SuperQuant-assisted comparative proteome analysis of glioblastoma subpopulations allows for identification of potential novel therapeutic targets and cell markers

Thiago Verano-Braga¹,²,*, Vladimir Gorshkov¹,*, Sune Munthe³,⁴, Mia D. Sørensen³,⁵, Bjarne W. Kristensen³,⁵ and Frank Kjeldsen¹

¹Department of Biochemistry and Molecular Biology, University of Southern Denmark, Odense, Denmark
²Department of Physiology and Biophysics, Federal University of Minas Gerais, Belo Horizonte, MG, Brazil
³Department of Pathology, Odense University Hospital, Odense, Denmark
⁴Department of Neurosurgery, Odense University Hospital, Odense, Denmark
⁵Department of Clinical Research, University of Southern Denmark, Odense, Denmark

*These authors contributed equally to this work

Correspondence to: Thiago Verano-Braga, email: tvbraga@icb.ufmg.br
Frank Kjeldsen, email: frankk@bmb.sdu.dk

Keywords: proteomics; glioblastoma; cancer stem cell; cancer proteome; migration

Received: May 30, 2017 Accepted: January 22, 2018 Published: January 25, 2018

ABSTRACT

Glioblastoma (GBM) is a highly aggressive brain cancer with poor prognosis and low survival rate. Invasive cancer stem-like cells (CSCs) are responsible for tumor recurrence because they escape current treatments. Our main goal was to study the proteome of three GBM subpopulations to identify key molecules behind GBM cell phenotypes and potential cell markers for migrating cells. We used SuperQuant—an enhanced quantitative proteome approach—to increase proteome coverage. We found 148 proteins differentially regulated in migrating CSCs and 199 proteins differentially regulated in differentiated cells. We used Ingenuity Pathway Analysis (IPA) to predict upstream regulators, downstream effects and canonical pathways associated with regulated proteins. IPA analysis predicted activation of integrin-linked kinase (ILK) signaling, actin cytoskeleton signaling, and lysine demethylase 5B (KDM5B) in CSC migration. Moreover, our data suggested that microRNA-122 (miR-122) is a potential upstream regulator of GBM phenotypes as miR-122 activation was predicted for differentiated cells while its inhibition was predicted for migrating CSCs. Finally, we validated transferrin (TF) and procollagen-lysine 2-oxoglutarate 5-dioxygenase 2 (PLOD2) as potential markers for migrating cells.

INTRODUCTION

Glioblastoma (GBM), a WHO grade IV glioma, is the most aggressive primary brain cancer [1] with a median survival rate of only 15 months [2]. Surgical resection combined with radiotherapy and chemotherapy is the standard treatment for GBM [2]. However, tumor recurrence is inevitable because GBM cells can invade surrounding tissues thus preventing full removal of tumor cells by resection. Moreover, subpopulations of cancer stem-like cells (CSCs), which are resistant to radiotherapy and chemotherapy, are present in tumor environment [3, 4]. CSCs self-renew and recapitulate the original tumor upon cerebral implantation in mice [5]. Evidences suggest that CSCs have developed a number of molecular mechanisms to escape and survive treatments. For example, radiation resistance seems to be due to increased activity of the DNA damage response machinery in CSCs.
Radiotherapy with concomitant temozolomide (TMZ) has increased the survival rate by several months [2], however, recurrence is still observed after treatment.

Understanding the regulatory system of CSCs is critical to design next-generation drugs to ultimately imbalance CSC homeostasis. Regulation of CSCs in gliomas involves extrinsic (microenvironment, niche factors and host immune system) and intrinsic (genetic, epigenetic, and metabolic) factors [6]. For example, CSCs maintain an undifferentiated state by aberrant activation of common signaling pathways such as Notch, NF-kB, BMP, Wnt, and PI3K-AKT signaling [6].

A number of biomarkers are available for glioma cells, such as mutation of isocitrate dehydrogenase 1 (mIDH1) [7, 8], methylation of MGMT gene (DNA repair enzyme O6-methylguanine-DNA methyltransferase) [9], as well as standard cell markers including, but not restricted to, CD133, Musashi-1, Bmi-1, Sox-2 and Nestin [10, 11]. Some CSCs are able to migrate from the tumor core, representing a significant challenge for treatment as they are left behind after tumor resection [9, 12]. Although standard stem cell markers can be used to identify migrating CSCs, these markers cannot be used to select the migrating phenotype from other CSCs.

Gene expression analysis is a valuable method in cancer research. However, gene counts do not necessarily correlate with protein abundances in cells due to mRNA regulation and protein turnover. During the past decade, mass spectrometry (MS)-based proteomics has become the leading technology to study biological systems. Our group developed the Complementary Finder software. This tool uses high-confidence complementary fragment ions to derive individual and unique peptide parent masses, deconvoluting mixture spectra and increasing proteome coverage [13]. More recently, we implemented the SuperQuant method to infer quantitative information to the output data of Complementary Finder using MS1 spectrum-based quantitation. On average, SuperQuant allows the identification and quantitation of 10% more proteins per study enabling a more insightful proteome interpretation [14].

