Cureus
Part of SPRINGER NATURE

Open Access Review Article

Proteogenomic Signatures in Complex Diseases: Implications for Targeted Interventions

Preeti Sharma [1], Rachna Sharma [2], Upasana Uniya [3], Eragam Siva Rami Reddy [4], Amit Nampalliwar [5], Karthika Padmavathy [6], Rahima Malek [7]

1. Department of Biochemistry, Amar Shaheed Jodha Singh Attaiya Thakur Dariyao Singh (ASJSATDS) Medical College and Hospital, Fatehpur, IND 2. Department of Biochemistry, Government Medical College, Karauli, IND 3. Department of Pathology, Gandhi Medical College, Bhopal, IND 4. Department of Medicine, Sri Adi Siva Sadguru Alli Saheb Sivaaryula Ayurvedic Medical College, Guntakal, IND 5. Department of Roga Nidan and Vikriti Vigyan (Diagnostic Procedure and Pathology), Government Ayurved College and Hospital, Bilaspur, IND 6. Department of Pathology, Sri Lalithambigai Medical College and Hospital, Dr. MGR Educational and Research Institute, Chennai, IND 7. Department of Biochemistry, Dr. N.D. Desai Faculty of Medical Science and Research, Dharmsinh Desai University, Nadiad, IND

Corresponding author: Preeti Sharma, ps9713182@gmail.com

Review began 08/22/2026
Review ended 09/06/2026
Published 10/22/2026

© **Copyright** 2026
Sharma et al. This is an open access article distributed under the terms of the Creative Commons Attribution License CC-BY 4.0., which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

DOI: 10.7759/cureus.118351

Abstract

Complex diseases arise from interactions among genetic susceptibility, epigenetic regulation, transcriptomic variation, protein expression, environmental exposure, and clinical phenotype. Background evidence indicates that genomic information alone may not fully explain disease progression, molecular heterogeneity, or therapeutic response, as functional disease mechanisms are frequently executed at the protein level. The objective of this review is to synthesize current knowledge on proteogenomic signatures in complex diseases and evaluate their implications for biomarker discovery, risk stratification, clinical decision-making, and targeted interventions. The methodology involved an integrative literature review of English-language studies published between 2015 and 2026, identified through PubMed, Scopus, Web of Science, ScienceDirect, and Google Scholar using terms related to proteogenomics, multi-omics, complex diseases, precision medicine, biomarkers, and targeted therapy. Eligible studies included original articles, reviews, meta-analyses, and consensus reports addressing human complex diseases with proteomic integration. The review highlights proteogenomic applications in cancer, cardiometabolic disorders, neurodegenerative diseases, autoimmune conditions, and inflammatory disorders. Integrated molecular profiling supports identification of active disease pathways, clinically relevant biomarkers, therapeutic vulnerabilities, and molecular subtypes beyond conventional diagnostic categories. In conclusion, proteogenomic signatures offer a promising route toward mechanism-guided precision medicine, although clinical translation requires standardized workflows, diverse validation cohorts, ethical governance, and cost-effective integration into healthcare systems.

Categories: Genetics, Pathology, Oncology
Keywords: biomarkers, complex diseases, precision medicine, proteogenomics, targeted therapy

Introduction And Background

Multifactorial diseases, characterized by different clinical manifestations and responses to therapy, pose a challenge in modern medicine [1]. Diseases such as cancer, cardiovascular diseases, diabetes mellitus, neurodegenerative diseases, autoimmune diseases, chronic kidney diseases, and inflammatory syndromes result from interactions among inherited susceptibility, somatic variation, epigenetic regulation, environmental factors, metabolic status, immune system activity, and tissue-specific molecular adaptation [2]. The classic approach to disease classification has been mainly based on clinical phenotype, histopathology, biochemical markers, and single genes. While these strategies continue to play a key role in diagnosis and treatment, these models can fail to explain the initiation of disease, disease progression, disease severity, disease recurrence, or treatment resistance [3].

Genomic technologies have revolutionized the knowledge of complex disease mechanisms through the identification of pathogenic variants, susceptibility loci, copy number variations (CNVs), gene fusions, mutational signatures, and inherited risk patterns [4]. Molecular characterization has been expanded and is now possible genome-wide using genome-wide association studies, next-generation sequencing, and transcriptomic profiling in a variety of disease states [5]. Despite these advances, the abundance of proteins, their function, activation of pathways, or therapeutic vulnerability cannot be consistently predicted by DNA- or RNA-based data [2]. A genetic variant might be of no biological effect, have a context-dependent effect, or affect the behavior of the disease through a molecular network in the downstream parts of the molecular system that is not completely inferable from the sequence [6]. The final functional phenotype of disease is achieved at the protein level by enzymes, receptors, transcription factors, signaling mediators, immune regulators, transporters, structural proteins, and post-translationally modified molecular complexes [7].

Proteogenomics has become an interdisciplinary field of research that integrates genomic and transcriptomic information with protein expression, protein isoforms, post-translational modifications, and

protein pathway activity [4]. This strategy offers a more functional understanding of disease biology by linking molecular variation with cell function and clinical phenotype [1]. In oncology, integrated proteogenomic profiling can contribute to the characterization of tumor heterogeneity, immune evasion, oncogenic signaling, neoantigen formation, and mechanisms of drug resistance [8]. Proteomic signatures related to inflammation, endothelial dysfunction, lipid metabolism, insulin resistance, and myocardial remodeling are correlated with cardiometabolic disorders and could be useful for risk stratification [4]. Proteomic patterns and proteogenomic signatures where genomic features are explicitly integrated with protein-level evidence may aid early diagnosis, mechanistic classification, and identification of therapeutically relevant pathways in neurodegenerative and autoimmune diseases [9].

Closely related reviews have addressed complementary aspects of multi-omics precision medicine. Abdelaziz et al. reviewed computational pipelines for multi-omics data integration, including preprocessing, integration strategies, analytical methods, and translational applications in precision medicine [10]. Nam et al. examined artificial intelligence-enabled multimodal omics integration and its potential to combine heterogeneous molecular, clinical, and phenotypic information [11]. Akhoundova and Rubin focused on advanced multi-omics tumor profiling, including genomic, transcriptomic, proteomic, spatial, and liquid-biopsy approaches for precision oncology [12]. Fathi et al. highlighted the role of molecular imaging in gastrointestinal cancers for diagnosis, prognosis, treatment planning, and longitudinal assessment [13]. In contrast, the present review places proteogenomic signatures at its core and extends the discussion beyond oncology to cardiometabolic, neurodegenerative, autoimmune, and inflammatory diseases, with emphasis on linking genomic alterations and protein-level consequences, biomarker development, risk stratification, and targeted interventions. Tissue-based proteogenomic information may also complement quantitative molecular imaging by providing biological context for disease characterization, treatment selection, and longitudinal monitoring. The clinical relevance of proteogenomics lies in its potential to connect molecular variation with functional disease mechanisms and clinically meaningful phenotypes. Integrated molecular signatures may support biomarker discovery, molecular subclassification, risk assessment, therapeutic target identification, and treatment-response evaluation. This approach is particularly relevant to diseases characterized by overlapping phenotypes, heterogeneous progression, and variable therapeutic response, where mechanism-based interpretation may complement conventional phenotype-based classification.

Still, major gaps remain in clinical translation. Proteogenomic datasets are often disease-specific, population-limited, technically heterogeneous, and difficult to compare across platforms. Standardized workflows for sample processing, mass spectrometry, sequence integration, bioinformatics, validation, and reporting are still evolving. Limited population diversity, sparse longitudinal data, variable biomarker reproducibility, high infrastructure costs, and uncertain reimbursement also restrict implementation. These barriers are particularly important in emerging and resource-limited countries, where limited access to advanced technologies, specialized expertise, and sustainable funding may widen disparities in proteogenomic care. Concerns regarding genomic privacy, incidental findings, consent, algorithmic bias, and equitable access must also be addressed before widespread clinical adoption.

