Research progress of proteomics in screening tumor markers of gastric cancer

Dongkai Wu\textsuperscript{1,a}, Chen Ling\textsuperscript{1,b}, Qiang Zhao\textsuperscript{2,c,*}

\textsuperscript{1}Changzhi Medical College, Changzhi, Shanxi, 046000, China
\textsuperscript{2}Heping Hospital Affiliated to Changzhi Medical College, Changzhi, Shanxi, 046000, China
\textsuperscript{a}wdk18403551908@163.com, \textsuperscript{b}1768081949@qq.com, \textsuperscript{c}2238982818@qq.com

*Corresponding author

Keywords: Gastric cancer, proteomics technology, protein interactions, tumor markers

Abstract: Gastric cancer is a significant disease that endangers human life and health, ranking 3rd in the spectrum of disease incidence and death in China. It is mostly locally progressive and advanced, with fewer early-stage patients and an overall survival rate and quality of survival significantly lower than abroad. Now, early diagnosis and treatment, extending the survival time, and improving the quality of survival have become hot spots in treating gastric cancer. Protein is the performer of life activities. The protein of gastric cancer is different from that of healthy people. The understanding of its protein will be of great significance for disease prevention, diagnosis, and treatment, but the latest research reports are few; This article reviews the Progress of proteomics in screening tumor markers of gastric cancer.

1. Introduction

Gastric cancer (GC) is one of the malignant tumors with the highest morbidity and mortality in the world. According to the Global Cancer Statistics Report 2020\cite{1}, stomach cancer ranks 5th in the spectrum of cancer incidence and 4th in the spectrum of causes of death worldwide. In China, stomach cancer accounts for 43.9\% of new cases and 48.6\% of deaths worldwide, ranking 3rd in both incidence and death spectrum. It is a medium to significant disease that seriously endangers people's life and health. The completion of the Human Tumor Genome Project has guided people's understanding of GC from the macro level to the micro level and also promoted the change of GC treatment from the traditional surgical treatment to the individualized comprehensive treatment with surgery as the primary treatment and radiotherapy, chemotherapy and immunotherapy as supplements. The retrospective study has formally shown that it has achieved significant results. Scholars have turned their attention to proteins at this stage, and the performers of life activities and related research have now become a frontier area for tumor attacks.

2. The research content and significance of proteomics

Protein is the result of gene expression, and since Williams and Wilkins first introduced the concept of "proteome," the study has evolved to include all the proteins expressed by a cell or
organism. The corresponding research techniques are called proteomics techniques. Proteomics can analyze the dynamics of protein composition, expression level, and modification status in the organism from a holistic perspective by comparing different proteins, understanding protein-protein interaction (PPI) and linkage, and revealing the pattern of protein function and life activity. Essential for understanding the development, prevention, diagnosis, and treatment of cancer [2].

3. Latest research techniques in proteomics

There are five kinds of proteomic detection technologies with different detection principles.

3.1. Electrophoresis technique

2-Dimensional gel-electrophoresis technology (2-DE) is one of the classical core techniques in proteomics. However, its disadvantages of low resolution, poor repeatability, and cumbersome operation limit its development. Later, two-dimension difference gel electrophoresis (2D-DIGE), combined with mass spectrometry detection technology, became the core technology of proteomics research at this stage; Electrophoresis technology is constantly updating. Capillary electrophoresis and multicellular electrophoresis have promoted the development of agricultural science and medicine. It also has excellent potential for development in other disciplines.

3.2. Mass spectrometry

Mass spectrometry detection technology has high resolution and high precision characteristics, which plays an essential role in many fields. At present, the commonly used chromatography-mass spectrometry (CMS), matrix-assisted laser desorption ionization (MALDI), and Electrospray ionization (ESI) can ensure the integrity of the sample to a certain extent.

Based on this, liquid chromatography-mass spectrometry (LC-MS) technology is in time, which can detect the enzymolysis products of complex samples. With the development of technology, single-cell analysis has become an essential tool to analyze tumor heterogeneity, which can analyze multiple markers of a single tumor cell. Single-cell analysis has become an essential tool to analyze tumor heterogeneity [3]; Mass Cytometry (MC) converts cells into molecular data for phenotype analysis of single cells, which can avoid crosstalk, help identify drug-resistant tumor cell subsets, and guide the development of targeted therapy [4-6].

