کارگاه‌های آموزشی مرکز اطلاعات علمی

مقاله نویسی علوم انسانی

اصول تنظیم قرارداده

آموزش مهارت های کاربردی در تدوین و چاپ مقاله
Cluster and Principal Component Analysis of Human Glioblastoma Multiforme (GBM) Tumor Proteome

Mehdi Pooladi¹,2, Mostafa Rezaei-Tavirani¹, Mehrdad Hashemi³, Saeed Hesami-Tackallou⁴, Solmaz Khaghanı-Razi-Abad¹,2, Afshin Moradi⁵, Ali Reza Zali⁶, Masoumeh Mousavi¹, Leila Firozi-Dalvand², Azadeh Rakhshan⁵, Mona Zamanian Azodi¹

Abstract

Background: Glioblastoma Multiforme (GBM) or grade IV astrocytoma is the most common and lethal adult malignant brain tumor. Several of the molecular alterations detected in gliomas may have diagnostic and/or prognostic implications. Proteomics has been widely applied in various areas of science, ranging from the deciphering of molecular pathogen nests of discuses.

Methods: In this study proteins were extracted from the tumor and normal brain tissues and then the protein purity was evaluated by Bradford test and spectrophotometry. In this study, proteins were separated by 2-Dimensional Gel (2DG) electrophoresis method and the spots were then analyzed and compared using statistical data and specific software. Protein clustering analysis was performed on the list of proteins deemed significantly altered in glioblastoma tumors (t-test and one-way ANOVA; P< 0.05).

Results: The 2D gel showed totally 876 spots. We reported, 172 spots were exhibited differently in expression level (fold > 2) for glioblastoma. On each analytical 2D gel, an average of 876 spots was observed. In this study, 188 spots exhibited up regulation of expression level, whereas the remaining 232 spots were decreased in glioblastoma tumor relative to normal tissue. Results demonstrate that functional clustering (up and down regulated) and Principal Component Analysis (PCA) has considerable merits in aiding the interpretation of proteomic data.

Conclusion: 2D gel electrophoresis is the core of proteomics which permitted the separation of thousands of proteins. High resolution 2DE can resolve up to 5,000 proteins simultaneously. Using cluster analysis, we can also form groups of related variables, similar to what is practiced in factor analysis.

Keywords: Glioblastoma; Glioma; Proteomics; Cluster; 2DG electrophoresis


Introduction

Glioblastoma Multiforme (GBM) or grade IV astrocytoma is the most common and lethal adult malignant brain tumor [1-3]. Malignant gliomas are the most common human primary brain tumors, GBM being the most aggressive and lethal form [4, 5], with a median survival of only 14 months even with the best available treatments. GBMs are characterized by their resistance to radiotherapy and chemotherapy, as well as their abundant and aberrant vasculature which is the most aggressive and devastating grade [6-8]. This was a remarkable observation at that time, as late as 1979; the World Health Organization (WHO) did not consider the glioblastoma as an astrocytic tumor, listing it instead in a group of poorly different and embryonic tumors [9]. With the introduction of immunohistochemistry, the glioblastoma was firmly categorized as astrocytic neoplasm [10], but the separation of primary and secondary glioblastoma remained conceptual [11, 12]. Growing tumor reaches a point at which the existing blood supply can no longer

Vol 7, No 2, Spring 2014
support the needs of the tumor leading to areas of hypoxia. In response to this hypoxia, GBMs undergo an antigenic switch, and increase secretion of various growth factors to promote new blood vessel formation [13-15]. In most cases diagnosis is only done after clinical symptoms become apparent and depends on histological investigations of tumor samples obtained by biopsy or resection [16-19].

Understanding the different biological characteristics of these heterogeneous cell subpopulations within a single tumor is invaluable because it may provide not only clues regarding tumor pathogenesis but also potential targets for treatment [20-23]. Knowledge about the molecular biology of cancer, including CNS tumors, continues to increase [24]. Several of the molecular alterations detected in glioblastoma have diagnostic and/or prognostic implications, as they are associated with histologically defined tumor type or malignancy grade. However, for most of the molecular changes this does not justify a designation as glioma biomarker [25, 26]. In fact, proteomics has been widely applied in various areas of science, ranging from the deciphering of molecular pathogen nests of discuses, to the discovery of potential diagnostic and prognostic biomarkers, where the technology is able to identify and quantify proteins associated with a particular disease by means of their altered levels of expression and post translational modification between the control and disease states [26-32].

