

Advancements in proteogenomics for preclinical targeted cancer therapy research

Yuying Suo^{1,2}, Yuanli Song¹, Yuqiu Wang^{1,4}, Qian Liu¹, Henry Rodriguez³✉, Hu Zhou¹✉

¹ Department of Analytical Chemistry, State Key Laboratory of Drug Research, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai 201203, China

² University of Chinese Academy of Sciences, Beijing 100049, China

³ Office of Cancer Clinical Proteomics Research, National Cancer Institute, National Institutes of Health, Rockville, MD 20850, USA

⁴ Department of Otolaryngology, Eye & ENT Hospital, Fudan University, Shanghai 200031, China

Received: 23 October 2024 / Accepted: 3 December 2024

Abstract Advancements in molecular characterization technologies have accelerated targeted cancer therapy research at unprecedented resolution and dimensionality. Integrating comprehensive multi-omic molecular profiling of a tumor, proteogenomics, marks a transformative milestone for preclinical cancer research. In this paper, we initially provided an overview of proteogenomics in cancer research, spanning genomics, transcriptomics, and proteomics. Subsequently, the applications were introduced and examined from different perspectives, including but not limited to genetic alterations, molecular quantifications, single-cell patterns, different post-translational modification levels, subtype signatures, and immune landscape. We also paid attention to the combined multi-omics data analysis and pan-cancer analysis. This paper highlights the crucial role of proteogenomics in preclinical targeted cancer therapy research, including but not limited to elucidating the mechanisms of tumorigenesis, discovering effective therapeutic targets and promising biomarkers, and developing subtype-specific therapies.

Keywords Proteogenomics, Targeted cancer therapy, Integrated omics analysis

INTRODUCTION

Cancer remains a significant health problem worldwide. It is imperative to study tumorigenesis and improve cancer therapy. Traditional drug research is costly, tedious, time-consuming, and fraught with uncertainty. The next-generation sequencing (NGS) has revolutionized cancer therapy research (Wishart 2022). Since then, many cancer therapy strategies have been “rationally designed”, meaning that the target is first identified and then strategies developed. Successful examples include using BCR-ABL- and c-kit-targeting agents on chronic myeloid leukemia (Gonzalez-Fierro

and Dueñas-González 2021). However, these strategies provide only limited treatment outcomes in clinical results. Gene mutations alone do not consistently predict prognosis or the response to therapies (Gerlinger *et al.* 2012; Simon and Roychowdhury 2013).

Proteogenomics is an interdisciplinary field that integrates proteomics and genomics, enabling researchers to incorporate genetic information into the interpretation of protein-level data. Proteogenomics was first introduced in 2004 and was initially applied to improve the annotation of genomes. Then the term has been comprehensively mapped to different aspects of biological activity, primarily focusing on genomes, transcriptomes, and proteomes (Jaffe *et al.* 2004; Tsang and Wong 2021). Early proteogenomic development was limited by proteomic technology. However, over the

✉ Correspondence: rodriguez@nih.gov (H. Rodriguez), zhouhu@simmm.ac.cn (H. Zhou)

past few decades, advancements in proteomics, particularly liquid chromatography coupled with tandem mass spectrometry (LC-MS/MS), have been rapidly developed. One of the key advantages of proteogenomics is that proteogenomics provides a comprehensive and multi-dimensional approach that helps unravel the complexities of cancer by bridging the gap between genotype and phenotype, offering a holistic understanding of the cancer (Menyhart and Gyorffy 2021; Neagu *et al.* 2023; Zhang *et al.* 2020a).

A typical proteogenomic study involves multiple stages, ranging from sample collection to data analysis and interpretation. The main processes include: (1) Sample collection and preparation, which involves extracting nucleic acids (DNA and RNA) and proteins from biological samples. (2) Genomics and transcriptomics data generation, through DNA and RNA sequencing. (3) Proteomics data generation, primarily using LC-MS/MS analysis. (4) Data processing: raw sequencing and MS data are preprocessed (*e.g.*, quality control, normalization, filtering) to reduce noise and enhance signal clarity. The genomic and transcriptomic data undergo variant calling and functional annotations. (5) Statistical analysis and discovery. The combined data are interpreted to identify key biological mechanisms, such as mutated genes that lead to aberrant protein expression or altered cellular pathways that drive cancer progression. (6) Validation and clinical application. Findings from the proteogenomic analysis (such as new biomarkers or drug targets) are validated through *in vitro* or *in vivo* experiments. The results are strongly recommended for inclusion in clinical assays or therapies, paving the way for precision medicine applications.

In this review, we first briefly introduce proteogenomics, including genomics, transcriptomics, proteomics, and post-translational modifications (PTM) omics, and then explore its various applications for pre-clinical cancer research. These data resources offer novel insights into tumors from various perspectives, further advancing personalized targeted cancer therapies (Fig. 1). We further introduce the multi-omics data analysis and pan-cancer analysis, which are popular strategies nowadays, emphasizing the role of proteogenomics in preclinical targeted cancer therapy.

OVERVIEW OF PROTEOGENOMICS

Advancements in technology have expanded the scope of proteogenomics to encompass various areas (Wang *et al.* 2023). Still, the typical proteogenomic data

include genomic, transcriptomic, and MS-based proteomics data.

Genomics and transcriptomics

Genomics focuses on studying entire genomes. Transcriptomics examines RNA levels genome-wide, including qualitatively (identification of present transcripts, splice sites, and RNA editing sites) and quantitatively (detecting the expression of each transcript) (Hasin *et al.* 2017). Large-scale genome sequencing to characterize various types of cancers began in earnest as early as 2005. The most typical project is The Cancer Genome Atlas (TCGA) (Tomczak *et al.* 2015). Since 2013, The Cancer Genome Project has compiled approximately five million somatic alterations for over 7000 primary cancers (Alexandrov *et al.* 2013).

It has become increasingly feasible and cost-effective to perform genomics and transcriptomics studies (Bustamante *et al.* 2011). NGS enables high-throughput sequencing of DNA and RNA fragments (Kamps *et al.* 2017). The commonly used methods in proteogenomic studies for DNA sequencing are whole genome sequencing (WGS) and whole exome sequencing (WES). WGS can sequence the total DNA content and WES means sequencing of all protein-coding regions (*i.e.*, the exome). The context and complexity of alterations such as point mutations, insertion-deletion mutations, and copy number variations (CNVs) can be identified (Ang *et al.* 2019; Petersen *et al.* 2017). The widely used method for transcriptomics is RNA sequencing (RNA-seq). It measures the total RNA from cells within bulk tissues and organs, providing an average expression level with mature technology and high throughput.

Conventional cancer omics research can only produce the average profile of cancer, regardless of heterogeneous cell types. Nowadays, single-cell omics have become popular for characterizing cells across different molecular levels. It enables the elucidation of cellular mechanisms and the deciphering of cell heterogeneity and diversity (Sun *et al.* 2024a). There are numerous single-cell omics technologies, among which single-cell transcriptomics has developed relatively quickly and paired with other omics (Baysoy *et al.* 2023; Bian *et al.* 2018; Wang *et al.* 2022). The popular tactics include genome-wide single-cell RNA sequencing (scRNA-seq) and single-nucleus RNA-Seq (snRNA-Seq) (Luecken and Theis 2019; Slyper *et al.* 2020; Sun *et al.* 2021). By uncovering cellular heterogeneity and building cell atlas resources, our understanding of the mechanisms of oncology can be improved.

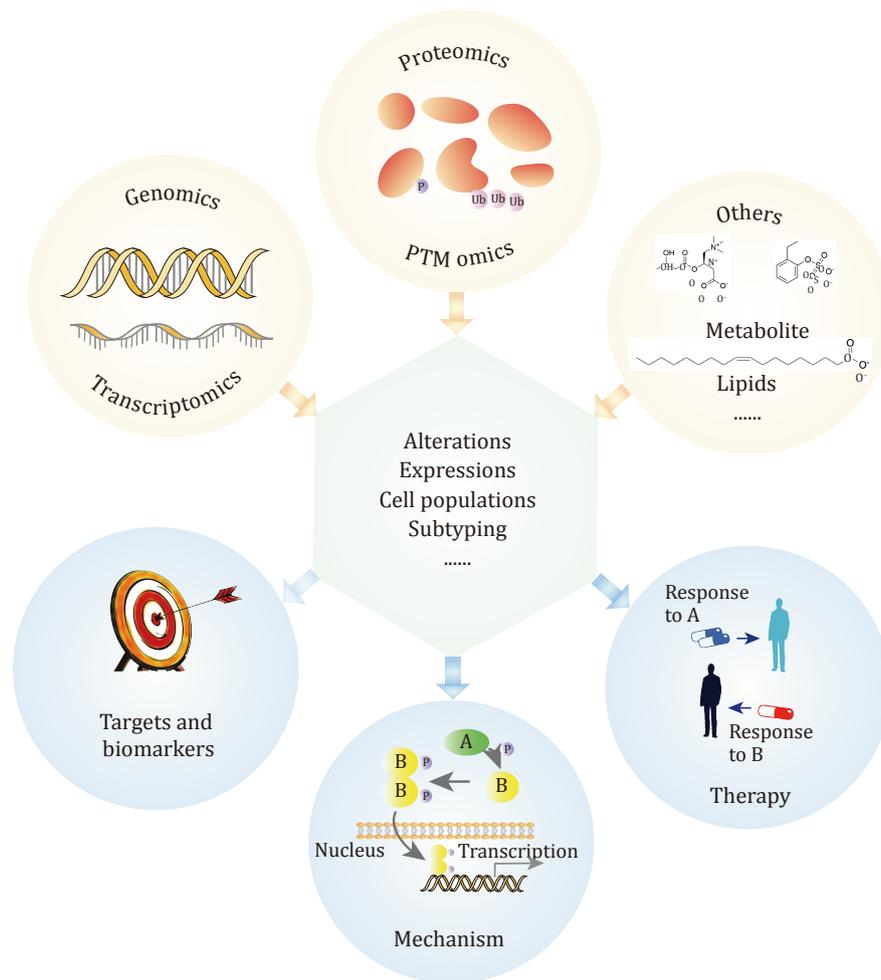


Fig. 1 Proteogenomics and their applications in target cancer therapy research. The application of proteogenomics (mainly including genomics, transcriptomics, and proteomics) has become an indispensable tool in the diagnosis and treatment of various types of cancer. Through the assessment of alterations, molecular expressions, cell populations, sample subtyping, and other aspects, proteogenomics can provide invaluable insights into preclinical target cancer therapy research. This includes unraveling the mechanism of tumorigenesis, identifying potential targets and specific biomarkers, and predicting subtype-specific therapy

Proteomics and post-translational modifications omics

Proteins serve as the fundamental building blocks of life, driving essential biological processes. The vast majority of signal transduction pathways are governed by dynamic protein PTMs, which include phosphorylation, acetylation, glycosylation, *etc.* These processes contribute to regulating cellular biological signaling, providing a wealth of potential targets for cancer diagnosis and therapeutic strategies. Genomics and transcriptomics are unable to capture phenomena such as protein accumulation or degradation and the intricacies of PTM processes (Doll *et al.* 2019).

Proteomics has developed methods to directly measure and annotate protein abundance, interaction,

localization, and modification states, offering new insights into tumor pathophysiology (Meissner *et al.* 2022; Suo *et al.* 2024; Xiao *et al.* 2020). There are several powerful tools for accurately assessing protein, including protein arrays (Akbari *et al.* 2014; Coarfa *et al.* 2021) and LC-MS/MS (Meissner *et al.* 2022). The latter method stands out as the most common method. Proteins are digested into peptides and sent to an LC instrument, which allows for the separation of complex peptide mixtures. The peptides are then ionized by the MS instrument, and the ions are manipulated using electromagnetic fields to measure their mass-to-charge (m/z) ratio. Quantitative proteomics focuses on measuring the abundance of proteins within a biological sample, encompassing both label-based and label-free methodologies (Meissner *et al.* 2022). The tandem mass tags (TMT) method and label-free quantification based

on data-independent acquisition are of special concern (Ludwig *et al.* 2018; Zhang *et al.* 2020b).

PTMs refer to the chemical changes occurring at protein terminuses or amino acid side chains (Pan and Chen 2022). To identify these modified peptides, matrices with appropriate physical properties (such as charge and hydrophobicity) or antibodies are widely used for enrichment before LC-MS/MS detection (Silva *et al.* 2013). Take phosphoproteomics and glycoproteomics as examples. Phosphorylation is the most common type of PTM and is frequently studied (Silva *et al.* 2013). The classic enrichment strategies include immobilized metal affinity chromatography (IMAC) and metal oxide affinity chromatography (MOAC) (Xie *et al.* 2022). Glycobiology has garnered increasing attention for its significant role in tumorigenesis (Pan and Chen 2022). Yet glycosylation is a complex process and the intricate structure of glycoproteins constitutes a major technical difficulty. Some technologies, such as enzymes and probes, have facilitated the detection of glycans and glycosylation enzymes (Luo *et al.* 2023; Narimatsu *et al.* 2018; Tian *et al.* 2022).

The single-cell proteomic methods are still limited (Dwivedi *et al.* 2019). Before the availability of single-cell proteomics by mass spectrometry (SCP), most single-cell proteomics methods were antibody-based. For example, CyTOF (Cytometry by time-of-flight) utilizes rare metal isotopes for antibody labeling and can quantify labeled targets of single cells (Bandura *et al.* 2009; Li *et al.* 2018). SCP aims to identify and quantify the proteome of single cells using LC-MS/MS without antibodies. Since 2018, there has been an increasing boom of customized sample preparation platforms using SCP (Brunner *et al.* 2021; Tan *et al.* 2024). Nowadays, SCP requires further optimization in aspects such as instrumentation, sample preparation, and MS methods.

ANALYSIS STRATEGIES AND PROMISING APPLICATIONS OF PROTEOGENOMICS IN PRECLINICAL TARGET CANCER THERAPY RESEARCH

To date, most cancer proteogenomic studies have been exploratory, hypothesis-generating, so-called 'landscape' studies, aiming at producing the omics dataset (*i.e.* resource) of representative samples. Leveraging these datasets, proteogenomic applications primarily delve into fascinating aspects such as genetic alterations, differential expression, single-cell information, patterns of PTMs, subtype characteristics, and immune landscape (Fig. 2). Although the following descriptions are separated, they are often interactive during the analysis.

Depiction of cancer proteogenomic landscape as valuable resources for subsequent exploration

The Clinical Proteomic Tumor Analysis Consortium (CPTAC) (Ellis *et al.* 2013) was established in 2011, which has developed a series of collaborating groups, to achieve proteomic analysis of different cancers and store large amounts of proteogenomic data. In order to remove the multiple sources of variability, CPTAC provides a standard and uniform proteomics analysis platform for all of the collaborators, including experimental design, MS methods, and database searching, making it possible to compare the data between different samples, cancer types, and institutions (Rudnick *et al.* 2016).

In proteogenomic studies, the selection of clinical samples is the crucial step. Tumors and their corresponding normal adjacent tissues (NATs) are the preferred samples, which directly represent the characteristics of the tumors. Biofluids are also highly sought after because they contain biomarkers that can help diagnose and monitor various medical conditions. WGS, WES, RNA-seq, miRNA-seq, scRNA-seq, proteomics, and PTM omics analysis are widely performed. These techniques provide insights into genetic and transcript alterations, RNA and protein quantification, levels of PTMs, and other molecular characteristics (Cao *et al.* 2021; Clark *et al.* 2019; Vasaikar *et al.* 2019; Wang *et al.* 2021a). Clinical information is also very important, especially survival information. By examining survival differences across different groups, researchers can determine whether the factor is key to clinical relevance (Gao *et al.* 2019; Liu *et al.* 2024). Patient characteristics such as gender, race, age, and tumor grade, are also documented.

Researchers have conducted comprehensive profiling across various cancer types over the past decade (Cao *et al.* 2021; Clark *et al.* 2019; Gao *et al.* 2019; Huang *et al.* 2021; Jayavelu *et al.* 2022; Kramer *et al.* 2022; Krug *et al.* 2020; Lehtio *et al.* 2021; Li *et al.* 2023b; Liu *et al.* 2024; Vasaikar *et al.* 2019; Wang *et al.* 2021a; Xiao *et al.* 2022). As data production increases, there are numerous public repositories besides TCGA and CPTAC (Clark and Lillard 2024; Subramanian *et al.* 2020). Multi-omics data provide the landscape of cancer characterization from various dimensions and perspectives, providing a data foundation for cancer research.

