Vol 3, No 4, Dec, 2017, P 159-169
DOI, 10.15562/tcp.59

CANCER PRESS
Vol. 3, No. 4, Dec, 2017

Review Article

Classification of Proteins Expression in some Popular Cancers for Protein Biomarkers Identification

Ali Farid Mohamadzadegan[1,*], Kazem Porbadakhshan[2]

ABSTRACT

Recognition of the source and stage of cancer has always been one of the Issues of interest to scientists. On the other hand, cancer is the second leading cause of death worldwide after cardiovascular disease. According to the Global Burden of Disease Cancer Report In 2015, there were 17.5 million cancer cases worldwide and over 8.7 million cancer deaths. Based on the same report, breast cancer, TBL (tracheal, bronchus, and lung) cancer and colorectal cancer were the most common incidents. From another perspective, one of the requirements for the treatment of different cancers is early diagnosis in the early stages. With the end of the human genome project, molecular medicine moved to a step beyond the genome called "proteomics". Proteomic ideas play an important role in discovering cancer biomarkers for early diagnosis of disease, prediction and prognosis, identifying new drug goals, monitoring the effectiveness of treatment and personal therapy. Nowadays with new developments in mass spectrometry and bioinformatics, new biomarkers can be identified for different cancers. To analyze a cancer, identifying only one biomarker does not provide enough information for that cancer, but paying attention to changes in the level of expression of various proteins is valuable. In this paper, effective proteins for breast, lung and colorectal cancers, have been identified and classified. Biomarkers sparse in different articles are combined using Text Mining and reviewing articles that introduced a cancer biomarker. In fact, by examining changes in the expression of proteins in the cancerous tissue and considering their significant changes, they are referred to as cancer marker candidates for early diagnosis or even prediction of future illness. This research offers a text mining algorithms to collect cancer biomarker's.

Keywords: Proteomics, Biomarkers, Breast Cancer, Lung Cancer, Colorectal Cancer, Classification

Author Information

1. Master student of control engineering, Department of Electrical Engineering, Quchan University of Technology

2. Assistant Professor, Department of Electrical Engineering, Quchan University of Technology

Submitted: 28-10-2017
Accepted : 28-11-2017
Published : 29-12-2017

Cancer has always been one of the main problems of human society and is usually caused by functional impairment of regulatory mechanisms of cell growth and division, Therefore, cancer occurs when the mechanisms responsible for stabilizing the growth process of the cells are in impaired (1, 2).

References

1. Zamanian-Azodi M, Rezaie-Tavirani M, Heydari-Kashal S, Kalantari S, Dailian S, Zali H. Proteomics analysis of MKN45 cell line before and after treatment with Lavender aqueous extract. Gastroenterol Hepatol Bed Bench. 2012;5 (1):35-42.

1) Cancer and Global Statistics

After cardiovascular disease, cancer is the second leading cause of death in the world. According to global statistics in 2015, there were 17.5 million cancer cases worldwide and 8.7 million deaths. Between 2005 and 2015, cancer cases increased by 33%. At the global level, the odds of developing cancer during a lifetime (age 0-79 years) differed between the sexes: they were 1 in 3 for men and 1 in 4 for women (3). For men, the most common cancer globally was prostate cancer (1.6 million cases). After that, (TBL) and colorectal are in the next ranks. The most common causes of cancer deaths for men were TBL, liver, and stomach cancer with 1.21 million, 577000, and 535000 deaths, respectively. For women in 2015, the most common incident cancers were breast, colorectal, and TBL cancer, with 2.4 million, 733000 and 640000 respectively. These cancers were responsible for 46% of all incident cases among women. The leading causes of cancer Deaths were breast, TBL, and colorectal cancer,523000,517000, and 376000 deaths, respectively (3).

