Principal component analysis on LC-MS/MS and 2DE-MALDI-TOF in glioblastoma cell lines reveals that mitochondria act as organelle sensors of the metabolic state in glioblastoma

LEOPOLDO GÓMEZ-CAUDILLO¹, ARIADNA J. ORTEGA-LOZANO¹, ÁNGEL G. MARTÍNEZ-BATALLAR¹, HAYDEE ROSAS-VARGAS², FERNANDO MINAURO-SANMIGUEL² and SERGIO ENCARNACIÓN-GUEVARA¹

¹Center for Genomic Sciences, National Autonomous University of Mexico, Cuernavaca, Morelos 62210;
²Medical Research Unit in Human Genetics, Hospital of Pediatrics, National Medical Center XXI Century, Mexican Social Security Institute, Mexico City 06720, Mexico

Received February 8, 2019; Accepted April 8, 2020
DOI: 10.3892/or.2020.7625

Abstract. Glioblastoma is a difficult disease to diagnose. Proteomic techniques are commonly applied in biomedical research, and can be useful for early detection, making an accurate diagnosis and reducing mortality. The relevance of mitochondria in brain development and function is well known; therefore, mitochondria may influence the development of glioblastoma. The T98G (with oxidative metabolism) and U87MG (with glycolytic metabolism) cell lines are considered to be useful glioblastoma models for studying these tumors and the role of mitochondria in key aspects of this disease, such as prognosis, metastasis and apoptosis. In the present study, principal component analysis of protein abundance data identified by liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS) and matrix-assisted laser desorption/ionization-time of flight mass spectrometry (MALDI-TOF) from 2D gels indicated that representative mitochondrial proteins were associated with glioblastoma. The selected proteins were organized into T98G- and U87MG-specific protein-protein interaction networks to demonstrate the representativeness of both proteomic techniques. Gene Ontology overrepresentation analysis based on the relevant proteins revealed that mitochondrial processes were associated with metabolic changes, invasion and metastasis in glioblastoma, along with other non-mitochondrial processes, such as DNA translation, chaperone responses and autophagy. Despite the lower resolution of 2D electrophoresis, principal component analysis yielded information of comparable quality to that of LC-MS/MS. The present analysis pipeline described a specific and more complete metabolic status for each cell line, defined a clear mitochondrial performance for distinct glioblastoma tumors, and introduced a useful strategy to understand the heterogeneity of glioblastoma.

Introduction

Pediatric solid brain tumors are the most common central nervous system neoplasia in children and the second most common in individuals <20 years old (1). Glioblastoma multiforme (GBM), or grade IV astrocytoma, is the most common and lethal adult malignant brain tumor, whereas it only occurs in 8-12% of the pediatric population (2). Nevertheless, glioblastoma in both populations is characterized by an aggressive medical behavior, as well as high mortality and morbidity rates, with an incidence of 3.19 cases per 100,000 individuals and a 5-year survival rate of 5% (3,4). GBM has a high diversity in terms of morphology, localization and genetic alterations; therefore, GBM has been poorly characterized, which makes glioblastoma difficult to diagnose (4). Understanding glioblastoma heterogeneity should be a priority for developing improved therapies and searching for novel biomarkers (5).

Mitochondria, which are the 'power houses' of the cell, are abundant in the brain. Biogenesis, mitophagy, migration and
of mitochondria in GBM. Therefore, mitochondria can affect the susceptibility of the brain to injury, and they serve a role in innate immunity, in inflammation in response to infection and acute damage, and in antiviral and antibacterial defense (6,8). As mitochondria serve critical roles in numerous bioenergetic, anabolic and cell biochemical pathways (9,10), their genetic and metabolic alterations have been suggested to be a pathogenic cause of, or contributing factor to, a broad range of human diseases, including cancer (11,12). Several common tumor cell features can result from mitochondrial dysregulation, because, the biology of mitochondria supports cell transformation during carcinogenesis (11,13,14), which suggests that the mitochondrial proteome is versatile and can sense the spatial and temporal dynamics of cell biological processes from the onset to the end of cancer. However, the specific role of mitochondria in cancer has not been completely uncovered, mainly due to the large amount of information regarding mitochondrial processes in cancer not having been properly integrated.