Identification of genetic and transcript alterations for uncovering the mechanisms of tumorigenesis and new therapeutic targets

Among the numerous alterations observed, only a select few — known as driver mutations — have the

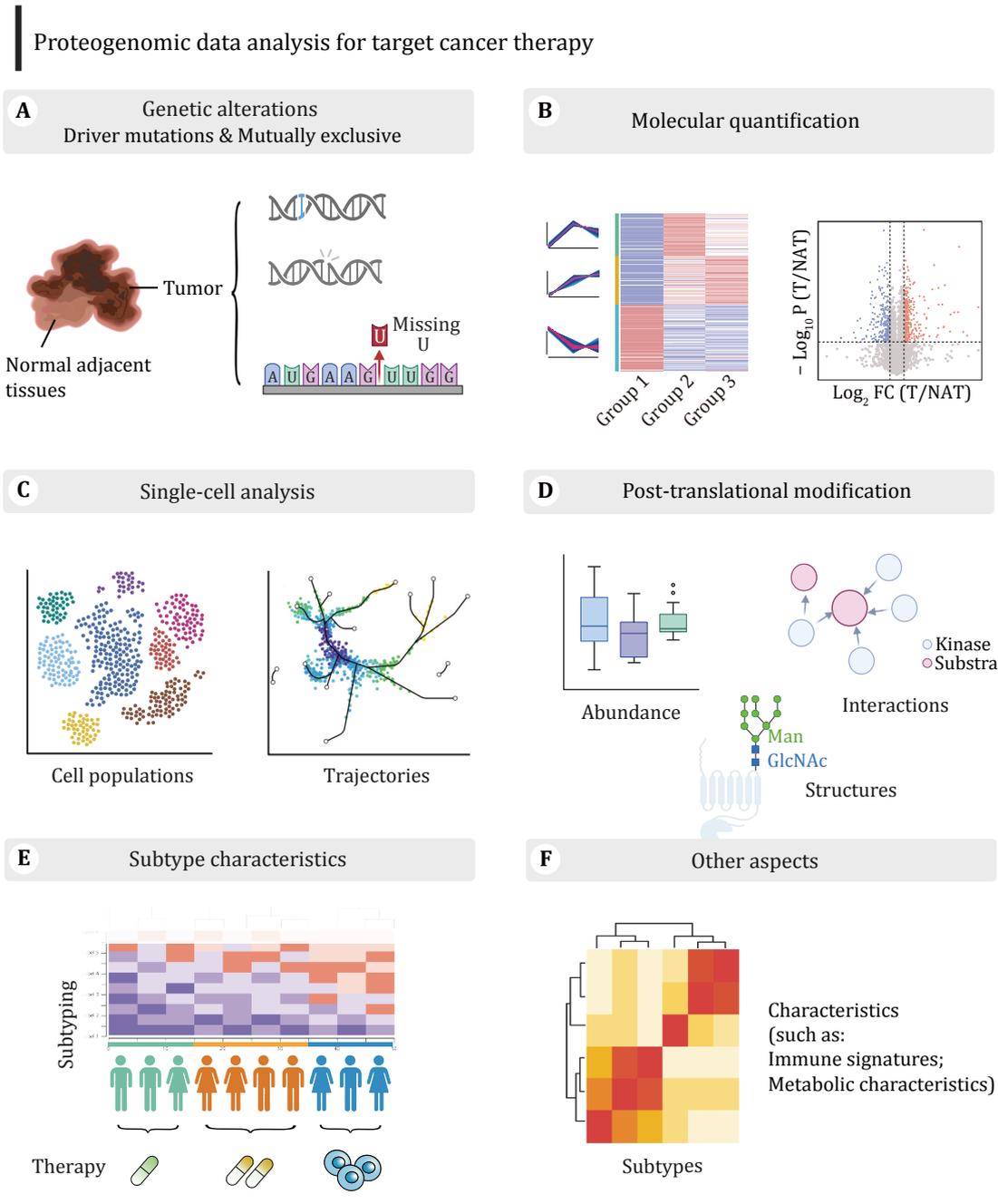


Fig. 2 The proteogenomic data offers different perspectives for target cancer therapy research. Proteogenomics data provide comprehensive insights into various cancer samples (such as tumors, normal tissues, metastatic tissues, *etc.*), revealing their molecular biases and providing the basis for clinical precision medicine. The popular analytical aspects include, but are not limited to, identifying the genetic alterations and single-cell populations in tumor samples, comparing the RNA and protein expression levels among different kinds of samples, depicting the PTM patterns, clustering patients into distinct subtypes, analyzing immune landscapes, *etc.*

potential to initiate and propel cancer progression (Ostroverkhova *et al.* 2023). They are associated with clinical phenotypes and different risk factors, and some of them can be regarded as therapeutic targets and biomarkers (Cancer Genome Atlas Research 2012;

Schulze *et al.* 2015). Since the 1970s, the list of driver genes has been growing gradually, and there are now more than 500 cancer genes (Martínez-Jiménez *et al.* 2020). Somatic changes in cancer genomes often result in transcript alterations. Various forms of RNA

alterations, such as altered splicing and gene fusions, have been described in cancer (Calabrese *et al.* 2020).

By identifying the genomic drivers, we can enhance our understanding of tumorigenesis and identify potential therapy targets. One group applied exome sequencing to hepatocellular carcinoma (HCC), identifying mutational signatures associated with distinct risk factors. They found that 28% of patients harbored at least one alteration potentially targeted by an FDA-approved drug, which was potential treatment chances (Schulze *et al.* 2015). Another group carried out a proteogenomic analysis of HCC across different clinical stages and etiologies. Authors identified the driver gene list, including UBQLN4 and RRB2M, and suggested them as therapeutic targets (Ng *et al.* 2022). Alterations not only can be directly inhibited by targeted drugs but also might indirectly influence drug sensitivity, such as the relevance between NQO1 expression and HSP90 inhibitors (Schulze *et al.* 2015).

A series of filtering criteria can be used to identify putative key drivers (Cao *et al.* 2021; Huang *et al.* 2021; Vasaikar *et al.* 2019). For instance, Vasaikar *et al.* performed a systematic proteogenomic analysis of colon cancer and prioritized genetic alterations. By correlating copy number variations with mRNA and protein levels, and examining the consistent differences in protein expression between paired tumor and NAT samples, they prioritized 90 somatic CNV genes for further investigation. These genes and significantly mutated genes are enriched in cell proliferation, cell death, and Hippo signaling pathways (Vasaikar *et al.* 2019).

Many proteogenomic studies have observed mutually exclusive genomic alterations (Chen *et al.* 2020; Gao *et al.* 2019; Gillette *et al.* 2020; Huang *et al.* 2021). Two plausible situations may explain the mutually exclusive alterations. The first point is that alterations are functionally equivalent, meaning changes to other genes offer no additional benefit (Ciriello *et al.* 2012). Another is synthetic lethality (SL) interactions, where the combination of two genetic abnormalities leads to lethality, despite each being harmless on its own (Ciriello *et al.* 2012; Lee *et al.* 2021). One of its most significant applications is in the formulation of drug combination strategies. Zhang *et al.* developed the Synthetic Lethality Knowledge Graph (SLKG) to explore the SL and synthetic dosage lethality. They integrated genomics and drug datasets into a drugs-targets-indications knowledge graph. A computational model identified potential drug combinations, with the top-ranked SL gene pair CDKN2A + RRM2 validated through experimental methods, indicating that the RRM2 inhibitor may be a promising treatment for melanoma patients with CDKN2A mutations (Zhang *et al.* 2021a).

Quantification of mRNA and protein for revealing the cancer mechanisms and potential biomarkers

The levels of mRNA and proteins are extensively analyzed to detect gene expression levels and signaling pathway activity among different samples (such as comparison between tumors and NATs), thereby offering a distinct molecular foundation. Compiling lists of differentially expressed genes along with their enrichment analysis can highlight potential key factors in tumorigenesis and uncover mechanisms associated with cancer development (Li *et al.* 2023c; Xu *et al.* 2020).

In recent years, the combination of bulk RNA-seq and scRNA-seq analysis has been a popular strategy, offering diverse and complementary perspectives on biological research (Bao *et al.* 2021; Bischoff *et al.* 2021; Briggs *et al.* 2021; Liu *et al.* 2022a, b; Sun *et al.* 2024b). In prostate cancer studies, through mRNA expression analysis at both bulk and single-cell levels, researchers have identified the JAK/STAT signaling pathway as a crucial factor in lineage plasticity, providing a theoretical basis for therapeutic targets (Chan *et al.* 2022; Deng *et al.* 2022). Loss of function of TP53/RB1 and activation of SOX2 occur in lineage plasticity-driven resistance in metastatic castration-resistant prostate cancer (mCRPC). Deng *et al.* analyzed transcriptomic changes induced by TP53/RB1 deficiency and overexpression of SOX2 in cell lines. They found that JAK/STAT pathway was the only cancer-related pathway that was concomitantly altered with the dysfunction in mCRPC. Two major clusters of PCa cell subpopulations were identified by scRNA-seq, with increased expression of JAK/STAT signaling genes observed in the TP53/RB1-low subpopulation. These results indicated that JAK/STAT members may serve as promising targets for resensitizing resistant tumors to AR-targeted therapy (Deng *et al.* 2022).

It's worth noting that the weak correlation between mRNA and proteins suggests the presence of multiple post-transcriptional modifications (Chakraborty *et al.* 2018; Clark *et al.* 2019; Gao *et al.* 2019), underscoring the significance of quantifying proteins directly in biomedical research. Clear cell renal cell carcinoma (CCRCC) is classified as a metabolic disease (Wettersten *et al.* 2017). A study revealed a prominent inconsistency of oxidative phosphorylation (OXPHOS) mRNA and protein expression in CCRCC tumors. Proteins and mRNAs in the glycolysis pathway were upregulated, and the mRNA expression level of the OXPHOS components was not changed. However, proteins associated with OXPHOS were significantly downregulated, reflecting the regulation of OXPHOS primarily at the translational level (Clark *et al.* 2019).

The differences in molecular expression levels among different samples can be analyzed in conjunction with clinical characteristics. For example, smoking is an important factor contributing to lung cancer. One group defined a smoking score (SS) to study the relevant phenotypes in lung adenocarcinoma (LUAD). Pathways such as cell cycle and signaling by FGFR were reduced, while P53 pathway and apoptosis were higher in NATs with high SS, but these patterns were reversed in tumors. Notably, the abundance of ARHGEF5 and its phosphosite Y1370y were elevated in tumors from strict never-smokers, highlighting a potential oncogenic mechanism in never-smokers (Gillette *et al.* 2020).

Single-cell omics capturing tumoral and microenvironmental heterogeneity

The core application of single-cell omics is defining cell populations, shedding light on how specific cell types contribute to tumor characteristics. The application of single-cell omics to cell biology and preclinical research is still in its infancy, yet it holds significant potential for elucidating the cellular landscape (Baysoy *et al.* 2023). The tumor microenvironment (TME) is composed of various cellular components, which could provide growth and anti-apoptotic benefits to surviving cancer cells, decrease drug penetration, and alter disease characteristics while distorting clinical outcomes (Sun 2016). Single-cell omics has revolutionized our understanding of the TME, allowing for the exploration of the notable differences in its cell types.

Understanding the complex disturbance within the TME is crucial for understanding oncology (Cao *et al.* 2023; Kim *et al.* 2024; Luo *et al.* 2024; Ramberger *et al.* 2024; Sorin *et al.* 2023; Wu *et al.* 2023). Take renal cell carcinoma (RCC) as an example. Metastatic CCRCC is considered an immunotherapy sensitive tumor while others are not (Zhang *et al.* 2021c). There are some studies that reveal the immune-related heterogeneity of RCC through single-cell RNA-seq (Li *et al.* 2023b, 2024; Zhang *et al.* 2021c). Zhang *et al.* performed scRNA-seq on patients of RCCs and benign adjacent tissues. High immune cell infiltration was observed in CCRCC compared with classic chromophobe RCC, which is consistent with their responses to immunotherapy. They identified a list of genes that were associated with clinical tyrosine kinase inhibitors followed by nivolumab outcomes. It is notable that genes with a negative correlation with the response were primarily expressed in the major endothelial cluster in CCRCC. This indicated that metastatic CCRCC patients with

high endothelial cell levels may respond poorly to immunotherapy (Zhang *et al.* 2021c).

Single-cell sequencing can model dynamic cellular changes by analyzing expression pattern similarities to construct differentiation and growth trajectories. Comparing characteristics of different trajectories offers insights into cancer development and therapy response (Trapnell *et al.* 2014; Van den Berge *et al.* 2020). Prostate cancer (PCa) is an extensive heterogeneous cancer, with androgen-deprivation therapy (ADT) being the standard treatment for metastatic cases. Bian *et al.* performed a single-cell analysis of PCa tissues and clustered myeloid cells into seven cell types. They found that IL1B-NLRP3 macrophages were enriched in ADT-treated patients. Pseudotime analysis identified two evolutionary trajectories of macrophages. Key genes defining these branches were identified, suggesting that fate2, primarily composed of IL1B-NLRP3 subsets, represents a pro-inflammatory phenotype or immune activation state. These results indicated that ADT might promote the formation of IL1B-NLRP3 macrophages, which may have anticancer effects (Bian *et al.* 2024). Another research of PCa revealed FOXA2 as a neuroendocrine lineage-pioneering transcription factor by inferencing cell differentiation progression trajectory. Further studies indicate that FOXA2 can drive KIT pathway activation in neuroendocrine cells in PCa (Han *et al.* 2022).

Annotation of PTMs for uncovering molecular mechanisms and potential therapeutic targets and biomarker

PTMs play a crucial role in enhancing the functional diversity of proteins and providing a complex layer of regulation essential for cellular processes. Analyzing trends in PTM site expression helps researchers explore tumor mechanisms at the post-translational level. For example, Vasaikar *et al.* discovered that RB1 was amplified and protein was upregulated in colorectal cancer (CRC). This phenomenon contradicted the role of RB1 as a tumor suppressor gene (TSG). Then they explored the phosphoproteomic data and found that RB1 phosphorylation is upregulated. Through their analysis, they concluded that RB1 phosphorylation, rather than RB1 mutations, drives tumor growth and has an anti-apoptotic effect (Vasaikar *et al.* 2019).

Kinases are one of the most intensively pursued targets in antitumor drug research, due to their critical roles in cellular signaling (Wu *et al.* 2015). There is considerable interest in measuring phosphorylation abundance levels and kinase activity to identify potential

therapeutic targets (Huang *et al.* 2017; Krug *et al.* 2020). Liu *et al.* subtyped a cohort of small cell lung cancer (SCLC) into four clusters. By comparing the difference in phosphosite abundance among different samples, they identified a subtype characterized by elevated receptor tyrosine kinase (RTK) signaling and suggested RTK inhibitors as an effective strategy for these tumors (Liu *et al.* 2024). There has also been increased focus on inferring kinase activity, understanding kinase-substrate networks, and assessing their implications for drug responses (Savage and Zhang 2020). Several algorithms are helping to predict the kinase activity such as KSEA (kinase-substrate enrichment analysis) (Wiredja *et al.* 2017), IKAP (Inference of kinase activities from phosphoproteomics) (Mischnik *et al.* 2016), *etc.* (Piersma *et al.* 2024). McDermott *et al.* performed a proteogenomic study on ovarian high-grade serous cancer (HGSC). Using KSEA revealed significantly increased activity of multiple kinases associated with cell-cycle control in tumors, including CDKs and AURKA. These results suggest the possible therapeutic target for HGSC and prompted an investigation of mechanisms related to cell cycle regulation and replication stress in this tumor (McDermott *et al.* 2020).

Glycobiology has become increasingly prominent in cancer research, involving numerous vital biological processes such as inflammation, cell–cell adhesion, and cell–matrix interaction (Pinho and Reis 2015). Aberrant glycans and glycosylation enzymes have been identified as potential targets for cancer diagnosis and treatment (Pinho and Reis 2015; Tian *et al.* 2022). For example, Rolland *et al.* generated the landscape of N-glycoproteins from lymphoma tissue biopsies and cell lines. They focused on anaplastic lymphoma kinase-positive anaplastic large cell lymphoma (ALKALCL) and several proteins with cytokine receptor activity. Among them, IL-31R β was exclusively expressed in ALKALCL cell lines and regulated by ALK. Through functional verification by biological experiments such as RNA interference, IL-31R β is considered a therapeutic target for ALKALCL (Rolland *et al.* 2017).

Patient stratification for subgroup-specific therapeutics

Different pathogenic factors can activate diverse molecular mechanisms, leading to significant heterogeneity among patients. A popular strategy is classifying clinical samples and analyzing their distinct characteristics. Common methods for classifying include consensus clustering (Clark *et al.* 2019; Monti *et al.* 2003) and non-negative matrix factorization (NMF) (Cao *et al.* 2021; Huang *et al.* 2021; Rappoport and Shamir 2018).

By patient stratification, precise treatment plans for different subtypes can be formulated, aiding the selection of more effective treatments for individuals (Laurent-Puig *et al.* 2001; Menyhart and Györfy 2021; Nguyen *et al.* 2020; Smith *et al.* 2022).

Patients in different subtypes can exhibit significant differences in various characteristics, such as mutations, RNA or protein or PTM site abundance levels, and pathway activity. Through subtyping of patient cohorts, potential biomarkers or targets can be identified, aiding the selection of more effective treatments for individuals. Triple-negative breast cancer (TNBC) is known for its heterogeneity and treatment difficulties (Burstein *et al.* 2015; Gong *et al.* 2022; Stewart *et al.* 2020; Xiao *et al.* 2022). Burstein *et al.* carried out proteogenomic characterization on 198 TNBC tumors, confirming four distinct subtypes based on RNA expression profiling. They analyzed subtype specific characterization and identified subtype-specific targets. One of the subtypes exhibited activation of immunoregulation pathways and STAT mediated pathways, indicating that STAT inhibitors and cytokine/cytokine receptor antibodies might be suitable for this subtype (Burstein *et al.* 2015). Another study on TNBC in China divided another cohort into four subtypes and found that the JAK/STAT3 signaling pathway was noticed in the mesenchymal-like subtype, suggesting that STAT3 inhibitors could be a promising treatment strategy for this subgroup (Jiang *et al.* 2019).

Subtypes identified using different datasets may show discrepancies in their results. Several studies highlight the advantages of proteomic typing, such as providing a more accurate indication of clinical responses and patterns (Gao *et al.* 2019; Macklin *et al.* 2020; Xu *et al.* 2020). Gao *et al.* performed clustering of HCC based on differentially expressed proteins between tumor and NAT, and established several screening criteria to identify representative proteins for subtypes. They pinpointed PYCR2 and ADH1A exhibited significant differential expression across the proteomics subtypes and strongly correlated with patient survival. These proteins were identified as prognostic biomarkers with potential clinical utility (Gao *et al.* 2019). The same team conducted another research for intrahepatic cholangiocarcinoma and identified four proteomic subgroups (S1–S4) with divergent genomic and clinical features. They discovered that subtypes S1 and S2 exhibited greater sensitivity to EGFR inhibitors, while cell lines other than those in S1 and S3 demonstrated sensitivity to Gemcitabine and Paclitaxel (Dong *et al.* 2022).