Thermal proteome profiling [7] (TPP) analysis is a new frontier technology that uses the principle of protein dissolution curve shift under high temperatures for analysis, which can identify small molecular target proteins. However, it is still challenging to obtain reliable and repeatable data results due to the complex operation.

3.3 Gel-independent

3.3.1 Marker quantification techniques

In the past, the quantitative analysis of protein digestion results by using the peak intensity or peak area of mass spectrometry, which poor accuracy, promoted the generation of isotope labeling technology. Stable isotope labeling with amino acids in cell culture (SILAC) is generally considered the "gold standard" for isotope labeling methods. Isobaric tags for relative and absolute quantitation (iTRAQ) and isobaric tags for relative and absolute quantification/tandem mass tags (iTRAQ/TMT) can accurately identify and quantify proteins in the mixing system [8] and has significant advantages in repeatability and accuracy. The mobile liquid phase multi-protein quantitative
technology developed on this basis can simultaneously detect multiple proteins quickly for accurate qualitative and quantitative analysis. It has made Progress in immunology, drug toxicology, and other disciplines.

3.3.2 Non-marker quantitative techniques

Label-free shotgun and Data Independent Acquisition (DIA) quantitative techniques are independent of isotope labeling. The former has the characteristics of simple processing steps, low cost, and high authenticity of samples; It can use for quantitative analysis of multiple samples simultaneously. It is a conventional choice for large-scale experiments; The latter is an accurate panoramic mass spectrometry assay, can significantly shorten the detection cycle, and is a common choice for large-scale experiments.

3.4. Antibody-based

The protein microarray technology is now clinically used primarily for screening and finding tumor markers and drug target ligands for rapid mass screening. It has the characteristics of high throughput, high speed, high sensitivity, low sample requirement, and better reproducibility [9-10]. However, there are still shortcomings of high cost, poor stability, and false positives. In recent years, Surface plasmon resonance (SPR) protein microarray and optical protein microarray have solved the above problems to a certain extent. It can efficiently screen low-abundance tumor markers, which has good development potential in clinical medicine, translational medicine, and the biological industry.

3.5. Based on protein interactions

PPI includes the formation of stable protein complexes or transient interactions, mostly a dynamic process characterized by transience, weak interactions, and regulated by multiple factors. There are three types of research methods: biophysical methods, molecular biological methods, and genetic methods.

3.5.1 Biophysical methods

Including Co-Immunoprecipitation (Co-IP), binding assay in vitro, and Tandem affinity purification (TAP), Co-IP can identify intracellular interacting proteins. Chromatin immunoprecipitation assay (ChIP) can combine with microarray, sequencing, and other technologies, and remarkable achievements have been made in modifying transcription factors, histones, and other fields. Binding assay in vitro is easy to operate but easily affected by many factors, leading to false positive results. TAP can effectively reduce the binding of non-specific proteins, but it is challenging to detect partial transients and weak interactions.

3.5.2 Molecular biology methods

It includes yeast hybridization technology and phage display technology. The former can detect the weak and instantaneous interaction between proteins, but it has the disadvantages of false positives, false negatives, and low transformation efficiency. Proximity-dependent biotin identification (BioID) is a unique method for studying PPI in living cells. It has been successfully applied to the study of insoluble proteins and transient or weakly interacting proteins and gradually applied to disease pathogenesis and drug target screening [11]. Phage display technology obtains specific phages through antigen antibody-specific binding, which is essential in designing and
screening small molecule drugs. In China, scholars use phage display technology to screen specific nano antibodies against novel coronavirus, which will have broad prospects.

3.5.3 Genetic methods

Including synthetic lethality (SL) and extragenic suppressor. Two genes with SL interaction have bright prospects in digging potential therapeutic targets, improving combination or new drug regimens, and even reversing tumor immune escape. Extragenic suppressors make up for the mutation of the original gene by mutation of one gene. However, the research process is cumbersome, the cost is high, and it requires high-quality protein.

4. Application of proteomics in gastric cancer research

GC proteomics generally has five research directions: tissue, serum, body fluid, single cell, and microorganism, each of which has its characteristics.