Proteomic analysis can be applied in GBM research for early detection, for making a reliable diagnosis and for performing an accurate risk assessment. However, Petrak et al (15), Deighton et al (16) and Valledor and Jorrín (17) agree that, despite the utility of proteomic research to obtain insights into cancer-associated biological processes and the knowledge of neuro-oncology, the glioblastoma proteome studies performed to date have focused on how proteins are up- or downregulated, and these results have been generated without any specific approach to establish the existence of key proteins and/or specific signaling pathways in cancer development or regulation. To the best of our knowledge, most of the generated data lack reproducibility, validity and comparability, mainly due to methodological and analytical constraints. The identified proteins in these studies are diverse and make it hard to understand the nature of the disease or its background.

Mitochondria compose a biological system that interacts in, with or between other living systems, and that maintains physiological associations with other subsystems in cells, such as organelles, genes and proteins. In a complex disease, the mitochondrion and its environment are altered (14). Therefore, systemic questions in the context of cancer are: How do the mitochondria interact with other components and their environment? Additionally, what are the roles of the mitochondria within the cell? (18) In this regard, analyzing protein abundance data according to fold-change, univariate statistics (such as Student's t-test or ANOVA) or analog nonparametric tests do not answer these questions due to their nature; on the other hand, multivariate statistical analysis involves ≥2 variables in ≥2 conditions simultaneously for any possible association or empirical relationship, so the description of the interactions between cellular components can be improved.

Proteomic data enclose information about the whole cellular system, bringing an improved description through multivariate statistical approaches. Additionally, principal component analysis (PCA) applied to GBM mitochondrial proteomic data could reveal that mitochondria metabolism acts as a cellular sensor of specific isolated cancer states, which could result in reliable and useful information to help improve diagnosis and risk assessment, as well as to understand the role of mitochondria in GBM.

In the present study, a proteomic functional analysis based on PCA of liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS) and 2D isoelectric focusing (IEF)/SDS-PAGE intensity data was conducted. A specific mitochondrial proteomic landscape was obtained from the glioblastoma T98G and U87MG cell lines associated with biological processes that characterize ‘oxidative’ and ‘glycolytic’ types of tumor. Additionally, the present cell model resembles the metabolic transition from mitochondrial oxidative phosphorylation (OXPHOS) to glycolysis (a process known as the Warburg effect), as reported during tumorgenesis (19). Finally, protein-protein interaction networks (PPINs) and Gene Ontology (GO) overrepresentation based on IEF revealed that LC-MS/MS and 2D IEF/SDS-PAGE analysis were comparable and complementary with each other, indicating that mitochondria may act as key sensing organelles for GBM tumor characterization and serve as valuable tools for therapeutic targets.

Materials and methods

Cell culture. The T98G (CRL-1690™) and U87MG (glioblastoma of unknown origin; HTB-14™) cell lines were purchased from the American Type Culture Collection and cultured in 175-cm² plastic flasks at 37°C with 5% CO₂ in Eagle's Minimum Essential Medium (In Vitro S. A.) supplemented with 10% fetal bovine serum (Gibco; Thermo Fisher Scientific, Inc). Cells were grown to 80-90% confluence, harvested with trypsin, washed twice in PBS and used for mitochondrial isolation.

Mitochondrial isolation. Mitochondria were isolated by differential centrifugation. Cells were separately disrupted in 250 mM sucrose, 1 mM EGTA and 10 mM HEPES (pH 7.4) at 4°C and centrifuged for 10 min at 1,500 x g and 4°C to recover the supernatant. This step was repeated three times. Subsequently, all of the supernatants were pooled and centrifuged for 10 min at 12,000 x g and 4°C to obtain a mitochondrial pellet. The pellets were used immediately or kept at -80°C until use.

LC-MS/MS

Mitochondrial proteome extraction. A total of six mitochondrial pellets (3 biological replicates each from T98G and U87MG cells) from the same passages were lysed, incubated and sonicated at 4°C (5 cycles of 20 pulses) in lysis buffer (4% SDS, 0.1 M DTT and 0.1 M Tris pH 8.6). To reduce disulfide bridges, samples were incubated at 40°C for 30 min, and cysteine residues were alkylated with 100 mM iodoacetamide for 30 min in the dark. The protein content was estimated by 1D SDS-PAGE scanned in a GS-800 densitometer (Bio-Rad Laboratories, Inc.), stained overnight at room temperature with colloidal Coomassie brilliant blue R-250 and quantified using the Quantity One software v4.6.9 (Bio-Rad Laboratories, Inc.).