Analysis of the immune landscape suggesting the applicability of immunotherapy

Immunotherapy is one of the important methods of cancer therapy, and proteogenomic analysis aids in the research of immunotherapy from many aspects. The composition and density of immune cells in the TME, which significantly impact tumor progression and the effectiveness of anti-cancer treatments, can be deconstructed by various methods (Sturm *et al.* 2019). Besides, tumor-specific antigens can stimulate T cell-mediated antitumor immune responses and serve as available targets for cancer immunotherapy.

Based on bulk proteomic and gene expression data, there are many tools such as xCell (Aran *et al.* 2017) and BayesDeBulk (Petralia *et al.* 2023), that can analyze and classify the immune and stromal cell composition (Sturm *et al.* 2019). Tumors can be clustered based on cell signatures, such as being grouped into immune hot, warm, and cold clusters, allowing for the investigation of disparities among these groups, and providing insights into cancer immunotherapy (Cao *et al.* 2021; Gillette *et al.* 2020; Krug *et al.* 2020; Liu *et al.* 2024; Satpathy *et al.* 2021; Wang *et al.* 2021a). Pancreatic ductal adenocarcinoma (PDAC) is a highly heterogeneous and aggressive tumor, usually resistant to immune checkpoint inhibitors. One group of CPTAC profiled the proteogenomic characterization of PDAC and delineated the cellular composition of tumors by xCell. They classified the tumors into four clusters according to the cell composition and characterized the 'immune-cold' phenotype. These tumors showed reduced expression of endothelial adhesion proteins and increased activity of VEGF and hypoxia pathways, which were important to the remodeling of endothelial cells. Immune-cold samples also exhibited elevated levels of glycolysis. Researchers proposed that targeting glycolysis and endothelial cell remodeling, could be therapeutically leveraged to enhance antitumor immunity in immune-cold PDACs, which can be verified by clinical outcomes (Cao *et al.* 2021).

Single-cell omics data provide a more intuitive representation of the cellular composition within the TME (Bärthel *et al.* 2023; Cao *et al.* 2023; Christodoulou and Zaravinos 2023; Chung *et al.* 2017; Gessain *et al.* 2024). Zhang *et al.* carried out single-cell analysis for TNBC patients. A total of 22 patients were included in this cohort, with half receiving paclitaxel monotherapy and the other half receiving paclitaxel in combination with atezolizumab. Through measuring the correlations between cell proportion change with tumor size changes,

T cells notably exhibited a positive role in PD-L1 blockade therapy. Researchers identified two T cell clusters (*i.e.*, CD8-CXCL13 and CD4-CXCL13), which showed decreased levels in paclitaxel treatment responders but increased levels in responders following paclitaxel plus atezolizumab treatment. These results demonstrated the important role of CXCL13+ T cells in predicting the clinical outcomes of anti-PD-L1 therapy (Zhang *et al.* 2021b).

Neoepitopes are identified by the adaptive immune system and act as a mechanism for the immune system to distinguish cancer cells from normal cells, serving as the promising target. Somatic or passenger mutations in the tumor lead to the formation of new epitopes or neoepitopes (Brennick *et al.* 2017; Ren *et al.* 2024). MS is the primary technique for directly identifying neoepitopes displayed on tumor cells (Xiang *et al.* 2022). Dong *et al.* performed proteogenomic analysis for intrahepatic cholangiocarcinoma (iCCA). FGFR2 fusions were found in 11.9% (30/253) in this cohort and the common fusion partners included BICC1. For the two most common FGFR2::BICC1 fusion proteins, a set of fusion-derived peptides was built. Although the results showed a low frequency of neoepitope T cell reactivity, some peptides derived from FGFR2 fusion proteins exhibited strong immunogenicity and were potential targets for immune antigens (Dong *et al.* 2022).

INTEGRATIVE ANALYSIS

Combined analysis of omic data

Comprehensive proteogenomic studies treat the tumor as a system, offering insights from myriad layers, thereby promising a more systematic and holistic understanding of the intricate nature of cancer and drug interactions. Therefore, integrating proteogenomic data enhances the credibility and reliability of research outcomes, ultimately refining preventive measures and treatment strategies for cancer (Subramanian *et al.* 2020; Zhang *et al.* 2020a).

There are some common strategies to implement integrated analyses. The first strategy is data mapping, where genomic features are mapped to transcriptomic and proteomic data to identify relationships between genetic variations and their downstream effects. A specific CNV or mutation could be linked to gene expression changes, which then influence protein abundance or PTM levels. The second is comparing data from different omics based on shared identifiers (*e.g.*, gene symbol). Statistical methods (*e.g.*, statistical tests and

regression) can be used to identify relationships between multiple datasets, such as gene expression, protein abundance and PTM abundance data. The third strategy is the application of machine learning methods. For example, unsupervised learning is widely used for patient subtyping. Clustering algorithms (*e.g.*, k-means, hierarchical clustering, t-SNE) are used to identify patterns across multiple omics data types without predefined labels. Additionally, there are many tools and software available for multi-omics integration.

Combining the different levels of omics data can provide novel and comprehensive insights into cancer pathogenesis, further suggesting the identification of candidate biomarkers and therapeutic targets. A typical example was provided by Satpathy *et al.* They studied lung squamous cell carcinoma samples and found that NSD3 is an amplified cancer-related gene within a focal amplicon that includes FGFR1. Their proteomics data showed that NSD3, not FGFR1, was significantly upregulated in tumors with high FGFR1 amplification. These results suggest the limited effectiveness of FGFR1-targeted therapies and highlight NSD3 as a potential alternative treatment (Satpathy *et al.* 2021). Gillette *et al.* focused on therapeutic vulnerabilities in LUAD and integrated mutations and phosphosites data. They found that KRAS mutant tumors showed significant upregulation of SOS1 phosphorylation on S1161, while in EGFR mutant tumors was PTPN11/Shp2 at Y26. This may indicate new therapies and PTPN11/Shp2 inhibitors perform well in preclinical trials (Gillette *et al.* 2020).

There are numerous studies revealing that the median correlation between mRNA and protein abundance is weak (around 0.4), indicating post-transcriptional regulation. Further analysis indicates that the degree of correlation between mRNA and protein levels is related to the gene's biological function. The weaker correlations were observed in highly stable and abundant proteins (*e.g.* ribosomes, spliceosomes), while proteins that are more dynamically regulated in response to disturbances (such as a metabolic process) show stronger correlations (Clark *et al.* 2019; Gao *et al.* 2019; Gillette *et al.* 2020; Zhang *et al.* 2016). These findings suggest that discrepancies between mRNA and protein data are common and normal, reflecting the regulation at different biological levels. To address these issues, we can analyze the data in conjunction with the protein's function and select the appropriate data according to the specific research requirements.

Genetic variants can broadly impact the transcriptional, translational, and post-translational product. If

the effect is on the expression of itself, it is referred to as 'cis-effect'; if it affects other genes, it is termed 'trans-effect' (Clark *et al.* 2019; Li *et al.* 2023c). The correlation between RNA and CNV changes is usually more significant than that between protein and CNV changes (Gao *et al.* 2019; Gonçalves *et al.* 2017). Observing the dosage cascade regulation can provide novel insights into the mechanisms of cancer. A group conducted proteogenomic analysis on Chinese CCRCC patients. The data showed that the center of CNVs with cis-effect included 3q, where the copy number was negatively correlated with survival. The researchers prioritized genes located at 3p that showed significant cis-effects and evaluated their correlation with survival. They focused on the SLC4A7, which is crucial for cellular net acid extrusion and maintaining intracellular pH balance, indicating the significance of SLC4A7 in regulating cellular pH balance in tumors (Qu *et al.* 2022).

It is worth mentioning that a popular strategy is multi-omics clustering. The mutations, CNVs, mRNA expression, protein, and PTM site abundance can all be used as input features for data clustering (Krug *et al.* 2020; Wang *et al.* 2021a), providing the basis for subtype-specific therapeutic chances. Liu *et al.* subtype SCLC tumors into four nmf subtypes based on multi-omics data. The nmf1 subtype was characterized by high replication stress and neuroendocrine differentiation, suggesting the response to drugs that exacerbate genome instability; the nmf2 subtype showed a high abundance of DLL3, suggesting the response to anti-DLL3 therapies; the nmf3 subtype showed elevated RTK signaling, which may be response to RTK inhibitor treatment; the nmf4 subtype displayed high expression of MYC and POU2F3, indicating its benefit from AURK inhibitors (Liu *et al.* 2024).

There are also some designed statistical methods to integrate and analyze multiple types of omics data, such as Multi-Omics Factor Analysis (MOFA) (Argelaguet *et al.* 2018, 2020) and mixOmics (Rohart *et al.* 2017). MOFA infers a set of latent factors (LF) representing sources of variability from different omics data sets across the samples (Argelaguet *et al.* 2018). It helped researchers to reveal factors associated with clinically relevant biology (Herbst *et al.* 2022). Sharma *et al.* employed MOFA+ to integrate multi-omics data from breast cancer patients. LF1 and LF2 represented most of the heterogeneity showing a strong association with tumor grade. LF1 and LF13 were significantly associated with BC metastasis and death. Based on LF1, LF2, and LF13, patients were grouped into three clusters, providing a new basis for the characterization and understanding of BC (Sharma *et al.* 2024).

Pan-cancer analysis

Various cancers exhibit unique characteristics. Yet, overarching commonalities exist across an array of cancers, such as significant driver mutations, altered pathways, and immune signatures. These similarities highlight the value of a strategy called ‘pan-cancer’, which targets traits shared among various cancer types (Chen *et al.* 2021). Pan-cancer studies focus on the molecular similarities and differences across various cancer types, rather than solely on the tissue of origin.

After completing genomic investigation across 33 cancer types, TCGA demonstrated the significance of the pan-cancer study and published the project namely the ‘Pan-Cancer Atlas project’ (Bailey *et al.* 2018). Later in 2020, the Pan-Cancer Analysis of Whole Genomes (PCAWG) consortium was established, integrating 2,658 whole-cancer genome data shared by compute clouds (Consortium 2020). The comprehensive cancer genome analysis studies were carried out from many aspects, such as driver mutations (Consortium 2020), non-coding somatic drivers (Rheinbay *et al.* 2020), mutational signatures (Alexandrov *et al.* 2020), *etc.* PCAWG mainly concentrates its attention on the analysis of genomics data. CPTAC consortium has also published a series of proteogenomic pan-cancer studies and provided a comprehensive public data resource, profiling many aspects of pan-cancer, including the characterization of oncogenic drivers, the pattern of PTMs, DNA methylation, tumor immunity, and therapeutic targets (Geffen *et al.* 2023; Li *et al.* 2023a, c; Liang *et al.* 2023; Petralia *et al.* 2024; Savage *et al.* 2024). They generated the harmonized proteogenomic and clinic data over ten cancer types and also established a pipeline for pan-cancer sample collection and data harmonization (Li *et al.* 2023a).

Pan-cancer studies can facilitate the identification of potential cancer drivers and prognostic relevance across different cancer types (Petralia *et al.* 2024; Wang *et al.* 2021b; Zack *et al.* 2013). Savage *et al.* published a pan-cancer study to expand the landscape of therapeutic targets. Through analyzing the associations between alteration of TSGs and variation in protein abundance, they focused on the TP53-TOP2A pairs in EC. TOP2A can be targeted by chemotherapy agents such as doxorubicin, whose abundance was significantly higher in tumors harboring TO53 loss. These results suggested that TP53 loss may serve as a biomarker for selecting EC patients who are suitable for treatment with doxorubicin (Savage *et al.* 2024). Li *et al.* provided insights into shared oncogenic drivers across ten types of cancer. They identified that PLOD2, UBE2C, and

MARCKSL1 were commonly differentially expressed between NATs and tumors. Among these genes, PLOD2 abundance was associated with survival and can be regarded as a pan-cancer prognostic biomarker (Li *et al.* 2023c).

SUMMARY AND PERSPECTIVES

Proteogenomic technologies are revolutionizing pre-clinical research by objectively and comprehensively delineating the entire molecular landscape. Proteogenomic datasets contain information from various biological dimensions, helping to uncover the mechanisms of tumorigenesis, identify potential therapeutic targets and specific biomarkers, and drive the development of precision medicine. The ultimate goal of preclinical proteogenomic studies is to create strategies for patient-specific cancer therapy, taking into account the heterogeneity of different cancers.

Artificial intelligence (AI) has been rapidly applied in preclinical cancer research because of its predictive power and precision. A major area of AI research is centered on enabling machines to derive general rules from data, a process known as machine learning (ML), to make predictions about previously unseen samples. ML algorithms can identify hidden patterns and relationships within large datasets from diverse sources, which might not be evident through traditional methods (Rehman *et al.* 2024). AI-driven methods can uncover novel targets that were previously unrecognized or ignored and develop more robust classification approaches for clinically relevant cancer subtypes (Bonniolo *et al.* 2021). In the future, we look forward to AI, especially ML, playing a key role in large-scale data analysis, and further promoting the development of personalized therapies based on individual molecular profiles.

Despite significant progress, we still have a long way to go in fully understanding cancer and overcoming the various challenges. Firstly, advancements in proteomic technologies have not progressed as rapidly, especially for single-cell proteomic analysis (Olivier *et al.* 2019). There are numerous challenges such as detection sensitivity, dynamic range, analytical throughput, and equipment costs. It is worth mentioning that advanced protein sequencing techniques may provide more convenient and accessible platforms for proteomic research (Guo *et al.* 2024; Singh 2022). For example, nanopore technology can directly capture molecular nature as single molecules pass through nanopores, enabling the sequencing of intact protein (Motone *et al.*

2024). Secondly, a major challenge lies in integrating and interpreting the vast array of results from diverse sources. There are many public data resources and trying to combine data from different public resources is complicated. Different experimental and computational methods can yield inconsistent results, leading to potential false positives or negatives. Besides, there is difficulty trying to extract authentic knowledge out of large multi-omic data sets. With the impending of AI, machine learning and similar technologies may become especially valuable for interpreting large-scale data (Darbandi *et al.* 2024; Sun *et al.* 2022; Xiao *et al.* 2021). Thirdly, translating biological cancer research into clinical practice remains challenging. There are some reasons: (1) Tumor tissues exhibit significant heterogeneity, and samples used in proteogenomic studies can only represent a subset of tumor characteristics. (2) The patients represent only a subset of the population and do not reflect the entire patient population (Liu *et al.* 2024). Even so, studies have showcased the potential of current proteogenomic findings for clinical applications (Klomp *et al.* 2024; Liu *et al.* 2024).

Looking ahead, there are several future directions in proteogenomic research. Firstly, proteogenomics is expected to evolve into single-cell spatial multi-omics, three-dimensional spatial omics, and spatiotemporal multi-omics. These innovations will generate more detailed, high-resolution data, providing valuable insights into the dynamic alterations in tumor molecular structures throughout progression and treatment. Secondly, an increasing number of patients, including those from different regions, populations, and various cancer subtypes, should be included in research cohorts to enhance the generalizability and representativeness of the studies. Thirdly, diverse research perspectives are needed rather than a standardized, assembly-line approach to analysis, which may rely on the development of bioinformatics tools and AI. For instance, we should conduct clinically-driven analyses, focusing on specific clinical issues such as drug treatment, and tailor the analysis accordingly. Fourthly, the ultimate goal of proteogenomic research is to advance disease treatment. If the identified targets or biomarkers are used in clinical trials and yield positive results, the study will attract widespread attention. Collaboration among experts from diverse backgrounds becomes increasingly crucial (Alseekh *et al.* 2021; Doll *et al.* 2019), for developing improved treatment strategies for patients and advancing precision medicine.

Acknowledgements This work was financially supported by the National Key Research and Development Program

(2022YFA1302902 to H.Z.) and the Shanghai Municipal Science and Technology Major Project. The work was supported by the Strategic Priority Research Program of the Chinese Academy of Sciences (XDB0830000), and the National Natural Science Foundation of China (22425703 to H.Z.). This illustrations were created with biorender.com.

Compliance with Ethical Standards

Conflict of interest Yuying Suo, Yuanli Song, Yuqiu Wang, Qian Liu, Henry Rodriguez, Hu Zhou declare that they have no conflict of interest.