Objective

This review aimed to synthesize current knowledge on proteogenomic signatures in complex diseases, with emphasis on molecular mechanisms, biomarker potential, and clinical translation. It also aimed to evaluate implications for targeted interventions, precision medicine, and future integration into diagnostic and therapeutic decision-making.

Methodology

We conducted an integrative literature review to synthesize evidence on proteogenomic signatures in complex diseases and their implications for biomarker discovery, risk stratification, clinical decision-making, and targeted interventions. Studies published in English from 2015 through 2026 and available at the time of the literature search were considered. Electronic searches were conducted using bibliographic databases and academic research platforms, including PubMed, Scopus, Web of Science, ScienceDirect, and Google Scholar. PubMed, Scopus, and Web of Science were used as bibliographic databases, whereas ScienceDirect and Google Scholar were used as supplementary academic search platforms. Search terms were combined using Boolean operators, including (“proteogenomics” OR “proteogenomic signatures” OR “multi-omics”) AND (“complex diseases” OR “cancer” OR “cardiometabolic disease” OR “neurodegenerative disease” OR “autoimmune disease”) AND (“precision medicine” OR “targeted therapy” OR “biomarkers”). We also screened reference lists of relevant publications to identify additional literature.

Studies were eligible if they focused on human complex diseases, integrated proteomic evidence with genomic, transcriptomic, or other molecular data, and addressed disease mechanisms, biomarker discovery, risk stratification, therapeutic targeting, or clinical translation. Consistent with the integrative review design, original human studies were the principal evidence base, while systematic reviews, narrative reviews, meta-analyses, and consensus statements from recognized scientific or professional organizations were used where relevant to provide contextual or interpretive evidence. Studies were excluded if they were non-English, unrelated to the review objective, limited to single-omics analysis without proteomic integration, restricted to non-human experimental models without direct clinical relevance, conference abstracts,

editorials, letters, or duplicate publications. We also excluded publications that did not provide sufficient information on the disease population or context, molecular methodology, or clinically relevant findings needed for interpretation in the review.

Review

Conceptual basis of proteogenomics in complex disease biology

Proteogenomics is an integrative approach that links genomic variation with proteomic evidence, often using transcriptomic data as an intermediate layer. Broader integration of genomics, transcriptomics, proteomics, metabolomics, epigenomics, and clinical data is more appropriately described as multi-omics integration unless genomic features are explicitly connected to protein-level findings [14]. Genomic profiling can detect variants, mutations, structural rearrangements, and copy number alterations; transcriptomic analysis can assess gene expression and splicing; proteomics can quantify protein abundance, isoform diversity, and post-translational modifications, while pathway activity is generally inferred through bioinformatic interpretation of these molecular measurements [14]. The combined use of genomic, transcriptomic, and proteomic information can help identify whether molecular alterations are expressed at the RNA and protein levels and whether they are associated with clinically relevant disease phenotypes, biomarkers, or therapeutic targets [14]. Changes in the genome are not always reflected at the protein level [15]. Single-nucleotide variants (SNVs) may be synonymous, missense, nonsense, splice-altering, regulatory, or located in untranslated regions; only some alter the amino acid sequence, and even missense variants do not necessarily affect protein stability or function [15]. Copy number alterations may influence gene dosage, but final protein abundance can also be modified by transcriptional regulation, RNA degradation, translation efficiency, and protein turnover [4]. Structural variants can generate fusion proteins, disrupt regulatory regions, or produce novel molecular products [16]. Proteogenomic analysis can demonstrate whether genomic alterations are transcribed and translated by detecting corresponding RNA and peptide evidence, but such detection alone does not establish that the resulting protein product is functionally or clinically pathogenic [9].

Another layer of complexity is that alternative splicing may produce more than one protein product from a single gene. These isoforms can be differently localized, bind differently, have different catalytic activity, be recognized differently by the immune system, or be differently sensitive to drugs. Neoantigens can be created due to aberrant splicing and somatic mutations in cancer, which can help in immunotherapy target discovery [17]. Changes in protein networks can signal inflammatory, metabolic, neurological, endothelial, and/or mitochondrial pathway disruption in autoimmune, neurodegenerative, cardiovascular, and metabolic diseases, even before clinical symptoms appear [6]. Proteogenomic interpretation is greatly dependent on post-translational modification. Activity changes in proteins can occur without alteration of DNA sequence, such as through the process of phosphorylation, acetylation, methylation, ubiquitination, glycosylation, and proteolytic cleavage. These changes affect processes like signal transduction, immune activation, cell-cycle regulation, cell death, metabolism, and cellular stress response [14]. These protein-level events can be mapped to identify active disease mechanisms and actionable therapeutic nodes.

Proteogenomics serves as a link between molecular variation and clinical decision-making. It can identify passive genetic signatures versus active drivers of disease, categorize molecular subtypes, aid in the discovery of biomarkers, and direct targeted intervention. For complex diseases, this paradigm offers a new interpretation - not just a sequence change, but how pathways function and, consequently, enhances understanding of mechanisms and their application in precision medicine [16]. Table *1* summarizes the major molecular components supporting proteogenomic interpretation.

Molecular feature	Genomic/transcriptomic alteration	Proteomic consequence	Clinical relevance	Example disease context	References
SNV	Single-nucleotide sequence change that may alter coding or regulatory regions	May produce an altered peptide sequence or affect protein abundance/stability when translated	Variant prioritization and potential target selection	Cancer, inherited cardiomyopathy	[16]
CNA	Increase or decrease in genomic copy number	May be associated with increased or reduced protein abundance, although this is not always proportional to gene dosage	Prognostic stratification and pathway assessment	Breast cancer, ovarian cancer	[17]
SV	Genomic rearrangement, including fusion or disruption of regulatory regions	May generate fusion proteins or altered protein products when expressed and translated	Potential fusion-guided therapeutic selection	Lung cancer, leukemia	[15]
AS	Transcript-level generation of alternative RNA isoforms	May produce distinct protein isoforms with altered localization, stability, or function	Isoform-based biomarker development and response assessment	Cancer, neurodegeneration	[15]
PTM	Protein-level modification rather than a genomic or transcriptomic alteration	Phosphorylation, acetylation, methylation, ubiquitination, glycosylation, or proteolytic processing	Pathway assessment, therapeutic targeting, and treatment monitoring	Cancer, autoimmune disease	[14]

TABLE 1: Genomic, transcriptomic, and proteomic features linking molecular alterations with clinical interpretation.

SNV: single-nucleotide variant; CNA: copy number alteration; SV: structural variant; AS: alternative splicing; PTM: post-translational modification

Molecular architecture of complex diseases: from genotype to phenotype

Complex diseases arise from multilayered interactions among genetic, molecular, environmental, and tissue-specific factors rather than from alterations in a single gene [18]. Unlike monogenic disorders, these disorders are multifaceted and influenced by many genetic variants, regulatory mechanisms, environmental exposures, physiological and cellular responses, and tissue-specific molecular context [18]. The contribution of each factor alone may be small, but their cumulative action can lead to substantial changes in disease susceptibility, progression, severity, and response to therapy [19]. This architecture contributes to the wide clinical diversity observed in cancer, type 2 diabetes, Alzheimer's disease, cardiovascular disease, inflammatory bowel disease, rheumatoid arthritis, and asthma [18,19]. One of the core aspects of complex disease biology is polygenic inheritance. Several susceptibility loci for metabolic, immune, cardiovascular, neurologic, and malignant disorders have been identified through genome-wide association studies [8]. Many variants occur in non-coding regions and can influence transcription factor binding, enhancer function, chromatin conformation, and gene expression without directly altering the encoded amino acid sequence. These regulatory effects may still modify downstream protein abundance, expression patterns, or pathway activity. These findings suggest that genomic risk may operate through regulatory networks rather than solely through alterations in protein structure [20]. Proteogenomic analysis extends genomic interpretation by linking genetic variation to corresponding changes in protein abundance and disease-relevant molecular phenotypes, while pathway activity is inferred from integrated molecular data [21].