Peptide separation and spectrometry. The peptide mixture was subjected to reverse phase chromatography on a Dionex Ultimate 3,000 RSLC nano-UPLC system (Thermo Fisher Scientific, Inc) in-line coupled to a Q-Exactive Plus high-resolution mass spectrometer (Thermo Fischer Scientific, Inc.). Peptides (2 µg) resuspended were first trapped on a precolumn (C18 PepMap...
100, 5 µm, 100 A, 300 µm inner diameter x 5 mm; Thermo Fisher Scientific, Inc.), then separated using an EASY-Spray PepMap RSLC C18 capillary column (2 µm, 15 cm x 50 µm; Thermo Fisher Scientific, Inc.) with a 250-min elution gradient at 250 nL/min. The mobile phases were: A) 2% acetonitrile and 0.1% formic acid in water; and B) 90:10 (v:v) acetonitrile:water and 0.1% formic acid in water. The mass spectrometer was operated in positive data-dependent acquisition mode and the full MS range was 300-1,800 m/z. A total of 10 of the most intense ions were isolated in the quadrupole and fragmented under higher-energy collisional dissociation with a normalized collision energy of 27%. Precursor ions were measured at a resolution of 70,000 (at 200 m/z) and the fragments were measured at 17,500. Only ions with charge states ≥2 were fragmented with an isolation window of 2 m/z.

**Protein identification and quantification.** Protein identification and label-free quantification were performed using MaxQuant v1.6.2.3 (20). The parameters included trypsin/P as the digestion enzyme, carbamidomethyl-cysteine as a fixed modification, and N-terminal protein acetylation and methionine oxidation as variable modifications. Proteins were identified with a 1% false discovery rate (FDR) based on the target-decoy strategy provided by MaxQuant, setting Arg-C as the digestion enzyme and carbamidomethylcysteine as a fixed modification. Proteins were identified with an FDR of 1% based on the target-decoy strategy provided by MaxQuant.

Protein identification was performed according to the human reference proteome UP000005640 from the UniProt repository downloaded on August 03, 2018 (https://www.uniprot.org/proteomes/UP000005640). Label-free quantification was performed using proteins with ≥2 razor-unique peptides identified by LC-MS/MS.

**Multivariate analysis of protein intensities.** Statistical analysis of protein abundances was performed only with proteins with ≥2 intensity values in each cell. The protein abundance was normalized between conditions and missing values were imputed with the Random Forest method. PCA was carried out on the protein intensity correlation matrix (FactoMiner v2.3; R package v3.5) (22) to generate a protein abundance pattern for the cell lines (15). To determine whether any component could distinguish between the cell lines, the sample scores for each component were plotted. After finding the component, the significant proteins with discriminatory capacity in that component were identified using the cos² of the correlation matrix between the components and the proteins (23).

To evaluate the PCA performance on LC-MS/MS data, a t-test (significance level, 0.05) and a fold-change analysis were conducted to compare the abundance spots of proteins between cell lines.

**2D SDS-PAGE**

**Mitochondrial proteome extraction.** T98G and U87MG mitochondrial-associated proteins were obtained according to Hurkman's protocol (24), which was modified as follows: Six mitochondrial pellets (three biological replicates each from T98G and U87MG cells) from the same passage were resuspended in 500 µl extraction buffer (0.7 M sucrose, 0.5 M Tris-Base, 0.1 M KCl, 0.03 M HCl, 0.05 M EDTA and 2% β-mercaptoethanol) and 500 µl saturated phenol, and incubated for 20 min at -20°C. Subsequently, mitochondrial samples were centrifuged for 10 min at 400 x g and 4°C, and the phenolic phase was recovered after the addition of 0.1 M ammonium acetate for 12-15 h at -20°C. Subsequently, mitochondrial samples were washed twice with 0.1 M ammonium acetate and centrifuged for 10 min at 4,000 x g and 4°C. Pellets with mitochondrial proteins were washed with 1 ml 80% acetone and centrifuged for 10 min at 4,000 x g and 4°C. The supernatants were discarded, and the pellets were resuspended in IEF buffer (7 M urea, 2 M thiourea, 0.06 M DTT, 2% amphotoles at pH 3-10 and 4% CHAPS) and centrifuged for 30 min at 8,000 x g and 4°C. The obtained supernatants were recovered and frozen at -80°C until use in 2D electrophoresis (2DE).