Human and animal rights and informed consent This article does not contain any studies with human or animal subjects performed by the any of the authors.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Akbani R, Ng PK, Werner HM, Shahmoradgoli M, Zhang F, Ju Z, Liu W, Yang JY, Yoshihara K, Li J, Ling S, Sevior EG, Ram PT, Minna JD, Diao L, Tong P, Heymach JV, Hill SM, Dondelinger F, Stadler N, Byers LA, Meric-Bernstam F, Weinstein JN, Broom BM, Verhaak RG, Liang H, Mukherjee S, Lu Y, Mills GB (2014) A pan-cancer proteomic perspective on The Cancer Genome Atlas. *Nat Commun* 5: 3887. <https://doi.org/10.1038/ncomms4887>
- Alexandrov LB, Kim J, Haradhvala NJ, Huang MN, Tian Ng AW, Wu Y, Boot A, Covington KR, Gordenin DA, Bergstrom EN, Islam SMA, Lopez-Bigas N, Klimczak LJ, McPherson JR, Morganella S, Sabarinathan R, Wheeler DA, Mustonen V, Group PMSW, Getz G, Rozen SG, Stratton MR, Consortium P (2020) The repertoire of mutational signatures in human cancer. *Nature* 578(7793): 94–101
- Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SAJR, Behjati S, Biankin AV, Bignell GR, Bolli N, Borg A, Børresen-Dale A-L, Boyault S, Burkhardt B, Butler AP, Caldas C, Davies HR, Desmedt C, Eils R, Eyfjörd JE, Foekens JA, Greaves M, Hosoda F, Hutter B, Ilicic T, Imbeaud S, Imielinski M, Jäger N, Jones DTW, Jones D, Knappskog S, Kool M, Lakhani SR, López-Otín C, Martin S, Munshi NC, Nakamura H, Northcott PA, Pajic M, Papaemmanuil E, Paradiso A, Pearson JV, Puente XS, Raine K, Ramakrishna M, Richardson AL, Richter J, Rosenstiel P, Schlesner M, Schumacher TN, Span PN, Teague JW, Totoki Y, Tutt ANJ, Valdés-Mas R, van Buuren MM, van 't Veer L, Vincent-Salomon A, Waddell N, Yates LR, Zucman-Rossi J, Andrew Futreal P, McDermott U, Lichter P, Meyerson M, Grimmond SM, Siebert R, Campo E, Shibata T, Pfister SM, Campbell PJ, Stratton

- MR, Australian Pancreatic Cancer Genome I, Consortium IBC, Consortium IM-S, PedBrain I (2013) Signatures of mutational processes in human cancer. *Nature* 500(7463): 415–421
- Alseikh S, Aharoni A, Brotman Y, Contrepolis K, D'Auria J, Ewald J, C. Ewald J, Fraser PD, Giavalisco P, Hall RD, Heinemann M, Link H, Luo J, Neumann S, Nielsen J, Perez de Souza L, Saito K, Sauer U, Schroeder FC, Schuster S, Siuzdak G, Skirycz A, Sumner LW, Snyder MP, Tang H, Tohge T, Wang Y, Wen W, Wu S, Xu G, Zamboni N, Fernie AR (2021) Mass spectrometry-based metabolomics: a guide for annotation, quantification and best reporting practices. *Nat Methods* 18(7): 747–756
- Ang MY, Low TY, Lee PY, Wan Mohamad Nazarie WF, Guryev V, Jamal R (2019) Proteogenomics: From next-generation sequencing (NGS) and mass spectrometry-based proteomics to precision medicine. *Clin Chim Acta* 498: 38–46
- Aran D, Hu Z, Butte AJ (2017) xCell: digitally portraying the tissue cellular heterogeneity landscape. *Genome Biol* 18(1): 220. <https://doi.org/10.1186/s13059-017-1349-1>
- Argelaguet R, Arnol D, Bredikhin D, Deloro Y, Velten B, Marioni JC, Stegle O (2020) MOFA+: a statistical framework for comprehensive integration of multi-modal single-cell data. *Genome Biol* 21(1): 111. <https://doi.org/10.1186/s13059-020-02015-1>
- Argelaguet R, Velten B, Arnol D, Dietrich S, Zenz T, Marioni JC, Buettner F, Huber W, Stegle O (2018) Multi-Omics Factor Analysis—a framework for unsupervised integration of multi-omics data sets. *Mol Syst Biol* 14(6): e8124. <https://doi.org/10.15252/msb.20178124>
- Bailey MH, Tokheim C, Porta-Pardo E, Sengupta S, Bertrand D, Weerasinghe A, Colaprico A, Wendl MC, Kim J, Reardon B, Ng PK-S, Jeong KJ, Cao S, Wang Z, Gao J, Gao Q, Wang F, Liu EM, Mularoni L, Rubio-Perez C, Nagarajan N, Cortés-Ciriano I, Zhou DC, Liang W-W, Hess JM, Yellapantula VD, Tamborero D, Gonzalez-Perez A, Suphavitai C, Ko JY, Khurana E, Park PJ, Van Allen EM, Liang H, Group TMW, Network TCGAR, Lawrence MS, Godzik A, Lopez-Bigas N, Stuart J, Wheeler D, Getz G, Chen K, Lazar AJ, Mills GB, Karchin R, Ding L (2018) Comprehensive characterization of cancer driver genes and mutations. *Cell* 173(2): 371–385
- Bandura DR, Baranov VI, Ornatsky OI, Antonov A, Kinach R, Lou X, Pavlov S, Vorobiev S, Dick JE, Tanner SD (2009) Mass cytometry: technique for real time single cell multitarget immunoassay based on inductively coupled plasma time-of-flight mass spectrometry. *Anal Chem* 81(16): 6813–6822
- Bao X, Shi R, Zhao T, Wang Y, Anastasov N, Rosemann M, Fang W (2021) Integrated analysis of single-cell RNA-seq and bulk RNA-seq unravels tumour heterogeneity plus M2-like tumour-associated macrophage infiltration and aggressiveness in TNBC. *Cancer Immunol, Immunother* 70(1): 189–202
- Bärthel S, Falcomatà C, Rad R, Theis FJ, Saur D (2023) Single-cell profiling to explore pancreatic cancer heterogeneity, plasticity and response to therapy. *Nat Cancer* 4(4): 454–467
- Baysou A, Bai Z, Satija R, Fan R (2023) The technological landscape and applications of single-cell multi-omics. *Nat Rev Mol Cell Biol* 4(10): 695–713
- Bian S, Hou Y, Zhou X, Li X, Yong J, Wang Y, Wang W, Yan J, Hu B, Guo H, Wang J, Gao S, Mao Y, Dong J, Zhu P, Xiu D, Yan L, Wen L, Qiao J, Tang F, Fu W (2018) Single-cell multiomics sequencing and analyses of human colorectal cancer. *Science* 362(6418): 1060–1063
- Bian X, Wang W, Abudurexiti M, Zhang X, Ma W, Shi G, Du L, Xu M, Wang X, Tan C, Sun H, He X, Zhang C, Zhu Y, Zhang M, Ye D, Wang J (2024) Integration analysis of single-cell multi-omics reveals prostate cancer heterogeneity. *Adv Sci* 11(18): 2305724. <https://doi.org/10.1002/adv.202305724>
- Bischoff P, Trinks A, Obermayer B, Pett JP, Wiederspahn J, Uhlitz F, Liang X, Lehmann A, Jurmeister P, Elsner A (2021) Single-cell RNA sequencing reveals distinct tumor microenvironmental patterns in lung adenocarcinoma. *Oncogene* 40(50): 6748–6758
- Boniolo F, Dorigatti E, Ohnmacht AJ, Saur D, Schubert B, Menden MP (2021) Artificial intelligence in early drug discovery enabling precision medicine. *Expert Opin Drug Discov* 16(9): 991–1007
- Brennick CA, George MM, Corwin WL, Srivastava PK, Ebrahimi-Nik H (2017) Neoepitopes as cancer immunotherapy targets: key challenges and opportunities. *Immunotherapy* 9(4): 361–371
- Briggs EM, Rojas F, McCulloch R, Matthews KR, Otto TD (2021) Single-cell transcriptomic analysis of bloodstream *Trypanosoma brucei* reconstructs cell cycle progression and developmental quorum sensing. *Nat Commun* 12(1): 1–15
- Brunner A-D, Thielert M, Vasilopoulou CG, Ammar C, Coscia F, Mund A, Hoerning OB, Bache N, Apalategui A, Lubeck M, Richter S, Fischer DS, Raether O, Park MA, Meier F, Theis FJ, Mann M (2021) Ultra-high sensitivity mass spectrometry quantifies single-cell proteome changes upon perturbation. *bioRxiv*. <https://doi.org/10.1101/2020.12.22.423933>
- Burstein MD, Tsimelzon A, Poage GM, Covington KR, Contreras A, Fuqua SAW, Savage MI, Osborne CK, Hilsenbeck SG, Chang JC, Mills GB, Lau CC, Brown PH (2015) Comprehensive genomic analysis identifies novel subtypes and targets of triple-negative breast cancer. *Clin Cancer Res* 21(7): 1688–1698
- Bustamante CD, De La Vega FM, Burchard EG (2011) Genomics for the world. *Nature* 475(7355): 163–165
- Calabrese C, Davidson NR, Demircioğlu D, Fonseca NA, He Y, Kahles A, Lehmann K-V, Liu F, Shiraishi Y, Soulette CM, Urban L, Calabrese C, Davidson NR, Demircioğlu D, Fonseca NA, He Y, Kahles A, Lehmann K-V, Liu F, Shiraishi Y, Soulette CM, Urban L, Greger L, Li S, Liu D, Perry MD, Xiang Q, Zhang F, Zhang J, Bailey P, Erkek S, Hoadley KA, Hou Y, Huska MR, Kilpinen H, Korbel JO, Marin MG, Markowski J, Nandi T, Pan-Hammarström Q, Pedamallu CS, Siebert R, Stark SG, Su H, Tan P, Waszak SM, Yung C, Zhu S, Awadalla P, Creighton CJ, Meyerson M, Ouellette BFF, Wu K, Yang H, Fonseca NA, Kahles A, Lehmann K-V, Urban L, Soulette CM, Shiraishi Y, Liu F, He Y, Demircioğlu D, Davidson NR, Calabrese C, Zhang J, Perry MD, Xiang Q, Greger L, Li S, Liu D, Stark SG, Zhang F, Amin SB, Bailey P, Chateigner A, Cortés-Ciriano I, Craft B, Erkek S, Frenkel-Morgenstern M, Goldman M, Hoadley KA, Hou Y, Huska MR, Khurana E, Kilpinen H, Korbel JO, Lamaze FC, Li C, Li X, Li X, Liu X, Marin MG, Markowski J, Nandi T, Nielsen MM, Ojesina AI, Pan-Hammarström Q, Park PJ, Pedamallu CS, Pedersen JS, Siebert R, Su H, Tan P, Teh BT, Wang J, Waszak SM, Xiong H, Yackneen S, Ye C, Yung C, Zhang X, Zheng L, Zhu J, Zhu S, Awadalla P, Creighton CJ, Meyerson M, Ouellette BFF, Wu K, Yang H, Göke J, Schwarz RF, Stegle O, Zhang Z, Brazma A, Rättsch G, Brooks AN, Brazma A, Brooks AN, Göke J, Rättsch G, Schwarz RF, Stegle O, Zhang Z, Aaltonen LA, Abascal F, Abeshouse A, Aburatani H, Adams DJ, Agrawal N, Ahn KS, Ahn S-M, Aikata H, Akbani R, Akdemir KC, Al-Ahmadie H, Al-Sedairy ST, Al-Shahrour F, Alawi M, Albert M, Aldape K, Alexandrov LB, Ally A, Alsop K, Alvarez EG, Amary F, Amin SB, Aminou B, Ammerpohl O, Anderson MJ, Ang Y, Antonello D, Anur P, Aparicio S, Appelbaum EL, Arai Y, Aretz A, Arihiro K, Ariizumi S-i, Armenia J, Arnold L, Asa S, Assenov Y, Atwal G, Aukema S, Auman JT, Aure MRR, Awadalla P, Aymerich M, Bader GD, Baez-Ortega A, Bailey MH, Bailey PJ, Balasundaram M, Balu S, Bandopadhyay P, Banks RE, Barbi S, Barbour AP, Barenboim J, Barnholtz-Sloan J, Barr H, Barrera E, Bartlett J, Bartolome J, Bassi C, Bathe OF, Baumhoer D, Bavi P, Baylin SB, Bazant W, Beardsmore D, Beck TA, Behjati S, Behren A, Niu B, Bell C, Beltran S, Benz C, Berchuck A, Bergmann AK, Bergstrom

- EN, Berman BP, Berney DM, Bernhart SH, Beroukhim R, Berrios M, Bersani S, Bertl J, Betancourt M, Bhandari V, Bhosle SG, Biankin AV, Bieg M, Bigner D, Binder H, Birney E, Birrer M, Biswas NK, Bjerkehaugen B, Bodenheimer T, Boice L, Bonizzato G, De Bono JS, Boot A, Bootwalla MS, Borg A, Borkhardt A, Boroevich KA, Borozan I, Borst C, Bosenberg M, Bosio M, Boultonwood J, Bourque G, Boutros PC, Bova GS, Bowen DT, Bowlby R, Bowtell DDL, Boyault S, Boyce R, Boyd J, Brazma A, Brennan P, Brewer DS, Brinkman AB, Bristow RG, Broaddus RR, Brock JE, Brock M, Broeks A, Brooks AN, Brooks D, Brors B, Brunak S, Bruxner TJC, Bruzos AL, Buchanan A, Buchhalter I, Buchholz C, Bullman S, Burke H, Burkhardt B, Burns KH, Busanovich J, Bustamante CD, Butler AP, Butte AJ, Byrne NJ, Børresen-Dale A-L, Caesar-Johnson SJ, Cafferkey A, Cahill D, Calabrese C, Caldas C, Calvo F, Camacho N, Campbell PJ, Campo E, Cantù C, Cao S, Carey TE, Carlevaro-Fita J, Carlsen R, Cataldo I, Group PTC, Group PTW, Consortium P (2020) Genomic basis for RNA alterations in cancer. *Nature* 578(7793): 129–136
- Cancer Genome Atlas Research N (2012) Comprehensive genomic characterization of squamous cell lung cancers. *Nature* 489(7417): 519–525
- Cao G, Yue J, Ruan Y, Han Y, Zhi Y, Lu J, Liu M, Xu X, Wang J, Gu Q, Wen X, Gao J, Zhang Q, Kang J, Wang C, Li F (2023) Single-cell dissection of cervical cancer reveals key subsets of the tumor immune microenvironment. *EMBO J* 42(16): e110757. <https://doi.org/10.15252/embo.2022110757>
- Cao L, Huang C, Cui Zhou D, Hu Y, Lih TM, Savage SR, Krug K, Clark DJ, Schnaubelt M, Chen L, da Veiga Leprevost F, Eguez RV, Yang W, Pan J, Wen B, Dou Y, Jiang W, Liao Y, Shi Z, Terekhanova NV, Cao S, Lu RJ-H, Li Y, Liu R, Zhu H, Ronning P, Wu Y, Wyczalkowski MA, Easwaran H, Danilova L, Mer AS, Yoo S, Wang JM, Liu W, Haibe-Kains B, Thiagarajan M, Jewell SD, Hostetter G, Newton CJ, Li QK, Roehrl MH, Fenyö D, Wang P, Nesvizhskii AI, Mani DR, Omenn GS, Boja ES, Mesri M, Robles AI, Rodriguez H, Bathe OF, Chan DW, Hruban RH, Ding L, Zhang B, Zhang H, Amin M, An E, Ayad C, Bauer T, Birger C, Birrer MJ, Boca SM, Bocik W, Borucki M, Cai S, Carr SA, Cerda S, Chen H, Chen S, Chesla D, Chinnaiyan AM, Colaprico A, Cottingham S, Derejska M, Dhanasekaran SM, Domagalski MJ, Druker BJ, Duffy E, Dyer MA, Edwards NJ, Ellis MJ, Eschbacher J, Francis A, Francis J, Gabriel S, Gabrovski N, Gardner J, Getz G, Gillette MA, Goldthwaite CA, Grady P, Guo S, Hariharan P, Hiltke T, Hindenach B, Hoadley KA, Huang J, Jones CD, Ketchum KA, Kinsinger CR, Koziak JM, Kusnierz K, Liu T, Long J, Mallery D, Mareedu S, Matteotti R, Maunganidze N, McGarvey PB, Minoo P, Paklina OV, Paulovich AG, Payne SH, Potapova O, Pruetz B, Qi L, Roche N, Rodland KD, Rohrer DC, Schadt EE, Shabunin AV, Shelton T, Shutack Y, Singh S, Smith M, Smith RD, Sokoll LJ, Suh J, Thangudu RR, Tsang SX, Um KS, Valley DR, Vatanian N, Wang W, Wilson GD, Wiznerowicz M, Zhang Z, Zhao G (2021) Proteogenomic characterization of pancreatic ductal adenocarcinoma. *Cell* 184(19): 5031–5052
- Chakraborty S, Hosen MI, Ahmed M, Shekhar HU (2018) Onco-multi-OMICS approach: a new frontier in cancer research. *Biomed Res Int* 2018: 9836256. <https://doi.org/10.1155/2018/9836256>
- Chan JM, Zaidi S, Love JR, Zhao JL, Setty M, Wadosky KM, Gopalan A, Choo Z-N, Persad S, Choi J, LaClair J, Lawrence KE, Chaudhary O, Xu T, Masilionis I, Linkov I, Wang S, Lee C, Barlas A, Morris MJ, Mazutis L, Chaligne R, Chen Y, Goodrich DW, Karthaus WR, Pe'er D, Sawyers CL (2022) Lineage plasticity in prostate cancer depends on JAK/STAT inflammatory signaling. *Science* 377(6611): 1180–1191
- Chen F, Wendl MC, Wyczalkowski MA, Bailey MH, Li Y, Ding L (2021) Moving pan-cancer studies from basic research toward the clinic. *Nat Cancer* 2(9): 879–890
- Chen Y-J, Roumeliotis TI, Chang Y-H, Chen C-T, Han C-L, Lin M-H, Chen H-W, Chang G-C, Chang Y-L, Wu C-T, Lin M-W, Hsieh M-S, Wang Y-T, Chen Y-R, Jonassen I, Ghavidel FZ, Lin Z-S, Lin K-T, Chen C-W, Sheu P-Y, Hung C-T, Huang K-C, Yang H-C, Lin P-Y, Yen T-C, Lin Y-W, Wang J-H, Raghav L, Lin C-Y, Chen Y-S, Wu P-S, Lai C-T, Weng S-H, Su K-Y, Chang W-H, Tsai P-Y, Robles AI, Rodriguez H, Hsiao Y-J, Chang W-H, Sung T-Y, Chen J-S, Yu S-L, Choudhary JS, Chen H-Y, Yang P-C, Chen Y-J (2020) Proteogenomics of non-smoking lung cancer in east asia delineates molecular signatures of pathogenesis and progression. *Cell* 182(1): 226–244
- Christodoulou M-I, Zaravinos A (2023) Single-cell analysis in immuno-oncology. *Int J Mol Sci* 24(9): 8422. <https://doi.org/10.3390/ijms24098422>
- Chung W, Eum HH, Lee H-O, Lee K-M, Lee H-B, Kim K-T, Ryu HS, Kim S, Lee JE, Park YH, Kan Z, Han W, Park W-Y (2017) Single-cell RNA-seq enables comprehensive tumour and immune cell profiling in primary breast cancer. *Nat Commun* 8(1): 15081. <https://doi.org/10.1038/ncomms15081>
- Ciriello G, Cerami E, Sander C, Schultz N (2012) Mutual exclusivity analysis identifies oncogenic network modules. *Genome Res* 22(2): 398–406
- Clark AJ, Lillard JW (2024) A Comprehensive review of bioinformatics tools for genomic biomarker discovery driving precision oncology. *Genes* 15(8): 1036. <https://doi.org/10.3390/genes15081036>
- Clark DJ, Dhanasekaran SM, Petralia F, Pan J, Song X, Hu Y, da Veiga Leprevost F, Reva B, Lih T-SM, Chang H-Y, Ma W, Huang C, Ricketts CJ, Chen L, Krek A, Li Y, Rykunov D, Li QK, Chen LS, Ozbek U, Vasaikar S, Wu Y, Yoo S, Chowdhury S, Wyczalkowski MA, Ji J, Schnaubelt M, Kong A, Sethuraman S, Avtonomov DM, Ao M, Colaprico A, Cao S, Cho K-C, Kalayci S, Ma S, Liu W, Ruggles K, Calinawan A, Gümüş ZH, Geiszler D, Kawaler E, Teo GC, Wen B, Zhang Y, Keegan S, Li K, Chen F, Edwards N, Pierorazio PM, Chen XS, Pavlovich CP, Hakimi AA, Brominski G, Hsieh JJ, Antczak A, Omelchenko T, Lubinski J, Wiznerowicz M, Linehan WM, Kinsinger CR, Thiagarajan M, Boja ES, Mesri M, Hiltke T, Robles AI, Rodriguez H, Qian J, Fenyö D, Zhang B, Ding L, Schadt E, Chinnaiyan AM, Zhang Z, Omenn GS, Cieslik M, Chan DW, Nesvizhskii AI, Wang P, Zhang H, Clinical Proteomic Tumor Analysis C (2019) Integrated proteogenomic characterization of clear cell renal cell carcinoma. *Cell* 179(4): 964–983
- Coarfa C, Grimm SL, Rajapakshe K, Perera D, Lu HY, Wang X, Christensen KR, Mo Q, Edwards DP, Huang S (2021) Reverse-phase protein array: technology, application, data processing, and integration. *J Biomol Tech* 32(1): 15–29
- Consortium TITP-CAoWG (2020) Pan-cancer analysis of whole genomes. *Nature* 578(7793): 82–93
- Darbandi MR, Darbandi M, Darbandi S, Bado I, Hadizadeh M, Khorram Khorshid HR (2024) Artificial intelligence breakthroughs in pioneering early diagnosis and precision treatment of breast cancer: a multimethod study. *Eur J Cancer* 209: 114227. <https://doi.org/10.1016/j.ejca.2024.114227>
- Deng S, Wang C, Wang Y, Xu Y, Li X, Johnson NA, Mukherji A, Lo UG, Xu L, Gonzalez J, Metang LA, Ye J, Tirado CR, Rodarte K, Zhou Y, Xie Z, Arana C, Annamalai V, Liu X, Vander Griend DJ, Strand D, Hsieh J-T, Li B, Raj G, Wang T, Mu P (2022) Ectopic JAK-STAT activation enables the transition to a stem-like and multilineage state conferring AR-targeted therapy resistance. *Nat Cancer* 3(9): 1071–1087
- Doll S, Gnad F, Mann M (2019) The case for proteomics and phospho-proteomics in personalized cancer medicine. *Proteomics Clin Appl* 13(2): 1800113. <https://doi.org/10.1002/prca.201800113>