2) 3 most common cancer in the world

This section introduces 3 most common cancer, in which the statistical community consists of both male and female groups.

a) Breast cancer: According to the latest published statistics, breast cancer is the most common cancer among all the people in the world. The number of people with this cancer was 2.44 million, the vast majority occurred in women, with 2.4 million Cases vs 44000 cases in men. It was the cause of death for 523000 women and 10000 men. One in 14 women and 1 in 603 men developed breast cancer between birth and age79 years at the global level. It is the most common cancer for women in 183 countries or territories and the most common cause of cancer deaths in women in 115 countries or territories (3).

b) Tracheal, Bronchus, and Lung Cancer (TBL): In 2015, there were 2 million incident cases of TBL cancer and 1.7 million deaths. 1in18 men and 1in 45 women developing TBL cancer between birth and age 79 years. Tracheal, bronchus, and lung cancer was the cause of the most incident cases for men in 38 countries and the most common cause for cancer deaths in 113 countries or territories for women, TBL cancer was the most common cause of cancer deaths in 20countries and territories (3).

C) Colon and Rectum Cancer: In2015, therewere1.7million incident cases of colon and rectum cancer, and it caused 832000deaths. The odds of developing colon and rectum cancer before age79 years at the global level was higher for men than for women (1 in 28 men, 1 in 43 women). Colon and rectum cancer was the cancer with the highest incidence in 2015 for men in 6 countries. For women, colon and rectum cancer was the most common cause of cancer deaths in 5 countries (3).

References

2. Zali H, Rezaei-Tavirani M, Azodi M. Gastric cancer: prevention, risk factors and treatment. Gastroenterol Hepatol Bed bench. 2011;4(4):175-85.

3. Fitzmaurice C, Dicker D, Pain A, Hamavid H, Moradi-Lakeh M, MacIntyre MF, et al. The global burden of cancer 2013. JAMA oncology. 2015;1(4):505-27.

This paper is organized as follows: In sectionII, Proteomics analysis are presented. biomarkers applications in cancer are presented, in sectionIII. In SectionIV, Text mining and it's rolls in cancer research are presented. Collected biomarkers Were shown in sectionV and the final part is conclusion.

Proteomics and its role in cancer analysis

Proteomics simply means many or groups of proteins. Proteomics is about understanding the structure and function of all proteins. Defective proteins are of the main causes of cancer. Therefore, they are important indications for cancer diagnosis and treatment. In addition, proteins are the main target for most medicines and are also a prior nature of designing different medicines. As a result, proteomics studies are highly important in order to recognize their role in cancer formation and control (4, 5). By identifying cancer-related proteins, they can be targeted with drugs that are designed by computer software, for example if a specific protein is involved in cancer, by knowing protein 3D structure, proteomics can be designed to counteract it effectively and prevent protein activity (6, 7). Since different individuals have different genetic data, they will have different proteins. Hence, one of the proteomics applications is to design a more specific drug for each person's treatment by determining the proteome of each individual (8, 9). Many of the experimental techniques used in proteomics, such as gel electrophoresis, liquid chromato-graphy and mass spectrometry (MS), have been around for several decades, yet it is only in the last 10–15 years that approaches have been developed allowing many proteins to be detected or quantified simultaneously, coming closer to “global” methods of proteome analysis (10).

1) Proteomics applications in cancer:

Proteomic based approaches are continuing to play a major role in studying the natural history and treatment of cancer. For example, proteomics can facilitate discovery of new molecular targets for therapy, biomarkers for early detection, and new endpoints for therapeutic efficacy and toxicity (11 ,

References

4. Karin D. Rodland,"Proteomics and cancer diagnosis: the potential of mass spectrometry", Clinical Biochemistry 37 (2004) 579–583.

5. Hayat MA, "Methods of cancer diagnosis therapy and prognosis: liver cancer",1th:ed Springer Verlag. 2009.P.14

6. Pooladi M, Sobhi S, Abad SK, Hashemi M, Moradi A, Zali AR, et al. "The Investigation of Heat Shock Protein (HSP70) Expression Change in Human Brain Asterocytoma Tumor", Iran J Cancer Prevent. 2013;6:6-11.

7. Rozek W, Ciborowski PS, "Proteomics and Genomics Neuroimmune Pharmacology",2thed,Springer; 2008.P.725-41.

8. Gottfries J, Sjögren M, Holmberg B, Rosengren L, Davidsson P, Blennow K. "Proteomics for drug target discovery", Chemom Intell Lab Sys. 2004;73:47-53.

