Computational and Structural Biotechnology Journal 23 (2024) 1725–1739

Contents lists available at ScienceDirect

Computational and Structural Biotechnology Journal

journal homepage: www.elsevier.com/locate/csbj

Review article

Proteomics appending a complementary dimension to precision oncotherapy

Zhaokai Zhou [a,b,1], Ruiqi Zhang [a,1], Aoyang Zhou [a,1], Jinxiang Lv [c], Shuang Chen [d], Haijiao Zou [d], Ge Zhang [e], Ting Lin [a], Zhan Wang [b], Yuyuan Zhang [a], Siyuan Weng [a], Xinwei Han [a,f,g,*], Zaoqu Liu [a,f,g,h,*]

[a] Department of Interventional Radiology, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan 450052, China
[b] Department of Urology, The First Affiliated Hospital of Zhengzhou University, Henan 450052, China
[c] Department of Gastroenterology, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan 450052, China
[d] Center of Reproductive Medicine, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan 450052, China
[e] Department of Cardiology, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan 450052, China
[f] Interventional Institute of Zhengzhou University, Zhengzhou, Henan 450052, China
[g] Interventional Treatment and Clinical Research Center of Henan Province, Zhengzhou, Henan 450052, China
[h] Institute of Basic Medical Sciences, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100730, China

ARTICLE INFO

Keywords:
Proteomics
Proteomics analysis technology
Data integration
Protein network
Precision oncotherapy
Molecular typing

ABSTRACT

Recent advances in high-throughput proteomic profiling technologies have facilitated the precise quantification of numerous proteins across multiple specimens concurrently. Researchers have the opportunity to comprehensively analyze the molecular signatures in plentiful medical specimens or disease pattern cell lines. Along with advances in data analysis and integration, proteomics data could be efficiently consolidated and employed to recognize precise elementary molecular mechanisms and decode individual biomarkers, guiding the precision treatment of tumors. Herein, we review a broad array of proteomics technologies and the progress and methods for the integration of proteomics data and further discuss how to better merge proteomics in precision medicine and clinical settings.

1. Introduction

In 2003, the International Human Genome Sequencing Consortium successfully concluded a monumental scientific endeavor, the Human Genome Project, yielding a comprehensive catalog of nucleotide sequences serving as a reference for over 20,000 protein-coding genes inherent in the human genome[1,2]. Subsequently, a nascent medical paradigm has emerged, namely Genomic medicine[3]. The genomic era revolutionized the screening and treatment of human diseases. Numerous targeted medicines have been added to the arsenal to treat cancer, particularly lung and breast adenocarcinomas. However, there are essential limitations of genomic and transcriptomic analyses, which only provide indirect measures of cellular status and cannot accurately reflect the corresponding protein abundance and variations such as post-translational modifications (PTMs). As such, exclusive reliance on genomic data could not offer a complete picture of the underlying disease mechanisms[4]. Nevertheless, human genome sequencing propelled the field in proteomics, enabling quantitative and qualitative

Abbreviations: CCA, cholangiocarcinoma; ccRCC, clear cell renal cell carcinoma; ce6, chlorin e6; CRC, colorectal cancer; CEA, carcinoembryonic antigen; CYFRA 21-1, cytokeratin 19 fragment; DGC, diffuse-type gastric cancer; DNA, DeoxyriboNucleic Acid; DVP, deep visual proteomics; EGFR, epidermal growth factor receptor; ErbB, erythroblastic oncogene B; ESCC, esophageal squamous carcinoma; EV, extracellular vesicles; HCC, hepatocellular carcinoma; HPLC, high-performance LC; IL-8, interleukin (IL)-8; IL-6R, interleukin-6 receptor; LC, liquid chromatography; MAP4Ks, mitogen-activated protein kinase kinase kinase kinases; MAPK, mitogen-activated protein kinase; MS, Mass Spectrometry; NGS, next-generation sequencing; NMF, non-negative matrix decomposition; NSCLC, non-small cell lung cancer; PCA, principal component analysis; PCR, Polymerase Chain Reaction; PEA, Proximity Extension Assay; RNA, RiboNucleic Acid; PTMs, post-translational modifications; RPPA, Reverse-phase protein array; SVM, support vector machine; TieDIE, Tied Diffusion Through Interacting Events.

* Corresponding authors at: Department of Interventional Radiology, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan 450052, China.
E-mail addresses: fcchanxw@zzu.edu.cn (X. Han), liuzaoqu@163.com (Z. Liu).

[1] These authors have contributed equally to this work and share the first authorship

https://doi.org/10.1016/j.csbj.2024.04.044
Received 6 February 2024; Received in revised form 11 April 2024; Accepted 17 April 2024
Available online 20 April 2024

analysis of protein arrays that characterize cells, tissues, and organisms. Proteomics aimes to integrate proteomic experiments with data-based network consolidation analysis to investigate the composition, structure, expression, and modification status of proteins as well as the interactions and connections of proteins-to-proteins in organisms. It has opened promising avenues for life science exploration in the post-genomic era, providing insights into the cellular and molecular mechanisms and dynamic regulation of disease[5,6].

Employing advanced technologies to tailor diagnosis, treatment, and prevention strategies according to the unique characteristics of each individual, such as genomics, proteomics, and metabolomics could promote precision medicine. Over the past decade, it has been increasingly evident that even patients with the same cancer respond differently to homogeneous regimens such as chemotherapy and radiotherapy. Precision oncology leverages the molecular characteristics of individual patient tumors to assess the likelihood of benefit or toxicity of personalized therapeutic interventions. The approach relies on the assumption that the molecular target state of the drug in a patient's tumor matching the molecular mechanism of action with a therapeutic agent will optimize treatment outcomes. In recent years, proteomics which provides unique information relevant to cancer biology has become a research hotspot in the field of cancer precision therapy. Highly diverse and spatio-temporally dynamic proteins are involved in a myriad of cell functions, ranging from the transmission of genetic information to metabolic processes. Proteomic analysis could detect changes in the expression of different proteins, providing a more holistic tumor signature. The utilization of histology datasets in precision oncology is limited by elongated sample run durations, sophistication and ambulatory scope of proteomics samples, scarcity of interlaboratory reproductive capacity, variation in quantitative methods, and other confounding factors[7,8]. Fortunately, accomplishments in instrumentation, sample preparation, and data analysis, coupled with standardization initiatives, have boosted the accessibility of high-level, reproducible, and holistic proteomics and phosphoproteomics datasets and agreements[9–11]. Sharma *et al.* developed a rigorous procedural and computational workflow to plot over 50,000 distinct phosphorylated peptides in a single tumor cell line[12]. The Clinical Proteomic Tumor Analysis Consortium also amassed numerous targeted and global proteomics datasets to assist cancer research[13].

Regarding the escalating threat of malignancy, proteomics could provide novel prospects into the tumor molecular landscape, contributing to a more accurate diagnosis and personalized therapies targeted at the specific tumor biomarkers. We searched for relevant literature in PubMed and Web of Science by using keywords including proteomics and precision oncotherapy. In this review, we have selected mainstream and cutting-edge techniques each with the analytical scope of characterizing hundreds to thousands of protein targets used in proteomics for oncology applications. Afterward, we highlighted bioinformatics approaches for integrating proteomics data. Ultimately, we explored the potential applications of proteomics in the context of precision oncology.

2. Proteomics analysis technology

2.1. Mass spectrometry (MS)

Since proteomics qualitative analysis techniques could not satisfy the requirements of clinical investigation, the focus gradually shifted to quantitative analysis, with an increasing demand for MS. Quantitative proteomics strategies based on MS, including global and targeted approaches, can directly measure the protein quantity in a cell. Global proteomics methods indiscriminately quantify and identify all proteins from a given sample, having the advantage of not requiring any hypotheses other than measurable differences in one or more protein species between the samples.[14]. In contrast, the targeted proteomics strategy narrows down the monitoring to a limited number of features [15]. Following the acquisition of specific proteins or biomarkers of interest in global proteomics, these can be further validated and examined in more samples using targeted proteomics, which requires knowledge of the target proteins' amino acid sequence to set the appropriate mass parameters[16]. The selected reaction monitoring methods primarily encompassed three stages: 1) Primary MS scanning to screen out parent ions consistent with the specificity of the target molecule; 2) Collisional fragmentation of the parent ions to remove interfering ions; and 3) Acquisition of mass spectral signals from selected specific ions only[17,18]. Compared to global proteomics, targeted methods are optimized for higher sensitivity and throughput across a large number of samples.

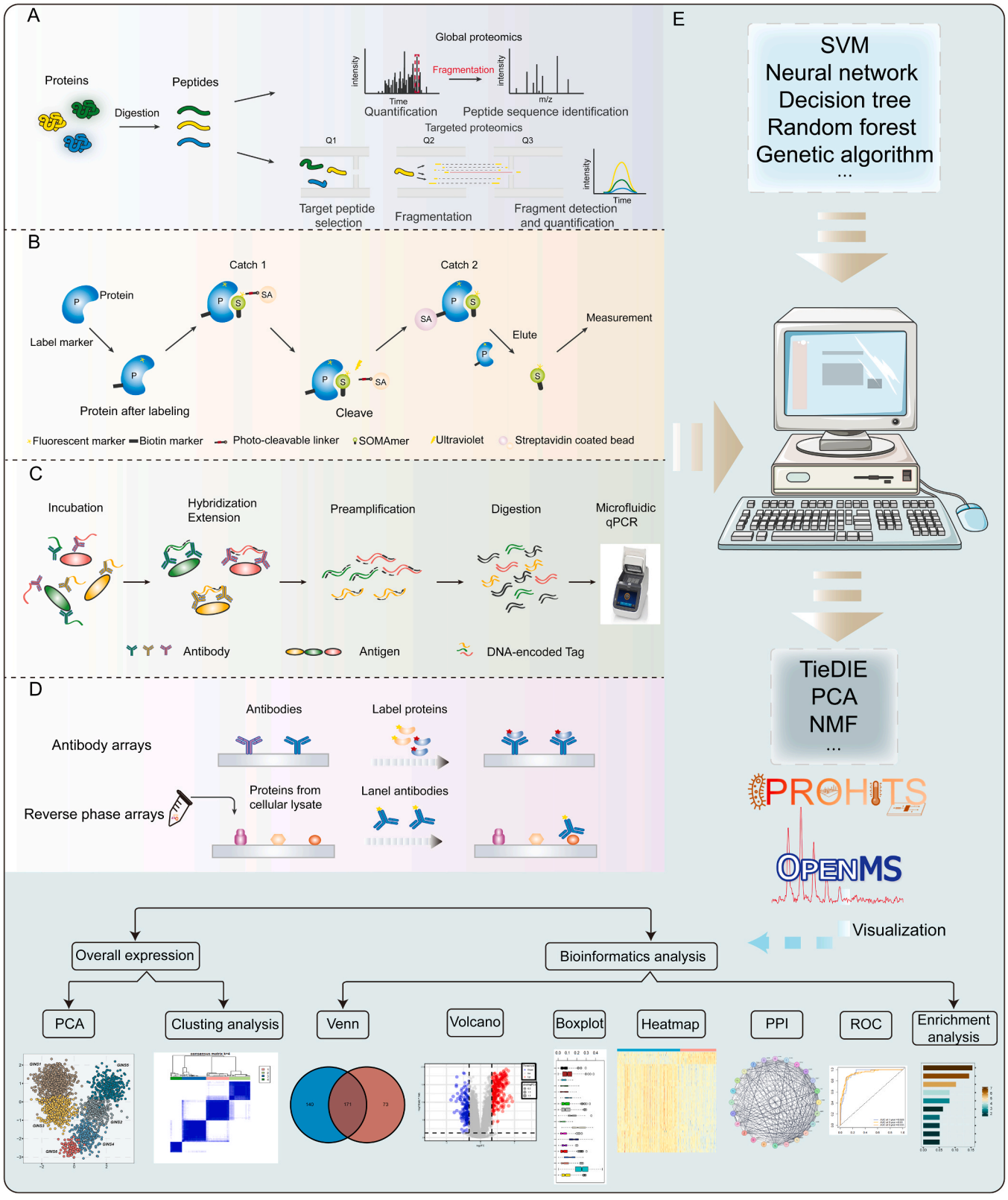
The quintessential solution for MS-based protein analytics is bottom-up proteomics **(**Fig. 1**A)**[14]. In this approach, proteins extracted from plasma, cells, or tissue specimens are first digested by enzymes into peptides. Subsequently, these peptides are typically isolated using liquid chromatography (LC) and then processed for analysis via electrospray ionization MS. Ultimately, the LC electrospray ionization tandem MS provides detailed information on the abundance and sequence profile of the peptides in the sample[14,19].