- Dong L, Lu D, Chen R, Lin Y, Zhu H, Zhang Z, Cai S, Cui P, Song G, Rao D, Yi X, Wu Y, Song N, Liu F, Zou Y, Zhang S, Zhang X, Wang X, Qiu S, Zhou J, Wang S, Zhang X, Shi Y, Figeys D, Ding L, Wang P, Zhang B, Rodriguez H, Gao Q, Gao D, Zhou H, Fan J (2022) Proteogenomic characterization identifies clinically relevant subgroups of intrahepatic cholangiocarcinoma. *Cancer Cell* 40(1): 70–87
- Dwivedi S, Purohit P, Misra R, Lingeswaran M, Vishnoi JR, Pareek P, Misra S, Sharma P (2019) Single cell omics of breast cancer: an update on characterization and diagnosis. *Indian J Clin Biochem* 34(1): 3–18
- Ellis MJ, Gillette M, Carr SA, Paulovich AG, Smith RD, Rodland KK, Townsend RR, Kinsinger C, Mesri M, Rodriguez H, Liebler DC, Consortium obotCPTA (2013) Connecting genomic alterations to cancer biology with proteomics: the NCI Clinical Proteomic Tumor Analysis Consortium. *Cancer Discov* 3(10): 1108–1112
- Gao Q, Zhu H, Dong L, Shi W, Chen R, Song Z, Huang C, Li J, Dong X, Zhou Y, Liu Q, Ma L, Wang X, Zhou J, Liu Y, Boja E, Robles AI, Ma W, Wang P, Li Y, Ding L, Wen B, Zhang B, Rodriguez H, Gao D, Zhou H, Fan J (2019) Integrated proteogenomic characterization of HBV-related hepatocellular carcinoma. *Cell* 179(2): 561–577
- Geffen Y, Anand S, Akiyama Y, Yaron TM, Song Y, Johnson JL, Govindan A, Babur Ö, Li Y, Huntsman E, Wang L-B, Birger C, Heiman DI, Zhang Q, Miller M, Maruvka YE, Haradhvala NJ, Calinawan A, Belkin S, Kerelsky A, Clauser KR, Krug K, Satpathy S, Payne SH, Mani DR, Gillette MA, Dhanasekaran SM, Thiagarajan M, Mesri M, Rodriguez H, Robles AI, Carr SA, Lazar AJ, Aguet F, Cantley LC, Ding L, Getz G, An E, Anurag M, Bavarva J, Birrer MJ, Babur Ö, Cao S, Ceccarelli M, Chan DW, Chinnaiyan AM, Cho H, Chowdhury S, Cieslik MP, Colaprico A, Carr SA, da Veiga Leprevost F, Day C, Domagalski MJ, Dou Y, Druker BJ, Edwards N, Ellis MJ, Fenyö D, Foltz SM, Francis A, Gonzalez Robles TJ, Gosline SJ, Gümüş ZH, Hiltke T, Hong R, Hostetter G, Hu Y, Huang C, Iavarone A, Jaehnig EJ, Jewell SD, Ji J, Jiang W, Katsnelson L, Ketchum KA, Kolodziejczak I, Kumar-Sinha C, Krug K, Lei JT, Liang W-W, Liao Y, Lindgren CM, Liu T, Liu W, Ma W, McKerrow W, Mesri M, Mani DR, Nesvizhskii AI, Newton C, Oldroyd R, Omenn GS, Paulovich AG, Petralia F, Pugliese P, Reva B, Rodland KD, Ruggles KV, Rykunov D, Rodrigues FM, Savage SR, Schadt EE, Schnaubelt M, Schraink T, Shi Z, Smith RD, Song X, Stathias V, Storrs EP, Schürer S, Selvan ME, Tan J, Terekhanova NV, Thangudu RR, Tignor N, Thiagarajan M, Wang JM, Wang P, Wang Y, Wen B, Wiznerowicz M, Wu Y, Wyczalkowski MA, Yao L, Yi X, Yao L, Zhang B, Zhang H, Zhang X, Zhang Z, Zhou DC (2023) Pan-cancer analysis of post-translational modifications reveals shared patterns of protein regulation. *Cell* 186(18): 3945–3967
- Gerlinger M, Rowan AJ, Horswell S, Larkin J, Endesfelder D, Gronroos E, Martinez P, Matthews N, Stewart A, Tarpey P, Varela I, Phillimore B, Begum S, McDonald NQ, Butler A, Jones D, Raine K, Latimer C, Santos CR, Nohadani M, Eklund AC, Spencer-Dene B, Clark G, Pickering L, Stamp G, Gore M, Szallasi Z, Downward J, Futreal PA, Swanton C (2012) Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N Engl J Med* 366(10): 883–892
- Gessain G, Anzali A-A, Lerousseau M, Mulder K, Bied M, Auperin A, Stockholm D, Signolle N, Sassi F, Marques Da Costa ME, Marchais A, Sayadi A, Weidner D, Uderhardt S, Blampey Q, Nakkireddy SR, Broutin S, Dutertre C-A, Busson P, Walter T, Marhic A, Moya-Plana A, Guerlain J, Breuskin I, Casiraghi O, Gorphe P, Classe M, Scoazec J-Y, Bleriot C, Ginhoux F (2024) Trem2-expressing multinucleated giant macrophages are a biomarker of good prognosis in head and neck squamous cell carcinoma. *Cancer Discov* 14(12): 2352–2366
- Gillette MA, Satpathy S, Cao S, Dhanasekaran SM, Vasaikar SV, Krug K, Petralia F, Li Y, Liang W-W, Reva B, Krek A, Ji J, Song X, Liu W, Hong R, Yao L, Blumenberg L, Savage SR, Wendl MC, Wen B, Li K, Tang LC, MacMullan MA, Avanesian SC, Kane MH, Newton CJ, Cornwell M, Kothadia RB, Ma W, Yoo S, Mannan R, Vats P, Kumar-Sinha C, Kawaler EA, Omelchenko T, Colaprico A, Geffen Y, Maruvka YE, da Veiga Leprevost F, Wiznerowicz M, Gümüş ZH, Veluswamy RR, Hostetter G, Heiman DI, Wyczalkowski MA, Hiltke T, Mesri M, Kinsinger CR, Boja ES, Omenn GS, Chinnaiyan AM, Rodriguez H, Li QK, Jewell SD, Thiagarajan M, Getz G, Zhang B, Fenyö D, Ruggles KV, Cieslik MP, Robles AI, Clauser KR, Govindan R, Wang P, Nesvizhskii AI, Ding L, Mani DR, Carr SA, Clinical Proteomic Tumor Analysis C (2020) Proteogenomic characterization reveals therapeutic vulnerabilities in lung adenocarcinoma. *Cell* 182(1): 200–225
- Gonçalves E, Fragoulis A, Garcia-Alonso L, Cramer T, Saez-Rodriguez J, Beltrao P (2017) Widespread post-transcriptional attenuation of genomic copy-number variation in cancer. *Cell Syst* 5(4): 386–398
- Gong TQ, Jiang YZ, Shao C, Peng WT, Liu MW, Li DQ, Zhang BY, Du P, Huang Y, Li FF, Li MY, Han ZL, Jin X, Ma D, Xiao Y, Yang PY, Qin J, Shao ZM, Zhu W (2022) Proteome-centric cross-omics characterization and integrated network analyses of triple-negative breast cancer. *Cell Rep* 38(9): 110460. <https://doi.org/10.1016/j.celrep.2022.110460>
- Gonzalez-Fierro A, Dueñas-González A (2021) Drug repurposing for cancer therapy, easier said than done. *Semin Cancer Biol* 68: 123–131
- Guo W, Liu Y, Han Y, Tang H, Fan X, Wang C, Chen PR (2024) Amplifiable protein identification via residue-resolved barcoding and composition code counting. *Natl Sci Rev* 11(7): nwae183. <https://doi.org/10.1093/nsr/nwae183>
- Han M, Li F, Zhang Y, Dai P, He J, Li Y, Zhu Y, Zheng J, Huang H, Bai F (2022) FOXA2 drives lineage plasticity and KIT pathway activation in neuroendocrine prostate cancer. *Cancer Cell* 40(11): 1306–1323
- Hasin Y, Seldin M, Lusic A (2017) Multi-omics approaches to disease. *Genome Biol* 18(1): 83. <https://doi.org/10.1186/s13059-017-1215-1>
- Herbst SA, Vesterlund M, Helmboldt AJ, Jafari R, Siavelis I, Stahl M, Schitter EC, Liebers N, Brinkmann BJ, Czernilofsky F, Roeder T, Bruch P-M, Iskar M, Kittai A, Huang Y, Lu J, Richter S, Mermelekas G, Umer HM, Knoll M, Kolb C, Lenze A, Cao X, Österholm C, Wahnschaffe L, Herling C, Scheinost S, Ganzinger M, Mansouri L, Kriegsmann K, Kriegsmann M, Anders S, Zapatka M, Del Poeta G, Zucchetto A, Bomben R, Gattei V, Dreger P, Woyach J, Herling M, Müller-Tidow C, Rosenquist R, Stilgenbauer S, Zenz T, Huber W, Tausch E, Lehtiö J, Dietrich S (2022) Proteogenomics refines the molecular classification of chronic lymphocytic leukemia. *Nat Commun* 13(1): 6226. <https://doi.org/10.1038/s41467-022-33385-8>
- Huang C, Chen L, Savage SR, Eguez RV, Dou Y, Li Y, da Veiga Leprevost F, Jaehnig EJ, Lei JT, Wen B, Schnaubelt M, Krug K, Song X, Cieslik M, Chang H-Y, Wyczalkowski MA, Li K, Colaprico A, Li QK, Clark DJ, Hu Y, Cao L, Pan J, Wang Y, Cho K-C, Shi Z, Liao Y, Jiang W, Anurag M, Ji J, Yoo S, Zhou DC, Liang W-W, Wendl M, Vats P, Carr SA, Mani DR, Zhang Z, Qian J, Chen XS, Pico AR, Wang P, Chinnaiyan AM, Ketchum KA, Kinsinger CR, Robles AI, An E, Hiltke T, Mesri M, Thiagarajan M, Weaver AM, Sikora AG, Lubiński J, Wierzbicka M, Wiznerowicz M, Satpathy S, Gillette MA, Miles G, Ellis MJ, Omenn GS, Rodriguez H, Boja ES, Dhanasekaran SM, Ding L, Nesvizhskii AI, El-Naggar AK, Chan DW, Zhang H, Zhang B, Agarwal A, Anderson ML, Avanesian SC, Avtonomov D, Bathe OF, Birger C, Birrer MJ, Blumenberg L, Bocik WE, Borate U, Borucki M, Burke MC, Cai S,