In the present study, we investigated a change in protein expression in human brain glioblastoma tumor to get an understanding of data and specific software molecular diagnosis of glioblastoma. Here, proteins of tumoral and normal brain tissues were extracted and evaluated by proteomic tools (2DG). After providing cluster and PCA of spots, their alterations are monitored using statistical data and specific software. Using different proteomic approaches, multiple differentially expressed glioblastoma proteins were identified, a few of which could be investigated further as potential surrogate markers for glioblastoma tumors.

Figure 1. Red spots indicate Up Regulation and blue spots indicate Down Regulation in glioblastoma tumor relative to normal tissue. Two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) of glioblastoma tumor; 2D-PAGE was performed to separate proteins by isoelectric focusing (IEF) (pH: 3–10) in the first dimension and SDS-PAGE (SDS: Sodium dodecyl sulfate) (MW 10–100 kDa) in the second dimension; the gel was stained with Coomassie Blue dye.

Figure 2. It shows chart up-regulation spots marked in figure 1.
Materials and Methods

Patient samples

Tissues were obtained, with informed consent and institutional review board approval, from patients undergoing tumor resection. For this study, all individuals filled a written informed consent form and glioblastoma tumors were surgically removed at Shohada Tajrish Hospital. The tumors were classified by a team of neuropathologists according to the guidelines of the WHO classification of tumors of the central nervous system. Non-tumoral brain tissues were obtained from normal areas (either grey or white matter) of brain tissues removed from patients undergoing non-tumor epileptic surgery.

Tissue and samples preparation

Tissue samples of both tumoral and normal brain tissues were snap-frozen immediately after operation in liquid nitrogen and stored at -80°C until used for proteomic analysis. To obtain tissue extracts, the samples were broken into suitable pieces and were homogenized in lysis buffer II consisting of lysis buffer I {7 M urea, 2 M thiourea, 4% 3-(3-cholamidopropyl) dimethylammonio]-1-propanesulfonate (CHAPS), 0.2% 100×Bio-Lyte 3/10}, dithiothreitol (DTT), and 1 mM ampholyte and protease inhibitor on ice. Cell lysis was completed by subsequent sonication (4×30 pulses). The samples were then centrifuged at 20,000 g at 4°C for 30 minutes to remove insoluble debris. The supernatants were combined with 100% acetone and centrifuged at 15,000 g, and then the supernatants were decanted and removed (3 times). Acetone 100% was added to the protein precipitant and kept at -20°C overnight. The samples were then centrifuged again at 15,000 g and the precipitant was incubated 1 hour at room temperature. The protein samples were dissolved in rehydration buffer [8 M urea, 1% CHAPS, DTT, ampholyte pH (4) and protease inhibitor]. Protein concentrations were determined using the Bradford test and spectrophotometry method, and the protein extracts were then separated and used for 2D gel electrophoresis.

2DG electrophoresis

The isoelectric focusing for first-dimensional electrophoresis was performed using 18 cm, pH 3–10 Immobilized PH Gradient (IPG) strips. The samples were diluted in a solution containing rehydration buffer, IPG buffer, and DTT to reach a final protein amount of 500µg per strip. The strips were subsequently subjected to voltage gradient as described in the instructions of the manufacturer. Once focused, the IPG strips were equilibrated twice for 15 minutes in equilibration buffer I [50 mM Tris-HCl (pH: 8.8), 6 M urea, 30% glycerol, 2% Sodium Dodecyl Sulfate (SDS), and DTT] and equilibration buffer II. The second-dimension SDS-PAGE was carried out using 12% PAGEs. Following SDS-PAGE, the gels were stained using the Coomassie Blue method overnight (Figure 1).

Image analysis

The gel images were analyzed by Progenesis Same Spots software to identify spots differentially expressed between tumor and control samples based on their volume and density. The spots were carefully matched individually and only spots that showed a definite difference were defined as altered.

Data statistics analysis

The Student’s t-test was used to rank proteins found altered in glioblastoma tumor compared to normal tissue according to statistical probability. The t-test was chosen to create a hierarchy because it is easily understood by many different target audiences and is currently a common practice in the majority of proteomics analysis. Protein clustering analysis was performed on the list of proteins deemed significantly altered in glioblastoma tumors (p<0.05).