9. Scarano EM, Fiorita AN, Picciotti PM, Passali GC, Calo L, Cabras T, et al. "Proteomics of saliva:personal experience", Acta Otorhinolaryngol Ital. 2010;30:125.

10. Simon J, Andrew R. "Proteome Bioinformatics", Methods in Molecular Biology, vol. 604:2010.

11. Bichsel VE, Liotta LA. "Cancer proteomics: from biomarker discovery to signal pathway profiling", Cancer J. 2001;7(1):69-78.

12). Identification of specific and sensitive markers of cancer is an important public health concern. Currently, there are no effective screening options available for many cancer patients (13, 14). As mentioned, proteomics applications are the identification of biomarkers in the early diagnosis of diseases In this regard, we can compare this information with the status of cancer by examining changes in the isoforms and the various changes that occur in protein molecules (15- 17).

2) Some limitations on the use of proteomics in diagnosis:

To use proteomics technique, a lot of protein is needed (18). Low protein amount is an unreliable source for electrophoresis, in addition, healthy tissues and tumors are a mixture of different cells. To find biomarkers and research in this field, there is a need for pure samples to achieve reliable results.

Biomarkers:

Over recent years lots of attention has been paid to the role of biomarkers in the diagnosis of cancer in clinical studies (19). Biomarkers are important tools for tracking and studying cancer. The use of methods such as proteomics can help for the early diagnosis of cancer and also identification of protein biomarkers (20). This process will be considered as a fundamental way to treatment of cancer. Today, with advances in proteomics, biomarkers can be identified in various cancers, as well as the design and function of proteins. The advancement of proteomics and genomics has led to the identification of a wide range of clinically valuable biomarkers. Biomarkers' cognition is helpful in determining the (stage) of the disease and its specific treatment (21). Biomarkers are really important in timely diagnosis of cancer, the progression of disease, effective treatment by

References

12. Clarke W, Zhang Z, Chan DW. "The Application of Clinical Proteomics to Cancer and other Diseases", Clin Chem Lab Med. 2003;41(12):1562–70.
13. Borchers C, Chen T, Neamati N. "Application of Proteomics in Basic Biological Sciences and Cancer", Molecular Carcinogenesis and the Molecular Biology of Human Cancer,chapter12,26pages.
14. Ardekani AM, Liotta LA, Petricoin EF. "Clinical potential of proteomics in the diagnosis of ovarian cancer". Expert Rev Mol Diagn. 2002;2(4):312-20.
15. Maes E, Mertens I, Valkenborg D, Pauwels P, Rolfo C, Baggerman G. "Proteomics In cancer research:Are We ready for clinical practice? ", 2015;96(3):437-48.
16. Krüger T, Lautenschläger J, Grosskreutz J, Rhode H. "Proteome analysis of body fluids for amyotrophic lateral sclerosis biomarker discovery", Proteomics Clin Appl. 2013;7(1-2):123-35.
17. Robert Grützmann,"Application of Proteomics in Cancer Biomarker DiscoveryCancer,Gene Profiling: Methods and Protocols", Methods in Molecular Biology, vol. 1381.
18. Leth-Larsen R, Lund RR, Ditzel HJ. "Plasma membrane proteomics and its application in clinical cancer biomarker discovery", Mol Cell Proteom 2010;9:1369-82.
19. Honda K, Ono M, Shitashige M, Masuda M, Kamita M, Miura N, et al. "Proteomic approaches to the discovery of cancer biomarkers for early detection and personalized medicine",Japan J Clin Oncol. 2013;43:103-9.
20. Wulfkuhle JD, Liotta LA, Petricoin EF,"Proteomic applications for the early detection of cancer", Nat Rev Cancer. 2003;3:267-75.
21. Ludwig JA, Weinstein JN. "Biomarkers in cancer staging, prognosis and treatment selection", Nat Rev Cancer. 2005;5:845-56.