- Calinawan AP, Cerda S, Charamut A, Chen LS, Chowdhury S, Clauser KR, Culpepper H, Czernicki T, D'Angelo F, Day J, De Young S, Demir E, Ding F, Domagalski MJ, Dort JC, Druker B, Duffy E, Dyer M, Edwards NJ, Elburn K, Ermakova TS, Fenyó D, Ferrarotto R, Francis A, Gabriel S, Garofano L, Geffen Y, Getz G, Goldthwaite CA, Hannick LI, Hariharan P, Hayes DN, Heiman D, Hindenach B, Hoadley KA, Hostetter G, Hyrcza M, Jewell SD, Jones CD, Kane MH, Karz A, Kothadia RB, Krek A, Kumar-Sinha C, Liu T, Liu H, Ma W, Malc E, Malovannaya A, Mareedu S, Markey SP, Marrero-Oliveras A, Maunganidze N, McDermott JE, McGarvey PB, McGee J, Mieczkowski P, Migliozzi S, Montgomery R, Newton CJ, Ozbek U, Paulovich AG, Payne SH, Pazardzhikiev DD, Perou AM, Petralia F, Petrenko L, Piehowski PD, Placantonakis D, Polonskaya L, Ponomareva EV, Potapova O, Qi L, Qu N, Ramkissoon S, Reva B, Richey S, Robinson K, Roche N, Rodland K, Rohrer DC, Rykunov D, Schadt EE, Shi Y, Shutack Y, Singh S, Skelly T, Smith R, Sokoll LJ, Stawicki J, Stein SE, Suh J, Szopa W, Tabor D, Tan D, Tansil D, Teo GC, Thangudu RR, Tognon C, Traer E, Tsang S, Tyner J, Um KS, Valley DR, Vasilev LV, Vatanian N, Velvulou U, Vernon M, Westbrook TF, Whiteaker JR, Wu Y, Xu M, Yao L, Yi X, Yu F, Zaalishvili K, Zakhartsev Y, Zelt R, Zhao G, Zhu J (2021) Proteogenomic insights into the biology and treatment of HPV-negative head and neck squamous cell carcinoma. *Cancer Cell* 39(3): 361–379
- Huang KL, Li S, Mertins P, Cao S, Gunawardena HP, Ruggles KV, Mani DR, Clauser KR, Tanioka M, Usary J, Kavuri SM, Xie L, Yoon C, Qiao JW, Wrobel J, Wyczalkowski MA, Erdmann-Gilmore P, Snider JE, Hoog J, Singh P, Niu B, Guo Z, Sun SQ, Sanati S, Kawaler E, Wang X, Scott A, Ye K, McLellan MD, Wendl MC, Malovannaya A, Held JM, Gillette MA, Fenyó D, Kinsinger CR, Mesri M, Rodriguez H, Davies SR, Perou CM, Ma C, Reid Townsend R, Chen X, Carr SA, Ellis MJ, Ding L (2017) Proteogenomic integration reveals therapeutic targets in breast cancer xenografts. *Nat Commun* 8: 14864. <https://doi.org/10.1038/ncomms14864>
- Jaffe JD, Berg HC, Church GM (2004) Proteogenomic mapping as a complementary method to perform genome annotation. *Proteomics* 4(1): 59–77
- Jayavelu AK, Wolf S, Buettner F, Alexe G, Häupl B, Comoglio F, Schneider C, Doebele C, Fuhrmann DC, Wagner S, Donato E, Andresen C, Wilke AC, Zindel A, Jahn D, Spletstoesser B, Plessmann U, Münch S, Abou-El-Ardat K, Makowka P, Acker F, Enssle JC, Cremer A, Schnütgen F, Kurrle N, Chapuy B, Löber J, Hartmann S, Wild PJ, Wittig I, Hübschmann D, Kaderali L, Cox J, Brüne B, Röhlig C, Steffen B, Bornhäuser M, Trumpp A, Urlaub H, Stegmaier K, Serve H, Mann M, Oellerich T (2022) The proteogenomic subtypes of acute myeloid leukemia. *Cancer Cell* 40(3): 301–317
- Jiang YZ, Ma D, Suo C, Shi J, Xue M, Hu X, Xiao Y, Yu KD, Liu YR, Yu Y, Zheng Y, Li X, Zhang C, Hu P, Zhang J, Hua Q, Zhang J, Hou W, Ren L, Bao D, Li B, Yang J, Yao L, Zuo WJ, Zhao S, Gong Y, Ren YX, Zhao YX, Yang YS, Niu Z, Cao ZG, Stover DG, Verschraegen C, Kaklamani V, Daemen A, Benson JR, Takabe K, Bai F, Li DQ, Wang P, Shi L, Huang W, Shao ZM (2019) Genomic and transcriptomic landscape of triple-negative breast cancers: subtypes and treatment strategies. *Cancer Cell* 35(3): 428–440
- Kamps R, Brandão RD, Bosch BJvd, Paulussen ADC, Xanthoulea S, Blok MJ, Romano A (2017) Next-generation sequencing in oncology: genetic diagnosis, risk prediction and cancer classification. *Int J Mol Sci* 18(2): 308. <https://doi.org/10.3390/ijms18020308>
- Kim K-H, Migliozzi S, Koo H, Hong J-H, Park SM, Kim S, Kwon HJ, Ha S, Garofano L, Oh YT, D'Angelo F, Kim CI, Kim S, Lee JY, Kim J, Hong J, Jang E-H, Mathon B, Di Stefano A-L, Bielle F, Laurence A, Nesvizhskii AI, Hur E-M, Yin J, Shi B, Kim Y, Moon K-S, Kwon JT, Lee SH, Lee SH, Gwak HS, Lasorella A, Yoo H, Sanson M, Sa JK, Park C-K, Nam D-H, Iavarone A, Park JB (2024) Integrated proteogenomic characterization of glioblastoma evolution. *Cancer Cell* 42(3): 358–377
- Klomp JA, Klomp JE, Stalneckner CA, Bryant KL, Edwards AC, Drizyte-Miller K, Hibshman PS, Diehl JN, Lee YS, Morales AJ, Taylor KE, Peng S, Tran NL, Herring LE, Prevatte AW, Barker NK, Hover LD, Hallin J, Chowdhury S, Coker O, Lee HM, Goodwin CM, Gautam P, Olson P, Christensen JG, Shen JP, Kopetz S, Graves LM, Lim K-H, Wang-Gillam A, Wennerberg K, Cox AD, Der CJ (2024) Defining the KRAS- and ERK-dependent transcriptome in KRAS-mutant cancers. *Science* 384(6700): eadk0775. <https://doi.org/10.1126/science.adk0775>
- Kramer MH, Zhang Q, Sprung R, Day RB, Erdmann-Gilmore P, Li Y, Xu Z, Helton NM, George DR, Mi Y, Westervelt P, Payton JE, Ramakrishnan SM, Miller CA, Link DC, DiPersio JF, Walter MJ, Townsend RR, Ley TJ (2022) Proteomic and phosphoproteomic landscapes of acute myeloid leukemia. *Blood* 140(13): 1533–1548
- Krug K, Jaehnig EJ, Satpathy S, Blumenberg L, Karpova A, Anurag M, Miles G, Mertins P, Geffen Y, Tang LC, Heiman DI, Cao S, Maruvka YE, Lei JT, Huang C, Kothadia RB, Colaprico A, Birger C, Wang J, Dou Y, Wen B, Shi Z, Liao Y, Wiznerowicz M, Wyczalkowski MA, Chen XS, Kennedy JJ, Paulovich AG, Thiagarajan M, Kinsinger CR, Hiltke T, Boja ES, Mesri M, Robles AI, Rodriguez H, Westbrook TF, Ding L, Getz G, Clauser KR, Fenyó D, Ruggles KV, Zhang B, Mani DR, Carr SA, Ellis MJ, Gillette MA, Clinical Proteomic Tumor Analysis C (2020) Proteogenomic landscape of breast cancer tumorigenesis and targeted therapy. *Cell* 183(5): 1436–1456
- Laurent-Puig P, Legoix P, Bluteau O, Belghiti J, Franco D, Binot F, Monges G, Thomas G, Bioulac-Sage P, Zucman-Rossi J (2001) Genetic alterations associated with hepatocellular carcinomas define distinct pathways of hepatocarcinogenesis. *Gastroenterology* 120(7): 1763–1773
- Lee JS, Nair NU, Dinstag G, Chapman L, Chung Y, Wang K, Sinha S, Cha H, Kim D, Schperberg AV, Srinivasan A, Lazar V, Rubin E, Hwang S, Berger R, Beker T, Ronai Z, Hannenhalli S, Gilbert MR, Kurzrock R, Lee SH, Aldape K, Ruppin E (2021) Synthetic lethality-mediated precision oncology via the tumor transcriptome. *Cell* 184(9): 2487–2502
- Lehtio J, Arslan T, Siavelis I, Pan Y, Soccirelli F, Berkovska O, Umer HM, Mermelekas G, Pirmoradian M, Jonsson M, Brunstrom H, Brustugun OT, Purohit KP, Cunningham R, Asl HF, Isaksson S, Arbajian E, Aine M, Karlsson A, Kotevska M, Hansen CG, Haakensen VD, Helland A, Tamborero D, Johansson HJ, Branca RM, Planck M, Staaf J, Orre LM (2021) Proteogenomics of non-small cell lung cancer reveals molecular subtypes associated with specific therapeutic targets and immune evasion mechanisms. *Nat Cancer* 2(11): 1224–1242
- Li GX, Chen H, Hsiao Y, Mannan R, Zhang Y, Luo J, Petralia F, Cho H, Hosseini N, Leprevost FdV, Calinawan A, Li Y, Anand S, Dagar A, Geffen Y, Kumar-Sinha C, Chugh S, Le A, Ponce S, Guo S, Zhang C, Schnaubelt M, Al Deen NN, Chen F, Caravan W, Houston A, Hopkins A, Newton CJ, Wang X, Polasky DA, Haynes S, Yu F, Jing X, Chen S, Robles AI, Mesri M, Thiagarajan M, An E, Getz GA, Linehan WM, Hostetter G, Jewell SD, Chan DW, Wang P, Omenn GS, Mehra R, Ricketts CJ, Ding L, Chinnaiyan AM, Cieslik MP, Dhanasekaran SM, Zhang H, Nesvizhskii AI, Lazar AJ, Paulovich AG, Antczak A, Green A, Ma'ayan A, Pruetz B, Zhang B, Reva B, Druker BJ, Goldthwaite CA, Birger C, Mani DR, Chesla D, Fenyó D, Schadt EE, Wilson G, Kołodziejczak I, John I, Hafron J, Vo J, Zaalishvili K, Ketchum KA, Rodland KD, Nyce K, Wiznerowicz M, Domagalski MJ, Anurag M, Borucki M, Gillette MA, Birrer MJ,

- Edwards NJ, Vatanian N, VanderKolk P, McGarvey PB, Dhir R, Thangudu RR, Crispen R, Smith RD, Payne SH, Cottingham S, Cai S, Carr SA, Liu T, Le T, Ma W, Zhang X, Lu Y, Shutack Y, Zhang Z (2024) Comprehensive proteogenomic characterization of rare kidney tumors. *Cell Rep Med* 5(5): 101547. <https://doi.org/10.1016/j.xcrim.2024.101547>
- Li L, Yan S, Lin B, Shi Q, Lu Y (2018) Chapter eight — Single-cell proteomics for cancer immunotherapy. In: Broome A-M (Eds). *Cancer nanotechnology*. Academic Press. pp 185-207
- Li Y, Dou Y, Da Veiga Leprevost F, Geffen Y, Calinawan AP, Aguet F, Akiyama Y, Anand S, Birger C, Cao S, Chaudhary R, Chilappagari P, Cieslik M, Colaprico A, Zhou DC, Day C, Domagalski MJ, Esai Selvan M, Fenyö D, Foltz SM, Francis A, Gonzalez-Robles T, Gümüş ZH, Heiman D, Holck M, Hong R, Hu Y, Jaehnig EJ, Ji J, Jiang W, Katsnelson L, Ketchum KA, Klein RJ, Lei JT, Liang W-W, Liao Y, Lindgren CM, Ma W, Ma L, MacCoss MJ, Martins Rodrigues F, McKerrow W, Nguyen N, Oldroyd R, Pillozzi A, Pugliese P, Reva B, Rudnick P, Ruggles KV, Rykunov D, Savage SR, Schnaubelt M, Schraink T, Shi Z, Singhal D, Song X, Storrs E, Terekhanova NV, Thangudu RR, Thiagarajan M, Wang L-B, Wang JM, Wang Y, Wen B, Wu Y, Wyczalkowski MA, Xin Y, Yao L, Yi X, Zhang H, Zhang Q, Zuhl M, Getz G, Ding L, Nesvizhskii AI, Wang P, Robles AI, Zhang B, Payne SH, Lazar AJ, Paulovich AG, Colaprico A, Iavarone A, Chinnaiyan AM, Druker BJ, Kumar-Sinha C, Newton CJ, Huang C, Mani DR, Smith RD, Huntsman E, Schadt EE, An E, Petralia F, Hostetter G, Omenn GS, Cho H, Rodriguez H, Zhang H, Kolodziejczak I, Johnson JL, Bavarva J, Tan J, Rodland KD, Clauser KR, Krug K, Cantley LC, Wiznerowicz M, Ellis MJ, Anurag M, Mesri M, Gillette MA, Birrer MJ, Ceccarelli M, Dhanasekaran SM, Edwards N, Tignor N, Babur Ö, Pugliese P, Gosline SJ, Jewell SD, Satpathy S, Chowdhury S, Schürer S, Carr SA, Liu T, Hiltke T, Yaron TM, Stathias V, Liu W, Zhang X, Song Y, Zhang Z, Chan DW (2023a) Proteogenomic data and resources for pan-cancer analysis. *Cancer Cell* 41(8): 1397–1406
- Li Y, Lih T-SM, Dhanasekaran SM, Mannan R, Chen L, Cieslik M, Wu Y, Lu RJ-H, Clark DJ, Kolodziejczak I, Hong R, Chen S, Zhao Y, Chugh S, Caravan W, Naser Al Deen N, Hosseini N, Newton CJ, Krug K, Xu Y, Cho K-C, Hu Y, Zhang Y, Kumar-Sinha C, Ma W, Calinawan A, Wyczalkowski MA, Wendl MC, Wang Y, Guo S, Zhang C, Le A, Dagar A, Hopkins A, Cho H, Leprevost FdV, Jing X, Teo GC, Liu W, Reimers MA, Pachynski R, Lazar AJ, Chinnaiyan AM, Van Tine BA, Zhang B, Rodland KD, Getz G, Mani DR, Wang P, Chen F, Hostetter G, Thiagarajan M, Linehan WM, Fenyö D, Jewell SD, Omenn GS, Mehra R, Wiznerowicz M, Robles AI, Mesri M, Hiltke T, An E, Rodriguez H, Chan DW, Ricketts CJ, Nesvizhskii AI, Zhang H, Ding L, Francis A, Paulovich AG, Antczak A, Green A, Colaprico A, Hakimi A, Pruetz B, Hindenach B, Yadav BK, Reva B, Fevrier-Sullivan B, Druker BJ, Szczylik C, Goldthwaite CA, Birger C, Jones CD, Rohrer DC, Tansil D, Chesla D, Heiman D, Duffy E, Schadt EE, Petralia F, Bromiński G, Quiroga-Garza GM, Wilson GD, Li GX, Zhao G, Hsiao Y, Hsieh J, Lubiński J, Bavarva J, Huang J, Hafron J, Eschbacher J, Hon J, Francis J, Freymann J, Vo J, Wang J, Kirby J, Zaalishvili K, Ketchum KA, Hoadley KA, Um KS, Qi L, Domagalski MJ, Tobin M, Dyer M, Anurag M, Borucki M, Gillette MA, Birrer MJ, Ittmann MM, Roehrl MH, Schnaubelt M, Smith M, Fam M, Roche N, Vatanian N, Maunganidze N, Potapova O, Paklina OV, VanderKolk P, Castro P, Kurzawa P, Hariharan P, Li Q, Li QK, Dhir R, Thangudu RR, Montgomery R, Smith RD, Mareedu S, Payne SH, Cerda S, Cottingham S, Haynes S, Satpathy S, Richey S, Singh S, Tsang SX, Cai S, Cao S, Gabriel S, Carr SA, Liu T, Bauer T, Le T, Chen XS, Zhang X, Shutack Y, Zhang Z (2023b) Histopathologic and proteogenomic heterogeneity reveals features of clear cell renal cell carcinoma aggressiveness. *Cancer Cell* 41(1): 139–163
- Li Y, Porta-Pardo E, Tokheim C, Bailey MH, Yaron TM, Stathias V, Geffen Y, Imbach KJ, Cao S, Anand S, Akiyama Y, Liu W, Wyczalkowski MA, Song Y, Storrs EP, Wendl MC, Zhang W, Sibai M, Ruiz-Serra V, Liang W-W, Terekhanova NV, Rodrigues FM, Clauser KR, Heiman DI, Zhang Q, Aguet F, Calinawan AP, Dhanasekaran SM, Birger C, Satpathy S, Zhou DC, Wang L-B, Baral J, Johnson JL, Huntsman EM, Pugliese P, Colaprico A, Iavarone A, Chheda MG, Ricketts CJ, Fenyö D, Payne SH, Rodriguez H, Robles AI, Gillette MA, Kumar-Sinha C, Lazar AJ, Cantley LC, Getz G, Ding L, An E, Anurag M, Bavarva J, Birrer MJ, Calinawan A, Ceccarelli M, Chan DW, Chinnaiyan AM, Cho H, Chowdhury S, Cieslik MP, da Veiga Leprevost F, Day C, Domagalski MJ, Dou Y, Druker BJ, Edwards N, Ellis MJ, Selvan ME, Foltz SM, Francis A, Gonzalez Robles TJ, Gosline SJ, Gümüş ZH, Hiltke T, Hong R, Hostetter G, Hu Y, Huang C, Huntsman E, Jaehnig EJ, Jewel SD, Ji J, Jiang W, Katsnelson L, Ketchum KA, Kolodziejczak I, Lei JT, Liao Y, Lindgren CM, Liu T, Ma W, McKerrow W, Nesvizhskii AI, Newton C, Oldroyd R, Omenn GS, Paulovich AG, Petralia F, Reva B, Rodland KD, Ruggles KV, Rykunov D, Savage SR, Schadt EE, Schnaubelt M, Schraink T, Shi Z, Smith RD, Song X, Tan J, Thangudu RR, Tignor N, Wang JM, Wang P, Wang Y, Wen B, Wiznerowicz M, Yi X, Zhang B, Zhang H, Zhang X, Zhang Z (2023c) Pan-cancer proteogenomics connects oncogenic drivers to functional states. *Cell* 186(18): 3921–3944
- Liang W-W, Lu RJ-H, Jayasinghe RG, Foltz SM, Porta-Pardo E, Geffen Y, Wendl MC, Lazcano R, Kolodziejczak I, Song Y, Govindan A, Demicco EG, Li X, Li Y, Sethuraman S, Payne SH, Fenyö D, Rodriguez H, Wiznerowicz M, Shen H, Mani DR, Rodland KD, Lazar AJ, Robles AI, Ding L, Aguet F, Akiyama Y, An E, Anand S, Anurag M, Babur O, Bavarva J, Birger C, Birrer M, Calinawan A, Cantley LC, Cao S, Carr S, Ceccarelli M, Chan D, Chinnaiyan A, Cho H, Chowdhury S, Cieslik M, Clauser K, Colaprico A, Zhou DC, da Veiga Leprevost F, Day C, Dhanasekaran M, Domagalski M, Dou Y, Druker B, Edwards N, Ellis M, Selvan ME, Francis A, Getz G, Gillette MA, Robles TG, Gosline S, Gümüş Z, Heiman D, Hiltke T, Hong R, Hostetter G, Hu Y, Huang C, Huntsman E, Iavarone A, Jaehnig E, Jewel S, Ji J, Jiang W, Lee Johnson J, Katsnelson L, Ketchum K, Krug K, Kumar-Sinha C, Lei J, Liao Y, Lindgren C, Liu T, Liu W, Ma W, Rodrigues FM, McKerrow W, Mesri M, Nesvizhskii AI, Newton C, Oldroyd R, Omenn G, Paulovich A, Petralia F, Pugliese P, Reva B, Ruggles K, Rykunov D, Satpathy S, Savage S, Schadt E, Schnaubelt M, Schraink T, Shi Z, Smith D, Song X, Stathias V, Storrs E, Tan J, Terekhanova N, Thangudu R, Thiagarajan M, Tignor N, Wang J, Wang L-B, Wang P, Wang Y, Wen B, Wu Y, Yao L, Yaron TM, Yi X, Zhang B, Zhang H, Zhang Q, Zhang X, Zhang Z, Chan DW, Dhanasekaran SM, Schürer S, Smith RD, Wyczalkowski MA (2023) Integrative multi-omic cancer profiling reveals DNA methylation patterns associated with therapeutic vulnerability and cell-of-origin. *Cancer Cell* 41(9): 1567–1585
- Liu F, Wang P, Sun W, Jiang Y, Gong Q (2022a) Identification of ligand-receptor pairs associated with tumour characteristics in clear cell renal cell carcinoma. *Front Immun* 13: 874056. <https://doi.org/10.3389/fimmu.2022.874056>
- Liu Q, Zhang J, Guo C, Wang M, Wang C, Yan Y, Sun L, Wang D, Zhang L, Yu H, Hou L, Wu C, Zhu Y, Jiang G, Zhu H, Zhou Y, Fang S, Zhang T, Hu L, Li J, Liu Y, Zhang H, Zhang B, Ding L, Robles AI, Rodriguez H, Gao D, Ji H, Zhou H, Zhang P (2024) Proteogenomic characterization of small cell lung cancer identifies biological insights and subtype-specific therapeutic strategies. *Cell* 187(1): 184–203
- Liu X, Xu X, Wu Z, Shan Q, Wang Z, Wu Z, Ding X, Huang W, Wang Z (2022b) Integrated single-cell RNA-seq analysis identifies