Given the poor prognosis of GBM, the identification of specific biomarkers and therapeutic targets for this disease are of imminent importance to improve the early diagnosis and treatment, including the development of personalized medicine [15]. In this study, our goals were to identify key regulators and cell markers from different phenotypes of GBM cells, focusing on the migrating CSC phenotype. We used SuperQuant to assess deeper into the GBM proteome of CSC spheroids, migrating CSCs and differentiated GBM cells. We identified and quantified the protein abundances of more than 2,817 proteins, of which 304 proteins were found differentially regulated. Among regulated proteins, we identified potential therapeutic targets and cell markers that could serve as subjects for further investigation.

**RESULTS AND DISCUSSION**

**Experimental design used to study proteins involved in the GBM migration and differentiation**

We used the proteomics pipeline represented in Figure 1 to compare the proteome of three GBM cell phenotypes. Tissue fragments from a grade IV GBM patient subjected to brain tumor resection were obtained in a study published previously [16]. Tissue fragments were dissociated to the single cell level and thereafter they formed spheroids in serum free medium. We decided to grow the tumor cells as organotypic tumor spheroids (OTS) because 3D culture preserves important features of the original tumor tissue, including cell-to-cell interactions, thus representing a more in vivo-like model [17, 18]. Migrating and differentiated cells were obtained from CSC spheroids. To select CSCs with migrating phenotype, we used the migration assay reported by Munthe et al. [19] and represented in the supporting information (Supplementary Figure 1). The migration assay was conducted with stem cell medium to maintain cells undifferentiated. On the other hand, to produce the differentiated cells we cultured the CSC spheroids with medium containing 10% (w/v) fetal calf serum (FCS) (Figure 1).

In vitro experiments performed with neural stem cells showed that these cells expand and form neurospheres when grown under serum-free conditions [20]. Lee et al. showed that GBM cells cultured under serum-free conditions more closely resembled primary GBMs and preserved CSC features [21]. Therefore serum-free medium more accurately reflect the cell environment in brain tissue compared to serum-containing medium for culturing CSCs.

Recently we reported a new post-acquisition method (named SuperQuant) that allows increased quantitative proteome depth by extracting co-isolated/co-fragmented peptides from MS/MS output files that are neglected in a normal shotgun proteomic analysis [14]. SuperQuant processing led to the identification and quantitation of 2,817 protein groups, representing an increment of 7% compared to unprocessed data (Supplementary Figure 2, supporting information). This improvement is comparable with our previous study reporting approximately 10% increase of quantified proteins using HeLa cells [14]. To evaluate quantitation reproducibility of our method, we calculated the coefficient of variance (CV) between replicates for each protein. The CVs distributions are presented in Supplementary Figure 3 (supporting information). The average CV between replicate experiments was 3.5%.

**The proteome of CSC spheroids and migrating CSCs are more similar than the proteome from differentiated cells**

Since migrating CSCs and differentiated cells were generated from CSC spheroids, we used the dataset of
CSC spheroids as reference (denominator) to obtain protein ratios. As consequence, upregulation of a given protein from the migrating CSCs dataset means that this protein is statistically more abundant in migrating CSCs than in CSC spheroids. Likewise, down-regulation of a given protein from differentiated cells means that this protein is differentially more abundant in CSC spheroids than in differentiated cells.

Out of 2,817 proteins detected and quantified in this study, 199 and 148 were differentially regulated in the differentiated and migrating CSCs datasets, respectively (Supplementary Table 1). Comparing the regulated proteins from both datasets, we observed that only 43 proteins (14%) were identified in both datasets (Supplementary Figure 4), indicating that the differential proteomes from migrating CSCs and differentiated cells are rather different.

Hierarchical clustering analysis shows that the proteome of CSC spheroids and migrating CSCs are more similar than the proteome of differentiated cells (Figure 2). Although the similarity is only marginal in relation to the proteome of the differentiated cells, this result is likely to reflect the stem-like phenotype of spheroids and migrating cells [19]. However, we would like to point out that even these CSCs have clusters containing proteins with opposing abundance profiles indicating genes with distinct regulations.

**Pathway analysis of migrating CSCs and differentiated cells**

Activation z-score in IPA was used to predict modulation patterns of the top-scoring canonical pathways. This algorithm uses the causal network to predict activation...
or inhibition of given pathways based on experimental data and expected expression levels of signaling effectors from a curated biological database (Ingenuity Knowledge Base) [22].

As shown in Figure 3, IPA predicted opposing modulations for all top-scoring canonical pathways from migrating CSCs and differentiated cells, except for the integrin-linked kinase (ILK) signaling, predicted to be activated in both datasets. ILK connects integrin receptors with the actin cytoskeleton to regulate many cellular processes including proliferation, angiogenesis, survival, differentiation, migration, and invasion [23]. The highest activation z-score was calculated for migrating cells suggesting that ILK signaling is more active in migrating CSCs. Data suggest that activation of ILK signaling is important for migration of CSCs and differentiation. Since GBM recurrence after treatment is thought to be due to CSCs [6], our data suggest that inhibition of ILK signaling can be a potential treatment for GBM. In fact, a recent study has shown that p53-wildtype GBM cells were more susceptible to radiotherapy after the knockdown of ILK and its signaling partners PINCH1 and ILKAP [24].

Activation of actin cytoskeleton signaling was predicted only for migrating CSCs (Figure 3). Actin cytoskeleton reorganization is a key process in migration of GBM cells. Migrating cells produce actin-rich protrusions of the plasma membrane, named invadopodia, which interact with the extracellular matrix (ECM) inducing its proteolysis and remodeling. The presence of invadopodia is usually associated with poor prognosis [25, 26].

