessed by comparing with flow cytometry data of the same samples.

**Conclusions:** Implementing multiple imputation and deconvolution methods our pipeline can serve as a valuable tool for working with blood MS proteomics data. This enables the characterization of protein abundance in immune cells without the need to isolate them or in samples where sample volume is a limiting factor.

 