**Image pre-processing.** The gels were scanned in a GS-800 densitometer (Bio-Rad Laboratories, Inc.), and six images were acquired, wrapped and overlapped with PdQuest v8.0.1 software (Bio-Rad Laboratories, Inc.). Subsequently, with all six images combined, a master gel was created by the default PdQuest algorithm from the intensity sum of all of the spots in the gel images.

**Random spot selection in the master gel.** To increase the protein representativeness of the cellular processes carried out in the T98G and U87MG cell lines, 400 spots (of the 1,274 detected by PdQuest) were randomly selected from the master gel regardless of their size, intensity or abundance differences between cell lines. This process ensures that every spot in the master gel had an equal chance of being selected and allowed to obtain a representative mitochondrial proteome sample (26,27). The generated spot sample was rematched in all gel images to allow for a more reliable abundance analysis (17).

**Multivariate analysis of spot intensities.** To select the spots to be identified, a spreadsheet with the normalized intensity of the 400 spots sampled was exported from PdQuest. The abundance of the spots was logarithmically transformed, and missing values were imputed with the Random Forest method (missForest; R package) (21) to perform multivariate analysis. Abundance analysis was performed using PCA of the spot intensity correlation matrix (FactoMinerR; R package) (22) to generate a spot abundance pattern for the cell line gels (17). To understand whether any component could distinguish between the cell lines, the gel scores for each component were plotted. Subsequently, significant spots were identified using the cos² of the correlation matrix between the components and the spots with discriminatory capacity (23).
To evaluate the PCA performance on 2DE-MALDI-TOF data, a t-test (significance level, 0.05) and a fold-change analysis were conducted to compare the abundance spots of proteins between cell lines.

**Mass spectrometry.** Each selected spot was cut from the gel, alkylated, reduced, digested and automatically transferred to a matrix-assisted laser desorption/ionization (MALDI) analysis target via a Proteineer SPII and SP robot using the SP control v3.1.48.0 software (Bruker Corporation) with the aid of a DP Chemicals 96 gel digestion kit (Bruker Corporation) and processed in a MALDI-time of flight (TOF) Autoflex spectrometer (Bruker Corporation) to obtain the peptide mass fingerprints. A total of 100 satisfactory shots in 20 short steps were performed, the peak resolution threshold was set at 1,500, the signal/noise ratio of tolerance was 6 and contaminants were not excluded. The spectrum was annotated using flexAnalysis 1.2 vSD1 Patch 2 (Bruker Corporation). The search engine Mascot (28) was used to compare the fingerprints against the SwissProt (29) 2016 release database with the following parameters: Taxon-Human, mass tolerance of ≤200 ppm, one missed cleavage allowed, and fixed modification of carbamidomethyl and oxidation of methionine as the variable modifications.

**PPIn construction.** A PPIn network for T98G and another for U87MG cell lines were built with overexpressed and specific proteins obtained from the LC-MS/MS abundance data employing the GeneMANIA application v3.4.1 (30) in Cytoscape v3.3 (31). The networks were constructed only using the experimental evidence of physical interactions between proteins without added nodes. Subsequently, to compare the consistency between the PCA results from the LC-MS/MS and the 2DE data, overexpressed and specific proteins obtained from 2DE were localized to the LC-MS/MS PPIn.

**Representative biological processes identification.** To further understand the critical biochemical processes taking place in each PPIn, only the connected proteins in the networks were taken into account for comparative overrepresentation analysis based on GO (32). Overrepresentation was performed online employing the Gene List Analysis tool on the PANTHER Classification system site (33) (http://www.pantherdb.org/). As inputs, the official gene symbols were uploaded as identifiers. The overrepresented biological processes were clustered with REVIGO web tool (http://revigo.irb.hr/) (34) and R software v3.4 (35).