employing the most effective techniques and also as a measurable factor between human populations (22, 23). In fact, mutation of genes and changes that occur in the transcription and translation of proteins can also be defined as cancer biomarkers. Even changes in serum proteins with the cancer process can be described as biomarkers of that cancer (24, 25). To analyze a cancer, identifying only one biomarker does not provide enough information for that cancer, but paying attention to changes in the level of expression of various proteins is valuable (20, 26, 27). One of the methods employed in the field of proteomic biomarkers is the two-dimensional electrophoresis and mass spectrometry. Although spectrometry is an ideal method for identifying biomarkers, But additional testing, along with electrophoresis, may also be helpful to verify the performance of mass spectrometry, including tests (LCM, Protein Arrays, Antibody Arrays and High density) (28, 11). By 2006, only 12 different types of biomarkers have been approved by the World Organization (FDA), The use of other techniques to determine the correct biomarkers from the protein of healthy, cancerous and malignant cells is necessary.

Here in, we discuss three major technologies presently available in proteome analysis. First the separation of protein mixtures by 2-DE, which is the only technique currently available that can reveal hundreds of proteins at a time. Second, characterization of every single protein for which mass spectrometric (MS) technique is the methodology of choice in this regard. Third, identification of the proteins or polypeptides by MS followed by database searching using numerous computational algorithms is nowadays becoming routinely accessible .Mass spectrometry plays a fundamental role in the majority of protein identification pipelines used in proteomics. Individual proteins can typically be “identified” either by a single stage of mass spectrometry, called peptide mass fingerprinting (PMF), or by two stages, called tandem MS or MS/MS. If a single stage is used, a protein is digested (for example with trypsin) and the pattern of peptide masses (PMF) is compared with a theoretical digest of database sequences to make an identification (13, 10).

Text mining

References

22. Srinivas PR, Verma M, Zhao Y, Srivastava S. "Proteomics for cancer biomarker discovery". Clin Chem. 2002;48:1160-9.

23. Phan JH. "Biomarker discovery and clinical outcome prediction using knowledge based bioinformatics", Georgia Institute of Technology Emory University May 2009

24. Chaerkady R, Pandey A. "Quantitative proteomic for identification of cancer biomarkers", Proteomics Clin Appl. 2007;1(9):1080-9.

25. Srinivas PR, Kramer BS, Srivastava S, "Trends in biomarker research for cancer detection. Lancet Oncol", 2001;2 (11):698-12.

26. Safaei A, Rezaei-Tavirani M, Sobhi S, Akbari ME. "Breast Cancer Biomarker Discovery: Proteomics and Genomics Approaches", Iran J Cancer Prevent. 2013;6:45-53

27. Rezaei-Tavirani M, Zali H, Jazii FR, Heidari MH, Hoseinzadeh-Salavati B, Daneshi-Mehr F, et al. "Introducing aldolase C as a differentiation biomarker:A proteomics approach", J Para med Sci 2010;1.33-39

28.Carol E. Parker, Christoph H. Borchers,"Mass spectrometry based biomarker discovery, verification, and validation e Quality assurance and control of protein biomarker assays",molecular oncology 8(2014)840-858

Text mining, also referred to as text data mining, roughly equivalent to text analytics, is the process of deriving high-quality information from text. High-quality information is typically derived through the devising of patterns and trends through means such as statistical pattern learning. Typical text mining tasks include text categorization, text clustering, concept/entity extraction, production of granular taxonomies, sentiment analysis, document summarization, and entity relation modeling. text mining seeks to extract useful information from non-structured text data by detecting and displaying patterns. In other words, text mining is a way of extracting knowledge from texts. The discovery of new and previously unknown information is based on the automatic extraction of information from various sources of writing. Text analysis involves information retrieval, lexical analysis to study word frequency distributions, pattern recognition, tagging/ annotation, information extraction, data mining techniques including link and association analysis, visualization, and predictive analytics. The overarching goal is, essentially, to turn text into data for analysis, via application of natural language processing (NLP) and analytical methods. A range of text mining applications in the biomedical literature has been describe (29).One online text mining application in the biomedical literature is Pub Gene that combines biomedical text mining with network visualization as an Internet service (30, 31). Go PubMed is a knowledge-based search engine for biomedical texts. unstructured biomedical text is of great value for cancer diagnostics, treatment, and prevention. The immense body and rapid growth of biomedical text on cancer has led to the appearance of a large number of text mining techniques aimed at extracting novel knowledge from scientific text. Biomedical text mining on cancer research is computationally automatic and high-throughput in nature. However, it is error-prone due to the complexity of natural language processing (32).

