Proteomics of gliomas: Initial biomarker discovery and evolution of technology

Juliya Kalinina, Junmin Peng, James C. Ritchie, and Erwin G. Van Meir

Laboratory of Molecular Neuro-Oncology, Departments of Neurosurgery, Hematology and Medical Oncology, School of Medicine, and Winship Cancer Institute, Emory University, Atlanta, GA 30322 (J.K., E.G.V.M.); Department of Human Genetics, Center for Neurodegenerative Diseases, Emory Proteomics Service Center, Emory University, Atlanta, GA 30322 (J.P.); Department of Pathology, Emory University School of Medicine, Atlanta, GA 30322 (J.C.R.)

Gliomas are a group of aggressive brain tumors that diffusely infiltrate adjacent brain tissues, rendering them largely incurable, even with multiple treatment modalities and agents. Mostly asymptomatic at early stages, they present in several subtypes with astrocytic or oligodendrocytic features and invariably progress to malignant forms. Gliomas are difficult to classify precisely because of interobserver variability during histopathologic grading. Identifying biological signatures of each glioma subtype through protein biomarker profiling of tumor or tumor-proximal fluids is therefore of high priority. Such profiling not only may provide clues regarding tumor classification but may identify clinical biomarkers and pathologic targets for the development of personalized treatments. In the past decade, differential proteomic profiling techniques have utilized tumor, cerebrospinal fluid, and plasma from glioma patients to identify the first candidate diagnostic, prognostic, predictive, and therapeutic response markers, highlighting the potential for glioma biomarker discovery. The number of markers identified, however, has been limited, their reproducibility between studies is unclear, and none have been validated for clinical use. Recent technological advancements in methodologies for high-throughput profiling, which provide easy access, rapid screening, low sample consumption, and accurate protein identification, are anticipated to accelerate brain tumor biomarker discovery. Reliable tools for biomarker verification forecast translation of the biomarkers into clinical diagnostics in the foreseeable future. Herein we update the reader on the recent trends and directions in glioma proteomics, including key findings and established and emerging technologies for analysis, together with challenges we are still facing in identifying and verifying potential glioma biomarkers.

Keywords: biomarker, glioma, proteomics.

Gliomas constitute more than 40% of all primary CNS neoplasms. Although all gliomas originate from neuroepithelial tissue, they vary considerably in morphology, location, genetic alterations, and response to therapy.1,2 The World Health Organization (WHO) classifies gliomas based on the different histological tumor types (astrocytic, oligodendrogial, mixed oligoastrocytic, and ependymal), as well as malignancy grades (I, II, III, and IV).3 The most malignant of gliomas (WHO grade IV) is glioblastoma multiforme (GBM), causing over 10 000 deaths each year in the United States alone.1

Presently, histopathology represents the gold standard for the typing and grading of gliomas and relies largely on particular architectural similarities of tumor cells with nonneoplastic glial cells.4,5 However, this histological classification remains unsatisfactory because of the lack of reproducibility and absence of precision in terms of prognosis, as evidenced by large interobserver variability.

Development of objective diagnostic, prognostic, and predictive markers for these deadly neoplasms is therefore a priority. Biomarkers could also provide unique information related to longitudinal response to therapy, and some may be disease mediators that need to be targeted for therapy.

The current review is intended to update the reader on the recent trends and directions in glioma proteomics, including important findings and established and emerging technologies as well as the challenges still faced in identifying and verifying potential glioma biomarkers.

Importance of Proteomics in the Fight Against Brain Cancer

Proteomic profiling represents the large-scale analysis of protein expression, post-translational modification, and
interactome. It is complementary to genomic and metabolomic analyses and has the ultimate goal of unifying the information into protein networks.8–10 Specifically, differential proteomic profiling of brain tumor vs. disease-free state allows for the detection and monitoring of pathology-related changes via the identification of potential CNS neoplasia diagnostic, prognostic, predictive, or treatment-assessing biomarkers. In addition, proteomic analysis of a proximal fluid is less invasive than biopsy while potentially informative into the causes of origin and progression of brain tumor pathologies.

