Prime Scholars Library

Advances in Life Science and Biotechnology

Vol.11 (1), pp.01-02, June, 2023

Commentary

Available online at https://primescholarslibrary.org/

Proteomic technology limitations for global protein isoform

Jeon Jongbum*

Department of Agricultural Biotechnology, Seoul National University, Seoul, Korea

Received: 30-Mar-2023, Manuscript No. ALSB-23-93599; **Editor assigned:** 03-Apr-2023, PreQC No. ALSB-23-93599 (PQ); **Reviewed:** 17-Apr-2023, QC No. ALSB-23-93599; **Revised:** 29-May-2023, Manuscript No. ALSB-23-93599 (R); **Published:** 05-Jun-2023, DOI**:** 10. 51268/2736-1837.23.11.081.

DESCRIPTION

Proteomics is a rapidly evolving field that is essential to their knowledge of biological systems. Despite recent advances in proteomic technologies, detecting protein diversity on a global scale remains a major challenge. Shotgun proteomics, the most frequently used method for large scale protein identification, has limitations in distinguishing between protein variants and isoforms. This study investigates the limitations of shotgun proteomics for the global detection of protein isoforms and discusses alternative methods for achieving proteoform level detection. The importance of alternative splicing, post-translational changes, and single amino acid polymorphisms in increasing proteome complexity is also emphasized. A typical shotgun proteomics experiment identifies 10,000 human proteins from a single sample. Individual proteins, on the other hand, are usually identified by peptide sequences that represent a tiny fraction of their total amino acid sequence. As a result, a typical shotgun experiment fails to distinguish between distinct protein variants and isoforms. For the global discovery of protein isoforms, deeper proteome sequencing is needed. They identify a million unique peptides from 17,717 protein groups using six different human cell lines, six proteases, deep fractionation, and three tandem mass spectrometry fragmentation methods, with a median sequence coverage of roughly 80%. Most nonsynonymous variants are translated, according to a direct match with RNA expression data. They also hypothesized that undetected variations are most likely caused by mutation induced protein instability. Furthermore, detection rates for exon-exon junction peptides representing constitutive and alternative splicing events are similar. Mass Spectrometry (MS) can identify nearly complete proteomes of simple organisms after only 1 hour of analysis. Over 10,000 proteins can be monitored in a day for more complicated organisms. Community based maps of the human proteome, created using extensive data from different tissues and cell types from labs around the world, have shown that more than 90% of annotated protein coding genes are translated. Despite the fact that the human genome includes approximately 20,000 protein coding genes, alternative splicing events, in which precursor messenger RNA sequences are combined in various arrangements, are thought to have the potential to significantly increase proteome diversity. According to reports based on RNA sequencing (RNA-seq) analysis of human organs, transcripts from more than 95% of multi exon genes experience alternative splicing. Furthermore, single cell transcriptome sequencing has revealed that true splice isoform complexity is likely to be higher than previously thought. Other forms of proteome variation, such as Single amino Acid Polymorphisms (SAPs),

alternative splicing, and posttranslational modifications, add to the complexity of the proteome.

Proteomic technology limitations have prevented the detection of protein diversity on a global basis. Shotgun proteomic methods typically identify the presence of an entire protein using a small number of peptide proxies-as few as two or three. As a result, in most cases, sequence coverage in a proteomics experiment is inadequate to completely characterize all protein states present in a sample. However, understanding biological systems requires the capacity to accurately monitor protein isoforms. Even the most comprehensive proteomic datasets currently lack of sufficient sequence data to define proteoform globally. Top-down MS, a technique that measures intact protein mass before dissociation for sequence determination using tandem mass spectrometry, is one approach to achieving proteoform level detection. The top-down approach is appealing because it ensures no loss of resolution.

However, practical problems with high mass proteins, sequence coverage, and detection of low abundance species limit its application.

CONCLUSION

Given the technological challenges of top-down proteomics, they reconsidered the shotgun approach. Shotgun proteomics depends heavily on trypsin to catalyse protein hydrolysis. Trypsin cleaves C-terminal to lysine and arginine residues, yielding peptides with the best length and charge distribution for MS. Even with extensive chromatographic separation, not all portions of the proteome are accessible from tryptic peptides, and many of the peptides produced are either too short or too long to be detected using modern Liquid Chromatography-Mass Spectrometry (LC-MS) technology. Because proteoform can vary by only a few amino acids, extensive sequence coverage is required to distinguish near identical variants.