As shown in Figure 3, IPA predicted inhibition of RhoGDI, PI3K/AKT, RhoA, insulin receptor, CDK5, Rho family GTPases, and protein kinase A (PKA) signaling pathways for differentiated cells. This data suggest that the inhibition of these signaling pathways is important for GBM differentiation. Regarding PI3K/AKT signaling, its activation is a key mechanism in glioma progression to higher-grade tumors. Increased activity of PI3K/AKT in CSCs induces resistance by enhancing expelling of drugs via ABC transporters [27]. Moreover, CSC is maintained undifferentiated by aberrant activation of PI3K/AKT [6].

**Prediction of upstream regulators for migrating CSC and differentiated cell proteomes**

Figure 4A shows the predicted upstream regulators. For the differentiated cells, IPA predicted activation of microRNA 122 (miR-122) (z-score = 2.998) and tumor antigen p53 (Tp53) (z-score = 2.425), and inhibition of tumor antigen p63 (Tp63) (z-score = -0.083). For migrating CSCs, prediction outcomes were activation of lysine demethylase 5B (KDM5B) (z-score = 0.447) and inhibition of...
of miR-122 (z-score = –0.061). Figure 4B–4F represents causal associations between the predicted upstream regulators and their protein targets (experimental data). Annotations retrieved from IPA are provided in Table 1 (differentiated dataset) and Table 2 (migrating dataset).

MicroRNAs (miRNAs) are noncoding RNA that suppress expression of target genes by degrading correspondent mRNAs or inhibiting their translation process [28]. Deregulation of some miRNAs has been reported in brain cancer including miR-124a, a miRNA potentially involved in GBM invasion [29]. IPA predicted miR-122 inhibition for migrating CSCs and miR-122 activation for GBM differentiation (Figure 4A). Thus, our data suggest that miR-122 is a potential regulator of GBM fate. By functioning as an “ON/OFF button”, miR-122 seems to control GBM migration and differentiation. IPA analysis correlated miR-122 inhibition in migrating CSCs with the experimentally observed upregulation of ARHGAP1 and LAMC1 (Figure 4B). Although, the relationship between ARHGAP1 and GBM migration is yet to be determined, LAMC1 is indeed involved in GBM migration [29]. IPA indicated inconsistent findings for ALDOA and VPS4B. Based on the Ingenuity Knowledge Base [22], inhibition of miR-122 correlates with upregulation of ALDOA and VPS4B, but we detected downregulation of these proteins in migrating CSCs (Figure 4B), which suggests involvement of other upstream regulators than miR-122 regulating the migrating CSCs proteome. As shown in Figure 4D, activation of miR-122 in differentiated correlated with downregulation of ALDOA, GNPDA1, GYS1, KHGRBS1, LAMC1, NPEPPS, SEPT2, SEPT9, and VPS4B indicating that these downregulation events are potentially related to GBM differentiation. Similarly, activation of Tp53 (Figure 4E) and inhibition of Tp63 (Figure 4F) may also have a potential role in GBM cell differentiation.

MiR-122 is a liver-specific miRNA [30]. However, this miRNA seems to be less tissue-specific in cancer as miR-122 deregulation has been associated with colorectal cancer [31, 32], breast cancer [33], and gliomas [34], although miR-122 plays a role in hepatocellular carcinoma [35].

IPA predicted activation of KDM5B for migrating CSCs highlighting its potential role in promoting CSC migration. KDM5 activation was consistent with EHD1 upregulation (Figure 4C). Interestingly, EHD1 has been reported as a potential inducer of cell migration and metastasis in non-small cell lung cancer [36]. A previous study showed high levels of KDM5B in glioma, inducing tumor growth via p21 downregulation [37]. Our data suggest a new role for KDM5B as a regulator of CSCs migration in GBM. In line with this observation, increased expression of KDM5B has been reported in metastatic breast cancer cells highlighting its potential involvement in cancer cell migration [38].

After shedding light on potential upstream proteome regulators from migrating CSC and differentiated cells, we used the IPA downstream effect analysis. The downstream effect “migration of tumor cell lines” was found significantly regulated in the migrating CSCs (p-value = 0.007) and differentiated cells (p-value = 0.0003), which was predicted to be activated in migrating CSCs (z-score = 0.04) and inhibited in differentiated cells (z-score = –0.158).
(Figure 5). Upregulation of CAV1, GNB2L1, RUVBL1, and ADAM10 as well as downregulation of PRDX2 were consistent with activation of “migration of tumor cell lines” (right panel, Figure 5). Although IPA analysis did not predict upregulation of CTBP1 as an activator of migration of cancer cells, a recent report showed that silencing of CTBP1 suppressed migration of human glioma cells, indicating the potential role of CTBP1 in GBM invasiveness [39]. Inhibition of “migration of tumor cell lines” was consistent with downregulation of FABP7, FSCN1, KHDRBS1, LGALS3, MAPK1, PRMT5, PTN, RUVBL2, SEPT9, and STMN1 (left panel, Figure 5). IPA reported “inconsistent findings with the state of downstream molecule” (yellow arrows) for both datasets indicating that IPA predicted opposing protein abundances. Importantly, we included “all human tissues, cells and cell lines” for this particular analysis. Thus, the term “migration of tumor cell lines” also includes other cancer cell lines than glial cells, which could explain the observed inconsistent findings as a given protein may favor migration of one type of cancer cell line and inhibition of other.