Arithmetic cluster analysis was performed for two groups. Arithmetic cluster analysis employs correlation analysis to define if alterations in the levels of one individual protein are associated with alterations in the levels of a second protein across all samples (glioblastoma and normal tissues). Arithmetic correlation algorithms are integral to the Progenesis Same Spots software (Nonlinear Dynamics v 3.0, 2008). Multiple areas on correlation coefficients between protein features were calculated by Progenesis Same Spots and the information visually represented in the form of a dendrogram.

Results

Two-dimensional gel was used to identify proteins expressed in glioblastoma tumor and non-tumor samples. The spots were separated according to their isoelectric pH and molecular weights. On each analytical 2D gel, an average of 876 spots corresponds to protein with nonlinear Progenesis Same Spots software. The representative set of overlaid 2D-DIGE (DIGE: Difference gel electrophoresis) images is given in Figure 1.
The first-dimension analysis was performed with a broad pH range (pH: 3–10) and IPG using strips of 18 cm. The total number of protein features was matched and analyzed between gels in the control group and tumor group; 420 spots (around 48% of the entire detected spots) were matched across all the gels. In software analysis, a total of 420 differentially expressed spots satisfied the statistical parameters (t-test and one-way ANOVA; p<0.05).

Among them, 172 spots exhibited difference in expression level (fold >2). A total of 420 spots showed statistically significant differences (student’s t-test; p<0.05) in gel, out of which 188 spots exhibited up regulation in expression level, whereas the remaining 232 spots were decreased in glioblastoma tumor relative to normal tissue. Up regulation is shown as red and down regulation as blue in imaging gel [Figures 1]. Of the 188 up regulated spots, 106 spots were between 1.1 and 2 fold, 49 spots were between 2 and 4 fold, and 33 spots exhibited over four fold increase in expression level (showed in Figure 2 and 3). Of the 232 down regulated spots, 142 spots were between 1.1 and 2 fold, 67 spots were between 2 and 4, and 23 spots exhibited over four fold reduction in expression level (showed in Figure 4 and 5).

**Cluster analysis**

The total number of protein features matched and analyzed between tumor and normal tissues.
Spots were analyzed by Nonlinear Prognosis Same Spots software. Arithmetic cluster analysis was performed on this list of 420 spot proteins. Arithmetic cluster analysis explores how one individual protein level correlates with a second individual protein level across different samples. Protein levels that are tightly correlated suggest that the proteins might be regulated or involved in the same biological pathway. A clear cluster analysis (dendrogram) with several distinct subgroups of proteins was generated (Figure 6).

Two main groups reflected the 188 spot proteins that increased (red) and 232 spot proteins that decreased (blue) in expression level in glioblastoma relative to normal tissue.

The total down-regulated protein spots showed two main subgroups (subgroups I and II); subgroup II involved two branches (blue in Figure 6), and the total up regulated spot proteins showed two main subgroups, (subgroup III and IV); subgroup IV involved two branches (red in Figure 6).

PCA was performed on all the spot proteins and it showed two main groups (up and down regulated) (Figure 7).

**Discussion**

Proteomics combines technologies from several disciplines in an attempt to explain the structural, functional and interactive proteins in cells, tissues and body fluids. The ultimate goal is to identify the interactive pathways of proteins for diagnostic purposes or as candidates for intervention [33-36]. Proteomics analyses are important since normal, up-regulated, down-regulated and mutated genes may not be transcribed for a large number of epigenetic reasons [37-39]. The considerable post-translation modification of proteins that alter both structure and function are revealed with proteomics add genomics [40-42]. Proteomics studies rely heavily on a number of different techniques (sample processing, 2D gel electrophoresis, MS, bioinformatics, and biostatistics) that enable the identification and
quantitation of thousands of proteins. Originally, at the core of proteomics was 2D gel electrophoresis which permitted the separation of thousands of proteins (spots) [42, 43].

High resolution 2DE can resolve up to 5,000 proteins simultaneously (~ 2000 proteins routinely), and detect and quantify < 1 ng of protein per spot. Today’s 2DE technology with IPGs has largely overcome the former limitations of carrier impolite based 2DE with respect to reproducibility, handling, resolution, and separation of very acidic or alkaline proteins [44, 45]. In our study, 91% of the proteins in the acid region, and 9% in the alkaline region, have been separated by two-dimensional electrophoresis.