- immune heterogeneity associated with KRAS/TP53 mutation status and tumor-sideness in colorectal cancers. *Front Immunol* 13: 961350. <https://doi.org/10.3389/fimmu.2022.961350>
- Ludwig C, Gillet L, Rosenberger G, Amon S, Collins BC, Aebersold R (2018) Data-independent acquisition-based SWATH-MS for quantitative proteomics: a tutorial. *Mol Syst Biol* 14(8): e8126. <https://doi.org/10.15252/msb.20178126>
- Luecken MD, Theis FJ (2019) Current best practices in single-cell RNA-seq analysis: a tutorial. *Mol Syst Biol* 15(6): e8746. <https://doi.org/10.15252/msb.20188746>
- Luo Y, Wang Y, Tian Y, Zhou H, Wen L (2023) “Two Birds One Stone” strategy for the site-specific analysis of core fucosylation and O-GlcNAcylation. *J Am Chem Soc* 145(29): 15879–15887
- Luo Y, Xia Y, Liu D, Li X, Li H, Liu J, Zhou D, Dong Y, Li X, Qian Y, Xu C, Tao K, Li G, Pan W, Zhong Q, Liu X, Xu S, Wang Z, Liu R, Zhang W, Shan W, Fang T, Wang S, Peng Z, Jin P, Jin N, Shi S, Chen Y, Wang M, Jiao X, Luo M, Gong W, Wang Y, Yao Y, Zhao Y, Huang X, Ji X, He Z, Zhao G, Liu R, Wu M, Chen G, Hong L, Ma D, Fang Y, Liang H, Gao Q (2024) Neoadjuvant PARPi or chemotherapy in ovarian cancer informs targeting effector Treg cells for homologous-recombination-deficient tumors. *Cell* 187(18): 4905–4925
- Macklin A, Khan S, Kislinger T (2020) Recent advances in mass spectrometry based clinical proteomics: applications to cancer research. *Clin Proteomics* 17(1): 17. <https://doi.org/10.1186/s12014-020-09283-w>
- Martínez-Jiménez F, Muiños F, Sentís I, Deu-Pons J, Reyes-Salazar I, Arnedo-Pac C, Mularoni L, Pich O, Bonet J, Kranas H, Gonzalez-Perez A, Lopez-Bigas N (2020) A compendium of mutational cancer driver genes. *Nat Rev Cancer* 20(10): 555–572
- McDermott JE, Arshad OA, Petyuk VA, Fu Y, Gritsenko MA, Clauss TR, Moore RJ, Schepmoes AA, Zhao R, Monroe ME, Schnaubelt M, Tsai C-F, Payne SH, Huang C, Wang L-B, Foltz S, Wyczalkowski M, Wu Y, Song E, Brewer MA, Thiagarajan M, Kinsinger CR, Robles AI, Boja ES, Rodriguez H, Chan DW, Zhang B, Zhang Z, Ding L, Smith RD, Liu T, Rodland KD (2020) Proteogenomic characterization of ovarian HGSC implicates mitotic kinases, replication stress in observed chromosomal instability. *Cell Rep Med* 1(1): 100004. <https://doi.org/10.1016/j.xcrm.2020.100004>
- Meissner F, Geddes-McAlister J, Mann M, Bantscheff M (2022) The emerging role of mass spectrometry-based proteomics in drug discovery. *Nat Rev Drug Discov* 21(9): 637–654
- Menyhart O, Gyorffy B (2021) Multi-omics approaches in cancer research with applications in tumor subtyping, prognosis, and diagnosis. *Comput Struct Biotechnol J* 19: 949–960
- Mischnik M, Sacco F, Cox J, Schneider HC, Schäfer M, Hendlich M, Crowther D, Mann M, Klabunde T (2016) IKAP: a heuristic framework for inference of kinase activities from Phosphoproteomics data. *Bioinformatics* 32(3): 424–431
- Monti S, Tamayo P, Mesirov J, Golub T (2003) Consensus clustering: a resampling-based method for class discovery and visualization of gene expression microarray data. *Mach Learn* 52(1): 91–118
- Motone K, Kontogiorgos-Heintz D, Wee J, Kurihara K, Yang S, Roote G, Fox OE, Fang Y, Queen M, Tolhurst M, Cardozo N, Jain M, Nivala J (2024) Multi-pass, single-molecule nanopore reading of long protein strands. *Nature* 633(8030): 662–669
- Narimatsu H, Kaji H, Vakhrushev SY, Clausen H, Zhang H, Noro E, Togayachi A, Nagai-Okatani C, Kuno A, Zou X, Cheng L, Tao S-C, Sun Y (2018) Current technologies for complex glycoproteomics and their applications to biology/disease-driven glycoproteomics. *J Proteome Res* 17(12): 4097–4112
- Neagu A-N, Whitham D, Bruno P, Morrissiey H, Darie CA, Darie CC (2023) Omics-based investigations of breast cancer. *Molecules* 28(12): 4768. <https://doi.org/10.3390/molecules28124768>
- Ng CKY, Dazert E, Boldanova T, Coto-Llerena M, Nuciforo S, Ercan C, Suslov A, Meier M-A, Bock T, Schmidt A, Ketterer S, Wang X, Wieland S, Matter MS, Colombi M, Piscuoglio S, Terracciano LM, Hall MN, Heim MH (2022) Integrative proteogenomic characterization of hepatocellular carcinoma across etiologies and stages. *Nat Commun* 13(1): 2436. <https://doi.org/10.1038/s41467-022-29960-8>
- Nguyen QH, Nguyen H, Nguyen T, Le DH (2020) Multi-omics analysis detects novel prognostic subgroups of breast cancer. *Front Genet* 11: 574661. <https://doi.org/10.3389/fgene.2020.574661>
- Olivier M, Asmis R, Hawkins GA, Howard TD, Cox LA (2019) The need for multi-omics biomarker signatures in precision medicine. *Int J Mol Sci* 20(19): 4781. <https://doi.org/10.3390/ijms20194781>
- Ostroverkhova D, Przytycka TM, Panchenko AR (2023) Cancer driver mutations: predictions and reality. *Trends Mol Med* 29(7): 554–566
- Pan S, Chen R (2022) Pathological implication of protein post-translational modifications in cancer. *Mol Aspects Med* 86: 101097. <https://doi.org/10.1016/j.mam.2022.101097>
- Petersen B-S, Fredrich B, Hoepfner MP, Ellinghaus D, Franke A (2017) Opportunities and challenges of whole-genome and -exome sequencing. *BMC Genet* 18(1): 14. <https://doi.org/10.1186/s12863-017-0479-5>
- Petralia F, Krek A, Calinawan AP, Charytonowicz D, Sebra R, Feng S, Gosline S, Pugliese P, Paulovich AG, Kennedy JJ, Ceccarelli M, Wang P (2023) BayesDeBulk: a flexible bayesian algorithm for the deconvolution of bulk tumor data. bioRxiv. <https://doi.org/10.1101/2021.06.25.449763>
- Petralia F, Ma W, Yaron TM, Caruso FP, Tignor N, Wang JM, Charytonowicz D, Johnson JL, Huntsman EM, Marino GB (2024) Pan-cancer proteogenomics characterization of tumor immunity. *Cell* 187(5): 1255–1277
- Piersma SR, Valles-Martí A, Rolfs F, Pham TV, Henneman AA, Jiménez CR (2024) Inferring kinase activity from phosphoproteomic data: tool comparison and recent applications. *Mass Spectrom Rev* 43(4): 725–751
- Pinho SS, Reis CA (2015) Glycosylation in cancer: mechanisms and clinical implications. *Nat Rev Cancer* 15(9): 540–555
- Qu Y, Feng J, Wu X, Bai L, Xu W, Zhu L, Liu Y, Xu F, Zhang X, Yang G, Lv J, Chen X, Shi G-H, Wang H-K, Cao D-L, Xiang H, Li L, Tan S, Gan H-L, Sun M-H, Qiu J, Zhang H, Zhao J-Y, Ye D, Ding C (2022) A proteogenomic analysis of clear cell renal cell carcinoma in a Chinese population. *Nat Commun* 13(1): 2052. <https://doi.org/10.1038/s41467-022-29577-x>
- Ramberger E, Sapozhnikova V, Ng YLD, Dolnik A, Ziehm M, Popp O, Sträng E, Kull M, Grünschlager F, Krüger J, Benary M, Müller S, Gao X, Murgai A, Haji M, Schmidt A, Lutz R, Nogai A, Braune J, Laue D, Langer C, Khandanpour C, Bassermann F, Döhner H, Engelhardt M, Straka C, Hundemer M, Beule D, Haas S, Keller U, Einsele H, Bullinger L, Knop S, Mertins P, Krönke J (2024) The proteogenomic landscape of multiple myeloma reveals insights into disease biology and therapeutic opportunities. *Nat Cancer* 5(8): 1267–1284
- Rappoport N, Shamir R (2018) Multi-omic and multi-view clustering algorithms: review and cancer benchmark. *Nucleic Acids Res* 46(20): 10546–10562
- Rehman AU, Li M, Wu B, Ali Y, Rasheed S, Shaheen S, Liu X, Luo R, Zhang J (2024) Role of artificial intelligence in revolutionizing drug discovery. *Fundam Res*. <https://doi.org/10.1016/j.fmre.2024.04.021>
- Ren Y, Yue Y, Li X, Weng S, Xu H, Liu L, Cheng Q, Luo P, Zhang T, Liu

- Z, Han X (2024) Proteogenomics offers a novel avenue in neoantigen identification for cancer immunotherapy. *Int Immunopharmacol* 142: 113147. <https://doi.org/10.1016/j.intimp.2024.113147>
- Rheinbay E, Nielsen MM, Abascal F, Wala JA, Shapira O, Tiao G, Hornshøj H, Hess JM, Juul RI, Lin Z, Feuerbach L, Sabarinathan R, Madsen T, Kim J, Mularoni L, Shuai S, Lanzós A, Herrmann C, Maruvka YE, Shen C, Amin SB, Bandopadhyay P, Bertl J, Boroevich KA, Group PDAFIW, Group PSVW, Consortium P (2020) Analyses of non-coding somatic drivers in 2, 658 cancer whole genomes. *Nature* 578(7793): 102–111
- Rohart F, Gautier B, Singh A, Lê Cao K-A (2017) mixOmics: an R package for 'omics feature selection and multiple data integration. *PLoS Comput Biol* 13(11): e1005752. <https://doi.org/10.1371/journal.pcbi.1005752>
- Rolland DCM, Basrur V, Jeon Y-K, McNeil-Schwalm C, Fermin D, Conlon KP, Zhou Y, Ng SY, Tsou C-C, Brown NA, Thomas DG, Bailey NG, Omenn GS, Nesvizhskii AI, Root DE, Weinstock DM, Faryabi RB, Lim MS, Elenitoba-Johnson KSJ (2017) Functional proteogenomics reveals biomarkers and therapeutic targets in lymphomas. *Proc Natl Acad Sci USA* 114(25): 6581–6586
- Rudnick PA, Markey SP, Roth J, Mirokhin Y, Yan X, Tchekhovskoi DV, Edwards NJ, Thangudu RR, Ketchum KA, Kinsinger CR, Mesri M, Rodriguez H, Stein SE (2016) A description of the clinical proteomic tumor analysis consortium (CPTAC) common data analysis pipeline. *J Proteome Res* 15(3): 1023–1032
- Satpathy S, Krug K, Jean Beltran PM, Savage SR, Petralia F, Kumar-Sinha C, Dou Y, Reva B, Kane MH, Avanesian SC, Vasaikar SV, Krek A, Lei JT, Jaehnig EJ, Omelchenko T, Geffen Y, Bergstrom EJ, Stathias V, Christianson KE, Heiman DI, Cieslik MP, Cao S, Song X, Ji J, Liu W, Li K, Wen B, Li Y, Gumus ZH, Selvan ME, Soundararajan R, Visal TH, Raso MG, Parra ER, Babur O, Vats P, Anand S, Schraink T, Cornwell M, Rodrigues FM, Zhu H, Mo CK, Zhang Y, da Veiga Leprevost F, Huang C, Chinnaiyan AM, Wyczalkowski MA, Omenn GS, Newton CJ, Schurer S, Ruggles KV, Fenyo D, Jewell SD, Thiagarajan M, Mesri M, Rodriguez H, Mani SA, Udeshi ND, Getz G, Suh J, Li QK, Hostetter G, Paik PK, Dhanasekaran SM, Govindan R, Ding L, Robles AI, Clauser KR, Nesvizhskii AI, Wang P, Carr SA, Zhang B, Mani DR, Gillette MA, Clinical Proteomic Tumor Analysis Consortium (2021) A proteogenomic portrait of lung squamous cell carcinoma. *Cell* 184(16): 4348–4371
- Savage SR, Yi X, Lei JT, Wen B, Zhao H, Liao Y, Jaehnig EJ, Somes LK, Shafer PW, Lee TD, Fu Z, Dou Y, Shi Z, Gao D, Hoyos V, Gao Q, Zhang B (2024) Pan-cancer proteogenomics expands the landscape of therapeutic targets. *Cell* 187(16): 4389–4407
- Savage SR, Zhang B (2020) Using phosphoproteomics data to understand cellular signaling: a comprehensive guide to bioinformatics resources. *Clin Proteomics* 17(1): 27. <https://doi.org/10.1186/s12014-020-09290-x>
- Schulze K, Imbeaud S, Letoué E, Alexandrov LB, Calderaro J, Rebouissou S, Couchy G, Meiller C, Shinde J, Soysouvanh F, Calatayud A-L, Pinyol R, Pelletier L, Balabaud C, Laurent A, Blanc J-F, Mazzaferro V, Calvo F, Villanueva A, Nault J-C, Bioulac-Sage P, Stratton MR, Llovet JM, Zucman-Rossi J (2015) Exome sequencing of hepatocellular carcinomas identifies new mutational signatures and potential therapeutic targets. *Nat Genetics* 47(5): 505–511
- Sharma A, Debik J, Naume B, Ohnstad HO, Sahlber KK, Borgen E, Børresen-Dale A-L, Engebråten O, Fritzman B, Garred Ø, Geisler J, Geitvik GA, Hofvind S, Kristensen VN, Kåresen R, Langerød A, Lingjærde OC, Mælandsmo GM, Russnes HG, Sauer T, Skjerven HK, Schlichting E, Sørli T, Bathen TF, Giskeødegård GF, Oslo Breast Cancer C (2024) Comprehensive multi-omics analysis of breast cancer reveals distinct long-term prognostic subtypes. *Oncogenesis* 13(1): 22. <https://doi.org/10.1038/s41389-024-00521-6>
- Silva AMN, Vitorino R, Domingues MRM, Spickett CM, Domingues P (2013) Post-translational modifications and mass spectrometry detection. *Free Radic Biol Med* 65: 925–941
- Simon R, Roychowdhury S (2013) Implementing personalized cancer genomics in clinical trials. *Nat Rev Drug Discov* 12(5): 358–369
- Singh A (2022) Single-molecule protein sequencing on a semiconductor device. *Nat Methods* 19(12): 1517–1517
- Slyper M, Porter CBM, Ashenberg O, Waldman J, Drokhyansky E, Wakiro I, Smillie C, Smith-Rosario G, Wu J, Dionne D, Vigneau S, Jané-Valbuena J, Tickle TL, Napolitano S, Su M-J, Patel AG, Karlstrom A, Gritsch S, Nomura M, Waghray A, Gohil SH, Tsankov AM, Jerby-Arnon L, Cohen O, Klughammer J, Rosen Y, Gould J, Nguyen L, Hofree M, Tramontozzi PJ, Li B, Wu CJ, Izar B, Haq R, Hodi FS, Yoon CH, Hata AN, Baker SJ, Suvà ML, Bueno R, Stover EH, Clay MR, Dyer MA, Collins NB, Matulonis UA, Wagle N, Johnson BE, Rotem A, Rozenblatt-Rosen O, Regev A (2020) A single-cell and single-nucleus RNA-Seq toolbox for fresh and frozen human tumors. *Nat Med* 26(5): 792–802
- Smith KS, Bihannic L, Gudenas BL, Haldipur P, Tao R, Gao Q, Li Y, Aldinger KA, Iskusnykh IY, Chizhikov VV, Scoggins M, Zhang S, Edwards A, Deng M, Glass IA, Overman LM, Millman J, Sjoboen AH, Hadley J, Golser J, Mankad K, Sheppard H, Onar-Thomas A, Gajjar A, Robinson GW, Hovestadt V, Orr BA, Patay Z, Millen KJ, Northcott PA (2022) Unified rhombic lip origins of group 3 and group 4 medulloblastoma. *Nature* 609(7929): 1012–1020
- Sorin M, Rezanejad M, Karimi E, Fiset B, Desharnais L, Perus LJM, Milete S, Yu MW, Maritan SM, Doré S, Pichette É, Enlow W, Gagné A, Wei Y, Orain M, Manem VSK, Rayes R, Siegel PM, Camilleri-Broët S, Fiset PO, Desmeules P, Spicer JD, Quail DF, Joubert P, Walsh LA (2023) Single-cell spatial landscapes of the lung tumour immune microenvironment. *Nature* 614(7948): 548–554
- Stewart RL, Matynia AP, Factor RE, Varley KE (2020) Spatially-resolved quantification of proteins in triple negative breast cancers reveals differences in the immune microenvironment associated with prognosis. *Sci Rep* 10(1): 6598. <https://doi.org/10.1038/s41598-020-63539-x>
- Sturm G, Finotello F, Petitprez F, Zhang JD, Baumbach J, Fridman WH, List M, Aneichyk T (2019) Comprehensive evaluation of transcriptome-based cell-type quantification methods for immuno-oncology. *Bioinformatics* 35(14): i436–i445
- Subramanian I, Verma S, Kumar S, Jere A, Anamika K (2020) Multi-omics data integration, interpretation, and its application. *Bioinform Biol Insights* 14: 1177932219899051. <https://doi.org/10.1177/1177932219899051>
- Sun F, Li H, Sun D, Fu S, Gu L, Shao X, Wang Q, Dong X, Duan B, Xing F, Wu J, Xiao M, Zhao F, Han J-DJ, Liu Q, Fan X, Li C, Wang C, Shi T (2024a) Single-cell omics: experimental workflow, data analyses and applications. *Sci China Life Sci*. <https://doi.org/10.1007/s11427-023-2561-0>
- Sun Y (2016) Tumor microenvironment and cancer therapy resistance. *Cancer Lett* 380(1): 205–215
- Sun Y, Selvarajan S, Zang Z, Liu W, Zhu Y, Zhang H, Chen W, Chen H, Li L, Cai X, Gao H, Wu Z, Zhao Y, Chen L, Teng X, Mantoo S, Lim TK-H, Hariraman B, Yeow S, Alkaff SMF, Lee SS, Ruan G, Zhang Q, Zhu T, Hu Y, Dong Z, Ge W, Xiao Q, Wang W, Wang G, Xiao J, He Y, Wang Z, Sun W, Qin Y, Zhu J, Zheng X, Wang L, Zheng X, Xu K, Shao Y, Zheng S, Liu K, Aebersold R, Guan H, Wu X, Luo D, Tian W, Li SZ, Kon OL, Iyer NG, Guo T (2022) Artificial intelligence defines protein-based classification of thyroid nodules. *Cell Discovery* 8(1): 85. <https://doi.org/10.1038/s41421-022-00442-x>