### Profiling potential phenotype-specific markers

Clinicians and researchers face a great challenge dealing with GBM due to its heterogeneity since cancer cells with different phenotypes populate the tumor mass and surrounding areas. Since CSCs are regarded as key players in tumor recurrence, it is important to identify this phenotype. Stem cell markers such as CD133 [10], Musashi-1 [10], Bmi-1 [10], Sox-2 [10], Nestin [11], and Glut-3 [40] (to review other CSCs markers, see [6]) can be used to identify CSCs as well. However, we lack markers.

### Table 1: Differentiated proteome: predicted upstream regulators and downstream effectors

<table>
<thead>
<tr>
<th>Symbol</th>
<th>ID</th>
<th>Name</th>
<th>Location</th>
<th>Type(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>miR-122</td>
<td>—</td>
<td>microRNA 122</td>
<td>Cytoplasm</td>
<td>microRNA</td>
</tr>
<tr>
<td>ALDOA</td>
<td>P04075</td>
<td>aldolase, fructose-bisphosphate A</td>
<td>Cytoplasm</td>
<td>Enzyme</td>
</tr>
<tr>
<td>GNPDA1</td>
<td>P46926</td>
<td>glucosamine-6-phosphate deaminase 1</td>
<td>Cytoplasm</td>
<td>Enzyme</td>
</tr>
<tr>
<td>GYS1</td>
<td>P13807</td>
<td>glycogen synthase 1</td>
<td>Cytoplasm</td>
<td>Enzyme</td>
</tr>
<tr>
<td>KHDRBS1</td>
<td>Q07666</td>
<td>KH RNA binding domain containing, signal transduction associated 1</td>
<td>Nucleus</td>
<td>transcription regulator</td>
</tr>
<tr>
<td>LAMC1</td>
<td>P11047</td>
<td>laminin subunit gamma 1</td>
<td>Extracellular Space</td>
<td>Other</td>
</tr>
<tr>
<td>NPEPPS</td>
<td>P55786</td>
<td>aminopeptidase puromycin sensitive</td>
<td>Cytoplasm</td>
<td>peptidase</td>
</tr>
<tr>
<td>SEPT2</td>
<td>Q15019</td>
<td>septin 2</td>
<td>Cytoplasm</td>
<td>Enzyme</td>
</tr>
<tr>
<td>SEPT9</td>
<td>Q9UHD8</td>
<td>septin 9</td>
<td>Cytoplasm</td>
<td>Enzyme</td>
</tr>
<tr>
<td>VPS4B</td>
<td>O75351</td>
<td>vacuolar protein sorting 4 homolog B</td>
<td>Cytoplasm</td>
<td>transporter</td>
</tr>
<tr>
<td>Tp53'</td>
<td>—</td>
<td>tumor protein p53</td>
<td>Nucleus</td>
<td>transcription regulator</td>
</tr>
<tr>
<td>CAV1</td>
<td>Q03135</td>
<td>caveolin 1</td>
<td>Plasma Membrane</td>
<td>transmembrane receptor</td>
</tr>
<tr>
<td>CRYAB</td>
<td>P02511</td>
<td>crystallin alpha B</td>
<td>Nucleus</td>
<td>Other</td>
</tr>
<tr>
<td>FKBP3</td>
<td>Q00688</td>
<td>FK506 binding protein 3</td>
<td>Nucleus</td>
<td>Enzyme</td>
</tr>
<tr>
<td>PCNA</td>
<td>P12004</td>
<td>proliferating cell nuclear antigen</td>
<td>Nucleus</td>
<td>Enzyme</td>
</tr>
<tr>
<td>STMN1</td>
<td>P16949</td>
<td>stathmin 1</td>
<td>Cytoplasm</td>
<td>Other</td>
</tr>
<tr>
<td>TJP1</td>
<td>Q07157</td>
<td>tight junction protein 1</td>
<td>Plasma Membrane</td>
<td>Other</td>
</tr>
<tr>
<td>XPO1</td>
<td>O14980</td>
<td>exportin 1</td>
<td>Nucleus</td>
<td>transporter</td>
</tr>
<tr>
<td>Tp63'</td>
<td>—</td>
<td>tumor protein p63</td>
<td>Nucleus</td>
<td>transcription regulator</td>
</tr>
<tr>
<td>COL4A1</td>
<td>P02462</td>
<td>collagen type IV alpha 1 chain</td>
<td>Extracellular Space</td>
<td>Other</td>
</tr>
<tr>
<td>FUBP1</td>
<td>Q96AE4</td>
<td>far upstream element binding protein 1</td>
<td>Nucleus</td>
<td>transcription regulator</td>
</tr>
<tr>
<td>ITGB8</td>
<td>P26012</td>
<td>integrin subunit beta 8</td>
<td>Plasma Membrane</td>
<td>Other</td>
</tr>
<tr>
<td>PCNA</td>
<td>P12004</td>
<td>proliferating cell nuclear antigen</td>
<td>Nucleus</td>
<td>enzyme</td>
</tr>
<tr>
<td>TNC</td>
<td>P24821</td>
<td>tenascin C</td>
<td>Extracellular Space</td>
<td>Other</td>
</tr>
</tbody>
</table>

Molecule annotation was retrieved from IPA. (*) indicates the predicted upstream regulator.
Table 2: Migrating proteome: predicted upstream regulators and downstream effectors