4.1 Proteomic screening of tissues for differentially expressed proteins

Tissue samples have high protein abundance and are the most direct and accurate sample type. However, sample preparation is invasive, and the tissue samples are primarily mixtures. Guodong Lian [12] collected 140 tissue samples from clinical stage III patients who underwent GC eradication and assessed the protein expression in tumor tissues using a total protein pathway array. Six differential proteins, including PLK1 and DACH1, were finally screened as independent risk factors for drug resistance and developed related prediction models. Xiaotian Ni [13] collected 82 samples from different anatomical regions of the normal stomach and 58 mucosal samples from GC tumors and adjacent tissues and drew a reference map of the mucosal protein profile. Which can provide a range of protein expressions covering individual variation and a reference for disease and health. Xue Li [14] collected 324 cases of Tissue samples of GC and precancerous lesions for tissue proteome analysis. Propose four molecular subtypes of GC, and the subtype - S4 with the highest risk of progression was determined. Proposed that 104 positive and 113 harmful related proteins are related to early GC, and APOA1BP and PGC are related to GC progress, suggesting that APOA1BP, and PGC, were associated with GC progression. Xiaoxia Yuan [15] collected four pairs of GC tumors and tumor-adjacent tissues. Through Liquid chromatography-electrospray ionization tandem mass spectrometry (LC-ESI-MS/MS) and immunohistochemical analysis, believed that 6-phosphogluconolactonase (PGLS) could be an essential target for early diagnosis and follow-up monitoring of GC.

4.2 Proteomic screening of serum differentially expressed proteins

Serum samples have the advantages of being easy to obtain, low cost, less invasive, reproducible, providing a large amount of clinical information, and providing the earliest representation of disease onset, progression, and regression [16], which are the preferred study samples. Quijin Shen [17] collected 100 preoperative and postoperative serum samples from GC patients and compared them with standard serum samples (N=50). Nineteen kinds of serum proteins, such as CEACAM5 and CA9, were identified and considered to have better sensitivity and specificity in diagnosing patients with TNM stage I-II and patients with high microsatellite instability. Jiajia Shen [18] collected preoperative and 4-week postoperative serum specimens from patients with advanced GC and used liquid chromatography-mass spectrometry to identify differential proteins and validate the potential of SOX3 as a marker for GC. Zong Cao Liu [19] collected information from 400 subjects with GC and precancerous lesions to explore the pathological changes in the stomach and the lipoic
characteristics of GC. Finally, eight kinds of phospholipids and three kinds of free fatty acids were potentially crucial in the progression of gastric diseases to early GC, which had the transformation significance of GC prevention. Shahid Aziz [20] studied the sera of 219 patients diagnosed with GC, gastritis, and ulcers, searching GC-dependent proteins. They found ten groups of GC-specific serum protein markers and 29 groups of early GC prospective markers.

4.3 Proteomic screening of differentially expressed proteins in body fluids

Body fluid samples have the advantages of low acquisition cost and continuous collection but low protein concentration, poor stability, and easy contamination. Jonghwa Jin [21], constructing a deep proteomics study of ascites, identified two candidate proteins, gastrin, and osteopontin, which could differentiate malignant ascites from benign ascites in GC patients. Lei Huang [22] reported concerning a study of 1870 patients showing that the maximum fluorescence intensity of the first peak at 288 nm (P1 FI) in the gastric fluid was significantly higher in GC patients than in patients with benign disease. Clinical validation showed significantly higher mean P1 FI fluorescence intensity values in gastric fluid in the advanced GC (92.1). Early GC (90.8) groups than in the standard mucosa/chronic non-atrophic gastritis group (55.7), suggesting that gastric fluid fluorescence spectroscopy may be a potential screening method for early GC. Jian Shu [23] found significantly altered 15 candidate lectins by saliva microarray and lectin blotting of 64 GC, 30 gastritis patients, and 30 healthy salivae for salivary glycan pattern difference analysis. They demonstrated that salivary glycans have great potential as a biomarker for GC diagnosis. Takaya Shimura [24] identified urinary proteins by proteomic analysis of 282 urine samples from GC patients and healthy controls. They proposed the potential of urinary levels of TFF1 and ADAM12 as independent tumor markers for GC.