Relationships between genotype and phenotype are further complicated by the interaction of genes and the environment. Molecular pathways associated with inherited susceptibility can be altered by diet, smoking, infection, exposure to medications, pollutants, composition of the microbiome, psychosocial stress, physical activity, and aging [19]. Genetic susceptibility in type 2 diabetes is linked to obesity, insulin resistance, adipose inflammation, and mitochondrial dysfunction [17]. In cardiovascular disease, genetic susceptibility interacts with alterations in lipid metabolism, endothelial dysfunction, immune activation, and vascular remodeling. Environmental factors affect the epithelial barrier integrity, immune tolerance, cytokine signaling, and microbial-host interactions in asthma and inflammatory bowel disease [14]. The disease phenotype can also be influenced by epigenetic regulation, which does not involve changes in the DNA sequence. Mechanisms that regulate gene expression in a cell-specific and exposure-sensitive manner include DNA methylation, histone modification, chromatin remodeling, and non-coding RNA activity [15]. They are important in the pathogenesis of rheumatoid arthritis and may also contribute to disease mechanisms in cancer and neurodegeneration, where chronic stimulation can create stable pathogenic molecular states [20].

Tissue specificity and pathway redundancy further complicate the situation. The same genetic change can affect various tissues of the brain, liver, heart, immune cells, intestinal epithelium, or the tumor itself because of the different transcription and protein environments in each tissue [21]. Multiple molecular pathways may be able to act as compensatory mechanisms, thereby modulating disease manifestations and contributing to therapeutic resistance [18]. Even within the same clinical diagnosis, there is molecular heterogeneity that restricts uniform responses to treatment. Proteogenomic profiling does not overcome this heterogeneity but can help characterize it by integrating genomic, transcriptomic, and proteomic information into a more functionally informed interpretation of disease biology [18]. Figure *1* shows the interconnected genetic, molecular, environmental, tissue-specific, and proteogenomic factors that contribute to complex disease phenotypes.

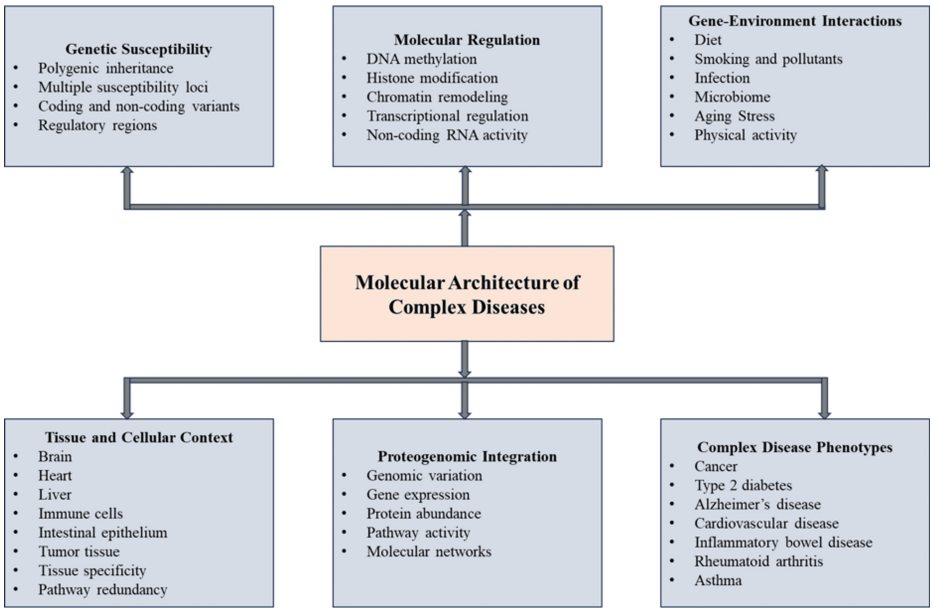


FIGURE 1: Molecular architecture of complex diseases.

This image was created by the authors of this study using Microsoft PowerPoint software (Redmond, WA: Microsoft Corp.), based on supporting literature [1,18,19].

Multi-omics integration strategies for proteogenomic profiling

Multi-omics integration offers a systematic approach to linking genomic variation to transcript expression, protein abundance, metabolic activity, epigenetic regulation, and clinical phenotype [22]. In proteogenomic workflows, next-generation sequencing (NGS) can identify inherited variants, somatic mutations, copy number alterations, structural rearrangements, and gene fusions, while mutational signatures are subsequently inferred computationally from patterns of somatic mutations [23]. The transcriptomic profile helps to characterize gene expression, alternative splicing, allele-specific expression, and non-coding RNA activity [24]. Proteogenomics integrates genomic and/or transcriptomic information with proteomic evidence, typically generated using mass spectrometry-based approaches. At the protein level, mass spectrometry-based proteomics can quantify protein abundance, detect variant peptides, identify protein isoforms, and characterize post-translational modifications, thereby providing protein-level evidence that complements molecular findings at the DNA and RNA levels [24].

Sample preparation and analytical harmonization are important components of proteogenomic integration and include standardized sampling, extraction procedures, sequencing quality control, protein digestion, peptide identification, and spectral matching [25]. Custom protein sequence databases, or sample-specific protein databases derived from genomic and transcriptomic information, can improve detection of variant peptides, fusion proteins, neoantigens, and disease-specific isoforms [26]. Proteomic findings are then integrated with genomic and transcriptomic data to determine whether molecular alterations are supported at the RNA and protein levels. This approach increases the likelihood that clinically relevant molecular drivers will be prioritized and can reduce, although not eliminate, uncertainty in the interpretation of variants with limited biological evidence [20]. Epigenomic and metabolomic data can provide additional regulatory and functional context. Differences in gene-expression patterns between physiological and pathological states may be associated with DNA methylation, histone modification, chromatin accessibility, and enhancer activity [17]. Metabolomic profiling can further characterize downstream biochemical processes involving lipid metabolism, oxidative stress, mitochondrial function, amino acid metabolism, and inflammatory mediators. Integrating these datasets with proteomic evidence facilitates reconstruction of disease pathways across regulatory, structural, enzymatic, and metabolic levels [19].

Single-cell omics can provide cellular-level resolution across genomic, transcriptomic, and, when protein-level assays are included, proteomic dimensions, thereby revealing heterogeneity that may be masked by bulk tissue analysis. Single-cell sequencing and protein-level assays can identify rare disease-driving clones, immune-cell states, stromal interactions, treatment-resistant populations, and lineage-specific molecular programs [22]. Spatial proteomics and spatial transcriptomics preserve tissue architecture, allowing identification of molecular signatures in tumor niches, inflamed tissues, vascular compartments, synaptic regions, or fibrotic microenvironments. This spatial information is particularly relevant in diseases influenced by cellular neighborhoods, immune-cell infiltration, and tissue remodeling [18]. Key challenges for bioinformatics pipelines include normalization, batch correction, feature alignment, differential expression analysis, variant peptide detection, pathway enrichment, and network modeling [24,25]. Pathway-based interpretation connects molecular variation to biological processes such as immune activation, kinase signaling, apoptosis, extracellular matrix remodeling, angiogenesis, proteostasis, and mitochondrial dysfunction. Computational network and pathway analyses can be applied to integrated molecular datasets to identify hub proteins, regulatory modules, and potentially druggable nodes [22]. Machine-learning approaches may further support biomarker discovery, molecular subtype classification, prognosis prediction, and modeling of therapeutic response. Robust validation, transparent analytical methods, external cohorts, and clinically interpretable outputs remain important for translation to precision medicine [18].