To enhance the accuracy and efficiency of protein/peptide identification, MS incorporated with various separation and pre-fractionation techniques offers a promising approach[20]. While two-dimensional polyacrylamide gel electrophoresis, a standard proteomics assay based on charge and molecular weight, is labor-intensive and time-consuming, making it unsuitable for high-throughput proteomics. Nevertheless, LC employed polarity, charge, and protein molecular weight for protein separation. LC or high-performance LC (HPLC) could be coupled with MS as LC/HPLC-MS to facilitate the sequential separation of thousands of proteins from complex mixtures[21,22]. Dai *et al.* analyzing serum specimens via LCMS/MS found that the sensitivity of diagnosis was enhanced from 35.1% to 84.0% when alpha-enolase (ENO1) cooperated with tumor protein biomarkers such as carcinoembryonic antigen (CEA) and cytokeratin 19 fragment (CYFRA 21-1) in non-small cell lung cancer (NSCLC)[23]. The synergistic detection of the proteins paraoxonase/arylesterase 1 and alpha-1-antichymotrypsin in serum specimens had a sensitivity of 94.4% and a specificity of 90.2% for early diagnosis of NSCLC[24]. In addition, LC-MS established for the first time a combination of protein markers, which could differentiate primary squamous lung cancer from head and neck squamous lung metastatic cancer, leading to a trustworthy foundation for classifying unknown origin squamous lung cancer[25].

2.2. Aptamer-based detection

Aptamers, single-stranded deoxyribonucleic acid or ribonucleic acid molecules, could fold into specific tertiary structures to tie with elevated affinity and specificity to homologous protein targets in their natural state[26,27]. In slow off-rate modified aptamer scanning assays, biotin and fluorescent markers labeled the protein to be detected. The aptamer captured and bound to the specific protein to form a complex and attached to streptavidin beads. After releasing by ultraviolet-based cleavage and binding to another biotin bead, the complex survived the subsequent challenge and was purified, characterized, and measured by DeoxyriboNucleic Acid (DNA) hybridization techniques to reflect protein abundance within the system **(**Fig. 1**B)**[28]. Compared to antibodies, nucleic acid aptamers conferred advantages in enhanced affinity and specificity, facile in vitro synthesis and selection, and minimal batch-to-batch variation, presenting an appealing option for biomarker discovery[29].

In addition to demonstrating the ability to perform a versatile range of molecular target screens encompassing challenging-to-produce antibodies such as low-immunogenicity targets and toxins, aptamer screening was characterized by cost-effectiveness and efficiency, with screening results obtained within a week[30]. Recently, automated DNA/RiboNucleic Acid (RNA) synthesis technologies have further



(caption on next page)

streamlined the production and chemical modification of aptamers on a large scale with low- to high-volume variation[31]. The chemical structure of these straightforward aptamers remained complete conformational restoration even after thermal or chemical denaturation.

Moreover, using already-known molecular targets, like proteins, peptides, or small molecules, aptamer screening allowed the indication of aptamers for cancer-related biomarkers. Kruspe *et al.* bound photosensitizer chlorin e6 (ce6) to human interleukin-6 receptor (IL-6R)-binding

Fig. 1. Proteomics Technology Processing, data analysis and integration. A Mass Spectrometry: proteins are first digested enzymatically into peptide fragments, which then enter the mass spectrometer for detection. In global proteomics, a full spectrum of peptide ions is obtained and fragmentation is performed to identify the peptide sequence. In targeted proteomics, signals that match the target ion rules are acquired with triple quadrupole mass spectrometry, with Q1 selecting the parent ion, Q2 fragmenting the peptide ion, and Q3 detecting the fragmented ion signal. **B** Aptamer-based detection: the protein to be detected is labeled with biotin and fluorescent markers, and the aptamer is captured and attached to a specific protein binding to a streptavidin bead. After ultraviolet cleavage, they are released and bound to another biotin bead, eluted and recovered for quantification, and hybridized to a custom DNA microarray to react to the amount of protein in the sample. **C** Proximity Extension Assay: each paired antibody with a unique DNA tag binds to the target protein of the sample. After binding, the DNA tags are nearby and hybridize, generating a double strand that is digitally recognized, amplified, and detected by qPCR, which translates into data. **D** Protein microarrays: in forward protein microarrays, specific antibodies can be immobilized on the chip surface to capture target proteins in complex samples, and the unbound components are washed away by rinsing to measure the fluorescence intensity at various points on the chip. In reverse-phase protein microarrays, proteins are bound to the chip, which is then detected with an antibody against the target protein, rinsed to remove unbound components, and fluorescence scanned to determine the fluorescence intensity at various points on the chip. **E** Data generated by technologies requires computational algorithms to derive novel biological insights. After data integration using relevant methods (TieDIE, PCA and NMF) and platforms (ProHits and OpenMS), visualization figures are formed through standardized analysis.

RNA aptamer AIR-3A to produce AIR-3A-ce6, which was rapidly and specifically internalized by IL-6R-presenting cells. Under photo-irradiation, the target cells were selectively killed, while the free ce6 showed no toxic effect, doing its part in remedy of breast tumors and skin cancer[32].

The elongated shelf lifetime and nucleic acid nature granting direct sensitivity to antidotes were other properties of the aptamers, making them attractive candidates as contributors to targeted therapies[33,34]. It was reported that the selection of RNA aptamers for conjugating tumor-associated membrane antigens and the application of RNA aptamers to prostate-specific cellular markers[35]. In the clinical setting, these aptamers may function as N-acetyl-alpha-linked acid dipeptidase inhibitors or be modified to carry imaging and therapeutic agents targeting prostate cancer cells. Then, Wang and colleagues successfully certificated DNA aptamers to promote the development of novel targeted cancer detection, imaging, and therapy by identifying multiple epidermal growth factor receptor (EGFR)-expressing cancer cells but not augmenting EGFR-negative cells[36].

2.3. Proximity extension assay (PEA)

PEA is a unique proteomics method that uses antibodies linked to oligonucleotides for quantifying proteins with real-time Polymerase Chain Reaction (PCR). The technique had been shown to possess excellent readout specificity and sensitivity enabling high-volume multiplex analysis over a wide dynamic range as well as conserving minimal sample volume. By employing paired oligonucleotide-labeled antibodies coupled to their target antigens in a paired fashion, PEA could bring matched oligonucleotides into proximity. Upon antibody binding, the downstream procedure usually carried out by quantitative polymerase chain reaction entails the acquisition of DNA polymerase to create, amplify, detect, and quantify the PCR target sequence (Fig. 1**C**) [37,38]. To enhance the high-throughput screening of biological samples and expand assay libraries, researchers have automated, miniaturized, and adopted next-generation sequencing (NGS) technology to propel PEA to a frontier[39]. Up to now, NGS stands as the market leader in massively parallel short-read sequencing. The amalgamation of PEA technology with NGS reads constituted an impressive landmark in the new era of protein identification and quantification.

PEA remained unaffected when analyzing complex biological materials without the washing steps, becoming a panacea of liquid biopsy-based detection[38]. Target-specific antibody pairs were tethered to DNA strands and generated real-time PCR amplicons in a proximity-dependent manner by conjugating simultaneously to the target analyte in the presence of DNA polymerase. In the PEA, multiple antibody pairs of the target protein are aggregated, each labeled with a complementary DNA oligonucleotide sequence. High-fidelity discrimination hybridization could be achieved by several measures, including reducing pipetting steps, selecting superheat-resistant enzymes, introducing interplate controls, and novel standardized procedures[37]. As a superior vehicle of serological discovery, PEA was applicable to discriminate early-stage colorectal cancer (CRC) and identify potential amalgamations of plasma protein biomarkers, further extending to asymptomatic CRC individuals[40]. Pre-treatment plasma samples from anti-programmed death-ligand 1 therapy NSCLC patients were analyzed using PEA to quantify different immune oncology-related proteins[41]. By employing a similar strategy, Liu *et al.* deciphered the high content of organization pleiotrophin to be an independent predictive marker of chemotactic relapse and migration progress in low and moderate prostate cancer[42].

2.4. Protein microarrays

Protein microarrays concurrently measured a vast array of distinct proteins in a single experiment, permitting parallel evaluation of an extensive collection of proteins. There were generally two types of protein arrays: 1) Forward-phase protein arrays were arrayed for the detection of proteins and consisted of differentiated capturing substances; 2) Reverse-phase protein array (RPPA) detected molecules of interest such as proteins, drugs, and nucleic acids by the proteins on the array (Fig. 1**D**)[43].

For forward-phase protein microarrays, commonly antibody arrays, the components upon the array were capturing domains[44]. As trailblazing tools in targeted proteomics, antibody arrays parsed many proteins of an individual specimen to enable the discrimination of specific disease biomarkers. Antibodies, acting as capture or decoy probes, immobilized specific proteins on the revised plane-primed substrate through chemical interaction or physical encapsulation[45,46]. After a hybrids-like capture, the following phase was to inspect the components of the particular capture. There were two main classifications of assignment formats: label-based and sandwich-based. The former approach necessitated sample pre-treatment before incubating in the microarray. Analytes could be labeled directly with fluorescent clusters or quanta, or marked after capturing with secondarily labeling antibodies[47]. Nevertheless, the sandwich methodology did not require pre-labeling, but every targeted spot demanded two separate antibodies. The primary antibody captured the sample and the secondary antibody combined with another protein-conjugated domain of the target site for binding detection[48]. In high-complexity analysis, materials were predominantly branded with fluorometric, chemo-luminescence, or oligo tags to permit divergent signal magnification and monitoring. The protocol could virtually characterize more than 1000 proteins or their modification forms at minimum immunogenicity crossing correlation triggered by the antibody response mixture.

Antibody arrays overcome the sensitivity issues associated with non-targeted proteomics technologies on account of their excellent performance in knowledge-based biological mechanism exploration. It discerned, quantified, and monitored a broad spectrum of intracellular and serum proteins such as cytokines, hormones, chemokines, and intracellular signaling molecules to screen for disease markers, thereby elucidating mechanisms of disease onset and progression[49,50]. Cytokine antibody arrays became an awesome assistant in explaining the indispensable function of interleukin (IL)- 8 in breast carcinogenesis by measuring protein expression levels of different cytokines in breast

carcinomatosis[51]. Subsequently, leveraging antibody microarrays, some researchers analyzed the resistance mode of various breast carcinoma cell clusters to adriamycin, which improved the understanding of protein expression changes associated with adriamycin resistance and identified novel biomarkers[52]. Ingvarsson *et al.* accessed protein signatures relevant to a shorter life expectancy in the sera of pancreatic cancer patients by recombinant single-chain variable fragment antibody microarrays[53]. Taken together, antibody arrays offered valuable insights into tumor diagnosis, treatments, and prognosis.

Several decades ago, Paweletz *et al.* pioneered RPPA, a proteomic technique with a high degree of sensitivity, precision, and linearity, especially for quantifying the phosphorylated status of signal proteins in tumor cell subpopulations[54]. The platform sequence steps involved the immobilization of protein-containing samples on nitrocellulose-coated slides, antibody recognition of target proteins, signal amplification chemistry to detect protein-antibody complexes, and quantification of spot intensity. Subsequently, each RPPA slide was assayed with catalytic signal amplification chemistry and a primary antibody, a secondary antibody, and fluorescent or colorimetric dyes. Nowadays, this technology has been graduated for application in clinical trials with higher sensitivity and precision[55].

