Lecture Title and Date

DATA - Proteomics I - 1/15

Objectives of the Lecture

By the end of this lecture, students should be able to:

- know the central dogma of molecular biology and understand the flow of information translation, and how it relates to proteomics.
- know the major challenges prevent complete proteome analysis
- know the definition of proteomics and protein-protein interactions
- Understand the principles of the Mass Spectrometry
- Understand the fundamental mechanisms and work-flow of using Mass Spectrometry to detect/sequence proteins
- know the basics of protein-protein interaction
- know the basics of single cell proteomics & Spatial Proteomics

Key Concepts and Definitions

- Central dogma of molecular biology: DNA → RNA → PROTEIN
- The definition of **proteomics**: the study of the expression, location, modification, interaction, function, and structure of all the proteins in a given cell, organelle, tissue, organ, or whole organism
- Shotgun Proteomics: Bottom-up proteomics approach involving protein digestion followed by LC-MS/MS analysis

Main Content/Topics

- major challenges prevent complete proteome analysis
    - Proteomics is a **sample limited** (even till now it is still not comprehensive at all, unlike DNA samples that can be amplified powered by DNA polymerases), proteomics has no "polymerase" or amplification method, thus it must content with natural abundance
    - Mass spectrometry has limitations, it can not provide full amino-acid resolution of a proteome yet
- Mass Spectrometry

- - Principles: Measure m/z (mass-to-charge) in a mass spectrum; measure peptide m/z for proteins; sample must be ionizable in order to be analyzed
  - Ionization techniques enabled direct measurement and sequencing of intact peptides & MS based Proteomics -> success of mass spectrometry in the life science
    - Two ionization technique: Electrospray Ionization (ESI) & Matrix Assisted Laser Desorption Ionization (MALDI)

Suggest references for many of the key concepts

- Wilhelm, Mathias, et al. "Mass-spectrometry-based draft of the human proteome." *Nature* 509.7502 (2014): 582-587.
- ProteomicsDB: https://www.proteomicsdb.org/
- PeptideAtlas: http://www.peptideatlas.org/docs/protein_ident_terms.php

Additional suggested reference:

- Schilling, Oliver, et al. "Proteome-wide analysis of protein carboxy termini: C terminomics." *Nature methods* 7.7 (2010): 508-511.

Discussion/Comments

- The lecture discussed the role of databases in organizing and interpreting proteomics data. The use of databases like ProteomicsDB and PeptideAtlas helps in the real-time analysis and integration of massive proteomics datasets, which is crucial for advancing our understanding of protein functions and interactions.
- The current pace of proteomics is behind genomic technologies, highlighting the need for advances in mass spectrometry and protein analysis techniques.