<table>
<thead>
<tr>
<th>Symbol</th>
<th>ID</th>
<th>Name</th>
<th>Location</th>
<th>Type(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>miR-122’</td>
<td>—</td>
<td>microRNA 122</td>
<td>Cytoplasm</td>
<td>microRNA</td>
</tr>
<tr>
<td>ALDOA</td>
<td>P04075</td>
<td>aldolase, fructose-bisphosphate A</td>
<td>Cytoplasm</td>
<td>enzyme</td>
</tr>
<tr>
<td>ARHGAP1</td>
<td>Q07960</td>
<td>Rho GTPase activating protein 1</td>
<td>Cytoplasm</td>
<td>other</td>
</tr>
<tr>
<td>LAMC1</td>
<td>P11047</td>
<td>laminin subunit gamma 1</td>
<td>Extracellular Space</td>
<td>other</td>
</tr>
<tr>
<td>VPS4B</td>
<td>O75351</td>
<td>vacuolar protein sorting 4 homolog B</td>
<td>Cytoplasm</td>
<td>transporter</td>
</tr>
<tr>
<td>KDM5B’</td>
<td>—</td>
<td>lysine demethylase 5B</td>
<td>Nucleus</td>
<td>transcription regulator</td>
</tr>
<tr>
<td>CAV1</td>
<td>Q03135</td>
<td>caveolin 1</td>
<td>Plasma Membrane</td>
<td>transmembrane receptor</td>
</tr>
<tr>
<td>EHD1</td>
<td>Q9H4M9</td>
<td>EH domain containing 1</td>
<td>Cytoplasm</td>
<td>other</td>
</tr>
<tr>
<td>FABP5</td>
<td>Q01469</td>
<td>fatty acid binding protein 5</td>
<td>Cytoplasm</td>
<td>transporter</td>
</tr>
<tr>
<td>IARS2</td>
<td>Q9NSE4</td>
<td>isoleucyl-tRNA synthetase 2, mitochondrial</td>
<td>Cytoplasm</td>
<td>enzyme</td>
</tr>
<tr>
<td>PRPS1</td>
<td>P60891</td>
<td>phosphoribosyl pyrophosphate synthetase 1</td>
<td>Cytoplasm</td>
<td>kinase</td>
</tr>
</tbody>
</table>

Molecule annotation was retrieved from IPA. (‘) indicates the predicted upstream regulator.

Figure 4: Potential upstream regulators of migrating CSC and differentiated cancer cell proteomes. (A) Heat map with activation z-scores of predicted upstream regulators. (B–F) Causal connections between predicted upstream regulators and their targets from migrating (B–C) and differentiated (D–F) proteomes. Experimental log-transformed ratios and q-values are represented below each protein associated with upstream regulators. More information related to upstream regulators and their associated proteins from differentiated and migrating proteomes are provided in Table 1 and Table 2, respectively.
to identify different CSCs phenotypes as, for example, migrating and static CSCs [9, 19].

To select potential markers for the migrating CSC phenotype, we focused the analysis on the top-50 proteins with lowest q-values from the migrating dataset. Figure 6 represents the selected top-50 proteins including their cellular localization (i.e., extracellular, plasma membrane, and intracellular). Based on their abundance profile, proteins were divided into 32 upregulated and 18 downregulated, representing proteins more abundant in migrating CSCs or more abundant in CSC spheroids, respectively. To select potential markers for migrating CSCs, we focused our

Figure 5: Profiling the proteins potentially involved in the migration of GBM cells. Regulated proteins potentially involved in the enriched downstream effect “migration of tumor cell lines” are shown.

Figure 6: Potential markers for migrating CSCs. The top-50 most confident regulated proteins (lowest q-values) from migrating CSCs are plotted. Numbers represent the log-transformed ratios. Numbers in the left column are from the differentiated dataset while numbers in the right column are from the migrating CSC dataset. Downregulation (q-value < 0.05) = green; upregulation (q-value < 0.05) = red; non-regulation (q-value > 0.05) = white. Cell localizations were retrieved from Ingenuity Pathway Analysis (IPA). Diff, differentiated; Migr, migrating CSCs; Spher, spheroid CSCs.
analysis on the 32-upregulated proteins (Figure 6). We selected candidates for further validation based on antibody availability and novelty (not previously used as GBM migration marker). By using these criteria, we selected VCL and TF for further analysis. Although we observed upregulation of PLOD2 in the differentiated dataset as well, we selected PLOD2 for further validation because this enzyme has already been highlighted as a potential migration marker in other cancers, such as bladder cancer [41] and renal cell carcinoma [42]. Although PTN was not found differentially regulated in the migrating dataset (p-value = 0.485; Supplementary Table 1), we also selected this protein for validation because of reports indicating its role in glioma cell migration [43, 44].

We performed in situ immunohistochemistry using the same–but non-cultured–tumor tissue from which we obtained the spheroids to validate the selected proteins identified in our SuperQuant-assisted proteomic approach. A total of 14 areas (7 tumor core and 7 tumor periphery) were evaluated by manual tumor cell counting.

Figure 7A–7D contains representative pictures of tissue immunostaining of TF expression in tumor core and periphery. We provided in the supporting information (Supplementary Figure 5) the immunostaining pictures of VCL, PTN, and PLOD2.