Proteogenomic signatures in cancer and precision oncology

Cancer represents one of the most advanced areas of clinical application for proteogenomics because somatic mutations, altered protein expression, dysregulated signaling, immune evasion, and therapeutic vulnerabilities are closely interconnected in tumor biology [26]. Tumor genomes contain driver mutations, copy number alterations, gene fusions, chromosomal instability, and mutational signatures associated with malignant transformation [23]. Proteogenomic profiling complements genomic information by determining whether genomic alterations are supported by corresponding transcriptomic and protein-level evidence and whether associated proteins show altered abundance or post-translational modification in tumor tissue. Not all genomic alterations produce detectable protein-level consequences or contribute equally to tumor progression [27]. Proteogenomic studies can therefore help distinguish genomic alterations from the downstream molecular pathways and protein-level changes associated with tumor behavior [22]. In breast cancer, integrated profiling can identify hormone receptor signaling, HER2 pathway activation, PI3K-AKT-mTOR pathway activity, and immune-enriched subgroups [28]. Proteogenomic analysis can also support interpretation of EGFR, ALK, ROS1, MET, Kirsten rat sarcoma virus (KRAS), and immune-related features in lung cancer. In colorectal cancer, mismatch repair status, WNT signaling, MAPK activation, and immune infiltration are relevant to prognosis and treatment decisions [17].

Another important application of proteogenomics is the identification of tumor-derived variant peptides. However, conventional whole-cell proteomics and HLA immunopeptidomics should be distinguished as follows: detection of a mutated protein or peptide by proteomics does not by itself establish that the peptide is processed, loaded onto HLA molecules, presented on the cell surface, or immunogenic [27]. HLA immunopeptidomics can provide more direct evidence of peptide presentation and may contribute to neoantigen prioritization, but its use for selecting patients for immune checkpoint inhibitors, cancer vaccines, adoptive cell therapies, or personalized immunotherapies remains investigational and should be interpreted alongside other genomic, immunologic, and clinical factors [21]. Immune evasion can also be studied through integrated molecular profiling of antigen-presentation machinery, checkpoint ligand expression, interferon signaling, and tumor-microenvironment remodeling [28]. Precision oncology remains constrained by tumor heterogeneity and drug resistance. Tumor cells within the same tumor may differ in genomic alterations, protein networks, and pathway activity [29]. Resistant populations may emerge through secondary mutations, bypass signaling, altered protein degradation, lineage plasticity, or immune escape under treatment pressure. Proteogenomic profiling can help characterize resistance-associated pathways and may inform molecularly guided combination strategies involving kinase inhibitors, immune checkpoint blockade, antibody-drug conjugates, poly(ADP-ribose) polymerase (PARP) inhibitors, and other pathway-specific therapies [18]. Although routine clinical implementation of proteogenomics remains limited, integrated genomic and proteomic information may provide complementary evidence for multidisciplinary molecular tumor boards when interpreted together with pathology, radiology, and treatment history [27]. Table *2* presents a disease-specific summary of genomic alterations, proteomic/pathway characteristics, and clinical status.

Cancer type	Genomic or molecular alteration	Proteomic/pathway characteristic	Clinical biomarker status	Potential therapeutic relevance	References
Breast cancer	ERBB2 amplification	HER2 protein expression, PI3K pathway activity, ER signaling	ERBB2/HER2-related biomarkers are clinically established; broader integrated proteogenomic signatures remain investigational	HER2-targeted therapy, PI3K inhibition, endocrine therapy	[26]
Lung cancer	EGFR, ALK, ROS1, KRAS, and MET alterations	Kinase signaling and immune-related pathway activity	Several genomic driver alterations are clinically established; integrated proteomic signatures remain investigational	TKI, ICI, combination strategies	[23]
Colorectal cancer	MSI and associated molecular alterations	WNT and MAPK pathway activity, immune infiltration	MSI has established clinical utility; broader pathway-level proteogenomic profiles remain investigational	ICI, EGFR-directed and MAPK-directed approaches	[27]
Glioblastoma	EGFR amplification and IDH status	RTK signaling, hypoxia, and resistance-associated pathways	Selected molecular markers have established diagnostic/prognostic relevance; integrated proteomic signatures remain investigational	RTK-directed and investigational combination strategies	[16]
Ovarian cancer	HRD and TP53 alterations	DNA-repair pathway dysfunction	HRD-related biomarkers have clinical relevance; broader proteogenomic signatures remain under investigation	PARP inhibition, platinum-based therapy, ADC strategies	[28]
Melanoma	BRAF and NRAS alterations	MAPK pathway activity and immune-checkpoint-related signaling	BRAF status is clinically established; integrated proteomic and immune signatures remain investigational	BRAF-MEK inhibition and ICI	[29]

TABLE 2: Genomic alterations, proteomic/pathway characteristics, and clinical biomarker status in precision oncology.

HER2: human epidermal growth factor receptor 2; PI3K: phosphoinositide 3-kinase; ER: estrogen receptor; ADC: antibody-drug conjugate; EGFR: epidermal growth factor receptor; ALK: anaplastic lymphoma kinase; ROS1: ROS proto-oncogene 1 receptor tyrosine kinase; KRAS: Kirsten rat sarcoma viral oncogene homolog; MET: mesenchymal-epithelial transition factor; TKI: tyrosine kinase inhibitor; ICI: immune checkpoint inhibitor; MSI: microsatellite instability; WNT: wingless/integrated signaling pathway; MAPK: mitogen-activated protein kinase; IDH: isocitrate dehydrogenase; RTK: receptor tyrosine kinase; HRD: homologous recombination deficiency; TP53: tumor protein p53; PARP: poly (ADP-ribose) polymerase; BRAF: B-Raf proto-oncogene; NRAS: neuroblastoma RAS viral oncogene homolog; MEK: mitogen-activated protein kinase

Neurodegenerative and neuropsychiatric disease signatures

Cellular diversity, regional specialization of the brain, slowly evolving molecular injury, and limited accessibility of affected tissue during early disease stages contribute to the proteogenomic complexity of neurodegenerative and neuropsychiatric disorders [30]. Alterations in protein homeostasis and related molecular pathways have been reported in several neurodegenerative and neuropsychiatric conditions, although their nature and extent vary across diseases [31]. Traditional diagnostic assessment relies on clinical examination, neuroimaging, cognitive evaluation, electrophysiology, and selected fluid biomarkers, while some molecular changes may precede overt clinical manifestations [23]. Proteogenomic signatures may help clarify relationships among amyloid precursor protein processing, tau phosphorylation, synaptic loss, complement activation, microglial response, lipid metabolism, and neuronal injury in Alzheimer's disease [32]. In Alzheimer's disease, cerebrospinal fluid biomarkers such as amyloid beta and phosphorylated tau have contributed to biological classification and disease staging, while other markers, including total tau, neurofilament light chain, and inflammatory proteins, may provide complementary information rather than serving uniformly as diagnostic markers [26]. In combination with genomic risk factors such as apolipoprotein E genotype, plasma proteomics may support identification of signatures associated with neurodegeneration and offer a less invasive alternative to tissue-based approaches [33].

Parkinson's disease is associated with degeneration and loss of dopaminergic neurons, aggregation of α-synuclein, mitochondrial dysfunction, lysosomal impairment, and neuroinflammatory activity [19]. Proteogenomic profiling may help identify genetically influenced subtypes involving LRRK2, GBA1, PRKN, and PINK1 pathways and distinguish them from idiopathic disease [26]. This classification could support

targeted drug development involving kinase inhibitors, lysosomal modulators, mitochondrial protective strategies, and α-synuclein-directed therapies [34]. The molecular heterogeneity of amyotrophic lateral sclerosis includes RNA-binding protein dysfunction, impaired protein clearance, cytoskeletal disruption, excitotoxicity, oxidative stress, and neuroimmune activation [27]. Modeling of protein networks in cerebrospinal fluid, plasma, spinal cord tissue, and motor neurons may reveal molecular signatures associated with disease progression and therapeutic sensitivity [23]. Multiple sclerosis is characterized by immune-mediated demyelination, blood-brain barrier disruption, astrocyte activation, microglial signaling, axonal injury, and impaired remyelination [16]. Protein signatures associated with complement activity, cytokine signaling, myelin injury, and neurofilament release may help monitor disease activity and treatment response [8].