A number of reports have emerged over the past decade reporting the analysis of the glioma proteome using human tissues and biofluids, as well as cell lines and animal models. A comprehensive overview of these studies was recently published.10 Herein, we focus on the established methodologies, current technologies, future directions, and challenges in glioma proteomics.

Biomarker Discovery

Sources of Biospecimens

Over the past decade, proteomic analysis has undergone many technological advances allowing for the discovery of comprehensive proteomic profiles using glioma patient biopsies, proximal fluids (cerebrospinal fluid [CSF] and cyst fluid), plasma, glioma cell lines, and animal models. Table 1 outlines a comprehensive list of important findings in the field of glioma proteomics.

Biological Fluids. Bodily fluids are an excellent source of disease markers, as they are readily accessible and represent drainage products from involved tissues, including proteins secreted by normal, diseased, and dying cells such as inflammatory markers, immunoglobulins or cytokines, vasoactive compounds, and mediators of cell signaling and/or growth pathways.11 The ability to follow such changes throughout the course of disease provides a unique opportunity to diagnose, prognose, predict, and monitor the presence and progression of a specific pathologic process.12,13

Plasma/ Serum. Plasma and serum have been valuable sample sources for human glioma biomarker research in the last 5 years.14–16 Plasma is a yellowish liquid component of whole blood that has been cleared of blood cells by whole blood centrifugation. It represents the largest component of the whole blood, making up approximately 55% of blood’s total volume. It is composed of mostly water (90% by volume), as well as an intricate mixture of proteins (with concentrations spanning 12 orders of magnitude), glucose, clotting factors, mineral ions, hormones, and carbon dioxide. The only difference between plasma and serum is the absence of clotting factors in the latter.

In 2005 Liu et al utilized 105 serum samples from glioma, benign brain tumors, and healthy individuals for screening and evaluation of potential serum biomarkers for the detection of astrocytomas (WHO grades I–IV). Specifically, using bioinformatics tools, they established the discriminate cluster analysis between GBM and healthy samples, as well as between GBM and low grades of astrocytoma samples, allowing identification of gliomas using panels of selected biomarkers.14 Two years later, another group reported 7 serum biomarkers specifically deregulated in astrocytomas (WHO grades I–IV) compared with the normal control group.15 In 2008, Petrik et al discovered and validated α-2-Heremans-Schmid glycoprotein as a predictive and prognostic biomarker for GBM patients identified in whole serum samples.16 Finally, another recent intriguing finding by Skog et al that tumor-derived microvesicles (exosomes) carrying mutant/variant epidermal growth factor receptor–vIII messenger RNA, microRNA, and proteins could be detected in serum provided impetus for utilizing the contents of exosomes as glioma biomarkers in the future.17

While undoubtedly valuable, several limitations associated with these fluids have restricted their use in the biomarker discovery for gliomas. These include: (1) low biomarker concentrations in comparison with the most abundant proteins (immunoglobulins and albumins), (2) a large dilution of CNS biomarkers in the peripheral circulation, (3) the existence of the blood–brain barrier, and (4) the ubiquitous collection of other proteins in plasma from throughout the body (not exclusively from the CNS).18,19 These detractors are, of course, offset by the ease of collection of plasma/serum.

Cerebrospinal fluid. To circumvent some of these weaknesses, another extracellular fluid, CSF, has emerged as a novel key source for biomarkers in glioma. CSF is a clear, colorless, and largely cell-free liquid circulating within the ventricles, in the spinal column, and around the brain in a cavity called the subarachnoid space.11 It takes approximately 7–8 h for CSF to be actively secreted and reabsorbed into the venous circulation, with 300 mL of new CSF produced each day.11,20