In line with the SuperQuant-assisted proteomics data, both TF (p-value = 0.032) and PLOD2 (p-value = 0.044) were significantly more abundant in the tumor periphery than the tumor core (Figure 7E). TF is involved in the main iron uptake route in animals, and hence one of the most abundant proteins in plasma. Although TF is primarily synthesized in the liver, human brain (mainly oligodendrocytes) synthesizes TF as well [45, 46]. We found a strong upregulation of TF in migrating CSCs (Figure 6), which could be validated by immunohistochemistry (Figure 7D and 7E). Schonberg et al. [47] have recently shown that the TF gene is highly expressed in GBM CSCs while its expression is lower in normal neurons and glial progenitors. Furthermore, the authors observed a trend towards worse survival in patients with high levels of TF.

The elevation of PLOD2 in migrating cells is well in-line with other reports showing that PLOD2 is involved in cancer cells migration. Eisinger-Mathason et al. [48] showed that hypoxia in the tumor microenvironment and the expression of hypoxia-inducible factor-1α (HIF-1α) led to upregulation of PLOD2 in sarcoma. In addition, Chang et al. [49] observed high expression level of PLOD2 in invasive ductal carcinomas of breast cancer. PLOD2 increases collagen stability by hydroxylation of lysine residues in procollagen [50] leading to crosslinked collagen fibers that induce ECM stiffness favoring cell migration, invasion, and metastasis [51]. Furthermore, high expression levels of PLOD2 correlated with poor prognosis in patients with hepatocellular carcinoma [52] and breast carcinoma [49]. Based on mRNA sequencing obtained from laser-capture microdissection of whole tumors, Dong et al. [53] suggested PLOD2 as a prognostic biomarker in patients with GBM. Authors observed differential upregulation of PLOD2 in perinecrotic palisading cells–thought to be the most aggressive tumor cell type–compared to non-palisading cells. Moreover, a lower survival rate was associated with patients expressing higher levels of PLOD2 [53]. Our data add an additional level of information to this study as we detected PLOD2 predominantly upregulated in migrating CSCs and to a lesser extent in differentiated cells. Data indicate both TF and PLOD2 as potential markers for migrating GBM cells. However, our results indicated that PLOD2 should not be used as a CSC marker since PLOD2 was also detected upregulated in the differentiated cells (Figure 6 and Supplementary Table 1).

VCL abundance was significantly lower in tumor periphery compared to tumor core. According to the immunostaining, VCL was approximately 33% less abundant in peripheral cells than central cells (Figure 7E). This stands in contrast to our proteomic data showing > 4-fold upregulation of VCL in migrating CSCs compared to CSC spheroids (Figure 6 and supporting information, Supplementary Table 1). Higher VCL-staining in tumor core may be a reflection of the recruitment of VCL-positive microglial cells towards the tumor site [54]. Ubiquitously expressed, VCL is involved in cell-cell interaction and cell-matrix adhesion [55]. According to literature, there is diverging information of the role of VCL in cancer biology. VCL has been reported as tumor suppressor, illustrated by VCL downregulation in highly metastatic colorectal cancer [56], breast cancer [57], squamous carcinoma [58], and rhabdomyosarcoma [59]. In contrast, VCL has been found significantly upregulated in pancreatic cancer tissue [60] as well as in human basal and squamous cell tumors [58]. The diverging results suggest that VCL behavior in cancer cells is potentially tissue-specific.

Secreted PTN is a heparin-binding growth factor [61] and has an important role during brain development reaching its maximum expression level around birth (see [62] for review). PTN abundance was not found significantly regulated in the migrating CSCs dataset (Supplementary Table 1). Likewise, no significant change in PTN abundance was detected in tumor tissue although a trend towards increased levels was observed in the periphery (Figure 7E). These data stand in contrast to the literature reporting that PTN induces migration of some GBM cell lines via the receptor tyrosine phosphatase PTPRZ1 [43, 44]. We believe that tumor heterogeneity and/or the subtype of GBM used in the studies [63] may be responsible for such diverging results.

CONCLUSIONS

Subjecting GBM subpopulations to a comprehensive quantitative proteome analysis revealed 304 differentially
regulated proteins. These proteins represent potential biomarker that may assist researchers and clinicians to accurately select and identify different GBM cell phenotypes, holding potential as next-generation brain cancer diagnostics and therapeutics. Our study indicated miR-122, KDM5B, Tp53, and Tp63 as potential regulators of GBM cell phenotypes.

MATERIALS AND METHODS

Reagents

If not explicitly stated, all common solutions and reagents were purchased from Sigma-Aldrich and were of HPLC gradient grade or proteomics grade (when applicable).

Patient tissue and cell culture

Organotypic tumor spheroids (OTS) were obtained from fresh tumor tissue from a WHO grade IV GBM patient (referred to as T78) who underwent surgery in 2009 as previously described [16]. Cells were cultured as free-floating spheroids in serum-free neural stem cell medium [16] at 36°C in a humidified incubator with 5% CO₂. We established this protocol to obtain CSCs in our laboratory [17] in accordance with the Regional Scientific Ethical Committee (approval number S-VF-20040102). These cells have the ability to form new spheroids at clonal density. They have a karyotype typical of GBMs, and the ability to form highly invasive tumors upon orthotopic xenografting. Moreover, they differentiate into cells expressing neuronal, astrocytic and oligodendrocyte markers upon culturing in serum-containing medium. The GBM spheroids from T78 have a

![Image](https://example.com/image1.png)