In recent years, molecular network disruption has gained attention as a framework for understanding neuropsychiatric disorders such as major depressive disorder and schizophrenia beyond symptom-based categories [33]. Proteogenomic studies have implicated processes including synaptic plasticity, immune signaling, stress-response pathways, mitochondrial energetics, glutamatergic transmission, and neurodevelopmental regulation [30]. Molecular signatures may also vary across brain regions and among different cell types within the same region, underscoring the importance of spatial and cellular context [24]. Combining cerebrospinal fluid markers, plasma proteomics, neuroimaging findings, genetic risk scores, and clinical phenotypes may support earlier detection, biologically informed classification, and development of disease-modifying strategies for neurological and psychiatric disorders [34]. Figure *2* illustrates the role of proteogenomics in linking molecular pathway disruption, biomarker discovery, and clinical applications in neurodegenerative and neuropsychiatric disorders.

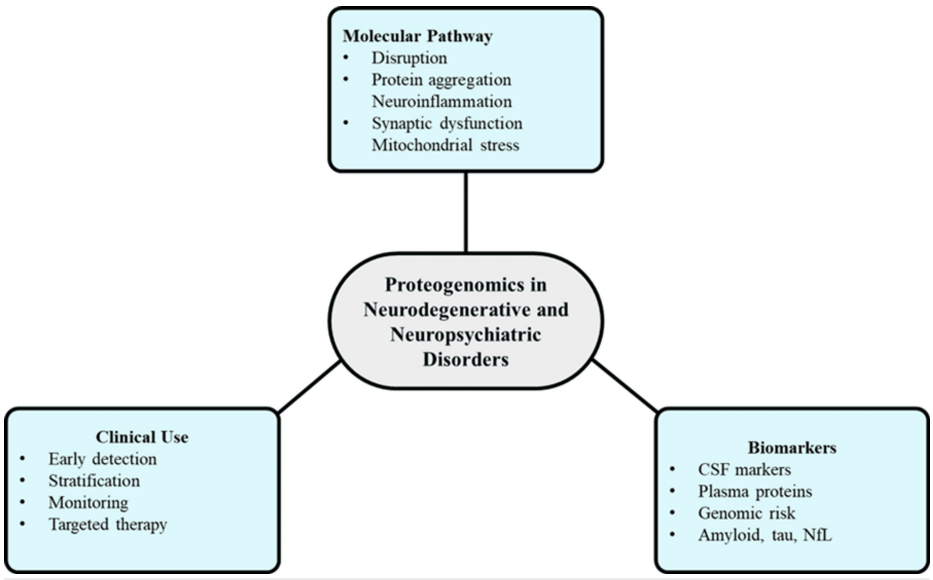


FIGURE 2: Proteogenomics in neurodegenerative and neuropsychiatric disorders.

This image was created by the authors of this study using Microsoft PowerPoint software (Redmond, WA: Microsoft Corp.), based on supporting literature [3,5,6,30,33].

Autoimmune and inflammatory disease proteogenomics

Autoimmune and inflammatory diseases can involve varying combinations of impaired immune recognition, loss of immune tolerance, persistent inflammatory signaling, and tissue injury, although the dominant mechanisms differ among individual conditions [35]. Rheumatoid arthritis, systemic lupus erythematosus, inflammatory bowel disease, psoriasis, asthma, and multiple sclerosis are clinically heterogeneous, with variable disease activity and therapeutic response. Proteogenomic profiling can integrate inherited genetic susceptibility, immune-cell transcriptional profiles, cytokine activity, autoantibody patterns, and protein-level pathway changes with clinical phenotypes [36]. This approach is particularly relevant to immune-mediated diseases, in which interacting genetic, epigenetic, microbial, environmental, and cellular factors may contribute to inflammatory activity [32].

Human leukocyte antigen (HLA) associations remain important determinants of autoimmune susceptibility. Certain HLA alleles influence antigen presentation, T-cell activation, autoantibody production, and tissue-specific immune responses [25]. In rheumatoid arthritis (RA), HLA-associated susceptibility interacts with immune responses to citrullinated antigens, synovial inflammation, complement activation, and cytokine amplification. Proteogenomic assessment may identify protein signatures associated with joint erosion,

systemic inflammation, and response to targeted immunomodulatory agents [14]. In systemic lupus erythematosus (SLE), type I interferon activity, immune-complex formation, complement consumption, B-cell dysfunction, and autoantibody diversity represent molecular features that may assist disease stratification and flare prediction [37]. Inflammatory bowel disease is characterized by epithelial barrier dysfunction, microbial dysbiosis, innate immune activation, T-cell polarization, cytokine release, and stromal remodeling [38]. Proteogenomic signatures may help distinguish Crohn's disease from ulcerative colitis, identify fibrosis-related pathways, and support prediction of response to anti-tumor necrosis factor (TNF), anti-integrin, IL-12/23, and Janus kinase (JAK)-targeted therapies [27]. Psoriasis is strongly associated with keratinocyte activation, dendritic-cell signaling, IL-23/IL-17 axis activity, and Th17-mediated inflammation. Mapping these pathways at the protein level may help clarify mechanisms underlying disease severity and inform selection of IL-17- or IL-23-targeted therapies [19].

Asthma comprises diverse clinical phenotypes that can be further characterized by underlying molecular endotypes, which complement rather than replace phenotype-based classification. Therapeutic variability is influenced by type 2 inflammation, eosinophilic activity, epithelial alarmins, IgE signaling, airway remodeling, and neutrophilic inflammation [31]. Proteogenomic and related molecular profiling may help refine biologic selection by identifying pathways involving IgE, IL-5, IL-4 receptor signaling, or TSLP. Multiple sclerosis (MS) involves immune-mediated demyelination, altered antigen presentation, microglial activation, complement-mediated injury, and neuroaxonal damage. Integrated molecular signatures may aid disease monitoring and assessment of treatment response [36]. Table *3* summarizes major proteogenomic patterns associated with immune-mediated disease mechanisms and targeted treatment selection.

Disease condition	Dominant proteogenomic pattern	Immune pathway relevance	Targeted intervention relevance	Clinical implication	References
RA	HLA risk, ACPA response, TNF and IL-6 activity	Synovial inflammation and joint destruction	Anti-TNF, IL-6 inhibitor, JAK inhibitor	Response prediction and erosion-risk stratification	[21]
SLE	IFN signature, complement alteration, autoantibody diversity	Immune-complex injury and B-cell activation	B-cell therapy, IFN pathway blockade	Flare prediction and organ-risk assessment	[36]
IBD	Barrier proteins, TNF activity, IL-12/23 signaling	Mucosal inflammation and fibrosis	Anti-TNF, anti-integrin, JAK inhibitor	Disease subtype refinement and treatment selection	[24]
Psoriasis	IL-23/IL-17 axis, keratinocyte activation	Th17 inflammation and epidermal hyperplasia	IL-17 inhibitor, IL-23 inhibitor	Severity assessment and biologic matching	[37]
Asthma	IgE activity, IL-5 signal, epithelial alarmins	Type 2 inflammation and airway remodeling	Anti-IgE, anti-IL-5, TSLP blockade	Endotype classification and biologic selection	[24]
MS	HLA association, complement activity, neurofilament release	Demyelination and neuroaxonal injury	B-cell therapy, immune modulation	Activity monitoring and progression assessment	[30]

TABLE 3: Immune-mediated disease signatures with targeted therapeutic relevance.