**Western blot analysis.** For OXPHOS system comparison, NADH dehydrogenase [ubiquinone] 1 OXPHOS complex, which is one of the most affected systems: at 4℃) purchased from Abcam against each subunit of the temperature, incubated with primary antibodies (overnight 100 V for 1 h (36,37). The membrane was blocked with 5% sodium deoxycholate, 0.1% SDS and 140 mM NaC) and protein concentration was determined using the Lowry method (38). As GAPDH and β-ATPase exhibited differences in abundance, western blotting was performed using both 10 and 20 µg of total protein to avoid artefacts due to protein quantity. Total protein was separated and transferred in the same way as mitochondrial extracts. The membrane was blocked, incubated and bands detected using the same protocol as for OXPHOS western blot analysis with primary antibodies purchased from Abcam against GAPDH (cat. no. ab8245; 1:1,000) and β-subunit ATP synthase (1:1,000), with an anti-mouse secondary antibody used (1:5,000). The experimental conditions, reagents, equipment and software used are those mentioned for mitochondrial extracts.

**Results**

**PCA of LC-MS/MS identified proteins.** The LC-MS/MS process identified 1,805 proteins, and 1,069 proteins were identified with ≥2 unique peptides and had ≥2 intensity values for each cell line (Table S1 and Fig. S1). However, 161 proteins were specific for T98G and 82 proteins for U87MG; these proteins were considered for the overrepresentation analysis (Fig. S1). Additionally, three proteins were eliminated during the imputation data process as they were identified as outliers. PCA was performed on 823 shared proteins (Table SII and Fig. S1). The total protein abundance variation was explained via five principal components (PCs; Fig. 1A). PC1 embraces 57% of the whole abundance variability, while the other four components only explain 43% of the remaining variability. The association of proteins and samples with PC1 and PC2 (Fig. 1B) revealed 235 proteins (black dots), with negative PC1 values, associated with T98G samples, and 308 proteins (black squares) with positive PC1 values, associated with U87MG samples. These 543 proteins (dark blue on the component scale in Fig. 1C) had a homogeneous intensity pattern within the cell lines (blocks 2 and 4 on the heatmap; Fig. 1C) and significant (r<0.5 or r>0.5) correlation values with PC1 (light red and light green in the mean correlation scale in Fig. 1C and Table SII), and their contribution to explain this intensity pattern was significant (white color in the contribution scale in Fig. 1C and Table SII). The proteins grouped on the left had a greater abundance in T98G and a lower abundance in U87MG (block 4 on the heatmap; Fig. 1C), and those grouped on the right were more abundant in U87MG (block 2 on the heatmap; Fig. 1C). The proteins in the center of the circular biplot (grey triangles; Fig. 1B) had heterogeneous intensity values within the cell lines (blocks 1, 3 and 5 on the heatmap; Fig. 1C), as well as low correlation and contribution values (Fig. 1C and Table SII).
**PCA for 2DE gel spots.** Overall, 400 protein spots were selected across all gel surfaces regardless of size, intensity or difference in abundance between the cell lines. A total of three spots did not pass quality control, and 161 protein spots were specific for either T98G or U87MG (Fig. S1). Finally, PCA was performed on 236 spots shared by both cell lines (Table SIII and Fig. S1). The 2DE-PCA results were close to those obtained via LC-MS/MS, as five PCs explained the gel intensity behavior (Fig. 2A). Similarly, to LC-MS/MS, PC1 accounted for 63% of the whole explained variance, while...
the other four components only explained 37%. The circular biplot (Fig. 2B) exhibits the same spatial arrangement between gels (proteins visualized as spots) as for the LC-MS/MS data. A total of 165 spots strongly correlated with PC1; 51 spots (black dots), associated with T98G gels, had negative values, and 114 spots (black squares) associated with U87MG gels, had positive PC1 values. Additionally, the 2DE intensity heatmap (Fig. 2C) replicated that obtained for LC-MS/MS. As expected, protein spots that differed between cell lines (dark blue on the component scale) had a homogeneous intensity within cell lines (blocks 1 and 3 on the heatmap; Fig. 2C), with significant correlation and contribution values (light red and light green in the mean correlation scale, and white in the contribution scale in Fig. 2C and Table SIII) with PC1. Likewise, the protein spots grouped on the left in the circular biplot (Fig. 2B), had a larger abundance in T98G cells and a lower abundance in U87MG cells (block 3 on the heatmap; Fig. 2C), and those grouped on the right were more abundant in U87MG cells (block 1 on the heatmap; Fig. 2C). Additionally, the spots in the center of the circular biplot (grey triangles in Fig. 2B) had very heterogeneous intensity values within the cell lines (blocks 2 and 4 on the heatmap; Fig. 2C), as well as low correlation and contribution values (Fig. 2C and Table SIII). According to the present results, these 165 spots, and 20 specific spots for T98G and 20 specific spots for U87MG cells (randomly selected) were selected for MALDI-TOF identification.