Protein quantified in cell lysates and PTMs in body fluids via RPPA could be harnessed for biomarker identification, protein pathway analysis, functional phenotyping, and drug discovery mechanisms of action, dramatically accelerating clinical management in individualized therapy. Murakoshi *et al.* validated candidate plasma biomarkers for early detection of CRC and age/sex-matched healthy individuals by RPPA for plasma protein analysis[56]. The protein expression data of RPPA had been certified to contain substantial predictive power for cancer dependencies[57]. In addition, Tibes and colleagues described the validation and robustness of RPPA in rare populations of normal and leukemic stem cells and pictured differences from bulk cell populations [58]. RPPA application in proteomic profiling of tumor-derived extracellular vesicles (EVs) has become another area of interest[59]. A preliminary study utilizing EV purification workflow followed by RPPA analysis of 276 cellular proteins, revealed seven protein biomarkers of breast cancer patients with diagnostic power[60]. Another RPPA study on EVs from prostate cancer patient sera also validated protein biomarkers with therapeutic monitoring value[61].

In contrast to antibody arrays consisting of ample antibodies deposited onto the substrate with the specimens as probes, protein lysates in RPPA were affixed onto a matrix, and a solitary primary antibody served as a probe molecule[54,62]. The reversed-phase modality facilitated simultaneous specimen quantification on a singular array, and the availability of multiple RPPA slides permitted for elaborate protein measurement, with each array acting as a comprehensive analysis. Accurate protein quantification relied on the identification of protein specimens in images acquired with high precision, followed by speckle analysis. Speckle analysis was capable of distinguishing speckles and measuring the pixel intensity of each speckle concerning the surrounding background[62,63].

2.5. Combined techniques in single cell proteomics

Off-target detection of proteins using mass spectrometry lacks sufficient sensitivity to be relevant for single-cell analysis. Therefore, single-cell proteomics studies have relied on selective target detection using antibodies or similar affinity reagents (e.g., affinity bodies or aptamers). Although the number of epitopes that can be detected simultaneously is limited by the spectral overlap between fluorescent moieties, flow cytometry has long been the workhorse for researchers seeking to assess protein expression at the single-cell resolution level. The advent of MS-flow cytometry, which involves the use of metal isotopes to label antibodies, has greatly expanded the number of markers that can be analyzed concurrently, allowing for high-dimensional detection of proteins expressed on single cells. Wagner *et al.* used a combination of 35 immune cell-centric and 38 tumor cell-centric antibody panels to generate a comprehensive single-cell atlas of breast cancer[64]. Another innovation based on single-cell barcode microarray technology is the IsoLight system, capable of detecting proteins secreted by single cells[65]. The system utilizes a microfluidic chip to capture information from 200 to 2000 single living cells and measures cytokines, chemokines, growth factors, and other secreted ligands using a sandwich ELISA-like assay over several hours of incubation. The system also has the flexibility to incorporate other analytes such as metabolites. The initial utility of the IsoLight System has been demonstrated in a Phase 1 trial in which patients with advanced melanoma underwent relay cell transfer supported by polyethylene glycolated IL-2, which resulted in enhanced levels of cytokines between polyfunctional T cells and NK cells [66].

Deep visual proteomics (DVP) enables single-cell analysis in a tissue context by first imaging the tissue at subcellular resolution using an optical microscope. The individual cellular or subcellular structures of each cell in the image are then automatically identified using sophisticated deep learning-based methods, followed by proteomic analysis by ultra-high sensitivity mass spectrometry using automated laser microdissection to target and capture the cell or region of interest. Based on the proteomics data, cells can be categorized into types or assigned functions. Based on this the proteome of cancer cell lines at different stages of the cell cycle characterized by fluorescence ubiquitination-based cell cycle indication (FUCCI) was determined[67]. In archived primary melanoma tissues, DVP identified spatially resolved proteomic changes as the transition of normal melanocytes to fully invasive melanoma, revealing pathways that change spatially as cancer progresses, such as dysregulation of mRNA splicing in metastatic vertical growth, which coincides with reduced interferon signaling and antigen presentation[68].

3. Proteomics data analysis and integration

The technologies mentioned above could generate vast amounts of data, which require computational algorithms to derive novel biological insights **(**Fig. 1**E)**. Biomarker identification faced the challenge of acquiring correlative and robust signatures from high-throughput proteomics data. Differential expression proteins between normal and disease states could address this problem[69,70]. For instance, an investigation demonstrated that α-enolase, α-catenin, 14–3-3 β, VDAC1, and calmodulin significantly expressed at least 2-fold between cancer and normal mice. Changes in these marker levels could provide insights into the physiologic aberrations and disease stages during pancreatic cancer progression[70]. Additionally, more elaborate techniques appeared to bridge this gap, like machine learning and network-based approaches, enabling the discernment of corollary biomarkers and the dissection of intricate relationships between proteins. Proteomics data analysis benefited from machine learning methods such as support vector machine (SVM)[71], neural network[72], decision tree[73], random forest[74], and genetic algorithm[75]. By comparing the average accuracy of random forest and SVM algorithms, Ahn *et al.* found that although SVM was talented in detecting small tumors, random forests generically outshined SVM no matter the phase or scale of the tumor. They eventually constructed a biomarker array platform and employed the random forest feature selection algorithm to screen candidate biomarkers in gastric adenocarcinoma[76]. These algorithms harvested enhanced sensitivity and specificity, precise diagnosis capabilities, and improved treatment outcomes. Frustratingly, The machine learning methods were weaker in the independent validation cohort, displaying the overfitting phenomenon[77]. Fortunately, to solve the deficiency of SVM in analyzing large datasets, a fast support vector classifier (FSVC) was proposed boasting free memory requirements and less time spent, showing promise for better data analysis[78].

Elucidating the underlying mechanisms of tumorigenesis and evolution paved the way for targeted therapeutic strategies. The researchers

typically attempted to recognize biological processes and functions derived from bioinformatic data analysis, such as gene ontology or pathway enrichment analysis for differential regulatory entities in each dataset[79–81]. Locard and colleagues performed a phosphoproteomic analysis and constructed a map of bidirectional signaling and interaction pathways by navigating identified phosphomodulatory proteins to the Kyoto Encyclopedia of Genes and Genomes pathway, characterized by alterations in phosphorylation upon tumor-endothelial interaction[82].

Evolving technologies to analyze protein interaction networks and pathway architectures could detect pathways disrupted in a patient-specific manner[83]. Among these techniques, the Signaling Pathway Impact Analysis algorithm was predicated on the location of genes in the pathway topology, incorporating the message of differential gene expression and its implications for the pathway[84]. Stemming from a network diffusion approach, Tied Diffusion Through Interacting Events (TieDIE) connected genomic perturbations to variations in gene expression underlying cancer subtype signatures, calculating sub-networks of protein-protein interactions, along with forecasted transcription factor-to-target connections[85]. Other methods included network propagation[86], clustering[87], network current flow[88], random wandering[89], and pathway models in determining disturbed function blocks or routes in the net and availed these as features to layer individuals or differentiate carcinogenic cellular lineages[90]. Nevertheless, these algorithms gravitated towards available interactome data and annotation pathway data, as these were presently fragmented and skewed toward heavily expressing proteins[91–93]. However, hampered by the instrumentation or technique application and the dynamical scope of abundance in the specimen, their coverage of the entire proteome was frequently much narrower than other correlative histological data[94]. The kinase-substrate enrichment analysis estimated kinase activity according to differences in the abundance of known substrates to accentuate specific signal networks active in each circumstance[95].

The methodologies amalgamating proteomics datasets with other omics or clinical data could be divided into homogeneous methods (datasets containing congruent data from diverse provenance) and heterogeneous methods (integrating several datasets with different database types). These methods could incrementally aggregate the data strata or products in the system integration pattern. Drake *et al.* adopted a stepwise method to consolidate genomics, transcriptomics, and phosphoproteomics datasets and implemented a diffusion-based algorithm, TieDIE, to discover differentially expressed master transcriptional regulators, functionally mutated genes, and differentially activated kinases[85,96]. The outcomes revealed clinically relevant pathway information as well as advanced prostate cancer patient stratification and targeted therapies, together with profiting from a global pathway-based reference for individual patient drug prioritization. The furthest prospective strategies to integrate datasets in a single step entail techniques predicated on principal component analysis (PCA) or factor analysis and non-negative matrix decomposition (NMF)[97–99]. These methodologies could integrate different large datasets and perform effective dimensionality reduction, thereby facilitating downstream network-base or machine learning analysis and constructing system-representative models[97,100]. Moreover, distinct supervised or unsupervised techniques could be implemented to select proper variables and contribute to the interpretation of the result[101,102]. These could embrace the implementation of linear discriminant analysis[102], Bayesian classifiers[103], SVM[104], and K-nearest neighbor methods [105]. Through factor analysis and linear discriminant analysis, combining microRNA, mRNA, and proteomics data and exploring the tumor molecular mechanisms, could reveal the related gene expression models and identify promising immunotherapeutic spots[103]. Matrix decomposition techniques such as NMF and variance traditionally analyzed genomics and transcriptomics data and had only recently been extended to proteomics datasets. Yuan *et al.* evaluated the prediction of patient survival in different molecular data types and described the potential prognostic and therapeutic relevance of multiple cancer types via NMF. Among the subgroups, they pinpointed a subgroup among other biomarkers and activation pathways profiting from mitogen-activated protein kinase (MAPK)-targeted therapy[106]. Another integration strategy utilized for proteomics datasets was predicted for co-inertial analysis of multilateral extensions. Through integrating cellular transcriptomic and proteomic profiles in NCI-60 cancer cells, it was uncovered that leukemia extravasation pathways were fundamental players in leukocyte migration and metastasis, but the same pathways were not observed in the analysis of individual datasets [107].

Inconsistent annotation and reporting of integrated data along with its dynamic nature brought about obstacles in models for investigating diseases or patients. With the evolvement of a standardized data collection and management process, it was critical to establish methods in synergy with extant techniques for genomic and other histological datasets and harness them to document relevant metadata. The growing sophisticated platforms and techniques such as ProHits and OpenMS, aimed at minimizing the variability associated with data acquisition [108,109]. ProHits, a software platform designed primarily for interaction proteomics, could deliver fancy options for data management and analysis to guarantee the analytical pipeline traceability[108]. Relatively, as an open-source framework tailored for high-throughput MS data analysis, OpenMS could serve as a transparent and scalable approach to implement various pipelines and analysis procedures[109]. Additionally, Compared to selected reaction monitoring, Sequential Window Acquisition of All Theoretical Mass Spectra (SWATH-MS), a mass spectrometric method that combines data-independent acquisition and targeted data analysis, could achieve the one-time, a priori generation of a specific measurement assay for each targeted protein, significantly broadening the production of targetable proteins[110]. TRIC, a software exploiting fragment-ion data, conducted cross-run alignment, consistent peak-picking, and quantification for high-throughput targeted proteomics. This minimized the identification error and corrected for highly nonlinear chromatographic effects, filling absent components in the pipeline of extensively parallel targeted proteomics datasets analysis[111].

The inherent variability of proteomics and phosphoproteomics datasets hampered data consolidation. The single-cell research demonstrated that activated cell-specific responses were modulated and coordinated with significantly less noise and sample variability compared to the resting state[112]. Where feasible, the acquiring of non-stationary data points assisted in mitigating data variations and boosting the signal-to-noise ratio. By combining microfluidic chips for all-in-one proteomic sample preparation and data-independent acquisition MS for proteomic analysis down to the single-cell level, the proteomics chips facilitated multiplexed and automated cell isolation/counting/imaging and sample processing in a single device[113]. Bendall *et al.* married the depth of single-cell MS flow cytometry with developed algorithms and expanded it to human B-cell lymphangiogenesis to construct a trajectory from hematopoietic stem cells to naïve B cells. They highlighted checkpoints where regulatory signals parallel changes in cellular state and laid the groundwork for introducing them into cancer development [114].

4. Proteomics in precision oncotherapy: administrating individualized clinical protocols of tumor treatment

Proteomics has become a research hotspot in the field of precision oncotherapy, which redefines disease subtypes, discerns the primary tumor cell, explores biomarkers and potential therapeutic targets, and ultimately discovers and validates tumor-related molecular regulatory mechanisms at the protein level for promoting therapeutic strategy.