The fluid is composed of water, glucose, salts, metabolites, nucleic acids, peptides, proteins, and enzymes. Although CSF shares many properties with serum, such as the presence of albumin and immunoglobulin as its largest constituents and maintaining similar levels of salt and ions, several key features make CSF unique in its potential application in glioma proteomics. These include: (1) CSF protein composition is at least an order of magnitude less complex than serum, with 0.2–0.5 mg/mL overall protein content in CSF vs. 55–80 mg/mL in serum; and (2) the protein profile of CSF originates chiefly from the CNS, reflecting the absence of a barrier at the CSF–brain interface. These characteristic features make CSF a protein reservoir into which CNS proteins and metabolites are secreted during normal brain function. Alterations in the protein content of CSF are measures of the changes occurring during a pathologic process in the brain; this
<table>
<thead>
<tr>
<th>#</th>
<th>Manuscript, Year</th>
<th>Source</th>
<th>Proteomic Method</th>
<th>Biomarkers Identified</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Hanash et al. 2002[97]</td>
<td>Biopsies: gliomas, grades I and II ((n = 11)) vs. GBM ((n = 12))</td>
<td>2DE</td>
<td>Identified 22 upregulated proteins, including oncoprotein 18, nucleoside diphosphate kinase A, GSTP, and moesin</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Zhang et al. 2003[98]</td>
<td>Cell lines: fetal human astrocytes (FHA), U87MG, U87MG(\Delta)EGFR</td>
<td>2DE, LC-MS-MS, MALDI-TOF MS</td>
<td>23 and 29 proteins were expressed exclusively in FHA and U87MG, respectively; 3 proteins were upregulated in U87MG(\Delta)EGFR relative to U87MG: ubiquitin, cystatin B, TTG and 4 proteins (Hsp27, major vault protein, TTG, and cystatin B) were further verified</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Hiratsuka et al. 2003[99]</td>
<td>Biopsies: gliomas, grades II and III and GBM; nontumoral tissue</td>
<td>2DE</td>
<td>11 proteins were upregulated in gliomas: CRMP4, albumin, hemopexin, copine I, prohibitin, phosphoserine phosphatase, TTR, apolipoprotein A-1, catechol-O-methyltransferase, fatty acid binding protein 7, and neurocalcin (\delta); 4 were downregulated in gliomas: SIRT2, profilin-2, UCH-L1, and PEA-15</td>
<td>SIRT2 was further verified</td>
</tr>
<tr>
<td>4</td>
<td>Zheng et al. 2003[26]</td>
<td>CSF: OD ((n = 2)), AOD ((n = 3)), AOA ((n = 2)), PA ((n = 2)), GBM ((n = 1)), medulloblastomas ((n = 2)) vs. normal controls ((n = 14))</td>
<td>2DE, MALDI-TOF MS</td>
<td>Identified 2 proteins that were found in primary brain tumors, but absent in controls, including N-myc and caldesmon (I-CaD); I-CaD was, in addition, absent in medulloblastomas samples</td>
<td>N-myc and I-CaD were further verified by IHC</td>
</tr>
<tr>
<td>5</td>
<td>Hobbs et al. 2003[100]</td>
<td>Biopsies: GBM</td>
<td>SELDI-TOF MS</td>
<td>Identified a panel of proteomic profiles from radiographically distinct regions of a tumor</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Furata et al. 2004[30]</td>
<td>Biopsies ((n = 30)): A ((n = 5)), AA ((n = 2)), GBM ((n = 13)); 6 primary, 7 secondary</td>
<td>2DE, NanoLC-MS/MS</td>
<td>Identified the following differentially expressed proteins in primary GBM: tenascin precursor, enolase-1, centrosome associated protein 350, EGFR; in secondary GBM: ERCC6, DUOX2, Wnt-11 precursor, cadherin-related tumor suppressor homolog precursor, ADAMTS-19, and hnRNP A3</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Schwartzs et al. 2004[11]</td>