4.4 Proteomic screening of single cell differentially expressed proteins

Cell samples have the advantages of fresh samples, good homogeneity, and easy accessibility, but they are more expensive to study and prone to contamination. Jai-Sing Yang [25] used iTRAQ to analyze the possible action network of Helicobacter pylori (H. pylori) targets in human GC cell AGS and found that HP-infected AGS cells may target EIF2 and BAG2 signaling pathways to regulate cell physiology. Furthermore. The expression of HSP70 and HSP105 was significantly lower than that of the control group. Shuanghuang Han [26] and others first applied CyTOF 2 flow cytometry technology to human GC cells AGS and MGC803 to radiotherapy and chemotherapy was detected. Present the specific clusters (44992, 44993) of differential expression had potential implications for evaluating the sensitivity of radiotherapy and chemotherapy and guiding individualized drug use. Jiaqi Lu [27] revealed the molecular mechanisms by which indooquinone inhibits the proliferation of GC cells AGS, HGC27, MGC803, and SGC7901 through neighborhood labeling techniques, quantitative proteomics, and biochemical studies. Their group found that indisulam enhanced the ubiquitination and degradation of RBM39 by promoting its interaction with DCAF15, thereby inhibiting the proliferation of gastric cancer cells. It may provide valuable information for drug discovery through protein hydrolysis targeting chimeras.

4.5 Proteomic screening for differentially expressed proteins within H. pylori

HP infection is now a significant cause of chronic gastritis, peptic ulcer disease, and gastric adenocarcinoma. It has been classified as a class I carcinogen by the International Agency for Research on Cancer. Malak Zoaiter [28] found that HP outer membrane vesicle (OMV) protein can
induce the secretion of oncostatin M (OSM) in AGS cells. Moreover, it guides type II receptor OSMR β (OSM/OSMR β) Overexpression. Its research group established HP OMVs and OSM/OSMR for the first time. Direct relationship between signal axes. Lusheng Song [29] used a programmable protein array to analyze the immune proteome, collected 50 GC cases and 50 healthy controls, and comprehensively evaluated GC's anti-HP humoral profile for the first time. Six anti-HP antibodies were found, including anti-HP1118/Ggt and anti-HP0516/Hs1U. The positive rate of GC serum was lower than that of the control group. Suggested that the change of specific anti-HP antibodies may have an immunoprotective effect on GC progress. Paradoxically. Yuka Koizumi [30] proposed that the prognosis of GC patients in the advanced stage of HP+was better than that of patients with HP -. They collected 491 patients' Information with HP infection and programmed death ligand 1 (PD-L1). They proposed that in the absence of gastric cancer immune escape mechanism, the host immune system regulated by hp might help to prolong the survival time. However, further studies are needed to investigate whether HP+ is a prognostic factor for postoperative chemotherapy in patients with advanced GC.

5. Conclusion

GC is a malignant disease worldwide; studying GC proteins has become a hot area. Some scholars tested the mRNA and protein expression levels of the same tissues. The consistency was less than 50%, proving that proteins are the actual manipulators of cell behavior. Later, foreign scholars proposed that compared with genomics, the verification and interpretation provided by proteomics research are closer to the phenotype, and almost all existing drug targets are proteins. In China, Ge [31], the study of "diffuse GC proteome" marked the "Chinese Human Proteome Project" to break out of the monopoly of the genome in the international arena.

With the continuous Progress of science and technology, single-cell and protein microarray technology has opened a new chapter for proteomics research and provided powerful support for understanding life's primary forms and core mechanisms. Proteomics interacts with bioinformatics analysis, analytical chemistry, statistics, and other disciplines. At the same time, it also set off a research upsurge in new research fields such as exocrine proteomics, thermal proteomics, and metabolomics. Proteomics technology has been widely used in clinical medicine, fisheries, animal husbandry, and other disciplines and has obtained significant disciplinary translation results.

At present, proteomics technology in the field of GC is an Active application to mine tumor-specific markers to achieve early detection, diagnosis, and treatment of GC. Individualized adjuvant therapy before surgery and detection of specific tumor markers after surgery to monitor tumor recurrence and drug resistance to improve the patient's life cycle and quality of life. Proteomics plays an irreplaceable role in this process.

At the same time, the complex spatial structure and composition of protein components have put forward higher requirements for science and technology, the author believes that the main problems faced by GC proteomics research at this stage are as follows: 1, protein separation capacity, repeatability, accuracy, anti-interference, etc.; with the Progress of science and technology, the above problems have been solved to a certain extent. However, high-throughput, high-precision and reproducible protein assay technologies are still necessary for the development of proteomics; 2) limitations in sample preparation; clinical samples are mostly mixtures, which are difficult to ensure the purity of the required proteins, and the components are easily contaminated or degraded, resulting in false positives and false negatives; 3) expensive research costs, which make it difficult to enter and implement at the grassroots level; looking ahead, proteomics research will undoubtedly develop a variety of complementary and advantageous new technologies and play a significant role in the road to overcome gastric cancer.
Acknowledgement

Thanks for Changzhi Medical College for this research support.

References