1) Text mining in cancer research: Many studies have been done on text mining in cancer research. We have reviewed the important research issues related to text mining in the biomedical field. Hui Li and Chunmei Liu (33) identified biomarkers using Text mining. In particular, DNA methylation is one of the hottest topics used for early diagnosis and treatment of cancer (34, 35). Different databases Including Pub Meth (36) and MeInfoText (32, 33)

References

29. Cohen, K. Bretonnel; Hunter, Lawrence, "Getting Started in Text Mining", PLoS Comput Biol. 2008 Jan;4(1):e20.
30. Jenssen TK, Lægreid A, Komorowski J, Hovig E,"A literature network of human genes for high-throughput analysis of gene expression", Nat Genet. 2001;28(1):21.
31. Masys DR. "Linking microarray data to the literature", Nat Genet. 2001;28(1):9.
32. Zhu F, Patumcharoenpol P, Zhang C, Yang Y, Chan J, Meechai A, et al. "Biomedical text mining and its applications in cancer research", J Biomed Inform. 2013;46(2):200-11.
33. Li H, Liu C. "Biomarker Identification Using Text Mining" 2012;2012:135780.
34. Fang YC, Huang HC, Juan HF,"MeInfoText: associated gene methylation and cancer information from text mining", BMC Bioinformatics. 2008;9(1):22.
35. Fang YC, Lai PT, Dai HJ, Hsu WL," MeInfoText 2.0: gene methylation and cancer relation extraction from biomedical literature", BMC Bioinformatics. 2011;12(1):471.
36. Ongenaert M, Van Neste L, De Meyer T, Menschaert G, Bekaert S, Van Criekinge W, "PubMeth: a cancer methylation database combining textmining and expert annotation", Nucleic Acids Res. 2008;36:D842–6.

have been developed for DNA methylation .Pub Meth (34) is a cancer methylation database with text mining tools and expert annotations. Associations among genes, methylation, and cancers in MeInfoText (34, 35) are extracted from a large body of biomedical literature. As a complex disease, cancer is related to a large number of genes and proteins. Biomedical researchers are interested in mining cancer-related genes and proteins from the literature to study cancer diagnostics, treatment, and prevention. Deng et al, (37) employed a text mining approach to identify prostate cancer-related genes as candidate genes and they used the OMIM (Online Mendelian Inheritance in Man) database to verify them. Natarajan et al, (38) also built gene–gene interaction networks for 72 genes using a text mining approach. Krallinger et al. implemented two cancer-related text mining applications (39). These two methods include 1-extraction of human gene mutations for predefined types of cancer from literatures and 2-breast cancer categorization and text-based breast cancer gene ranking. Korhonen et al, (40) applied biomedical text mining technology to cancer risk assessment. They extracted evidence from the literature as features and developed several classes for risk levels of the causes of cancer, from which researchers can acquire the risk levels. Guo et al, (41) developed classifiers for the automatic identification of schemes from abstracts to help cancer risk assessment to facilitate the study of cancer and other potential factors. Ben Abacha et al, (42) used a supervised machine learning approach to extract the relationships among medical problems, treatments, and tests. Lee et al, (43) developed a text mining-based system to discover the relationships among cancer and potential factors.

They mined relationships among diseases and potential factors in clinical medical records using self-organizing maps while they used SVM to evaluate them. There are several case studies in the context of text mining resources for cancer research. We introduced the general workflow of

References

37. Deng X, Geng H, Bastola DR, Ali HH. "Link test – a statistical method for finding prostate cancer biomarkers", Comput Biol Chem. 2006;30:425–33.

38. Natarajan J, Berrar D, Dubitzky W, Hack C, Zhang Y, DeSesa C, et al. "Text mining of full-text journal articles combined with gene expression analysis reveals a relationship between sphingosine-1-phosphate and invasiveness of a glioblastoma cell line", BMC Bioinformatics 2006;7:373.