- Sun Y, Wu L, Zhong Y, Zhou K, Hou Y, Wang Z, Zhang Z, Xie J, Wang C, Chen D, Huang Y, Wei X, Shi Y, Zhao Z, Li Y, Guo Z, Yu Q, Xu L, Volpe G, Qiu S, Zhou J, Ward C, Sun H, Yin Y, Xu X, Wang X, Esteban MA, Yang H, Wang J, Dean M, Zhang Y, Liu S, Yang X, Fan J (2021) Single-cell landscape of the ecosystem in early-relapse hepatocellular carcinoma. *Cell* 184(2): 404–421
- Sun Z, Wang J, Zhang Q, Meng X, Ma Z, Niu J, Guo R, Tran LJ, Zhang J, Liu Y, Ye F, Ma B (2024b) Coordinating single-cell and bulk RNA-seq in deciphering the intratumoral immune landscape and prognostic stratification of prostate cancer patients. *Environ Toxicol* 39(2): 657–668
- Suo Y, Du D, Chen C, Zhu H, Wang X, Song N, Lu D, Yang Y, Li J, Wang J, Luo Z, Zhou B, Luo C, Zhou H (2024) Uncovering PROTAC sensitivity and efficacy by multidimensional proteome profiling: a case for STAT3. *J Med Chem* 67(6): 4804–4818
- Tan YC, Low TY, Lee PY, Lim LC (2024) Single-cell proteomics by mass spectrometry: Advances and implications in cancer research. *Proteomics* 24(12-13): 2300210. <https://doi.org/10.1002/pmic.202300210>
- Tian Y, Wang Y, Yin H, Luo Y, Wei F, Zhou H, Wen L (2022) A sensitive and reversible labeling strategy enables global mapping of the core-fucosylated glycoproteome on cell surfaces. *Angew Chem Int Ed Engl* 61(49): e202206802. <https://doi.org/10.1002/anie.202206802>
- Tomczak K, Czerwińska P, Wiznerowicz M (2015) The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. *Contemp Oncol (Pozn)* 19(1A): A68–77
- Trapnell C, Cacchiarelli D, Grimsby J, Pokharel P, Li S, Morse M, Lennon NJ, Livak KJ, Mikkelsen TS, Rinn JL (2014) The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat Biotechnol* 32(4): 381–386
- Tsang O, Wong JWH (2021) Proteogenomic interrogation of cancer cell lines: an overview of the field. *Expert Rev Proteomics* 18(3): 221–232
- Van den Berge K, Roux de Bézieux H, Street K, Saelens W, Cannoodt R, Saeyns Y, Dudoit S, Clement L (2020) Trajectory-based differential expression analysis for single-cell sequencing data. *Nat Commun* 11(1): 1201. <https://doi.org/10.1038/s41467-020-14766-3>
- Vasaikar S, Huang C, Wang X, Petyuk VA, Savage SR, Wen B, Dou Y, Zhang Y, Shi Z, Arshad OA, Gritsenko MA, Zimmerman LJ, McDermott JE, Clauss TR, Moore RJ, Zhao R, Monroe ME, Wang YT, Chambers MC, Slebos RJC, Lau KS, Mo Q, Ding L, Ellis M, Thiagarajan M, Kinsinger CR, Rodriguez H, Smith RD, Rodland KD, Liebler DC, Liu T, Zhang B, Clinical Proteomic Tumor Analysis Consortium (2019) Proteogenomic analysis of human colon cancer reveals new therapeutic opportunities. *Cell* 177(4): 1035–1049
- Wang L-B, Karpova A, Gritsenko MA, Kyle JE, Cao S, Li Y, Rykunov D, Colaprico A, Rothstein JH, Hong R, Stathias V, Cornwell M, Petralia F, Wu Y, Reva B, Krug K, Pugliese P, Kawaler E, Olsen LK, Liang W-W, Song X, Dou Y, Wendl MC, Caravan W, Liu W, Cui Zhou D, Ji J, Tsai C-F, Petyuk VA, Moon J, Ma W, Chu RK, Weitz KK, Moore RJ, Monroe ME, Zhao R, Yang X, Yoo S, Krek A, Demopoulos A, Zhu H, Wyczalkowski MA, McMichael JF, Henderson BL, Lindgren CM, Boekweg H, Lu S, Baral J, Yao L, Stratton KG, Bramer LM, Zink E, Couvillion SP, Bloodsworth KJ, Satpathy S, Sieh W, Boca SM, Schürer S, Chen F, Wiznerowicz M, Ketchum KA, Boja ES, Kinsinger CR, Robles AI, Hiltke T, Thiagarajan M, Nesvizhskii AI, Zhang B, Mani DR, Ceccarelli M, Chen XS, Cottingham SL, Li QK, Kim AH, Fenyö D, Ruggles KV, Rodriguez H, Mesri M, Payne SH, Resnick AC, Wang P, Smith RD, Iavarone A, Chheda MG, Barnholtz-Sloan JS, Rodland KD, Liu T, Ding L, Agarwal A, Amin M, An E, Anderson ML, Andrews DW, Bauer T, Birger C, Birrer MJ, Blumenberg L, Bocik WE, Borate U, Borucki M, Burke MC, Cai S, Calinawan AP, Carr SA, Cerda S, Chan DW, Charamut A, Chen LS, Chesla D, Chinnaiyan AM, Chowdhury S, Ciešlik MP, Clark DJ, Culpepper H, Czernicki T, D'Angelo F, Day J, De Young S, Demir E, Dhanasekaran SM, Dhir R, Domagalski MJ, Druker B, Duffy E, Dyer M, Edwards NJ, Edwards R, Elburn K, Ellis MJ, Eschbacher J, Francis A, Gabriel S, Gabrovski N, Garofano L, Getz G, Gillette MA, Godwin AK, Golbin D, Hanhan Z, Hannick LI, Hariharan P, Hindenach B, Hoadley KA, Hostetter G, Huang C, Jaehnig E, Jewell SD, Ji N, Jones CD, Karz A, Kaspera W, Kim L, Kothadia RB, Kumar-Sinha C, Lei J, Leprevost FD, Li K, Liao Y, Lilly J, Liu H, Lubinski J, Madan R, Maggio W, Malc E, Malovannaya A, Mareedu S, Markey SP, Marrero-Oliveras A, Martinez N, Maunganidze N, McDermott JE, McGarvey PB, McGee J, Mieczkowski P, Migdizio S, Modugno F, Montgomery R, Newton CJ, Omenn GS, Ozbek U, Paklina OV, Paulovich AG, Perou AM, Pico AR, Piehowski PD, Placantonakis DG, Polonskaya L, Potapova O, Pruetz B, Qi L, Ramkissoon S, Resnick A, Richey S, Riggins G, Robinson K, Roche N, Rohrer DC, Rood BR, Rossell L, Savage SR, Schadt EE, Shi Y, Shi Z, Shutack Y, Singh S, Skelly T, Sokoll LJ, Stawicki J, Stein SE, Suh J, Szopa W, Tabor D, Tan D, Tansil D, Thangudu RR, Tognon C, Traer E, Tsang S, Tyner J, Um KS, Valley DR, Vasaikar S, Vatanian N, Velvulou U, Vernon M, Wan W, Wang J, Webster A, Wen B, Whiteaker JR, Wilson GD, Zakhartsev Y, Zelt R, Zhang H, Zhang L, Zhang Z, Zhao G, Zhu J (2021a) Proteogenomic and metabolomic characterization of human glioblastoma. *Cancer Cell* 39(4): 509–528
- Wang L, Liu Z, Liu L, Guo C, Jiao D, Li L, Zhao J, Han X, Sun Y (2021b) CELF2 is a candidate prognostic and immunotherapy biomarker in triple-negative breast cancer and lung squamous cell carcinoma: a pan-cancer analysis. *J Cell Mol Med* 25(15): 7559–7574
- Wang R, Peng G, Tam PP, Jing N (2022) Integration of computational analysis and spatial transcriptomics in single-cell study. *Genom Proteom Bioinf* 21(1): 13–21
- Wang X-Y, Xu Y-M, Lau ATY (2023) Proteogenomics in cancer: then and now. *J Proteome Res* 22(10): 3103–3122
- Wettersten HI, Aboud OA, Lara PN, Weiss RH (2017) Metabolic reprogramming in clear cell renal cell carcinoma. *Nat Rev Nephrol* 13(7): 410–419
- Wiredja DD, Koyutürk M, Chance MR (2017) The KSEA App: a web-based tool for kinase activity inference from quantitative phosphoproteomics. *Bioinformatics* 33(21): 3489–3491
- Wishart D (2022) Metabolomics and the multi-omics view of cancer. *Metabolites* 12(2): 154. <https://doi.org/10.3390/metabo12020154>
- Wu P, Nielsen TE, Clausen MH (2015) FDA-approved small-molecule kinase inhibitors. *Trends Pharmacol Sci* 36(7): 422–439
- Wu T, Zhang X, Liu X, Cai X, Shen T, Pan D, Liang R, Ding R, Hu R, Dong J, Li F, Li J, Xie L, Wang C, Geng S, Yang Z, Xing L, Li Y (2023) Single-cell sequencing reveals the immune microenvironment landscape related to anti-PD-1 resistance in metastatic colorectal cancer with high microsatellite instability. *BMC Med* 21(1): 161. <https://doi.org/10.1186/s12916-023-02866-y>
- Xiang H, Zhang L, Bu F, Guan X, Chen L, Zhang H, Zhao Y, Chen H, Zhang W, Li Y, Lee LJ, Mei Z, Rao Y, Gu Y, Hou Y, Mu F, Dong X (2022) A novel proteogenomic integration strategy expands the breadth of neo-epitope sources. *Cancers* 14(12): 3016. <https://doi.org/10.3390/cancers14123016>
- Xiao Q, Zhang F, Xu L, Yue L, Kon OL, Zhu Y, Guo T (2021) High-throughput proteomics and AI for cancer biomarker discovery. *Adv Drug Deliv Rev* 176: 113844. <https://doi.org/10.1016/j.addr.2021.113844>

- Xiao R, Ding C, Zhu H, Liu X, Gao J, Liu Q, Lu D, Zhang N, Zhang A, Zhou H (2020) Suppression of asparagine synthetase enhances the antitumor potency of ART and artemalogue SOMCL-14-221 in non-small cell lung cancer. *Cancer Lett* 475: 22–33
- Xiao Y, Ma D, Yang YS, Yang F, Ding JH, Gong Y, Jiang L, Ge LP, Wu SY, Yu Q, Zhang Q, Bertucci F, Sun Q, Hu X, Li DQ, Shao ZM, Jiang YZ (2022) Comprehensive metabolomics expands precision medicine for triple-negative breast cancer. *Cell Res* 32(5): 477–490
- Xie Z, Feng Q, Zhang S, Yan Y, Deng C, Ding C-F (2022) Advances in proteomics sample preparation and enrichment for phosphorylation and glycosylation analysis. *Proteomics* 22(23-24): 2200070. <https://doi.org/10.1002/pmic.202200070>
- Xu JY, Zhang C, Wang X, Zhai L, Ma Y, Mao Y, Qian K, Sun C, Liu Z, Jiang S, Wang M, Feng L, Zhao L, Liu P, Wang B, Zhao X, Xie H, Yang X, Zhao L, Chang Y, Jia J, Wang X, Zhang Y, Wang Y, Yang Y, Wu Z, Yang L, Liu B, Zhao T, Ren S, Sun A, Zhao Y, Ying W, Wang F, Wang G, Zhang Y, Cheng S, Qin J, Qian X, Wang Y, Li J, He F, Xiao T, Tan M (2020) Integrative proteomic characterization of human lung adenocarcinoma. *Cell* 182(1): 245–261
- Zack TI, Schumacher SE, Carter SL, Cherniack AD, Saksena G, Tabak B, Lawrence MS, Zhang C-Z, Wala J, Mermel CH, Sougnez C, Gabriel SB, Hernandez B, Shen H, Laird PW, Getz G, Meyerson M, Beroukhi R (2013) Pan-cancer patterns of somatic copy number alteration. *Nat Genetics* 45(10): 1134–1140
- Zhang B, Tang C, Yao Y, Chen X, Zhou C, Wei Z, Xing F, Chen L, Cai X, Zhang Z, Sun S, Liu Q (2021a) The tumor therapy landscape of synthetic lethality. *Nat Commun* 12(1): 1275. <https://doi.org/10.1038/s41467-021-21544-2>
- Zhang E, Zhang M, Shi C, Sun L, Shan L, Zhang H, Song Y (2020a) An overview of advances in multi-omics analysis in prostate cancer. *Life Sci* 260: 118376. <https://doi.org/10.1016/j.lfs.2020.118376>
- Zhang F, Ge W, Ruan G, Cai X, Guo T (2020b) Data-independent acquisition mass spectrometry-based proteomics and software tools: a glimpse in 2020. *Proteomics* 20(17-18): 1900276. <https://doi.org/10.1002/pmic.201900276>
- Zhang H, Liu T, Zhang Z, Payne SH, Zhang B, McDermott JE, Zhou J-Y, Petyuk VA, Chen L, Ray D, Sun S, Yang F, Chen L, Wang J, Shah P, Cha SW, Aiyetan P, Woo S, Tian Y, Gritsenko MA, Clauss TR, Choi C, Monroe ME, Thomas S, Nie S, Wu C, Moore RJ, Yu K-H, Tabb DL, Fenyo D, Bafna V, Wang Y, Rodriguez H, Boja ES, Hiltke T, Rivers RC, Sokoll L, Zhu H, Shih I-M, Cope L, Pandey A, Zhang B, Snyder MP, Levine DA, Smith RD, Chan DW, Rodland KD, Clinical Proteomic Tumor Analysis C (2016) Integrated proteogenomic characterization of human high-grade serous ovarian cancer. *Cell* 166(3): 755–765
- Zhang Y, Chen H, Mo H, Hu X, Gao R, Zhao Y, Liu B, Niu L, Sun X, Yu X, Wang Y, Chang Q, Gong T, Guan X, Hu T, Qian T, Xu B, Ma F, Zhang Z, Liu Z (2021b) Single-cell analyses reveal key immune cell subsets associated with response to PD-L1 blockade in triple-negative breast cancer. *Cancer Cell* 39(12): 1578–1593
- Zhang Y, Narayanan SP, Mannan R, Raskind G, Wang X, Vats P, Su F, Hosseini N, Cao X, Kumar-Sinha C, Ellison SJ, Giordano TJ, Morgan TM, Pitchiaya S, Alva A, Mehra R, Cieslik M, Dhanasekaran SM, Chinnaiyan AM (2021c) Single-cell analyses of renal cell cancers reveal insights into tumor microenvironment, cell of origin, and therapy response. *Proc Natl Acad Sci USA* 118(24): e2103240118. <https://doi.org/10.1073/pnas.2103240118>