MALDI-TOF protein identification. As a result of random sampling and PCA, 66 proteins exhibited a homogeneous distribution in the T98G and U87MG gels (Fig. 3A and B), which assured whole mitochondrial proteins were represented. The T98G cell line was represented by 33 proteins (4 specific and 29 upregulated), and the U87MG cell line was represented by 33 proteins (5 specific and 28 more abundant; Fig. 3).

PPIs from T98G and U87MG cell lines. According to LC-MS/MS data analysis, a specific T98G PPI (Fig. 4A) was built composed of 396 proteins (235 from PC1 and 161 specific). The U87MG PPI (Fig. 4B) included 389 proteins (307 from PC1 and 82 specifics). The T98G PPI had 24 no-interaction nodes, while U87MG had 27; their heterogeneity and connectivity were similar (Fig. 4A and B). For 2DE-identified proteins, 51 (77%) were found in PC1 from the LC-MS/MS data and mapped onto the T98G or U87MG LC-MS/MS PPI (Fig. 4). As shown in Fig. 4, 2DE proteins were distributed throughout the T98G (solid blue dots in Fig. 4A) and U87MG (solid red dots in Fig. 4B) LC-MS/MS networks. The present results suggested that the proteins obtained from 2DE-PCA from randomly selected spots were comparable with the proteins from PCA applied to the LC-MS/MS label-free data.

Overrepresentation analysis. To determine whether 2DE-MALDI-TOF and LC-MS/MS data were biologically comparable, GO overrepresentation analysis was performed.
Figure 4. T98G and U87MG PPIs. PPIs were built using GeneMANIA v3.4.1 with PC1-proteins from the liquid chromatography coupled to tandem mass spectrometry data without adding more nodes. The edges represent only experimentally tested physical interactions between proteins. The solid blue and red dots represent 2D electrophoresis-matrix-assisted laser desorption/ionization-time of flight PC1 weighted proteins in (A) T98G and (B) U87MG PPIs. PPIs, protein-protein interaction networks; PC, principal component.
Proteins identified in T98G cells exhibited typical mitochondrial functions (Fig. 5A): ‘Generation of precursor metabolites and energy process’ was the main enriched cellular process, followed by ‘primary metabolism’ and ‘cellular amino acid metabolism’, which together involve OXPHOS (UCR1, QCR1, QCR2, NUDS1 and NUDS3 proteins identified by MALDI-TOF), ‘β-oxidation’ (ECI1), and ‘tricarboxylic acid (TCA) cycle’ (ACON, PCKGM, SDHA, DHE3, SERA and 3HIDH). Notably, other cellular processes that are less reported for mitochondrial function were also present: ‘Protein metabolic process’, ‘protein folding’ (TCPQ, TCPB and HSP7C), ‘translation’ (EF2 and EF1G) and ‘cellular component organization or biogenesis’ (TOMM40, UQCRC2 and NDUFS1). The minor represented biological processes included ‘cell surface receptor signaling pathways’, ‘carbohydrate metabolism’ and ‘locomotion’. The present protein ranking provided a close picture of the mitochondrial function in T98G cells.

On the other hand, U87MG protein classification in cellular processes (Fig. 5B) was different compared with the T98G results. The more notable biological processes identified were associated with cancer. One of the most obvious was associated with energy metabolism change, ‘glycolytic process’ (ENO1A, PGAM1 and TPIS); however, this was ranked below mitochondrial organization issues such as ‘cytoskeleton organization’ (LMAN2, DYNC1L12, TPM3 and TPM4) and ‘cellular component morphogenesis’ (CCT2 and CCT8).
Notably, ‘exocytosis’ (ACTB, RAB1B and RAB2B) was another cancer-associated process identified in these cells, and a clear downregulation of the T98G cellular processes was evident (such as ‘primary metabolism’ and ‘cellular component organization or biogenesis’), indicating good biological congruence for the present analysis.

The aforementioned cellular processes were corroborated through protein spot identification by MALDI-TOF (Fig. 3), as the proteins associated with ‘glycolytic process’, ‘vesicle-mediated transport’, ‘protein translation and biomass’ (EF2 and EF1G) or molecular chaperones (HSP7C, TCPB and TPCQ) were upregulated. The present U87MG landscape presents mitochondria with modified cellular and metabolic functions, suggesting that mitochondria readjust their cellular processes according to different cancer states.