4.1. Identification and classification of cancer subtypes

Researchers discovered that patients with the same type of cancer showed substantial variations in the quality of livelihood and life expectancy, demanding imminent requirements for the clinical identification and classification of cancer[115,116]. Molecular typing represented an imperative dimension in precision cancer therapy, whereas molecular typing based on genetic mutations and transcriptional differences couldn't guarantee satisfactory guidance in clinical practice. Proteomic-based molecular typing gradually attracted attention in the research field. Leveraging MS-based characterization, ovarian cancers were typified with four proteomic clusters corresponding to mesenchymal, proliferative, immunoreactive, and differentiation subtypes as defined by genomic analysis. Specific protein acetylations associated with homologous recombination deficiency also provide stratified treatment for patients[117]. Yanovich and colleagues demonstrated intratumor heterogeneity of three breast cancer subtypes and normal tissue. One cluster of four proteomic clumps distinguished via unsupervised analysis manifested a novel luminal isoform characterized by increased PI3K signaling, which led to phosphorylation and elevated transcriptional activity of ER and represented tamoxifen resistance. This work might serve as a springboard for a targeted study of hormonal treatment response, and further research including clinical data is warranted[118]. Ge *et al.* classified diffuse-type gastric cancer (DGC) into three subtypes (PX1–3) according to solely altered proteomes. PX1 and PX2 exhibited disorders in the cell cycle with PX2 characterized by an additional epithelial-mesenchymal transition procedure. PX3 had the least favorable prognosis and was unresponsive to chemotherapy, while containing abundant immunoreactive proteins. Data acquisition decrypted four major vulnerabilities adaptable for therapeutic management and permitted the nominating of prospective immunotherapy targets for DGC individuals, especially those in PX3, indicating a prime target for immunotherapy (Fig. 2**A**)[119]. In addition, many other studies have been conducted, including brain cancer [120], breast cancer[121,122], lung cancer[123], clear cell renal cell carcinoma[124], and prostate cancer[125]. Overall, proteomic analysis might emerge as a promising superstar in cancer molecular subtyping followed by genomic and RNA sequencing, whose translation into clinical practice and care remains to be deciphered.

4.2. Discerning the primary cancer cell type

With the advent of the precision oncology era, proteomics delivered innovative insights into the mechanisms of cancer generation, empowering people to break through the constraints of previous investigations. Proteomic profiling unlocked that protein representation models were coincident with the origin of certain cancer cell types[126]. Though projecting the protein intensities onto a Uniform Manifold Approximation and Projection plot, Goncalves and colleagues identified tumor cell source type groupings including diverse populations of hematopoietic and lymphatic-like cells and further subdivided into different cell lineages. A robust and preserved cell type-specific procedure existed between transcription and translation. Ultimately, they also showed that protein expression characteristics correlate with cellular ancestry and

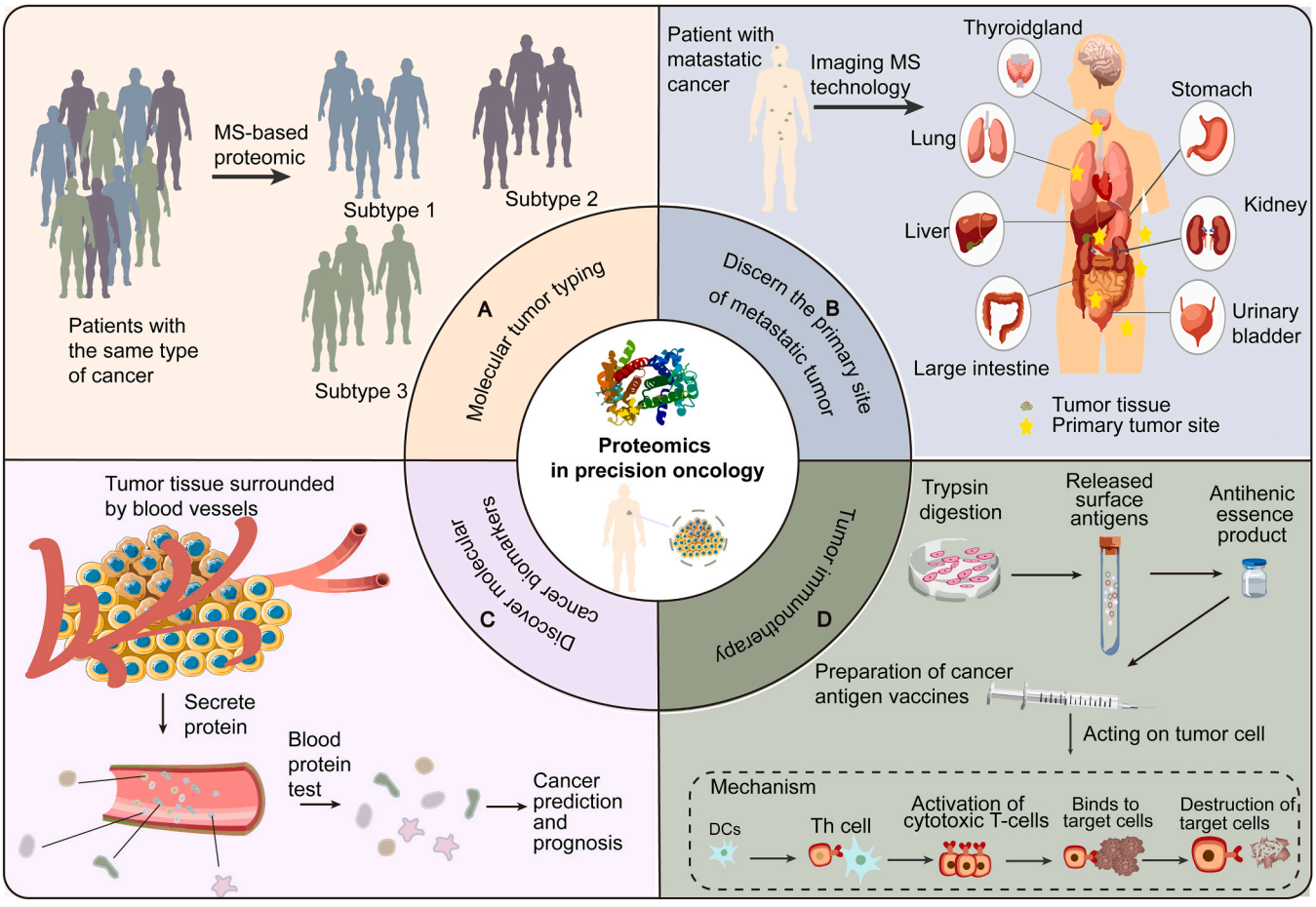


Fig. 2. Proteomics empowering precision oncology. A Patients with the same type of cancer could be further subdivided into different molecular subtypes based on the results of proteomic analysis. **B** For metastatic cancers with unknown primary foci, proteomic analysis assists in locating the primary foci. **C** Tumor tissues release proteins into the blood, and these proteins could be specific markers of tumor types and protein detection could determine cancer types. **D** Preparation of cancer antigen vaccines and their mechanism of action.

established a rigorous protein quantification strategy, thereby reducing high-dimensional complexity[127].

Given that various tumors and their subtypes necessitated distinct treatment modalities, accurately discerning the primary tumor site of metastatic cancer could guide clinicians in rendering pertinent medical decisions and evaluating expected outcomes, and overall prognosis (Fig. 2**B**). The diagnostic approach encompassed a detailed histopathological examination of formalin-fixed paraffin-embedded tissue stained with hematoxylin-eosin staining. This procedure was buttressed by histochemical reactions, immunohistochemical, and molecular biological methods[128]. It was noteworthy that the metastases differentiated moderately or poorly and lost the specific hallmarks of the primary tumor. Imaging MS technology could investigate a large number of proteins in association with morphological features, simultaneously maintaining their spatial integrity in tissue samples[129]. Matrix-assisted laser desorption ionization imaging constructed tumor proteomic signatures to classify adenocarcinoma entities located at different organ sites, which could discriminate from different organ sites and distinct tumor types in the same site[130]. Li and colleagues distinguished three CRC subtypes featured by diverse clinical prognostic and molecular profiles and successfully differentiated metastatic cases by phosphorylated proteomics. Phosphoproteomic data distinguished the primary tumors from metastatic CRC and non-metastatic CRC, resulting in the classification of six phosphoproteomic subtypes in which SC1, SC3, and SC5 were characteristic of the former, whilst SC2, SC4, and SC6 were enriched in the latter. Nevertheless, metastatic tissues rendered high similarity to original tumors at the hereditary level rather than the proteomic layer[131]. To summarize, proteomics emerged as a capable partner for clinicians to troubleshoot the conundrum of identifying the primary site of migratory cancers. Its integration with existing diagnostic methods was expected to refine the precision of cancer diagnosis.

4.3. Detection of biomarkers

Molecular cancer biomarkers undertook measurable indicators of cancer risk, genesis, or individual prognosis, and proteomics carried out a trustworthy strategy for the pursuit and validation of prognostic and predictive biomarkers (Fig. 2**C**). An international study conducted that high-throughput proteomics of serum EV identified early diagnostic markers for individuals at high risk of cholangiocarcinoma (CCA) and biomarkers for the differential diagnosis of intrahepatic CCA and hepatocellular carcinoma (HCC). Research has documented that in vitro CCA human cell-derived EVs contained a greater enrichment of oncogenic proteins versus EVs released from healthy human bile duct cells [132]. Wang *et al.* carried out profiling of pulmonary neuroendocrine carcinoma by quantitative tandem mass tag proteomics, characterized to subtype-specific secretory protein markers, and proved that insulin-like growth factor-binding protein 5 could emerge as a biomarker for oncogenic achaete-scute homolog 1 high expression small cell lung cancer subtype[133]. Indeed, the detection of low-abundance proteins presented a puzzle for tumor marker development, but Olink was making a difference, supplying assays and analytical data for cytokines, chemokines, and growth factors, among others. The research detected prediagnostic blood samples by the Olink96-inflammation panel and identified CDCP1, an early screening marker for lung adenocarcinoma, which in turn was demonstrated by mRNA-seq enrichment analysis to have a potential role in pathways associated with cell adhesion and migration[134]. Biomarker discoveries have also been reported in genomic studies of gastric[135], ovarian cancers[136,137], lung cancer [138], and brain cancer[139]. Protein-based and etiology-related logistic models with predictive and diagnostic capacities were emerging, moving a step forward into precision oncology.

4.4. Elucidating the molecular mechanisms of tumorigenesis

To investigate protein network regulation and cellular phenotypic modifications from the perspective of genomic variation remained extremely complicated, while direct analysis of protein co-regulation/covariation at the proteomic level could reveal the effects of relevant gene mutations on protein networks more accurately. Roumeliotis *et al.* demonstrated that gene loss-of-function mutations could influence protein levels to deliver protein co-regulatory networks through proteomic studies. The consequences of genomic variants could be transmitted and impinge on protein levels of other genes through protein reciprocal networks[140]. Proteogenomic research of malignancy tissues validated that copy-number variances in a large portion of the proteome could be reduced at the protein level, constituting a protein homeostasis network. Sousa and colleagues assembled genomic, proteomic, and structural data from malignant tissues and approved that the majority of proteins displayed a panorama of post-transcriptional regulatory traces and interaction-dependent control of protein abundance [141]. Failure to eliminate DNA lesions was associated with genome instability, a driving force in tumorigenesis. The maintenance of cellular functionality assured the protein homeostasis network under conditions of proteome instability, addressing the triage decision of protein fold, hold, or degrade. The outcome increased genomic instability due to reduced fidelity in processes like DNA replication or repair leading to cancer[142]. Overall, these findings provided evidence for the existence of an active protein homeostasis network, which was supported by direct proteome measurements.

Cancer signaling mechanisms have been investigated by diverse programs, and proteomics enables the visualization of gene-expressed proteins and shortens the research journey[143]. Proverbially, mitogen-activated protein kinase kinase kinase kinases (MAP4Ks) constituted a mammalian STE20-like serine/threonine kinase subfamily. The Hippo pathway, originally authenticated and elucidated in Drosophila, represented a conserved signaling pathway that regulated organ/body size by inhibiting cell proliferation and promoting apoptosis and bore considerable relevance in the suppression of cancer pathogenesis[144]. In the mammalian Hippo pathway, MAP4K1/2/3/4/6/7 operated simultaneously with two other STE20-like kinases, MST1/2, to phosphorylate and activate two ACG serine/threonine kinases, LATS1/2 phosphorylated and activated[145]. Upon phosphorylation, LATS1/2 and its splice protein MOB1A/B sequentially phosphorylated the downstream effector's YAP and TAZ. When the Hippo pathway was inhibited, unphosphorylated YAP/TAZ was released into the nucleus to form a complex with TEAD1–4 transcription factors that regulated the transcription of genes related to cell proliferation and survival[146]. Seo *et al.* harnessed proteomic analysis to obtain a MAP4KS interacting protein, STRN4, and correlation analysis of its expression with YAP in endometrial cancer suggested STRN4 as a putative oncogenic factor in endometrial cancer[147].