39. Krallinger M, Leitner F, Valencia A,"Analysis of biological processes and diseases using text mining approaches", Methods Mol Biol. 2010;593:341–82.

40. Korhonen A, Silins I, Sun L, Stenius U,"The first step in the development of TextMining technology for Cancer Risk Assessment: identifying and organizing scientific evidence in risk assessment literature", BMC Bioinformatics. 2009;10:303.

41. Guo Y, Korhonen A, Liakata M, Silins I, Hogberg J, Stenius U. A "comparison and user-based evaluation of models of textual information structure in the context of cancer risk assessment", BMC Bioinformatics. 2011;12:69.

42. Abacha AB, Zweigenbaum P,"Automatic extraction of semantic relations between medical entities: a rule based approach", J Biomed Semantics. 2011;2(Suppl. 5):S4.

43. Lee CH, Wu CH, Yang HC. "Text mining of clinical records for cancer diagnosis", Second international conference on innovative computing, informatio and control: IEEE computer society; 2007.

text mining to support cancer systems biology. We can see that text mining has been used widely in cancer research. In this study, we use the sentence and word search method, which is a text mining algorithm, to identify the biomarkers. finally, we collected the cancer associated biomarkers.

Biomarker Identification:

In this section according to numerous studies we have identified proteins with high potential that can be used in recognition process and becoming a biomarker. Therefore, the proteins shown in **Table1,2,3** are a set of proteins that express their expression significantly different in either healthy or cancerous tissue. This outstanding difference in the potential expression levels is the key factor to point out biomarkers. Identifying these effective proteins for cancer, and then their 3-D structure, can play a major role in drug design and early diagnosis of the disease. **Table1,2,3** list the biomarkers of breast, lung and colorectal cancer respectively.

Table 1: Identified biomarker candidates in Breast cancer

Type of cancer	Protein name	Condition in cancerous tissues	Reference
Breast cancer	ALDOA a-b; G3P a-e; KPYM a-b; PGK1 a-c; ENOG; S10A2; S10AB; S10AB a-b; S10AD; S10AG; PRS10; PRS6A; PRSA6B; PRS 7-8; PRSA 1-7; PSB3; PSB6; PSME1; TERA a -b; ANXA1 a; ANXA5; ANXA5 b; CATA a; CH60 a-c; COF1 a; COR1A; DDAH2; ENPL ; GDIR1; GRP78; HSPB1 c-f; LGUL ; PRDX2 a-b; PRDX3; RL40 a-b SODC b; SODM a-b; TCTP; THIO a-b; VDAC2;	Increase in the tumoral tissues	44
	GBR7	Over-expression in tumor	45
	PAK4		46
	HER-2		47,48,49
	P27	Reduced expression in tumor	47,48,49
	KI-67	only expresses in growing and proliferating cells	50
	BRCA1	Loss of expression is a marker of tumor aggressiveness	51
	GRHL3	expression appeared to decrease with tumor progression	52
	BIGH3	benign tissue had a 23-fold increase in expression compared to the infiltrating colloid carcinoma	53
	Serotransferrin	High levels in serum of cancerous	54,55
	Isoform 1 of Gelsolin	decreased in breast cancer	54,55
	vitronectin	lower in 2breast cancer	54,55
	DCC	decrease expression in tumor	56

Table 2: Identified biomarker candidates in TBL cancer

Type of cancer	Protein name	Condition in cancerous tissues	Reference
TBL cancer	Tubulin □-1B; laminin B1; COTL-1; γ-actin; Carbonic anhydrase1; ubiquitin-conjugating enzyme E2 - 25K; Ubiquitin carboxyl terminal esterase L1	Up-regulated more than 2fold in SCLC	57
	BCA-1; IGFBP-2; MAPK13; MMP-12; MMP-7; NAGK; VEGF; YES;	Up-regulated in NSLCC serum and tissue	58
	Cadherin-1	down-regulated in NSCLC serum and tissue	58
	Catalase; Endostatin; MRC1	Up-regulated in NSCLC serum and down-regulated in tissue	58
	RALA; Arf6	decreased expression in tumor	59
	SURVIVIN	increase in the tumoral tissues	59
	E2F1	higher in tumor samples than in non-tumor lung specimens	60
	RBBP6	۱,۶fold higher in Lung tumors compared with normal lung tissue	61
	P53	۲,۹fold lower in Lung tumors compared with normal lung tissue	61
	HE4	High levels in serum of lung cancer patients	62
	SIRT3	higher in NSCLC tissue than in adjacent tissue	63