**PCA versus fold-change and t-test.** Through selection of random 2DE spots followed by determination of their PCA abundance, 165 spots were obtained as candidates for identification (Fig. 6A). Ultimately, the log₂(fold-change) between -1 and 1 led to discard 31% of these 165 spots (black dots; Fig. 6A), and a t-test removed 30% (black dots below the horizontal dotted line; P=0.05 boundary). Following the log₂(fold-change) and the t-test, as commonly performed to select differentially expressed spots, 43% of the spots were discarded (black dots; Fig. 6A).

A similar behavior was observed for 89 MALDI-TOF-identified proteins (Fig. 6B), for which the log₂(fold-change) comparison removed ~35% of the spots, while the t-test removed 30%. When considering both tests, 44% of the identified proteins could be discarded (black dots; Fig. 6B). With respect to the 543 proteins obtained from LC-MS/MS PC1, there was a high number of removed proteins (black dots; Fig. 6C): The log₂(fold-change) and the t-test removed 80 and 35% of proteins, respectively. A total of 18% of proteins were significant according to the log₂(fold-change) and the t-test, indicating a substantial loss of information and representativeness.

---

**Figure 6.** Comparison of the significant proteins between PCA, t-test and log₂(fold-change). Random sampling and PCA show advantages over log₂(fold-change) and the t-test for the description of T98G and U87MG cells. Volcano plots for (A) PC1 significant 2DE spots, (B) 2DE-MALDI-TOF and (C) LC-MS/MS significant identified proteins. Graphs show spots and identified proteins discarded by log₂(fold-change) represented by black dots between -1 and 1 values of log₂(fold-change) (vertical dotted lines) and/or by t-test, represented by black dots under the horizontal dotted line (P=0.05 boundary). PCA, principal component analysis; 2DE, 2D electrophoresis; MALDI-TOF, matrix-assisted laser desorption/ionization-time of flight; LC-MS/MS, liquid chromatography coupled to tandem mass spectrometry.
Warburg metabolism. As the 'generation of precursor metabolites and energy' and 'metabolism' processes were amongst the most significantly enriched processes in the T98G (Fig. 5A) and U87MG (Fig. 5B) cell lines, respectively, the protein expression levels of OXPHOS complexes (CI-V and β-subunit ATP synthase) were verified on both cell lines via western blot analysis. The expression levels of the OXPHOS complexes in U87MG cells were significantly decreased compared with those in T98G cells (Fig. 7A and Fig. S2). Additionally, as the 'glycolytic process' was one of the most enhanced processes in U87MG (Fig. 5B), the bioenergetic signature (39) was investigated (Fig. 7B). The results were in accordance with Warburg's effect, as U87MG cells expressed more glycolytic proteins compared with T98G cells, where the OXPHOS system was predominant. The present results suggested that mitochondria may act as sensor organelles that change with biological states.

Discussion

Previous studies have stated that cancer proteomics results have an unclear association with various diseases (15-17). In regard to cancer, there can be a number of reasons for this unclear association; one of these may be associated with 'custom' data analysis focused on protein abundance fold-changes and/or univariate hypothesis tests. However, numerous proteins exhibit multiple or moonlighting functions and are involved in different biological pathways, or as bidirectional enzymes, they are involved in synthesizing or hydrolyzing according to cell duties. Therefore, a more sensible statistical approach consistent with the biological systems under study is required.

Unlike the log₂(fold-change) and t-test approach, PCA does not compare the protein mean abundance by multiple independent hypothesis tests between groups. Instead, the PCA summarizes the abundance behavior of whole proteins...
simultaneously in all samples to determine abundance patterns, i.e. proteins changing simultaneously in specific signaling pathways under certain conditions (40,41). On the other hand, random sampling on 2D gels is able to obtain enough representativeness to make reliable inferences (27,42).

LC and 2DE data analysis are in line with biochemical and proteomic evidence (15,16), providing an accurate description of T98G and U87MG cells with the representative biological processes according to each cell line (19,43‑45). In the present study, mitochondrial proteome PCA of LC‑MS/MS and 2DE random spot selection data revealed specific PPIns for each cell line in which 2D‑selected and identified proteins were included in a larger and improved LC PPIn.