4.5. Promoting cancer immunotherapy

Cancer survival and progression coincided with the ability of tumor cells to avoid immune recognition. Progress in the recognition of tumor immunity and mechanisms of tumor immune escape has permitted the evolution of immunotherapeutic approaches. However, primary and acquired resistance mechanisms limited the efficacy of immunotherapy. Further therapeutic advances required proteomics to understand the interplay between immune cells and tumors. Beck and colleagues analyzed high proteomic variability among metastatic sites and

identified context-specific cellular processes, with higher immune and mitochondrial activity in lung metastases compared to other metastases, highlighting potential tissue-based sensitivity to immunotherapy. Subsequently, they proposed that a combination of acute MAPKis treatment and immunotherapy had synergistic effects while a decrease in immune activity upon acquired resistance to MAPKis, which reinforced the significance of treatment regimen and timing[148].

The molecular composition of cancer cells approximated similarly to normal cells, and direct cancer vaccination frequently generated side effects and autoimmune reactions in the body, failing to eliminate cancer cells. Researchers uncovered that one should bypass the internal factories of cancer cells and focus on the surface of cancer cells (Fig. 2**D**) [149]. Recently, MS-based immunopeptidomics has allowed the exploration of noncanonical antigens-antigens generated from sequences beyond protein-coding regions or by noncanonical antigen-processing mechanisms. In combination with transcriptomics and ribosome profiling, it facilitated the authentication of noncanonical peptides, which could be detected exclusively in tumors[150]. The direct identification of mutated peptide ligands from primary tumor material by MS was proven to be possible and yielded true neoepitopes with high relevance for immunotherapeutic strategies in cancer[151]. Taken together, proteomics became a powerful weapon to break the limitations of previous tumor immunotherapy.

5. Future directions for proteomics pairing with other omics

Protein phosphorylation represented a reversible modification that operated on diverse kinases and phosphatases to achieve complex and precise biological regulation and functions in cell proliferation, survival, apoptosis, metabolism, transcription, and differentiation. However, abnormal phosphorylation of proteins ordinarily triggers disruption of cell proliferation, leading to tumorigenesis. For instance, there were four members in the erythroblastic oncogene B (ErbB) family of receptor tyrosine kinases: EGFR, ErbB2, ErbB3, and ErbB4. Epidermal growth factor-related peptides incorporated ErbB receptors and inducted the creation of distinct homodimers and heterodimers[152]. Receptor dimerization promoted internal kinase activation, resulting in the phosphorylation of specific tyrosines located in the cytoplasmic region of ErbB. Abnormal ErbB activity was significantly implicated in tumorigenesis and progression. Therefore, ErbB receptors have been proposed as promising treatment targets[153]. Besides, Yang and colleagues disclosed how phosphorylation modifications led to angiogenesis and sorafenib resistance in HCC cells[154]. Li *et al.* subjected esophageal squamous carcinoma (ESCC) to large-scale proteomic and phosphorylated proteomic assay to determine the molecular typing associated with clinical features and the potential therapeutic target of CDC-like kinase 1 for ESCC[155]. In studies of melanoma, investigators have also detected an essential role for protein phosphorylation in drug resistance and tumor transformation[156].

Understanding gene expression regulation requires deciphering DNA-protein interaction in chromatin. Recent developments in chromatin proteomics have enabled the analysis of entire chromatin or subfractions thereof. MS-based analysis of chromatin has emerged as an effective tool for identifying proteins involved in gene regulation through the ability to study protein function and protein complex formation in their in vivo chromatin-bound context[157]. Furthermore, the development of Chromatin Immuno Precipitation and Selective Isolation of Chromatin Associated Proteins (ChIP-SICAP) method allows the capture of specific chromatin domains facilitating the identification of unknown transcription factors interacting with them. Using ChIP-SICAP, Kiehlmeier and his colleague identify therapeutic targets of the hijacked super-enhancer complex in EVI1-rearranged leukemia [158].

Recent modifications in MS-based proteomics permitted direct interpretation of the genomic aberrations' impact, delivering a comprehensive and quantitative analysis of tumor tissue. The fusion of protein expression and PTMs with genomic, epigenomic, and transcriptomic data manifested a frontier area of proteogenomics, bringing fresh perspectives to the biology and diagnosis of cancer. Genomic and transcriptional analyses illustrated genomic alterations and their potential impact, while proteomics furnished immediate insights into protein regulation and signaling in response to these changes. Mertins *et al.* described genomically annotated quantitative MS-based proteomics and phosphoproteomics analyses of breast cancer. It elucidated the practical implications of somatic mutations, minimized the number of candidate nominations for driver genes within key deletion and amplification areas, and formulated therapeutic objectives[159]. To figure out the deregulatory function patterns driving clear cell renal cell carcinoma (ccRCC), Clark and colleagues undertook genomic, epigenomic, transcriptomic, proteomic, and phosphoproteomic signatures on therapeutic-naive ccRCC and paired-neighborhood organization (Fig. 3). A unique cohort of molecular subpopulations connected to genomic instability was evidenced by genomic analysis. The collaboration of proteogenomic assays distinctively determined protein dysregulation of cellular mechanisms influenced by genomic changes, which included oxidative phosphorylation-associated metabolism, protein translation procedures, and phospho-signaling modules[124].

6. Conclusions

The advantages and limitations of cutting-edge techniques for conducting highly intricate proteomics analyses for precision oncotherapy are discussed in Table 1. The maturation of proteomics technologies and data integration methodologies made extensive proteomic data analyses possible across all major cancer types. This enhanced our understanding of tumorigenesis mechanisms and the cancers' molecular composition at the protein level. Such knowledge spurred a more refined cancer classification in modern medicine. Furthermore, it has accelerated the development of prognostic and predictive biomarkers in novel clinical trials. These advancements have established potentially actionable therapeutic targets. Ultimately, these developments are facilitating the progress of individualized precision medicine for cancer patients. The consolidation of proteomics into clinical trials and patient care, as well as interdisciplinary collaboration, data sharing, and patient engagement, will encourage the virtuous circle of recognizing cancer subtypes that respond well or poorly to existing therapies. In this virtuous circle, the exploitation of inventive drugs and regimens that target the vulnerability and resistance mechanisms of recalcitrant cancer subtypes will transform the way we conceive of patients and treatments. However, translating preclinical findings into clinics remains numerous hard nuts to crack, application of proteomics in cancer research is summarized in Table 2.

In the domain of precision oncology, we are just beginning to generate comprehensive, unbiased, and truly multi-omics data[160]. Only by producing these disparate histological datasets from the same biological samples can we develop the necessary analytical and annotation instruments to assist us in interpreting these complex datasets and leverage the extraction of biologically and clinically relevant information. Initiatives to capture a more comprehensive study cohort of more extensive histological data are spreading. In addition to being a cumbersome and expensive process, multi-omics data counted on the availability of appropriate tissue specimens and biopsy material to satisfy the efficient analysis of the tissue genome, transcriptome, epigenome, proteome, and metabolome. With the harvest of multi-omics tactics, we have abundant reasons to believe in the bright future of precision oncology.

Ethics approval and consent to participate

Not applicable.

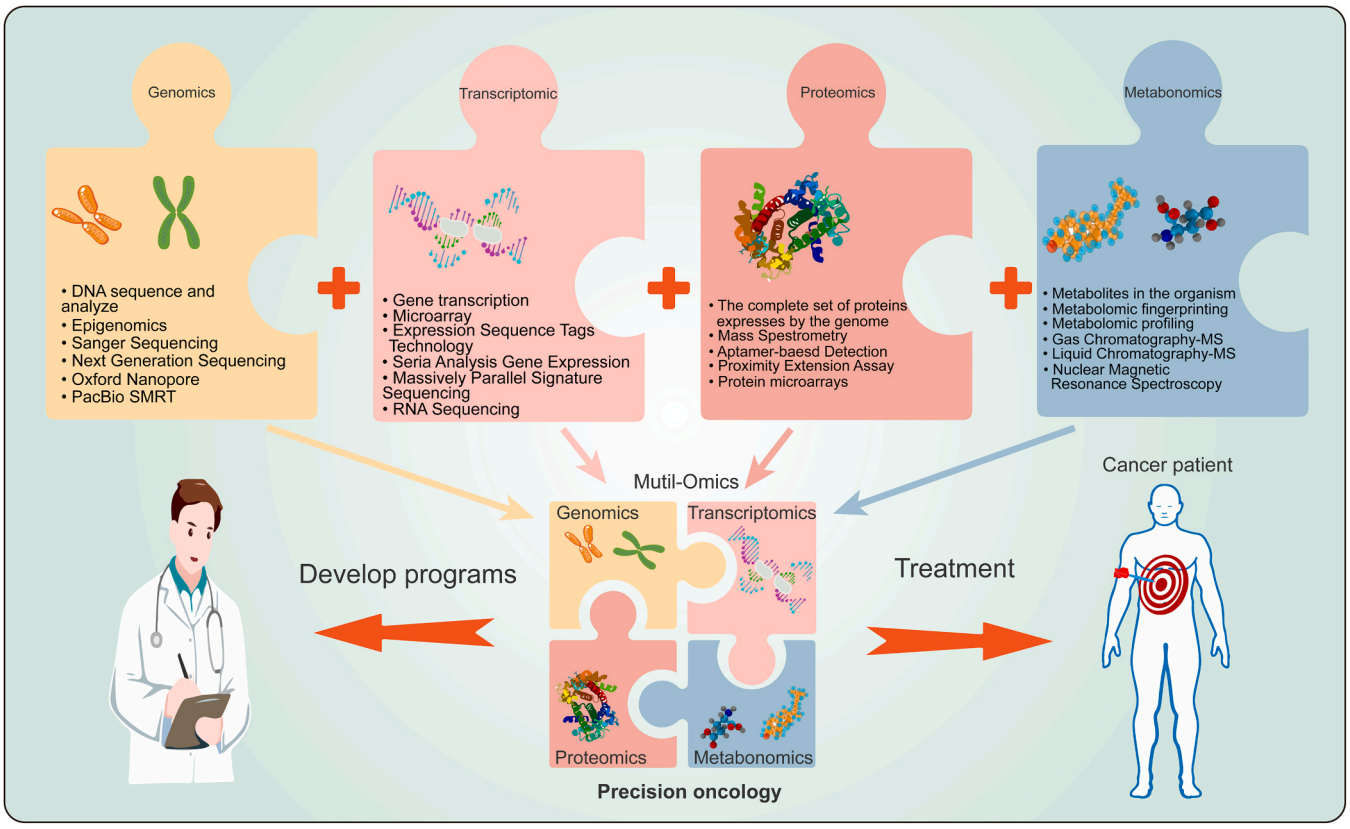


Fig. 3. Mutil-Omics for precision oncology. Genomics investigates gene structure and function, transcriptomics analyzes transcriptional processes, proteomics focuses on protein structure, function, and interactions, and metabolomics concentrates on metabolites and metabolic reactions. The clinicians obtain the different levels of histological information from the histology technology and integrate them like a "jigsaw puzzle" to acquire multi-omics data from which they can develop precise treatment plans for oncology patients.

Table 1
Cancer proteomics techniques application features.