RA: rheumatoid arthritis; HLA: human leukocyte antigen; ACPA: anti-citrullinated protein antibody; TNF: tumor necrosis factor; JAK: Janus kinase; SLE: systemic lupus erythematosus; IFN: interferon; IBD: inflammatory bowel disease; TSLP: thymic stromal lymphopoietin; MS: multiple sclerosis; Th17: T helper 17 cell

Biomarker discovery, risk stratification, and clinical decision support

Proteogenomic signatures offer a biology-based approach to transforming molecular complexity into clinically relevant biomarkers [38]. Depending on their intended use, biomarkers may support diagnosis, prognosis, therapeutic selection, treatment monitoring, or recurrence detection, provided that appropriate analytical validity, clinical validity, and clinical utility are established [39]. Within complex diseases, integration of genomic, transcriptomic, and proteomic data can uncover molecular signals associated with active disease pathways rather than genetic susceptibility alone. This is especially important in conditions with similar clinical phenotypes, limited diagnostic sensitivity, and variable treatment responses among patients with the same diagnosis [40]. Proteogenomic profiling-derived diagnostic biomarkers may support earlier and more precise disease detection. Protein signatures associated with tumor-specific mutations, inflammatory networks, endothelial injury, neurodegeneration, or metabolic dysfunction may indicate disease-related molecular changes before overt clinical manifestations [33]. Minimally invasive approaches can assess analytes such as circulating proteins, plasma peptides, extracellular vesicles, and autoantibody profiles, while liquid biopsy represents a broader sampling and analytical platform that may incorporate

several molecular analytes. In oncology, liquid biopsy combined with proteomic markers may help characterize tumor burden, molecular relapse, immune response, and resistance evolution [41]. Circulating molecular panels may also complement early risk stratification and disease subtyping in cardiovascular, autoimmune, and neurodegenerative conditions [27].

Prognostic biomarkers can help estimate disease course, severity, recurrence risk, or survival. Patterns involving DNA-repair defects, kinase activation, immune-cell infiltration, fibrosis, complement activation, mitochondrial dysfunction, or synaptic injury may help define clinically distinct risk groups [42]. Predictive biomarkers may inform therapeutic choice by identifying patients more likely to benefit from targeted agents, immunotherapies, biologics, metabolic treatments, or pathway-specific interventions. Pharmacodynamic biomarkers reflect molecular changes after treatment initiation, including reduced pathway activation, altered cytokine profiles, decreased tumor-derived proteins, or normalization of inflammatory mediators [43]. Some proteogenomic biomarkers may contribute to the development of companion diagnostics when they are specifically linked to the safe and effective use of a particular therapy, but this does not apply to all proteogenomic biomarkers [38]. Such biomarkers may support treatment selection when validated for a defined therapeutic context. Clinical decision-support systems can integrate biomarker profiles with imaging, pathological findings, laboratory data, comorbidities, medication exposure, and longitudinal outcomes. Algorithm-based interpretation may assist in identifying potentially actionable molecular features, prioritizing therapeutic options, and monitoring disease progression [42]. Laboratory variability, pre-analytical sample variation, batch effects, limited population diversity, and insufficient validation cohorts may compromise biomarker reproducibility, performance, and generalizability [40]. Before clinical application, relevant performance and utility criteria should be established, including sensitivity, specificity, predictive values, treatment relevance, and potential benefit to patient outcomes. Cost-effectiveness, funding, regulatory requirements, data security, and interpretability also require consideration [37].

Targeted interventions based on proteogenomic signatures

The concept of the proteogenomic signature is based on integrating molecular information to identify disease-relevant pathways that may inform targeted intervention [43]. Traditional drug selection is based on diagnosis, histology, and clinical severity. Proteogenomic analysis provides a multilayered molecular framework by integrating genomic changes, transcriptional signatures, protein expression levels, post-translational modifications, and signaling information associated with therapeutic response [44]. In oncology, proteogenomics may help identify targetable kinase activation, DNA-repair defects, immune-evasion mechanisms, metabolic dependencies, and resistance-associated protein networks [35]. Tumors with actionable kinase alterations may be candidates for kinase inhibitors, whereas homologous recombination deficiency may indicate potential sensitivity to PARP inhibitors or platinum-based therapy [45]. Identification of immune checkpoints, impaired antigen presentation, interferon-response patterns, and candidate neoantigens may contribute to evaluation of immunotherapeutic strategies [32]. Proteogenomics may also help distinguish likely biological drivers from passenger alterations [45]. Outside oncology, proteogenomic signatures may inform pathway-oriented investigation in cardiometabolic, autoimmune, inflammatory, neurological, and fibrotic diseases. Profiles characterized by inflammation, lipid dysregulation, endothelial activation, mitochondrial stress, or myocardial remodeling may identify biological processes potentially relevant to the evaluation of lipid-lowering, anti-inflammatory, metabolic, or cardioprotective therapeutic strategies [46]. In autoimmune disease, cytokine-dominant profiles may support evaluation of anti-TNF, interleukin, Janus kinase, B-cell, or complement-directed therapies [40]. In neurodegenerative disorders, signatures involving protein aggregation, lysosomal dysfunction, synaptic injury, neuroinflammation, and mitochondrial dysfunction may help identify candidate disease-modifying strategies and clinical trial opportunities [39].

A third translational application is drug repurposing based on network analysis of disease-associated proteins and pathways. Identification of approved drugs acting on relevant proteins or pathways may facilitate more rapid and potentially more cost-effective therapeutic discovery [23]. RNA-based therapies, including antisense oligonucleotides, small interfering RNA, and mRNA-based strategies, may enable targeting disease-associated RNA molecules or modulation of relevant molecular processes [43]. Gene therapies may also be considered in selected settings where molecular pathway disruption is supported by integrated molecular evidence [47]. Proteolysis-targeting chimeras and monoclonal antibody therapies provide additional approaches to targeted intervention. Proteolysis-targeting chimeras (PROTACs) may facilitate degradation of disease-promoting proteins that are difficult to inhibit using conventional drugs [45]. Monoclonal antibodies may inhibit cytokines, interfere with receptor signaling, deplete pathogenic immune-cell populations, or target tumor-associated antigens [43]. Combination approaches may be considered when parallel pathways, clonal heterogeneity, compensatory mechanisms, or immune escape contribute to treatment resistance. Proteogenomic profiles may assist in designing combination strategies intended to improve therapeutic response and address mechanisms of resistance [47].

Challenges and ethical considerations

Translational application of proteogenomic signatures continues to face challenges related to methodology, computation, economics, and ethics. Data complexity constitutes a significant hurdle since genomic,

transcriptomic, proteomic, epigenomic, metabolomic, and clinical data are inherently diverse in terms of magnitude, format, granularity, noise content, and biological relevance [48]. The unification of these disparate data types calls for effective bioinformatics infrastructure, reliable normalization protocols, replicable feature extraction techniques, and accurate statistical inference methods. Technical artifacts, such as batch effects, platform variation, specimen degradation, tissue heterogeneity, and pre-analytical factors, may affect molecular measurements and diminish comparability between laboratories [49]. Sample quality plays a critical role in proteogenomics research. Formalin-fixed tissues, fresh-frozen samples, blood plasma, serum, cerebrospinal fluid, and single-cell extracts vary in molecular stability and analytical fidelity [50]. Limited sample numbers, retrospective studies, incomplete metadata, and variable clinical annotation may compromise biomarker assessment. Numerous biomarkers do not advance into clinical practice due to inadequate external validation, poor analytical reproducibility, ambiguous clinical utility, or absence of well-defined thresholds for clinical decision-making [44]. Clinical utility should be demonstrated by showing how a molecular diagnostic tool surpasses existing clinical benchmarks in terms of diagnosis, prognostication, therapy guidance, or health benefits [49].

There are other impediments created by costs and infrastructure needs. Technologies such as high-throughput sequencing, mass spectrometry, specific storage, computational skills, and multidisciplinary analyses could be out of reach in resource-limited environments [37]. Disparities in access could exacerbate current healthcare inequalities, particularly because molecular databases would not have adequately sampled populations of various ancestral, geographical, and socioeconomic origins [48]. Inadequate sampling could limit the ability to generalize and make prediction algorithms prone to bias [50]. There are also issues relating to ethics. These involve genomic privacy, consent for omics analysis, secondary data utilization, unexpected results, family considerations, and potential misuse of predictive information [26]. There is a need for informed consent processes, proper data management, responsible data sharing, and appropriate disclosure strategies [48]. Interoperability, physician comprehension, patient privacy, auditability, and algorithmic discrimination also require consideration in embedding proteogenomics within electronic health records [50].