According to the current results, PPIns revealed that T98G protein groups belonged to well‑characterized cellular processes closer to those of typical mitochondria. The most represented protein groups according to GO overrepresentation were associated with ‘generation of precursor metabolites and energy’, followed by ‘chromatin organization’, ‘primary metabolism’, ‘protein folding’, ‘cellular component organization or biogenesis’ and the major group of proteins for ‘cellular amino acid metabolism’, where OXPHOS and TCA are implicitly represented. Additionally, some less reported mitochondrial processes (‘carbohydrate metabolism’ or ‘locomotion’) or non‑typical mitochondrial functions (‘cell surface receptor signaling pathway’ and ‘cell communication’) associated with cancer were identified. Therefore, PCA analysis appears to be sufficiently powerful to detect this ‘extra’ information. All the aforementioned processes were identified according to 2DE data, which demonstrated enough resolution power to build a reliable PPIn network skeleton; LC data detected the same pathways but in more detail.

U87MG cells displayed a more heterogeneous molecular landscape with more non‑mitochondrial processes detected. In these cells, mitochondrial organization processes (including ‘cytoskeleton organization’, ‘cellular component morphogenesis’, ‘protein complex assembly’ and ‘protein complex biogenesis’) occupied a central role. Additionally, the canonical energy metabolic shift was clear, as ‘glycolytic process’ proteins were well represented and OXPHOS was dysregulated. Notably, other non‑mitochondrial cancer‑associated processes were observed, including ‘exocytosis’ (such as ‘intracellular protein transport’ and ‘vesicle‑mediated transport’) and ‘protein folding’. Some of these cellular processes were also present in T98G cells but its commitment is different since in U87MG cells, i.e. the ‘generation of precursor metabolites and energy’ and ‘primary metabolism’ abundance pattern changes are represented mainly by ‘catabolic process’ or ‘protein synthesis for fueling other pathways, oligosynthesis, recycling (formation of metabolic precursors, RAB GTPases) and protein synthesis for fueling other pathways, such as the TCA cycle when basal or other metabolites are not available (53,54). In addition, there are proteins for amino acid and purine metabolism that could enable phagocytic structures such as phagosomes (55).

In addition to a metabolic shift, PCA can determine simultaneous cell processes, unlike other proteomic approaches based on proteins surpassing significant abundance changes (56,57). The present data analysis approach identified a specific proteomic landscape for T98G cells and another one for U87MG cells that defined concrete cell processes and temporality. This may allow the identification of targets or therapeutic tools that may result in reliable and useful information to help improve diagnosis and risk assessment.

The present data supported the hypothesis of mitochondria acting as dynamic organelles following and sensing the molecular events that take place during carcinogenesis (19,58‑60). In conclusion, PCA applied to LC‑MS/MS label‑free quantified data was able to describe the most relevant biological processes in each cell type. Similarly, random sampling of spots and their abundance PCA from 2DE before protein identification identified
proteins that exhibited the same information as LC, albeit with less resolution; with this information, a representative mitochondrial proteome landscape was built specifically for the T98G and U87MG cell lines, in which overrepresented biological processes were highlighted with the identified mitochondrial proteins.

Acknowledgements

Not applicable.

Funding

The present study was supported by the Mexican Social Security Institute (IMSS) and National Autonomous University of Mexico (UNAM), by a grant from CONACyT (grant no. 220790), by a fellowship from the National Council of Science and Technology (grant no. 375815) and by General Directorate of Academic Staff Affairs-Support Program for Research and Technological Innovation Projects (grant nos. IN-207519 and IN-213216).

Availability of data and materials

The datasets used and/or analyzed during the present study are available from the corresponding authors on reasonable request.

Authors' contributions

LGC, FMS and SEG conceptualized the research. LGC, FMS, and Principal Component Analysis of Human Glioblastoma Multiforme (GBM) Tumor Proteome. Iran J cancer Prev 7: 87‑95, 2014.


Valledor L and Jorrín J: Back to the basics: Maximizing the information obtained by quantitative two dimensional gel electrophoresis analyses by an appropriate experimental design and statistical analyses. J Proteomics 74: 1‑18, 2011.


49. Oba‑Shinjo S, Marie SK and Rosa JC: Changes in the expression of proteins associated with aerobic glycolysis and cell migration are involved in tumorigenic ability of two glioma cell lines. Proteome Sci 10: 53, 2012.