References

44. Pucci-Minafra I, Di Cara G, Musso R, Cancemi P, Albanese NN, Roz E, et al. "Retrospective Proteomic Screening of 100 Breast Cancer Tissues", Proteomes. 2017;5(3):15.

45. Ramsey B, Bai T, Newell AH, Troxell M, Park B, Olson S, et al. "GRB7 protein over-expression and clinical outcome in breast cancer", Breast Cancer Res Treat. 2011;127(3):659-69.

46. Minden A. "The Pak4 Protein Kinase in Breast Cancer", ISRN Oncol. ,2012;2012:694201.

47. Newman L, Xia W, Yang HY, Sahin A, Bondy M, Lukmanji F, et al. "Correlation of p27 Protein Expression With HER-2/neu Expression in Breast Cancer", Mol Carcinog. 2001;30(3):169-75.

48. Slamon DJ, Clark GM, Wong SG, Levin WJ, Ullrich A, McGuire WL. "Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu", oncogene Science. 1987;235(4785):177-82.

49. Porter PL, Malone KE, Heagerty PJ, Alexander GM, Gatti LA, Firpo EJ, et al. "Expression of cell-cycle regulators p27Kip1 and cyclin E,alone and in combination,correlate with survival in young breast cancer patients", Nat Med. 1997;3:222-5

50. Kheirandish S, Homaee F. "Ki67 protein:a proliferation index in breast cancer", Rev Clin Med. 2015;2(4):201-4.

Table 3: Identified biomarker candidates in Colorectal cancer

Type of cancer	Protein name	Condition in cancerous tissues	Reference
colorectal cancer	SPARC	low expression or absence of stromal SPARC in tumor	64
	nm23; P53; C-erbB-2; U-PA; VEGF;	higher expression in tumor	65
	$P33^{ING1B}$	lower in tumor samples compared with the normal adjacent samples	66
	UCP2	expression is increased in most human colon cancers	67
	PEBP4	higher expression in colorectal cancer than the normal peri carcinoma tissues	68
	UNC5C	lower in colorectal cancers Compared with the corresponding normal tissues	69
	CCL28	level in colon tumors lower than in normal tissue	70
	FABP5; RPS6; cyclophilin A; LDHA; HSC70; RPS6	Upregulated in cancer	71
	LDHB; galectin-١; GAPDH; cystatin-B; RAN; thioredoxin	Downregulation in cancer tissue	71

References

51. Madjd Z, Karimi A, Molanae S, Asadi-Lari M. BRCA1 protein expression level and CD44+ phenotype in breast cancer patients. Cell J. 2011;13(3):155.

52. Xu H, Liu C, Zhao Z, Gao N, Chen G, Wang Y, et al. "Clinical implications of GRHL3 protein expression in breast cancer", Tumor Biol. 2014,35:1827-31

53. Calaf GM, Echiburú-Chau C, Zhao YL, Hei TK, "BigH3 protein expression as a marker for breast cancer", Int J Mol Med. 2008;21(5):561-8.

54. Meiqun C, Zifan G, Kehuan S, Zhengzhi W," Application of iTRAQ Quantitative Proteomics in Identification of Serum Biomarkers in breast cancer", 2011 4th International Conference on Biomedical Engineering and Informatics,978-1-4244-9352-2011 IEEE

55. Asch HL, Head K, Dong Y, Natoli F, Winston JS, Connolly JL, et al. "Widespread loss of gelsolin in breast cancers of humans, mice, and rats" Cancer Res. 1996;56(21):4841-5.