Figure 7: Validation of some potential GBM population-specific marker candidates using immunohistochemical staining of histological sections. Transferrin (TF) expression was monitored in tumor core (A and B) and periphery (C and D). In A and C, tissue was stained with haematoxylin-eosin (HE) to define tumor core and tumor periphery. In B and D, tissue was immunohistochemically stained using anti-TF antibody. Few tumor cells expressed transferrin in the tumor core (B) while higher abundance of TF was found in the tumor periphery (D). Arrows highlight cells expressing TF and these areas are magnified in the inserts. (E) Fractions of cells positive for vinculin (VCL), pleiotrophin (PTN), transferrin (TF) and procollagen-lysine, 2-oxoglutarate 5-dioxygenase 2 (PLOD2) in tumor core and periphery.
Sample preparation for quantitative proteomics

We used three independent replicates per GBM subpopulation in this study. Cells (about 200000 for each cell type) were lysed and proteins were on-filter digested as previously published [65]. Briefly, GBM cells were lysed with a solution of 2% (w/v) sodium dodecyl sulfate (SDS), 20 mmol/L triethylammonium bicarbonate (TEAB), 0.1 mol/L dithiothreitol (DTT), phosphatase inhibitor (PhosSTOP, Roche, Switzerland), and protease inhibitor (cOmplete, Roche, Switzerland). Lysis was enhanced and DNA filaments sheared with tip sonication on ice. Protein concentration was measured using Qubit assay (Thermo Fisher Scientific, USA) as μg/μL. Proteins were loaded onto spin-filter units (Vivacon 500, 30,000 MWCO; Vivaproducts, USA) and the SDS-containing solution was washed out using an urea-containing solution (8 mol/L urea, 20 mmol/L TEAB). The lysate was loaded in three steps, 300 μL of urea solution was used for washing after each step and two washes with 600 μL of 1% (w/v) sodium deoxycholate (SDC), 20 mmol/L TEAB after all loadings. Alkylation of the reduced thiol groups was done with 50 mmol/L iodoacetamide, 1% (w/v) SDC, 20 mmol/L TEAB, followed by two times wash with 250 μL of 1% (w/v) SDC, 20 mmol/L TEAB, proteins were digested overnight with trypsin (1:50) (Promega, USA) in 1% (w/v) SDC, 20 mmol/L TEAB. Peptides were collected after centrifugation, and SDC was removed using ethyl acetate extraction from acidic solution (TFA (0.5% (v/v)).

Dimethyl labeling was performed according to a published protocol [66]. Briefly, 22.8 μg of peptides per labeling channel were dissolved in 100 μL of 0.1 mol/L TEAB. Peptide amount was measured by the amino acid analyzer (Biocrom 30, Biochrom, UK). Next, 4 μL of 4% (vol/vol) solution of CH₃O, CD₂O, or ¹³CD₂O was added and the samples were vortexed. Later, 4 μL of 0.6 mol/L NaBH₄CN or NaBD₄CN was added and the mixture was incubated for 60 min. at room temperature. The efficiency of labeling was monitored by liquid chromatography coupled to mass spectrometry (LC-MS) run before quenching the reaction. The reaction was quenched by adding 16 μL of 1% (vol/vol) ammonia solution and later 8 μL of 5% (vol/vol) formic acid. After quenching labelled peptides were mixed in 1:1:1 ratio resulting in 3 differently labeled samples (Table 3). The samples were dried in a SpeedVac and stored at −20°C until analyzed by LC-MS.

**LC-MS**

Peptides were separated using Dionex Ultimate 3000 nanoUPLC system coupled to Thermo Orbitrap Fusion mass spectrometer. Peptides were focused on the precolumn (PepMap C18 10 cm × 150 μm i.d., 5 μm) and eluted from analytical column (PepMap C18 50 cm × 75 μm i.d., 3 μm) with the gradient presented in Table 4. The mass spectrometer was configured to run in top speed mode with 3 seconds cycle duration. MS³ spectra were recorded in Orbitrap mass analyzer from 400 to 1200 Th, with 120000 resolution at 200 Th, AGC target value~5e5, maximum accumulation time~60 ms. Ions were isolated using quadrupole mass filter with 2 Th wide isolation window and fragmented using CID in the ion trap, spectra were acquired in Orbitrap, with 15,000 resolution at 200 Th, AGC target~1e4, maximum accumulation time~40 ms.

---

Table 3: Labeling scheme for GBM cells

<table>
<thead>
<tr>
<th>Replicate</th>
<th>Light (+ 28 Da)</th>
<th>Medium (+32 Da)</th>
<th>Heavy (+36 Da)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Replicate 1</td>
<td>Differentiated</td>
<td>Migrating</td>
<td>Spheroid</td>
</tr>
<tr>
<td>Replicate 2</td>
<td>Migrating</td>
<td>Spheroid</td>
<td>Differentiated</td>
</tr>
<tr>
<td>Replicate 3</td>
<td>Spheroid</td>
<td>Differentiated</td>
<td>Migrating</td>
</tr>
</tbody>
</table>

Table 4: LC gradient used for the analysis

<table>
<thead>
<tr>
<th>Time (min)</th>
<th>0</th>
<th>5</th>
<th>25</th>
<th>205</th>
<th>245</th>
<th>270</th>
<th>285</th>
<th>287</th>
<th>300</th>
</tr>
</thead>
<tbody>
<tr>
<td>% B</td>
<td>2</td>
<td>2</td>
<td>5</td>
<td>21</td>
<td>35</td>
<td>99</td>
<td>99</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

A: 0.1% formic acid; B: 80% acetonitrile in 0.1% formic acid.