Techniques	Basic principles	Readout	Data types	Advantages	Limitations
Mass spectrometry (MS)	Proteins are enzymatically digested to form peptide fragments, followed by peptide ionization, and MS detections.	Fragmented peptide sequences and spectrum counting	Mass charge ratio and abundance of ions	·Most commonly employed in proteomic studies; ·Non-hypothesis driven process for exploratory studies e.g. early phase of biomarker discovery.	·Low throughput; ·Relatively complex workflow ·limitations to analyze protein PTMs.
Aptamer-based detection	Short single-strand DNA or RNA folded into specific tertiary structures to be the aptamer and attached to a biotin bead of a specific protein, which is released and bound to another biotin bead after ultraviolet cleavage, eluted and recovered for quantification, and hybridized to a custom DNA microarray.	DNA microarray	Nucleotide sequence signaling	·High-plexity; ·Higher affinity and specificity than antibodies; ·Used for marking with therapeutic effect characteristic of elongated shelf life and sensitivity to antidotes.	·Limit on protein PTMs detection; ·DNA microarray readout dependence; ·Limited availability of aptamers requires high-quality development.
Proximity extension assay (PEA)	Each paired antibody with a unique DNA tag binds to the target protein of the sample followed by hybridization, generating a double strand that is digitally recognized, amplified, and detected by qPCR or NGS.	qPCR or NGS	Nucleotide sequence signaling	·Very little sample required with large dynamic ranges.	·Q-PCR/NGS readout dependence.
Forward-phase protein arrays	Antibodies are printed on solid phase substrates first capturing targeted proteins in samples and the signal intensity is measured by scanning devices.	Colorimetric assays/ fluorescence	Signal intensity	·Widely adopted approach with flexible experimental design and PTM profiling.	·Limit on inter-assay reproducibility, quantification, scalability, variation, sample labeling and costs of antibody arrays.
Reverse phase protein arrays (RPPA)	Samples are immobilized onto solid substrates first in the form of microarrays, and then proteins in the samples are detected by signal intensity with specific antibodies.	Colorimetric assays/ fluorescence	Signal intensity	·Large-scale parallel analysis for samples; ·Suitable for protein PTMs.	·Sophisticated experimental workflow.

Table 2
Application of proteomics in cancer research.

Cancer type	Methods	Application	Ref.
Brain cancer	LC-MS/MS Phosphopeptide Enrichment Consensus clustering	Refinement of cancer subtypes: (1) Cold-medullo; (2) Cold-mixed; (3) Neuronal; (4) Epithelial; (5) Hot.	[139]
Breast cancer	LC-MS/MS RPPA PCA	Refinement of cancer subtypes: CC1 is associated with luminal B, CC4 is associated with e triple-negative subtype, and CC2 and CC3 are associated with luminal A; CC2 is a novel luminal A breast cancer subtype with both triple-negative and ER+ signature that representing a group of more aggressive tumors.	[118]
	LC-MS/MS Unsupervised hierarchical clustering	Refinement of cancer subtypes: (1) Basal-like; (2) Luminal A; (3) Luminal B; (4) HER2; (5) Normal-like. Finding candidate immune target: lnc-AKAP14-1:3, lnc-CXorf36-3:1.	[121]
	LC-MS/MS NMF	Refinement of cancer subtypes: (1) NMF LumA-I; (2) NMF LumB-I; (3) NMF Basal-I; (4) NMF HER2-I.	[122]
Clear cell renal cell carcinoma	ESI-LC-MS/MS Hierarchical clustering PCA	Refinement of cancer subtypes: (1) ccRCC1: Upregulated adaptive immunity, N-glycosylation modification, OXPHOS and fatty acid metabolism-related proteins associated with high-grade and advanced tumors; (2) ccRCC2: Tumors overexpress proteins associated with natural immunity and platelet degranulation; (3) ccRCC3: Upregulated glycolysis, mTOR signaling, hypoxia-related proteins.	[124]
Colorectal cancer	RPPA	Detection biomarker for early detection: plasma C9	[56]
	PEA	Detection Potential prognostic biomarkers: operating characteristic (ROC) curve of Carcinoembryonic antigen (CEA), Transferrin Receptor-1 (TFRC), Macrophage migration inhibitory factor (MIF), Osteopontin (OPN/SPP1) and cancer antigen 242 (CA242).	[40]
	LC-MS/MS Phosphopeptide Enrichment Using IMAC	Rb phosphorylation as a driver and therapeutic target in colon cancer.	[161]
	LC-MS/MS Consensus clustering	Refinement of cancer subtypes: (1) CC1 is characterized by increased RNA processing and DNA mismatch repair (MMR); (2) CC2: the proteins upregulated in CC2 are enriched in extracellular matrix-receptor integration pathways and immune-related pathways. (3) CC3 is enriched in upregulated DNA replication and metabolic pathways. Differentiation of the primary tumors from metastatic CRC by six phosphoproteomic subtypes: metastatic CRC was represented by SC1, SC3, and SC5, whilst non-metastatic CRC was represented by SC2, SC4, and SC6.	[131]
Endometrial cancer	MS	Inferring a putative oncogenic factor: STRN4.	[147]
	LC-MS/MS	Insights into mechanisms of immune evasion.	[162]
Gastric cancer	LC-MS/MS Consensus clustering PCA	Classification of diffuse-type gastric cancer: (1) PX1 exhibits dysregulation in the cell cycle; (2) PX2 to PX1 features an EMT process; (3) PX3 is enriched in immune response proteins, has the worst survival, and is insensitive to chemotherapy.	[119]
	LC-MS/MS Consensus clustering Orthogonal NMF	Refinement of cancer subtypes: (1) represents cell proliferation processes; (2) represents immune response processes; (3) represents metabolism-related processes; (4) represents invasion-related processes.	[135]
Head-and-neck cancer	MS Hierarchical clustering	Inferring a proteomic signature for classifying squamous cell carcinomas as either lung or head-and-neck carcinomas.	[25]
Hepatocellular carcinoma	LC-MS/MS Consensus clustering PCA, NMF	Refinement of cancer subtypes: (1) metabolism subgroup (S-Mb); (2) microenvironment dysregulated subgroup (S-Me); (3) proliferation subgroup (S-Pf). Identification and validation of prognostic biomarkers: immunostaining of PYCR2 and ADH1A was significantly associated with patient survival.	[163]
Lungadenocarcinoma	LC-MS/MS Phosphopeptide Enrichment PCA, NMF	Phosphoproteomics identifies candidate ALK-fusion diagnostic markers and targets; Candidate drug targets: PTPN11(EGFR), SOS1 (KRAS), STK11.	[164]

(*continued on next page*)

Table 2 (*continued*)

Cancer type	Methods	Application	Ref.
	Nano-LC-MS/MS Consensus clustering PCA, NMF	Refinement of cancer subtypes: S-I; S-II; S-III. Detection Potential prognostic biomarkers: HSP 90 β.	[138]
Lung squamous cell carcinoma	Consensus clustering Spatial clustering RPPA PCA, NMF	Refinement of cancer subtypes: (1) basal-inclusive (B-I); (2) epithelial to mesenchymal transition-enriched (EMT-E); (3) classical; (4) inflamed-secretory (I-S); (5) proliferative-primitive (P-P). Detection of biomarkers: Rb phosphorylation is the biomarker of CDK4 /6 inhibitor test. Finding potential therapeutic targets: NSD3, survivin, LSD1, EZH2.	[123]
Small cell lung cancer	LC-MS/MS	Detection insulin-like growth factor-binding protein 5 as a biomarker for oncogenic achaete-scute homolog 1 high expression small cell lung cancer subtype.	[133]
Non-small cell lung cancer (NSCLC)	PEA Supervised cluster	Identifying protein signatures that could predict the response to anti-PD-(L)1 therapy.	[41]
Medulloblastoma	LC-MS/MS Consensus clustering, PCA	Differences in clinical features, histopathology, demographics, tumor imaging, and sensitivity to treatment among different subtypes of medulloblastoma were analyzed.	[120]
Melanoma	LC-MS/MS	The analysis of the high proteomic variability among metastatic sites contributes to an understanding of tissue-specific resistance mechanisms to prognosis and immunotherapy treatments.	[148]
Ovarian cancer	LC-MS/MS Phosphopeptide Enrichment Using IMAC	Identification of putative targets and biomarkers.	[136, 137]
Pancreatic cancer	LC-MS/MS NMF	Finding potential therapeutic targets: MAPK6, FLNA, BAD, MCM2, STAT3, CDK7, AKT1, PAK1, PAK2, and SRC.	[165]
Prostate cancer	PEA	Identifying an independent predictive marker of chemotactic relapse and migration progress.	[42]
	LC-MS/MS Consensus clustering	Refinement and extended characterization of cancer subtypes.	[125]

Note: LC-MS: liquid chromatography-mass spectrometry; NMF: non-negative matrix decomposition; PCA: principal component analysis; PEA: proximity extension assay; RPPA: reverse-phase protein array

Funding

Henan Provincial Science and Technology Research Project (221100310100).

CRediT authorship contribution statement

Yuyuan Zhang: Writing – review & editing. **Siyuan Weng:** Writing – review & editing. **Zaoqu Liu:** Funding acquisition, Writing – review & editing. **zhaokai Zhou:** Conceptualization, Data curation, Formal analysis, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. **Xinwei Han:** Supervision. **Ruiqi Zhang:** Visualization, Writing – original draft. **Aoyang Zhou:** Data curation, Writing – original draft. **Jinxiang Lv:** Writing – original draft. **Shuang Chen:** Writing – original draft, Writing – review & editing. **Haijiao Zou:** Writing – review & editing. **Ge Zhang:** Writing – review & editing. **Ting Lin:** Writing – review & editing. **Zhan Wang:** Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper: ZQL, XWH, and ZKZ provided direction and guidance throughout the preparation of this manuscript. ZKZ, RQZ, and AYZ wrote and edited the manuscript. ZKZ reviewed and made significant revisions to the manuscript. ZKZ drew all the figures. RQZ, AYZ, JXL, SC, GZ, HJZ, TL, ZW, YYZ, and SYW collected and prepared the related papers. All authors read and approved the final manuscript.

Acknowledgement

Not applicable.

Consent for publication

Not applicable.

Authors' contributions

ZQL, XWH, and ZKZ provided direction and guidance throughout the preparation of this manuscript. ZKZ, RQZ, and AYZ wrote and edited the manuscript. ZKZ reviewed and made significant revisions to the manuscript. ZKZ drew all the figures. RQZ, AYZ, JXL, SC, GZ, HJZ, TL, WZ, YYZ, and SYW collected and prepared the related papers. All authors read and approved the final manuscript.

References

[1] Shampo MA, Kyle RA, Craig Venter J. The human genome project. Mayo Clin Proc 2011;86(4):e26–7.
[2] Willyard C. New human gene tally reignites debate. Nature 2018;558(7710):354–5.
[3] Guttmacher AE, Collins FS. Welcome to the genomic era. N Engl J Med 2003;349(10):996–8.
[4] Wang K, Huang C, Nice E. Recent advances in proteomics: towards the human proteome. Biomed Chromatogr 2014;28(6):848–57.
[5] Zheng J, et al. Phenome-wide Mendelian randomization mapping the influence of the plasma proteome on complex diseases. Nat Genet 2020;52(10):1122–31.
[6] Digre A, Lindskog C. The Human Protein Atlas-Spatial localization of the human proteome in health and disease. Protein Sci 2021;30(1):218–33.
[7] Bell AW, et al. A HUPO test sample study reveals common problems in mass spectrometry-based proteomics. Nat Methods 2009;6(6):423–30.
[8] Nilsson T, et al. Mass spectrometry in high-throughput proteomics: ready for the big time. Nat Methods 2010;7(9):681–5.
[9] Zhou L, et al. Clinical proteomics-driven precision medicine for targeted cancer therapy: current overview and future perspectives. Expert Rev Proteom 2016;13(4):367–81.
[10] Guerin M, et al. How may targeted proteomics complement genomic data in breast cancer? Expert Rev Proteom 2017;14(1):43–54.