56. Koren R, Dekel Y, Sherman E, Weissman Y, Dreznik Z, Klein B, et al. "The expression of DCC protein in female breast cancer. Breast Cancer" Res Treat. 2003;80(2):215-20.

57. Knizhnik AV, Kovaleva OV, Laktionov KK, Mochalnikova VV, Komelkov AV, Tchevkina EM, et al. "Arf6,RalA and BIRC5 Protein Expression in Nonsmall Cell Lung Cancer", Mol Biol (Mosk). 2011;45(2):307-15.

58. Hung JJ, Hsueh CT, Chen KH, Hsu WH, Wu YC,"Clinical significance of E2F1 protein expression in non-small cell lung cancer", Exp Hematol Oncol. 2012,1:18

Conclusion

Because the identifying only one biomarker does not provide enough information for that cancer, we tried to collect a significant number of biomarkers. therefore, using intelligent techniques such as text mining, and we could collect biomarkers that have a protein base and have been introduced in various articles as biomarkers. In fact, effective proteins for breast, lung and colorectal cancers, have been identified and classified. By identifying several biomarkers in the patient's body, early diagnosis of cancer will be possible. Therefore, biomarkers play an important role in the treatment of cancer.

References

59. Knizhnik AV, Kovaleva OV, Laktionov KK, Mochalnikova VV, Komelkov AV, Tchevkina EM, et al. "Arf6,RalA and BIRC5 Protein Expression in Nonsmall Cell Lung Cancer", Mol Biol (Mosk). 2011;45(2):307-15.
60. Hung JJ, Hsueh CT, Chen KH, Hsu WH, Wu YC,"Clinical significance of E2F1 protein expression in non-small cell lung cancer", Exp Hematol Oncol. 2012,1:18
61. Motadi LR, Bhoola KD, Dlamini Z. "Expression and function of retinoblastoma binding protein 6(RBBP6)in human lung cancer", Immunobiology. 2011;216(10):1065-73.
62. Zhong H, Qian Y, Fang S, Yang L, Li L, Gu W. "HE4 expression in lung cancer,a meta-analysis", Clinica Chimica Acta. 2017;470:109-14.
63. Yang GC, Fu BC, Zhang DY, Sun L, Chen W, Bai L, et al. "The Expression and Related Clinical Significance of SIRT3 in NSCLC", Dis Markers. 2017;2017:8241953.
64. Liang JF, Wang HK, Xiao H, Li N, Cheng CX, Zhao YZ, et al. "Research Relationship and prognostic significance of SPARC and VEGF protein expression in colon cancer", J Exp Clin Cancer Res. 2010;29:71.
65. Berney CR, Fisher RJ, Yang JL, Russell PJ, Crowe PJ. "Protein Markers in Colorectal Cancer: Predictors of Liver Metastasis", Ann Surg. 1999;230(2):179-84.
66. Fallahnezhad S, Nikbakht M, Shokri S. "Expression of P33ing1b Protein in Colorectal Cancer", Middle East J Dig Dis. 2016;8(1):44-50.
67. Horimoto M, Resnick MB, Konkin TA, Routhier J, Wands JR, Baffy G. "Expression of Uncoupling Protein-2 in Human Colon Cancer", Clin Cancer Res. 2004;10(18):6203-7.
68. Liu H, Kong Q, Li B, He Y, Li P, Jia B. "Expression of PEBP4 protein correlates with the invasion and metastasis of colorectal cancer",Tumor Biol. 2012;33(1):267-73.
69. Wu J, Wang G, He B, Chen X, An Y. "Methylation of the UNC5C gene and its protein expression in colorectal cancer", Tumor Biol. 2017;39(4):1010428317697564.
70. Dimberg J, Hugander A, Wagsater D. "Protein expression of the chemokine,CCL28,in human colorectal cancer", Int J Oncol. 2006;28(2):315-9.
71. Koshiyama A, Ichibangase T, Imai K. "Comprehensive fl orogenic derivatization liquid chromatography/tandem mass spectrometry proteomic analysis of colorectal cancer cell to identify biomarker candidate", Biomed Chromatogr. 2013;27:440-50