This technique have some advances: complements and extends genomic analysis, protein expression rather than predicted message translation, posttranslational modification can be detected, ideal for target discovery and biomarker identification and amenable to automation [46, 47]. Until now, a pattern of protein expression using 2D-PAGE separation has provided the best repertoire of the cell or body fluid. However, this technology is limited in being low-throughput, labor intensive, time consuming, membrane proteins difficult to solubilize, high and low molecular mass proteins not well represented, statistics, dimensionality and also is problematic in detecting proteins that are basic in charge or smaller than 10,000 Da [48, 49].

Biostatistics is essential to ensure the collection of robust and meaningful data that results withstand the most rigorous statistical analyses at the level of the resulting clinical/analytical matrix [50, 51]. This includes the determination of both false positive and false negative rates, which are critical for evaluating the success of the biomarker [52, 53]. Better biomarkers are urgently needed for cancer (for example: GBM) detection, diagnosis and prognosis [27]. Our global protein difference was identified between normal tissue and GBM tumor samples. This study evaluates the effect of the global methodology on the clustering result of the GBM tumor. The effect of the clustering algorithm was also important, but a significant interaction was observed between clustering algorithm and data set [54, 56]. The term cluster analysis does not identify a particular statically method or model, as discriminant analysis, factor analysis, and regression do. Using cluster analysis, we can also form groups of related variables, similar to what is done in factor analysis. Interpretation has been restricted to assigning the proteins to a broad functional class (for example: brain tumor) with anecdotal discussion of how few of these proteins are pertinent to glioma. No systematic attempt has been made to determine if some or all of these proteins are altered as part of an interval response in glioma. Many proteins are implicated in numerous biological processes [57, 58]. For example, alteration of Hsp70 expression was detected in astrocytomas which has an important role in tumor survival and development, response cell stress and induces apoptosis in cells [59].

As a consequence it is naive to simply perform basic literature searches on individual proteins in a

Figure 7. PCA was performed on all the spot proteins; A: the PCA of 188 spot proteins was up-regulated, B: the PCA of 232 spot proteins were down-regulated in glioblastoma tumor.
protein list and then draw conclusions as to how these proteins contribute to the biological nature of the disease of interest [60, 61]. In analysis with normal variable distribution, PCA might provide a better solution to the data reduction issue. This technique examines underlying latent structures in a data set of variables. This may be particularly appropriate for proteomic research as the proteins under investigation are not expressed indecently from each other [62]. Power is crucial for understanding the statistical relevance of results generated and awareness of power should help increase the integrity of proteomic data reported in the literature [63].

Clustering analysis of the expression of proteins can be divided into two main clusters indicating that there are clusters of proteins with similar expression that these proteins can provide similar performance in terms of testing or indicating its presence in the same biological pathways. PCA analysis confirmed that the clustering results showed that the protein has been classified according to the test conditions. Finally, the results indicated that by using the statistical analysis software it is possible to quickly and easily demonstrate that a significant expression changes has been induced in the brain cancer (GBM) and Stages of malignancy on proteome level.

Acknowledgment
This study was supported by Proteomics Research Center, Faculty of Paramedical Sciences, Shahid Beheshti University of Medical Sciences, Tehran-Iran and contributed by Department of Pathology, Shohada Hospital, Shahid Beheshti University of Medical Science, Tehran, Iran and Department of Neurosurgery, Shohada Hospital, Shahid Beheshti Medical University, Tehran, Iran.

Authors’ Contribution
Mehdi Pooladi, Mostafa Rezaei-Tavirani, Mehrdad Hashemi, Saeed Hesami-Tackallou, and Solmaz Khaghani Razi Abad designed the study, colected and analyzed the data and wrote the paper. Masoumeh Mousavi, Leila Firozi Dalvand and Mona Zamanian Azodi contributed to study design. Afshin Moradi, Ali Reza Zali, Azadeh Rakhsheh and Mehdi Pooladi contributed to samples collection and indentation.

References


کارگاه‌های آموزشی مرکز اطلاعات علمی

مقاله نویسی علوم انسانی

اصول تنظیم قراردادها

آموزش مهارت های کاربردی در تدوین و چاپ مقاله