[11] Wang X. Statistical Assessment of QC Metrics on Raw LC-MS/MS Data. Methods Mol Biol 2017;1550:325–37.
[12] Sharma K, et al. Ultradeep human phosphoproteome reveals a distinct regulatory nature of Tyr and Ser/Thr-based signaling. Cell Rep 2014;8(5):1583–94.
[13] Edwards NJ, et al. The CPTAC data portal: a resource for cancer proteomics research. J Proteome Res 2015;14(6):2707–13.
[14] Aebersold R, Mann M. Mass-spectrometric exploration of proteome structure and function. Nature 2016;537(7620):347–55.
[15] Geyer PE, et al. Revisiting biomarker discovery by plasma proteomics. Mol Syst Biol 2017;13(9):942.
[16] Vidova V, Spacil Z. A review on mass spectrometry-based quantitative proteomics: Targeted and data independent acquisition. Anal Chim Acta 2017; 964:7–23.
[17] Matsumoto M, et al. A large-scale targeted proteomics assay resource based on an in vitro human proteome. Nat Methods 2017;14(3):251–8.
[18] Ludwig C, et al. Estimation of absolute protein quantities of unlabeled samples by selected reaction monitoring mass spectrometry. Mol Cell Proteom 2012;11(3). M111 013987.
[19] Olsen JV, et al. Higher-energy C-trap dissociation for peptide modification analysis. Nat Methods 2007;4(9):709–12.
[20] Peng L, et al. Tissue and plasma proteomics for early stage cancer detection. Mol Omics 2018;14(6):405–23.
[21] Corbett JR, Robinson DE, Patrie SM. Robustness and ruggedness of isoelectric focusing and superficially porous liquid chromatography with fourier transform mass spectrometry. J Am Soc Mass Spectrom 2021;32(1):346–54.
[22] Cupp-Sutton KA, Wu S. High-throughput quantitative top-down proteomics. Mol Omics 2020;16(2):91–9.
[23] Dai L, et al. Serological proteome analysis approach-based identification of ENO1 as a tumor-associated antigen and its autoantibody could enhance the sensitivity of CEA and CYFRA 21-1 in the detection of non-small cell lung cancer. Oncotarget 2017;8(22):36664–73.
[24] Jin Y, et al. Identification a novel clinical biomarker in early diagnosis of human non-small cell lung cancer. Glycoconj J 2019;36(1):57–68.
[25] Bohnenberger H, et al. Comparative proteomics reveals a diagnostic signature for pulmonary head-and-neck cancer metastasis. EMBO Mol Med 2018;10(9).
[26] Huang J, et al. Advances in aptamer-based biomarker discovery. Front Cell Dev Biol 2021;9:659760.
[27] Ellington AD, Szostak JW. In vitro selection of RNA molecules that bind specific ligands. Nature 1990;346(6287):818–22.
[28] Brody EN, et al. High-content affinity-based proteomics: unlocking protein biomarker discovery. Expert Rev Mol Diagn 2010;10(8):1013–22.
[29] Lollo B, Steele F, Gold L. Beyond antibodies: new affinity reagents to unlock the proteome. Proteomics 2014;14(6):638–44.
[30] Keefe AD, Pai S, Ellington A. Aptamers as therapeutics. Nat Rev Drug Discov 2010;9(7):537–50.
[31] Ma F, et al. Transcriptionally amplified synthesis of fluorogenic RNA aptamers for label-free DNA glycosylase assay. Chem Commun (Camb) 2022;58(73): 10229–32.
[32] Kruspe S, Meyer C, Hahn U. Chlorin e6 conjugated interleukin-6 receptor aptamers selectively kill target cells upon irradiation. Mol Ther Nucleic Acids 2014;3(1). e143.
[33] Rusconi CP, et al. Antidote-mediated control of an anticoagulant aptamer in vivo. Nat Biotechnol 2004;22(11):1423–8.
[34] Oney S, et al. Development of universal antidotes to control aptamer activity. Nat Med 2009;15(10):1224–8.
[35] Lupold SE, et al. Identification and characterization of nuclease-stabilized RNA molecules that bind human prostate cancer cells via the prostate-specific membrane antigen. Cancer Res 2002;62(14):4029–33.
[36] Wang DL, et al. Selection of DNA aptamers against epidermal growth factor receptor with high affinity and specificity. Biochem Biophys Res Commun 2014; 453(4):681–5.
[37] Assarsson E, et al. Homogenous 96-plex PEA immunoassay exhibiting high sensitivity, specificity, and excellent scalability. PLoS One 2014;9(4):e95192.
[38] Lundberg M, et al. Homogeneous antibody-based proximity extension assays provide sensitive and specific detection of low-abundant proteins in human blood. Nucleic Acids Res 2011;39(15):e102.
[39] Wik L, et al. Proximity Extension Assay in Combination with Next-Generation Sequencing for High-throughput Proteome-wide Analysis. Mol Cell Proteom 2021;20:100168.
[40] Thorsen SB, et al. Detection of serological biomarkers by proximity extension assay for detection of colorectal neoplasias in symptomatic individuals. J Transl Med 2013;11:253.
[41] Eltahir M, et al. Plasma Proteomic Analysis in Non-Small Cell Lung Cancer Patients Treated with PD-1/PD-L1 Blockade. Cancers (Basel) 2021;13(13).
[42] Liu S, et al. Discovery of PTN as a serum-based biomarker of pro-metastatic prostate cancer. Br J Cancer 2021;124(5):896–900.
[43] Liotta LA, et al. Protein microarrays: meeting analytical challenges for clinical applications. Cancer Cell 2003;3(4):317–25.
[44] Lueking A, Cahill DJ, Mullner S. Protein biochips: A new and versatile platform technology for molecular medicine. Drug Discov Today 2005;10(11):789–94.
[45] Haab BB. Antibody arrays in cancer research. Mol Cell Proteom 2005;4(4): 377–83.
[46] Reslova N, et al. xMAP Technology: Applications in Detection of Pathogens. Front Microbiol 2017;8:55.
[47] Shingyoji M, et al. Quantum dots-based reverse phase protein microarray. Talanta 2005;67(3):472–8.
[48] Tian S, et al. Facile synthesis of silver-rich Au/Ag bimetallic nanoparticles with highly active SERS properties. N J Chem 2019;43(37):14772–80.
[49] Yu X, Schneiderhan-Marra N, Joos TO. Protein microarrays for personalized medicine. Clin Chem 2010;56(3):376–87.
[50] Syu GD, Dunn J, Zhu H. Developments and Applications of Functional Protein Microarrays. Mol Cell Proteom 2020;19(6):916–27.
[51] Lin Y, et al. Identification of interleukin-8 as estrogen receptor-regulated factor involved in breast cancer invasion and angiogenesis by protein arrays. Int J Cancer 2004;109(4):507–15.
[52] Smith L, et al. The analysis of doxorubicin resistance in human breast cancer cells using antibody microarrays. Mol Cancer Ther 2006;5(8):2115–20.
[53] Ingvarsson J, et al. Detection of pancreatic cancer using antibody microarray-based serum protein profiling. Proteomics 2008;8(11):2211–9.
[54] Paweletz CP, et al. Reverse phase protein microarrays which capture disease progression show activation of pro-survival pathways at the cancer invasion front. Oncogene 2001;20(16):1981–9.
[55] Mueller C, Liotta LA, Espina V. Reverse phase protein microarrays advance to use in clinical trials. Mol Oncol 2010;4(6):461–81.
[56] Murakoshi Y, et al. Plasma biomarker discovery and validation for colorectal cancer by quantitative shotgun mass spectrometry and protein microarray. Cancer Sci 2011;102(3):630–8.
[57] Chen MM, et al. Predicting Cancer Cell Line Dependencies From the Protein Expression Data of Reverse-Phase Protein Arrays. JCO Clin Cancer Inf 2020;(4): 357–66.
[58] Tibes R, et al. Reverse phase protein array: validation of a novel proteomic technology and utility for analysis of primary leukemia specimens and hematopoietic stem cells. Mol Cancer Ther 2006;5(10):2512–21.
[59] Valencia K, Montuenga LM. Exosomes in Liquid Biopsy: The Nanometric World in the Pursuit of Precision Oncology. Cancers (Basel) 2021;13(9).
[60] Vinik Y, et al. Proteomic analysis of circulating extracellular vesicles identifies potential markers of breast cancer progression, recurrence, and response. Sci Adv 2020;6(40).
[61] Signore M, et al. Diagnostic and prognostic potential of the proteomic profiling of serum-derived extracellular vesicles in prostate cancer. Cell Death Dis 2021;12 (7):636.
[62] Gallagher RI, et al. Reverse phase protein microarrays: fluorometric and colorimetric detection. Methods Mol Biol 2011;723:275–301.
[63] Yeon S, et al. Dual-Color, multiplex analysis of protein microarrays for precision medicine. Methods Mol Biol 2017;1550:149–70.
[64] Wagner J, et al. A single-cell atlas of the tumor and immune ecosystem of human breast cancer. Cell 2019;177(5):1330–1345.e18.
[65] Liu D, et al. Single-Cell Multiplexed Proteomics on the IsoLight Resolves Cellular Functional Heterogeneity to Reveal Clinical Responses of Cancer Patients to Immunotherapies. Methods Mol Biol 2020;2055:413–31.
[66] Parisi G, et al. Persistence of adoptively transferred T cells with a kinetically engineered IL-2 receptor agonist. Nat Commun 2020;11(1):660.
[67] Lundberg E, Borner GHH. Spatial proteomics: a powerful discovery tool for cell biology. Nat Rev Mol Cell Biol 2019;20(5):285–302.
[68] Mund A, et al. Deep Visual Proteomics defines single-cell identity and heterogeneity. Nat Biotechnol 2022;40(8):1231–40.
[69] Kimura Y, et al. Identification of candidate diagnostic serum biomarkers for Kawasaki disease using proteomic analysis. Sci Rep 2017;7:43732.
[70] Kuo KK, et al. Quantitative proteomic analysis of differentially expressed protein profiles involved in pancreatic ductal adenocarcinoma. Pancreas 2016;45(1): 71–83.
[71] Siebert S, et al. Urinary proteomics can define distinct diagnostic inflammatory arthritis subgroups. Sci Rep 2017;7:40473.
[72] Zhang F, et al. A neural network approach to multi-biomarker panel discovery by high-throughput plasma proteomics profiling of breast cancer. BMC Proc 2013;7 (Suppl 7):S10.
[73] Chen YD, et al. Artificial neural networks analysis of surface-enhanced laser desorption/ionization mass spectra of serum protein pattern distinguishes colorectal cancer from healthy population. Clin Cancer Res 2004;10(24):8380–5.
[74] Ostroff RM, et al. Early detection of malignant pleural mesothelioma in asbestos-exposed individuals with a noninvasive proteomics-based surveillance tool. PLoS One 2012;7(10):e46091.
[75] Petricoin EF, et al. Use of proteomic patterns in serum to identify ovarian cancer. Lancet 2002;359(9306):572–7.
[76] Ahn HS, et al. Serum biomarker panels for the diagnosis of gastric adenocarcinoma. Br J Cancer 2012;106(4):733–9.
[77] Rogers MA, et al. Proteomic profiling of urinary proteins in renal cancer by surface enhanced laser desorption ionization and neural-network analysis: identification of key issues affecting potential clinical utility. Cancer Res 2003;63 (20):6971–83.
[78] Akram-Ali-Hammouri Z, et al. Fast Support Vector Classification for Large-Scale Problems. IEEE Trans Pattern Anal Mach Intell 2022;44(10):6184–95.
[79] Pozniak Y, et al. System-wide Clinical Proteomics of Breast Cancer Reveals Global Remodeling of Tissue Homeostasis. Cell Syst 2016;2(3):172–84.
[80] Ashburner M, et al. Gene Ontology: tool for the unification of biology. Nat Genet 2000;25(1):25–9.
[81] Subramanian A, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA 2005; 102(43):15545–50.
[82] Locard-Paulet M, et al. Phosphoproteomic analysis of interacting tumor and endothelial cells identifies regulatory mechanisms of transendothelial migration. Sci Signal 2016;9(414). p. ra15.