Limitations and future recommendations

This review has some important limitations. As an integrative review, the evidence synthesis was selective rather than based on a formal systematic-review framework, which may have resulted in uneven representation across disease areas. The available literature remains predominantly oncology-focused, with comparatively limited evidence in cardiometabolic, neurological, autoimmune, and inflammatory diseases. Many studies involve small discovery cohorts, limited ancestral and geographical diversity, and substantial heterogeneity in sequencing and proteomic platforms, specimen types, bioinformatic pipelines, analytical thresholds, and endpoint definitions. Independent replication of proposed proteogenomic signatures remains limited, and prospective studies evaluating clinical utility, patient outcomes, and cost-effectiveness are scarce.

Future studies should use harmonized protocols, clearly defined endpoints, larger and ancestrally diverse multicenter cohorts, and independent external validation. Prospective clinical-utility studies are needed to determine whether proteogenomic-guided decisions improve diagnosis, prognosis, treatment selection, and patient outcomes beyond conventional approaches. Health-economic evaluations should also assess cost-effectiveness, scalability, infrastructure requirements, and feasibility across different healthcare settings.

Conclusions

The use of proteogenomic signatures offers an important means of investigating complex diseases through the integration of genomic variation with protein expression, function, and disease-associated molecular pathways. This review's key contribution is a cross-disease synthesis of proteogenomic evidence linking molecular mechanisms to biomarker discovery, risk stratification, clinical decision support, and targeted interventions across oncologic and non-oncologic complex diseases. Rather than considering proteogenomics primarily within precision oncology, the review integrates evidence from cancer, cardiometabolic, neurodegenerative, autoimmune, and inflammatory disorders and distinguishes established clinically actionable biomarkers from broader proteogenomic signatures that remain investigational. This framework highlights how integrated molecular profiling may support biologically informed subclassification, treatment-response assessment, and identification of therapeutic vulnerabilities beyond conventional phenotype-based classification. Precision oncology currently represents the most developed clinical setting, whereas applications in non-oncologic diseases remain comparatively less validated. Clinical translation will require standardized specimen preparation, harmonized analytical workflows, independent replication, prospective clinical validation, diverse populations, cost-effectiveness evaluation, and interpretable decision-support systems. Issues involving consent, genomic privacy, equitable access, and algorithmic bias also require continued attention. The principal value of proteogenomics lies in its capacity to connect molecular alterations with functional protein-level consequences, providing a framework for moving from predominantly diagnosis-based treatment toward more mechanism-guided and patient-specific therapeutic strategies.

Additional Information

Author Contributions

All authors have reviewed the final version to be published and agreed to be accountable for all aspects of the work.

Concept and design: Preeti Sharma, Eragam Siva Rami Reddy, Amit Nampalliwar, Karthika Padmavathy, Rahima Malek, Rachna Sharma, Upasana Uniya

Acquisition, analysis, or interpretation of data: Preeti Sharma, Eragam Siva Rami Reddy, Amit Nampalliwar, Karthika Padmavathy, Rahima Malek, Rachna Sharma, Upasana Uniya

Drafting of the manuscript: Preeti Sharma, Eragam Siva Rami Reddy, Amit Nampalliwar, Karthika Padmavathy, Rahima Malek, Rachna Sharma, Upasana Uniya

Critical review of the manuscript for important intellectual content: Preeti Sharma, Eragam Siva Rami Reddy, Amit Nampalliwar, Karthika Padmavathy, Rahima Malek, Rachna Sharma, Upasana Uniya

Supervision: Karthika Padmavathy

Disclosures

Conflicts of interest: In compliance with the ICMJE uniform disclosure form, all authors declare the following: **Payment/services info:** All authors have declared that no financial support was received from any organization for the submitted work. **Financial relationships:** All authors have declared that they have no financial relationships at present or within the previous three years with any organizations that might have an interest in the submitted work. **Other relationships:** All authors have declared that there are no other relationships or activities that could appear to have influenced the submitted work.

References

1. Pietzner M, Wheeler E, Carrasco-Zanini J, et al.: Mapping the proteo-genomic convergence of human diseases. Science. 2021, 374:10.1126/science.abj1541
2. Koprulu M, Carrasco-Zanini J, Wheeler E, et al.: Proteogenomic links to human metabolic diseases. Nat Metab. 2023, 5:516-28. 10.1038/s42255-023-00753-7
3. Kaiser S, Zhang L, Mollenhauer B, et al.: A proteogenomic view of Parkinson's disease causality and heterogeneity. NPJ Parkinsons Dis. 2023, 9:10.1038/s41531-023-00461-9
4. Olin JW, Di Narzo AF, d'Escamard V, et al.: A plasma proteogenomic signature for fibromuscular dysplasia . Cardiovasc Res. 2020, 116:63-77. 10.1093/cvr/cvz219
5. Torshizi AD, Truong DT, Hou L, Smets B, Whelan CD, Li S: Proteogenomic network analysis reveals dysregulated mechanisms and potential mediators in Parkinson's disease. Nat Commun. 2024, 15:10.1038/s41467-024-50718-x
6. Diniz BS, Chen Z, Steffens DC, et al.: Proteogenomic signature of Alzheimer's disease and related dementia risk in individuals with major depressive disorder. Nat Ment Health. 2025, 3:879-88. 10.1038/s44220-025-00460-0
7. Sun BB, Suhre K, Gibson BW: Promises and challenges of populational proteomics in health and disease . Mol Cell Proteomics. 2024, 23:10.1016/j.mcpro.2024.100786
8. Carrasco-Zanini J, Pietzner M, Koprulu M, Wheeler E, Kerrison ND, Wareham NJ, Langenberg C: Proteomic prediction of diverse incident diseases: a machine learning-guided biomarker discovery study using data from a prospective cohort study. Lancet Digit Health. 2024, 6:470-9. 10.1016/S2589-7500(24)00087-6
9. Dick F, Johanson GA, Tysnes OB, Alves G, Dölle C, Tzoulis C: Brain proteome profiling reveals common and divergent signatures in Parkinson’s disease, multiple system atrophy, and progressive supranuclear palsy. Mol Neurobiol. 2025, 62:2801-16. 10.1007/s12035-024-04422-y
10. Abdelaziz EH, Ismail R, Mabrouk MS, Amin E: Multi-omics data integration and analysis pipeline for precision medicine: systematic review. Comput Biol Chem. 2024, 113:10.1016/j.compbiolchem.2024.108254
11. Nam Y, Kim J, Jung SH, et al.: Harnessing artificial intelligence in multimodal omics data integration: paving the path for the next frontier in precision medicine. Annu Rev Biomed Data Sci. 2024, 7:225-50. 10.1146/annurev-biodatasci-102523-103801
12. Akhoundova D, Rubin MA: Clinical application of advanced multi-omics tumor profiling: shaping precision oncology of the future. Cancer Cell. 2022, 40:920-38.
13. Fathi M, Taher HJ, Al-Rubiae SJ, et al.: Role of molecular imaging in prognosis, diagnosis, and treatment of gastrointestinal cancers: an update on new therapeutic methods. World J Methodol. 2024, 14:10.5662/wjm.v14.i4.93461
14. Monti C, Zilocchi M, Colugnat I, Alberio T: Proteomics turns functional. J Proteomics. 2019, 198:36-44. 10.1016/j.jprot.2018.12.012
15. Vitorino R, Choudhury M, Guedes S, et al.: Peptidomics and proteogenomics: background, challenges and future needs. Expert Rev Proteomics. 2021, 18:643-59. 10.1080/14789450.2021.1980388
16. Low TY, Mohtar MA, Ang MY, Jamal R: Connecting proteomics to next-generation sequencing: proteogenomics and its current applications in biology. Proteomics. 2019, 19:10.1002/pmic.201800235
17. Piana D, Iavarone F, De Paolis E, et al.: Phenotyping tumor heterogeneity through proteogenomics: study models and challenges. Int J Mol Sci. 2024, 25:10.3390/ijms25168830
18. Yong SY, Raben TG, Lello L, Hsu SD: Genetic architecture of complex traits and disease risk predictors . Sci Rep. 2020, 10:10.1038/s41598-020-68881-8