Migration assay

Geltrex (Gibco) and serum-free medium were mixed (1 + 49), and 1400 μL was added to each well in 12-well plates. Coated plates were incubated overnight at 36°C, and the following morning the supernatant was aspirated. One spheroid (100–200 μm) was aspirated into a 0.1–2.0 μl pipette and placed on the coating. After incubating the plate for 75 minutes at 36°C, 1000 μl serum-free medium was added. The migration was followed with time-lapse microscopy for 5 days. Migration speed was calculated to find the optimal time to isolate the migrating cells. At the highest migration speed, the migrating cells were isolated by removing the “central” residual spheroid with a micro-pipette. The migrating cells were then washed twice in phosphate buffered saline (PBS). Migrating cell pellets were stored at –80°C until further proteomic analysis was performed. For comparison with non-migrating cells, pellets from free-floating spheroids were obtained.
Data analysis

Data analysis was performed using Thermo Proteome Discoverer 2.0.0.644. Mascot 2.3 was used as the database search engine. SwissProt database (2014.04) restricted to *Homo sapiens* (20,340 protein sequences) combined with common contaminants database (231 protein sequences) was used. Search parameters were: parent ion mass tolerance – 5 ppm, fragment ion mass tolerance – 0.02 Th; fixed modifications – carboxamidomethylated cysteine, variable modifications – oxidized methionine and labeled N-terminal and lysine. Reversed decoy database was searched separately. For SuperQuant analysis, all MS2 spectra were processed using home-built deconvolution node to produce fragmentation spectra consisting only of 1+ fragments. Next, deconvoluted spectra were processed with ComplementaryFinder node before database search. A detailed explanation of SuperQuant data processing is published [14]. Database search results were evaluated using Percolator 2.05 [67], with the standard feature set. All PSMs with q-value < 0.01 were grouped together using sequence and theoretical mass, and the highest Percolator SVM score was used as the score for the group. Qvality 2.05 [68] was used for the estimation of q-value on the PSM group level, PSM groups were filtered by q-value < 0.01. Each PSM group gives rise to one peptide. Proteins related to the filtered peptides were grouped using maximum parsimony principle. Quantification of peptides and proteins was performed using standard settings provided by Proteome Discoverer. The mass spectrometry proteomics data were deposited to the ProteomeXchange Consortium (http://proteomexchange.org) via the PRIDE [69] partner repository with the dataset identifier PXD005245.

Peptides abundancies reported by Proteome Discoverer were log2 transformed, normalized by subtracting the mean, and protein rollup was performed using RRollup method from Damente package [70]. Log2-transformed intensities were used for hierarchical clustering and CV calculation prior to normalization. Significantly regulated proteins were detected using limma R package as described earlier [71, 72].

The software Ingenuity Pathway Analysis (IPA, Qiagen) was used to retrieve biologic and canonical functions. We used default settings except for the following instances: species = human; tissue & cell lines = astrocytes, microglia, neurons (all), stem cells (all), nervous system (all), CNS cell lines (all), neuroblastoma cell lines (all), tissues, primary cells and cells not otherwise specified; experimental p-value cutoff = 0.05.

Immunohistochemistry

Immunohistochemical staining was performed on tissue samples matching the patient from whom T78 was derived. Briefly, formaldehyde fixed and paraffin embedded tissue was cut into 3 μm serial cut sections and placed on glass slides. The first section was stained with haematoxylin-eosin to define tumor core and tumor periphery. Four sections were dewaxed, and heat-induced epitope retrieval followed by quenching of endogen peroxidase was performed. Sections were then incubated with primary antibodies against vinculin (VCL) (1 + 800), procollagen-lysine, 2-oxoglutarate 5-dioxygenase 2 (PLOD2) (1 + 200), pleiotrophin (PTN) (1 + 400), and transferrin (TF) (1 + 800). Detection and staining procedures were done using EnVision (Dako) and the AutostainerPlus (Dako), respectively.

Image analysis

Stained sections were evaluated using the Visiopharm software (Visiopharm, Hørsholm, Denmark). The five stainings were aligned using the TISSUEalign module (Visiopharm). Sample images were then collected for 14 regions (7 areas of tumor core and 7 areas of tumor periphery) using systematic uniform random sampling (meander fraction based) at a sampling fraction that resulted in approximately 5 images per region. To ensure optimal alignment, sample images were reviewed, and the two best-aligned images for each region were included for further analysis. Using the newCAST module (Visiopharm), tumor cells were counted manually to estimate the number of positive and negative tumor cells in each staining separately. Positive cell fractions were calculated on the basis of the total tumor cell number.

ACKNOWLEDGMENTS AND FUNDING

The authors gratefully acknowledge financial support by the European Research Council (ERC) under the European Union’s Horizon 2020 Research and Innovation Programme (grant agreement No 646603), the VILLUM Center for Bioanalytical Sciences at the University of Southern Denmark, the Odense University Hospital Research Funds, and the Brazilian Funding Agencies CNPq (TVB, 421021/2016-0) and FAPEMIG (TVB, APQ-03242-16). The authors acknowledge the expert technical assistance of Helle Wohlleben in the immunohistochemical stainings.

CONFLICTS OF INTEREST

None.

REFERENCES