[83] Creixell P, et al. Pathway and network analysis of cancer genomes. Nat Methods 2015;12(7):615–21.
[84] Tarca AL, et al. A novel signaling pathway impact analysis. Bioinformatics 2009;25(1):75–82.
[85] Paull EO, et al. Discovering causal pathways linking genomic events to transcriptional states using Tied Diffusion Through Interacting Events (TieDIE). Bioinformatics 2013;29(21):2757–64.
[86] Vanunu O, et al. Associating genes and protein complexes with disease via network propagation. PLoS Comput Biol 2010;6(1):e1000641.
[87] Hofree M, et al. Network-based stratification of tumor mutations. Nat Methods 2013;10(11):1108–15.
[88] Kim YA, Wuchty S, Przytycka TM. Identifying causal genes and dysregulated pathways in complex diseases. PLoS Comput Biol 2011;7(3):e1001095.
[89] Navlakha S, Kingsford C. The power of protein interaction networks for associating genes with diseases. Bioinformatics 2010;26(8):1057–63.
[90] Eduati F, et al. Drug resistance mechanisms in colorectal cancer dissected with cell type-specific dynamic logic models. Cancer Res 2017;77(12):3364–75.
[91] Hakes L, et al. Protein-protein interaction networks and biology–what's the connection? Nat Biotechnol 2008;26(1):69–72.
[92] Muller T, et al. Sense and nonsense of pathway analysis software in proteomics. J Proteome Res 2011;10(12):5398–408.
[93] Soh D, et al. Consistency, comprehensiveness, and compatibility of pathway databases. BMC Bioinforma 2010;11:449.
[94] Reinders J, et al. Challenges in mass spectrometry-based proteomics. Proteomics 2004;4(12):3686–703.
[95] Casado P, et al. Kinase-substrate enrichment analysis provides insights into the heterogeneity of signaling pathway activation in leukemia cells. Sci Signal 2013;6(268). p. rs6.
[96] Drake JM, et al. Phosphoproteome Integration Reveals Patient-Specific Networks in Prostate Cancer. Cell 2016;166(4):1041–54.
[97] Vaske CJ, et al. Inference of patient-specific pathway activities from multi-dimensional cancer genomics data using PARADIGM. Bioinformatics 2010;26(12):i237–45.
[98] Žitnik M, Zupan B. Data Fusion by Matrix Factorization. IEEE Trans Pattern Anal Mach Intell 2015;37(1):41–53.
[99] Metsalu T, Vilo J. ClustVis: a web tool for visualizing clustering of multivariate data using Principal Component Analysis and heatmap. Nucleic Acids Res 2015;43(W1):W566–70.
[100] Chen M, Li X. Robust Matrix Factorization With Spectral Embedding. IEEE Trans Neural Netw Learn Syst 2021;32(12):5698–707.
[101] Shen R, Olshen AB, Ladanyi M. Integrative clustering of multiple genomic data types using a joint latent variable model with application to breast and lung cancer subtype analysis. Bioinformatics 2009;25(22):2906–12.
[102] Liu Y, et al. Multilevel omic data integration in cancer cell lines: advanced annotation and emergent properties. BMC Syst Biol 2013;7:14.
[103] Persson O, et al. Microarray analysis of gliomas reveals chromosomal position-associated gene expression patterns and identifies potential immunotherapy targets. J Neurooncol 2007;85(1):11–24.
[104] Furey TS, et al. Support vector machine classification and validation of cancer tissue samples using microarray expression data. Bioinformatics 2000;16(10):906–14.
[105] Theilhaber J, et al. Finding genes in the C2C12 osteogenic pathway by k-nearest-neighbor classification of expression data. Genome Res 2002;12(1):165–76.
[106] Yuan Y, et al. Assessing the clinical utility of cancer genomic and proteomic data across tumor types. Nat Biotechnol 2014;32(7):644–52.
[107] Meng C, et al. A multivariate approach to the integration of multi-omics datasets. BMC Bioinforma 2014;15:162.
[108] Liu G, et al. ProHits: integrated software for mass spectrometry-based interaction proteomics. Nat Biotechnol 2010;28(10):1015–7.
[109] Pfeuffer J, et al. OpenMS - A platform for reproducible analysis of mass spectrometry data. J Biotechnol 2017;261:142–8.
[110] Rosenberger G, et al. A repository of assays to quantify 10,000 human proteins by SWATH-MS. Sci Data 2014;1:140031.
[111] Rost HL, et al. TRIC: an automated alignment strategy for reproducible protein quantification in targeted proteomics. Nat Methods 2016;13(9):777–83.
[112] Martinez-Jimenez CP, et al. Aging increases cell-to-cell transcriptional variability upon immune stimulation. Science 2017;355(6332):1433–6.
[113] Gebreyesus ST, et al. Streamlined single-cell proteomics by an integrated microfluidic chip and data-independent acquisition mass spectrometry. Nat Commun 2022;13(1):37.
[114] Bendall SC, et al. Single-cell trajectory detection uncovers progression and regulatory coordination in human B cell development. Cell 2014;157(3):714–25.
[115] Garderet L, et al. Pomalidomide, cyclophosphamide, and dexamethasone for relapsed multiple myeloma. Blood 2018;132(24):2555–63.
[116] Kümler I, Christiansen OG, Nielsen DL. A systematic review of bevacizumab efficacy in breast cancer. Cancer Treat Rev 2014;40(8):960–73.
[117] Zhang H, et al. Integrated Proteogenomic Characterization of Human High-Grade Serous Ovarian Cancer. Cell 2016;166(3):755–65.
[118] Yanovich G, et al. Clinical Proteomics of Breast Cancer Reveals a Novel Layer of Breast Cancer Classification. Cancer Res 2018;78(20):6001–10.
[119] Ge S, et al. A proteomic landscape of diffuse-type gastric cancer. Nat Commun 2018;9(1):1012.
[120] Archer TC, et al. Proteomics, post-translational modifications, and integrative analyses reveal molecular heterogeneity within medulloblastoma subgroups. Cancer Cell 2018;34(3):396–410. e8.
[121] Johansson HJ, et al. Breast cancer quantitative proteome and proteogenomic landscape. Nat Commun 2019;10(1):1600.
[122] Krug K, et al. Proteogenomic landscape of breast cancer tumorigenesis and targeted therapy. Cell 2020;183(5):1436–56. e31.
[123] Satpathy S, et al. A proteogenomic portrait of lung squamous cell carcinoma. Cell 2021;184(16):4348–71. e40.
[124] Clark DJ, et al. Integrated proteogenomic characterization of clear cell renal cell carcinoma. Cell 2019;179(4):964–83. e31.
[125] Sinha A, et al. The proteogenomic landscape of curable prostate cancer. Cancer Cell 2019;35(3):414–27. e6.
[126] Wang Y, Gao X, Wang J. Functional proteomic profiling analysis in four major types of gastrointestinal cancers. Biomolecules 2023;13(4).
[127] Goncalves E, et al. Pan-cancer proteomic map of 949 human cell lines. Cancer Cell 2022;40(8):835–49. e8.
[128] Schwamborn K, et al. Identifying prostate carcinoma by MALDI-Imaging. Int J Mol Med 2007;20(2):155–9.
[129] Kriegsmann M, et al. MALDI MS imaging as a powerful tool for investigating synovial tissue. Scand J Rheuma 2012;41(4):305–9.
[130] Meding S, et al. Tumor classification of six common cancer types based on proteomic profiling by MALDI imaging. J Proteome Res 2012;11(3):1996–2003.
[131] Li C, et al. Integrated Omics of Metastatic Colorectal Cancer. Cancer Cell 2020;38(5):734–47. e9.
[132] Arbelaiz A, et al. Serum extracellular vesicles contain protein biomarkers for primary sclerosing cholangitis and cholangiocarcinoma. Hepatology 2017;66(4):1125–43.
[133] Wang XD, et al. Subtype-specific secretomic characterization of pulmonary neuroendocrine tumor cells. Nat Commun 2019;10(1):3201.
[134] Dagnino S, et al. Prospective Identification of Elevated Circulating CDCP1 in Patients Years before Onset of Lung Cancer. Cancer Res 2021;81(13):3738–48.
[135] Mun DG, et al. Proteogenomic Characterization of Human Early-Onset Gastric Cancer. Cancer Cell 2019;35(1):111–24. e10.
[136] Hu Y, et al. Integrated Proteomic and Glycoproteomic Characterization of Human High-Grade Serous Ovarian Carcinoma. Cell Rep 2020;33(3):108276.
[137] McDermott JE, et al. Proteogenomic Characterization of Ovarian HGSC implicates mitotic kinases, replication stress in observed chromosomal instability. Cell Rep Med 2020;1(1).
[138] Xu JY, et al. Integrative proteomic characterization of human lung adenocarcinoma. Cell 2020;182(1):245–61. e17.
[139] Petralia F, et al. Integrated proteogenomic characterization across major histological types of pediatric brain cancer. Cell 2020;183(7):1962–85. e31.
[140] Roumeliotis TI, et al. Genomic determinants of protein abundance variation in colorectal cancer cells. Cell Rep 2017;20(9):2201–14.
[141] Sousa A, et al. Multi-omics Characterization of Interaction-mediated Control of Human Protein Abundance levels. Mol Cell Proteom 2019;18(8 suppl 1). S114-s125.
[142] Gumeni S, et al. Proteome stability as a key factor of genome integrity. Int J Mol Sci 2017;18(10).
[143] Al-Matouq J, et al. Post-transcriptional screen of cancer amplified genes identifies ERBB2/Her2 signaling as AU-rich mRNA stability-promoting pathway. Oncogenesis 2021;10(9):61.
[144] Zheng Y, Pan D. The hippo signaling pathway in development and disease. Dev Cell 2019;50(3):264–82.
[145] Xiang J, Jiang M, Du X. The role of Hippo pathway in ferroptosis. Front Oncol 2022;12:1107505.
[146] Maity S, Gridnev A, Misra JR. Assays Used for Discovering Small Molecule Inhibitors of YAP Activity in Cancers. Cancers (Basel) 2022;14(4).
[147] Seo G, et al. MAP4K interactome reveals STRN4 as a Key STRIPAK complex component in hippo pathway regulation. Cell Rep 2020;32(1):107860.
[148] Beck L, et al. Clinical proteomics of metastatic melanoma reveals profiles of organ specificity and treatment resistance. Clin Cancer Res 2021;27(7):2074–86.
[149] Lokhov PG, Balashova EE. Antigenic essence: upgrade of cellular cancer vaccines. Cancers (Basel) 2021;13(4).
[150] Chong C, Coukos G, Bassani-Sternberg M. Identification of tumor antigens with immunopeptidomics. Nat Biotechnol 2022;40(2):175–88.
[151] Bassani-Sternberg M, et al. Direct identification of clinically relevant neoepitopes presented on native human melanoma tissue by mass spectrometry. Nat Commun 2016;7:13404.
[152] Harris RC, Chung E, Coffey RJ. EGF receptor ligands. Exp Cell Res 2003;284(1):2–13.
[153] Holbro T, Hynes NE. ErbB receptors: directing key signaling networks throughout life. Annu Rev Pharm Toxicol 2004;44:195–217.
[154] Yang Y, et al. Involvement and targeted intervention of mortalin-regulated proteome phosphorylated-modification in hepatocellular Carcinoma. Front Oncol 2021;11:687871.
[155] Li Y, et al. Phosphoproteomics reveals therapeutic targets of esophageal squamous cell carcinoma. Signal Transduct Target Ther 2021;6(1):381.
[156] Li B, et al. Proteomics and phosphoproteomics profiling of drug-Addicted BRAFi-resistant melanoma cells. J Proteome Res 2021;20(9):4381–92.
[157] van Mierlo G, Vermeulen M. Chromatin proteomics to study epigenetics - challenges and opportunities. Mol Cell Proteom 2021;20:100056.
[158] Kiehlmeier S, et al. Identification of therapeutic targets of the hijacked super-enhancer complex in EVI1-rearranged leukemia. Leukemia 2021;35(11):3127–38.
[159] Mertins P, et al. Proteogenomics connects somatic mutations to signalling in breast cancer. Nature 2016;534(7605):55–62.
[160] Zhou Z, et al. Omics-based molecular classifications empowering in precision oncology. Cell Oncol (Dordr) 2024.

[161] Vasaikar S, et al. Proteogenomic analysis of human colon cancer reveals new therapeutic opportunities. Cell 2019;177(4):1035–49. e19.
[162] Dou Y, et al. Proteogenomic characterization of endometrial carcinoma. Cell 2020;180(4):729–48. e26.
[163] Gao Q, et al. Integrated proteogenomic characterization of hbv-related hepatocellular carcinoma. Cell 2019;179(2):561–77. e22.
[164] Gillette MA, et al. Proteogenomic characterization reveals therapeutic vulnerabilities in lung adenocarcinoma. Cell 2020;182(1):200–25. e35.
[165] Cao L, et al. Proteogenomic characterization of pancreatic ductal adenocarcinoma. Cell 2021;184(19):5031–52. e26.