Master thesis synopsis

Title: The housekeeping proteome: identifying proteins essential for human cells

Student
Name: Thi Huyen Mai Nguyen
Email: maihuyen.nha@gmail.com

Primary supervisor: Evelina Sjöstedt
Name: Dr Evelina Sjöstedt
Email: evelina.sjostedt@scilifelab.se

Co-supervisors: Mathias Uhlen, Åsa Sivertsson, Maria Bueno Alvez

Research group
The Human Protein Atlas – Uhlén Lab/Edfors Lab
KTH Royal Institute of Technology
School of Engineering Sciences in Chemistry, Biotechnology, and Health
Department of Protein Science – Division of Systems Biology
Situated at Science for Life Laboratory, Alfa 6

Project goal

To characterize the housekeeping proteome, a subset of proteins essential for human cell survival and function, as part of the Alpha Cell program by the Human Protein Atlas. The aim is to define the proteomic landscape of proteins ubiquitously expressed across dividing human cells, while distinguishing essential proteins from those that are ubiquitously expressed but non-essential due to functional redundancy (i.e., the cell can survive their knockout because other proteins compensate for their function).

Project outline

The project consists of four main phases:

1. **Defining criteria for housekeeping proteins.**
   Establish a robust framework to identify housekeeping proteins, focusing on proteins expressed across all dividing human cells. This phase includes literature review and data analysis to define essentiality.
2. **Identification of essential proteins.**
   Using proteomic data from the Human Protein Atlas (HPA) Alpha Cell program, identify proteins that meet the criteria for being essential. This includes addressing cases of functional redundancy where knockout (KO) of a protein does not affect survival due to compensation by other proteins.

3. **Validation and refinement of the housekeeping proteome.**
   Develop and apply filtering strategies to validate the list of housekeeping proteins. This involves analyzing data robustness and testing the hypothesis across diverse datasets to confirm essentiality and ubiquity.
4. **Integration and communication of findings.**
   Visualize and present the finalized housekeeping proteome in an accessible format, ensuring the results can be utilized effectively by users of the HPA resource. Create clear documentation to distinguish between truly essential proteins and ubiquitously expressed, non-essential proteins.

Time plan

- **Project planning:** Now – end of January
- **Phase 1 – Literature review and criteria definition:** February
- **Phase 2 – Identification of essential proteins:** February – March
- **Phase 3 – Validation and refinement:** March – April
- **Phase 4 – Visualization and reporting:** April – May
- **Writing:** May

Tools

- R
- RStudio
- Human Protein Atlas (HPA) resource
- UniProt and related protein databases