

Project log – week 46

Analysis of new Olink proteomics data from S3WP: toward multi-omics integration

Overview

The data exploration for both Olink Explore and Olink HT data sets has been going well, with the ongoing QC of the Olink HT. The platform comparison will be processed simultaneously.

Summary of activities and deliverables reached

- No meeting with my supervisor María Bueno Álvarez this week due to schedule conflicts, we maintain continuous contact via Slacks
 - o Discuss data set details and QC effectiveness
 - o Review project plan in case updates are needed
 - o Data transfer and handling
- Project activities
 - o Review the R code from previous publications (2, 3, 4)
 - o Data discovery: investigating metadata of participants, fixing protein name conflicts, combining/extracting relevant data, etc.
 - o QC programming for Olink HT data set
 - o Initial data comparison between Olink Explore and HT data sets

Summary of results

No results are available yet.

Reflection task

The central method of this project is computational plasma protein analysis, which includes differential expression, multivariate analyses, protein correlation across longitudinal samples, and pathway enrichment. With this analysis, we try to extract as much information from a complete plasma protein profile across different visits and changes in the lifestyle of participants. Hence, this method provides a comprehensive analysis of the longitudinal data, with the potential to discover important biological insights and understand the plasma proteomics profile. The drawback of this approach is that computational methods depend on certain hypotheses and data quality. There is no guarantee that the data will follow a normal distribution for some statistical tests.

Alternative methods for this project might be: nonparametric statistical methods for expression analysis, network-based analysis for protein-protein interaction or applied machine learning for automatic analysis of protein differential expression, protein correlation and pathway enrichment. We did not consider nonparametric statistical methods because the data normality was confirmed by the Olink QC report and our own QC and assessment. We chose to do R programming analysis for a good statistical analysis platform, and its effectiveness has been proven by prior publications.

Updated project plan

The project plan remains unchanged with corresponding deliverable results per step (table 1).

Table 1. Project plan

		Week										
Tasks	Deliverables	44	45	46	47	48	49	50	51	52	01	02
Literature search	Understand the research topic via prior publications											
Method exploration	Discover and test essential R packages for the analysis											
QC & pre-processing	QC and normalization for batch/technical bias											
Platform comparison	Compare data from Olink Explorer (3) and (new) Olink HT											
New proteomics analyses	Multivariate analyses, protein correlation across longitudinal samples, pathway enrichment											
Result interpretation	Visualization of the analysis result											
Presentation	Half-time and final presentation											
Report	Project report											

Reference

- (1) Bergström G, Berglund G, Blomberg A, Brandberg J, Engström G, Engvall J, et al. The Swedish CARDioPulmonary BioImage Study: objectives and design. *Journal of Internal Medicine* [Internet]. 2015 Jun 19;278(6):645–59. Available from: <https://doi.org/10.1111/joim.12384>
- (2) Tebani A, Gummeson A, Zhong W, Koistinen IS, Lakshmikanth T, Olsson LM, et al. Integration of molecular profiles in a longitudinal wellness profiling cohort. *Nature Communications* [Internet]. 2020 Sep 8;11(1). Available from: <https://doi.org/10.1038/s41467-020-18148-7>
- (3) Zhong W, Edfors F, Gummeson A, Bergström G, Fagerberg L, Uhlén M. Next generation plasma proteome profiling to monitor health and disease. *Nature Communications* [Internet]. 2021 May 3;12(1). Available from: <https://www.nature.com/articles/s41467-021-22767-z>
- (4) Wang J, Zenere A, Wang X, Bergström G, Edfors F, Uhlén M, et al. Longitudinal analysis of genetic and environmental interplay in human metabolic profiles and the implication for metabolic health. *medRxiv* (Cold Spring Harbor Laboratory) [Internet]. 2024 Sep 24; Available from: <https://www.medrxiv.org/content/10.1101/2024.09.23.24314199v1>