19. Lappalainen T, Li YI, Ramachandran S, Gusev A: Genetic and molecular architecture of complex traits . Cell. 2024, 187:1059-75. 10.1016/j.cell.2024.01.023
20. Claringbould A, De Klein N, Franke L: The genetic architecture of molecular traits . Curr Opin Syst Biol. 2017, 1:25-31. 10.1016/j.coisb.2017.01.002
21. Timpson NJ, Greenwood CM, Soranzo N, Lawson DJ, Richards JB: Genetic architecture: the shape of the genetic contribution to human traits and disease. Nat Rev Genet. 2018, 19:110-24. 10.1038/nrg.2017.101
22. Goddard ME, Kemper KE, MacLeod IM, Chamberlain AJ, Hayes BJ: Genetics of complex traits: prediction of phenotype, identification of causal polymorphisms and genetic architecture. Proc Biol Sci. 2016, 283:10.1098/rspb.2016.0569
23. Dimitrakopoulos L, Prassas I, Diamandis EP, Charames GS: Onco-proteogenomics: multi-omics level data integration for accurate phenotype prediction. Crit Rev Clin Lab Sci. 2017, 54:414-32. 10.1080/10408363.2017.1384446
24. Acharya D, Mukhopadhyay A: A comprehensive review of machine learning techniques for multi-omics data integration: challenges and applications in precision oncology. Brief Funct Genomics. 2024, 23:549-60. 10.1093/bfgp/elae013
25. Rajczewski AT, Jagtap PD, Griffin TJ: An overview of technologies for MS-based proteomics-centric multi-omics. Expert Rev Proteomics. 2022, 19:165-81. 10.1080/14789450.2022.2070476
26. Rodriguez H, Zenklusen JC, Staudt LM, Doroshow JH, Lowy DR: The next horizon in precision oncology: proteogenomics to inform cancer diagnosis and treatment. Cell. 2021, 184:1661-70. 10.1016/j.cell.2021.02.055
27. Blank-Landeshammer B, Richard VR, Mitsa G, et al.: Proteogenomics of colorectal cancer liver metastases: complementing precision oncology with phenotypic data. Cancers (Basel). 2019, 11:10.3390/cancers11121907
28. Satpathy S, Jaehnig EJ, Krug K, et al.: Microscaled proteogenomic methods for precision oncology . Nat Commun. 2020, 11:10.1038/s41467-020-14381-2
29. Ji S, Feng L, Fu Z, et al.: Pharmaco-proteogenomic characterization of liver cancer organoids for precision oncology. Sci Transl Med. 2023, 15:10.1126/scitranslmed.adg3358
30. Wingo TS, Liu Y, Gerasimov ES, et al.: Shared mechanisms across the major psychiatric and neurodegenerative diseases. Nat Commun. 2022, 13:10.1038/s41467-022-31873-5
31. Strauss A, Swann P, Kigar SL, et al.: Peripheral innate immunophenotype in neurodegenerative disease: blood-based profiles and links to survival. Mol Psychiatry. 2025, 30:1985-94. 10.1038/s41380-024-02809-w
32. Zhu K, Yang TH, Dorie V, Zheng T, Anastassiou D: Meta-analysis of expression and methylation signatures indicates a stress-related epigenetic mechanism in multiple neuropsychiatric disorders. Transl Psychiatry. 2019, 9:10.1038/s41398-018-0358-5
33. Cuttler K, Hassan M, Carr J, Cloete R, Bardien S: Emerging evidence implicating a role for neurexins in neurodegenerative and neuropsychiatric disorders. Open Biol. 2021, 11:10.1098/rsob.210091
34. Cummings J: The role of neuropsychiatric symptoms in research diagnostic criteria for neurodegenerative diseases. Am J Geriatr Psychiatry. 2021, 29:375-83. 10.1016/j.jagp.2020.07.011
35. Broeren MG, Wang JJ, Balzaretti G, et al.: Proteogenomic analysis of the autoreactive B cell repertoire in blood and tissues of patients with Sjögren's syndrome. Ann Rheum Dis. 2022, 81:644-52. 10.1136/annrheumdis-2021-221604
36. Li Y, Ma C, Liao S, et al.: Combined proteomics and single cell RNA-sequencing analysis to identify biomarkers of disease diagnosis and disease exacerbation for systemic lupus erythematosus. Front Immunol. 2022, 13:10.3389/fimmu.2022.969509
37. Zhang Y, Xu H, Xu L, Jiang S, Yu Y, Zhao H: Multi-omics-driven biomarker discovery in autoimmune diseases: a comprehensive review. Front Immunol. 2025, 16:10.3389/fimmu.2025.1652211
38. Tahir UA, Barber JL, Cruz DE, et al.: Proteogenomic analysis integrated with electronic health records data reveals disease-associated variants in Black Americans. J Clin Invest. 2024, 134:10.1172/JCI181802
39. Glaab E, Rauschenberger A, Banzi R, Gerardi C, Garcia P, Demotes J: Biomarker discovery studies for patient stratification using machine learning analysis of omics data: a scoping review. BMJ Open. 2021, 11:10.1136/bmjopen-2021-053674
40. Jørgensen JT: Predictive biomarkers and clinical evidence . Basic Clin Pharmacol Toxicol. 2021, 128:642-8. 10.1111/bcpt.13578
41. Wang J, Wang Y, Ren J, Li Z, Guo L, Lv J: Emerging applications of feature selection in osteoporosis research: from biomarker discovery to clinical decision support. J Bone Miner Res. 2025, 40:1106-13. 10.1093/jbmr/zjaf105
42. Kraus VB: Biomarkers as drug development tools: discovery, validation, qualification and use . Nat Rev Rheumatol. 2018, 14:354-62. 10.1038/s41584-018-0005-9
43. Susanto AP, Lyell D, Widyantoro B, Berkovsky S, Magrabi F: Effects of machine learning-based clinical decision support systems on decision-making, care delivery, and patient outcomes: a scoping review. J Am Med Inform Assoc. 2023, 30:2050-63. 10.1093/jamia/ocad180
44. Fu S, Liu X, Luo M, Xie K, Nice EC, Zhang H, Huang C: Proteogenomic studies on cancer drug resistance: towards biomarker discovery and target identification. Expert Rev Proteomics. 2017, 14:351-62. 10.1080/14789450.2017.1299006
45. Srivastava S, Jayaswal N, Kumar S, et al.: Unveiling the potential of proteomic and genetic signatures for precision therapeutics in lung cancer management. Cell Signal. 2024, 113:10.1016/j.cellsig.2023.110932
46. Shukla HD: Comprehensive analysis of cancer-proteogenome to identify biomarkers for the early diagnosis and prognosis of cancer. Proteomes. 2017, 5:10.3390/proteomes5040028
47. Krug K, Jaehnig EJ, Satpathy S, et al.: Proteogenomic landscape of breast cancer tumorigenesis and targeted therapy. Cell. 2020, 183:1436-56. 10.1016/j.cell.2020.10.036
48. Bittremieux W, Tabb DL, Impens F, Staes A, Timmerman E, Martens L, Laukens K: Quality control in mass spectrometry-based proteomics. Mass Spectrom Rev. 2018, 37:697-711. 10.1002/mas.21544
49. Lualdi M, Fasano M: Statistical analysis of proteomics data: a review on feature selection . J Proteomics. 2019, 198:18-26. 10.1016/j.jprot.2018.12.004

50. Yearby R: Racial disparities in health status and access to healthcare: the continuation of inequality in the United States due to structural racism. Am J Econ Sociol. 2018, 77:1113-52. 10.1111/ajes.12230