



Proteogenomic Analysis: Integrating Genetic and Protein Information for Advanced Molecular Profiling

Mei Lin *

Department of Molecular Medicine, Peking International University, Beijing, China

DESCRIPTION

Proteogenomic analysis combines genomic and proteomic information to examine biological systems from both genetic and protein perspectives. Genomics provides information about Deoxyribonucleic Acid (DNA) sequences, genetic variants, mutations and gene organization, while proteomics examines proteins, their abundance, molecular modifications, interactions and functional states. Bringing these data types together can provide a more complete view of how genetic information is expressed at the protein level and how molecular changes influence biological functions. DNA variation does not always produce a direct or predictable change in protein activity. A genetic variant may alter the amino acid sequence of a protein, affect gene expression, influence Ribonucleic Acid (RNA) processing, or have little measurable effect. Protein abundance can also be influenced by cellular signaling, degradation, environmental conditions and post-translational modifications. Proteogenomic analysis therefore connects genomic information with the actual protein molecules present in cells or tissues.

Genomic analysis commonly involves sequencing DNA to identify genetic alterations. These alterations may include single-nucleotide variants, insertions, deletions, copy-number changes and larger structural changes. Depending on the biological system, sequencing may focus on selected genes, coding regions, or the complete genome. The resulting information can identify molecular changes that may influence protein structure, expression, or function. Proteomic analysis provides information about proteins produced by cells. Mass spectrometry is widely used to identify and quantify proteins from complex biological samples. Proteins can be enzymatically digested into peptides before analysis and the resulting peptide signals can be matched with protein databases. Quantitative methods can compare protein abundance between different biological conditions and identify proteins that increase or decrease in association with particular molecular states.

One important application of proteogenomics is the identification of protein products associated with genomic variants. A DNA alteration that changes a coding sequence may generate a modified peptide containing a specific amino acid substitution. Mass spectrometry can sometimes detect such peptides directly. This provides evidence connecting a genomic alteration with its corresponding protein product. Cancer biology is a major application of proteogenomic analysis. Tumor cells can contain numerous genetic alterations, but not every alteration produces an active protein-level effect. Integrating DNA mutations with protein abundance and signaling information can help distinguish molecular changes that have functional consequences. This can support the characterization of tumor subtypes and provide information relevant to treatment selection.

Proteogenomic analysis can also examine signaling pathways. Many cellular pathways depend on protein phosphorylation and other post-translational modifications. A mutation may activate a signaling protein, but its functional state can only be fully understood by examining protein abundance and modification patterns. Phosphoproteomic analysis can therefore complement genomic information by showing which signaling pathways are active within a biological system. Protein isoforms provide another important area of analysis. Alternative RNA processing can produce different protein forms from the same gene. These isoforms may have different functions, cellular locations, or interactions with other molecules. Proteomic measurements can identify peptides that distinguish between protein isoforms, while genomic and transcriptomic information can provide information about the sequences and expression mechanisms responsible for them.

Proteogenomics can also improve genome annotation. Some predicted genes may lack direct evidence of protein production. Detection of peptides corresponding to predicted coding regions can provide additional evidence that a genomic region produces a protein. This can be particularly useful when examining

Correspondence to: Mei Lin, Department of Molecular Medicine, Peking International University, Beijing, China, E-mail: mei.lin@proteogen.cn

Received: 30-May-2025, Manuscript No. JPP-25-32592; **Editor assigned:** 02-Jun-2025, PreQC No. JPP-25-32592 (PQ); **Reviewed:** 16-Jun-2025, QC No. JPP-25-32592; **Revised:** 23-Jun-2025, Manuscript No. JPP-25-32592 (R); **Published:** 30-Jun-2025, DOI: 10.35248/2153-0645.25.16.138

Citation: Lin M (2025). Proteogenomic Analysis: Integrating Genetic and Protein Information for Advanced Molecular Profiling. J Pharmacogenom Pharmacoproteomics. 16:138.

Copyright: © 2025 Lin M. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution and reproduction in any medium, provided the original author and source are credited.

organisms with incomplete or newly generated genome annotations. In infectious disease biology, proteogenomic analysis can characterize genetic variation in microorganisms together with their protein expression. Pathogens can contain genomic changes that influence protein structure, metabolism, virulence-related functions, or antimicrobial sensitivity. Protein measurements can help determine whether these genetic changes result in altered cellular activity. Such information can contribute to improved understanding of microbial adaptation and treatment response.

Pharmacology also benefits from integrated genomic and proteomic analysis. Genetic variants can influence drug targets, metabolic enzymes, transport proteins and signaling pathways.

Protein measurements can determine whether the corresponding molecules are actually present and active. Combining these data can provide more complete information about drug response than genomic or proteomic information alone. Proteogenomic analysis requires extensive computational processing because genomic and proteomic datasets differ in structure and scale. DNA sequencing produces nucleotide sequences and variant information, whereas mass spectrometry generates peptide-level measurements. Specialized databases can connect genomic variants with predicted protein sequences and theoretical peptide fragments. Computational pipelines can then compare observed peptide signals with predicted sequences to identify variant-associated proteins.