<td>Biopsies ((n = 20)): gliomas, grade II ((n = 3)), grade III ((n = 3)), GBM ((n = 4)) vs. normal controls ((n = 5))</td>
<td>MALDI-MS</td>
<td>Utilized MALDI for direct proteomic tissue profiling</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Iwadate et al. 2004[29]</td>
<td>Biopsies ((n = 85)): A ((n = 10)), AA ((n = 13)), GBM ((52)) and normal brain tissues ((n = 10))</td>
<td>2DE, MALDI-TOF MS</td>
<td>Identified 37 differentially expressed proteins between low- and high-grade astrocytomas, including nucleolar GTP-binding protein, glutamate dehydrogenase, and CREB-1, Rac1, Rhoa, and enolase were verified further by IHC</td>
<td>CREB-1, Rac1, Rhoa, and enolase were verified further by IHC</td>
</tr>
<tr>
<td>9</td>
<td>Chumbalkar et al. 2005[51]</td>
<td>Biopsies ((n = 27)): A, grade I ((n = 2)), AA ((n = 1)), GBM ((n = 14)), and GBM ((n = 10))</td>
<td>2DE, MALDI-TOF</td>
<td>Identified 72 differentially expressed proteins in grade-specific manner. 29 of these were short-listed, including Rho-GDI, prohibitin, GFAP, HSP60, HSP70, vimentin, and (\alpha)-crystallin</td>
<td>Prohibitin and GFAP were further verified by WB</td>
</tr>
<tr>
<td>#</td>
<td>Manuscript, Year</td>
<td>Source</td>
<td>Proteomic Method</td>
<td>Biomarkers Identified</td>
<td>Notes</td>
</tr>
<tr>
<td>----</td>
<td>------------------</td>
<td>--------</td>
<td>------------------</td>
<td>----------------------</td>
<td>-------</td>
</tr>
<tr>
<td>10</td>
<td>Liu et al. 2005</td>
<td>Serum (n = 105); A, grades I, II, III and GBM, nontumoral controls</td>
<td>SELDI-TOF-MS</td>
<td>Provided an extensive bioinformatic analysis of nontumoral controls vs. GBM, and of low vs. high A grades</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Iwadate et al. 2005</td>
<td>Biopsies (n = 93); response to cytotoxic drugs</td>
<td>2DE, MALDI-TOF MS</td>
<td>Identified a list of 41 proteins that significantly affected the in vitro chemosensitivity to each category of anticancer agents tested</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Vogel et al. 2005</td>
<td>Biopsies and cell lines: U87, U251, U118, A172 compared with GBM biopsies (n = 8)</td>
<td>2DE</td>
<td>Identified 220 differentially expressed proteins, including DNA polymerase (\alpha)-catalytic subunit A, hHYD, LRP-1, CAPG, AKAP-9, Calgizzarin, Anx, LASP-1, PDI, PKM1/2, S100A6, GFAP</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Odreman et al. 2005</td>
<td>Biopsies: fibrillary A, grade II (n = 10) - GBM (n = 10)</td>
<td>2DE, LC-ESI-MS-MS</td>
<td>Identified 15 differentially expressed proteins, including peroxiredoxin 1/6, (\alpha)-internexin, BTF3, PDI A3, UCHL1, PKA</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Schwartz et al. 2005</td>
<td>Biopsies (n = 162) and cell line: nontumoral controls (n = 19), grade II (n = 29), grade III (n = 22), GBM (n = 57), and U118MG cell line</td>
<td>2D-LC, MALDI MS-MS, MALDI ESI-IT</td>
<td>Identified 6 potential biomarkers, including calcin, calpactin I light chain, tubuline-specific chaperone A, astrocytic phosphoprotein PEA 15, fatty acid binding protein 5, Dynein light chain 2 and calccin discriminated between 2 survivor groups</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Wang et al. 2005</td>
<td>Biopsies (n = 1): GBM (n = 1)</td>
<td>cIEF, ESI MS-MS</td>
<td>Identified 1820 unique proteins, including nestin and AnxA1, which were unambiguously identified</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Fiore et al. 2005</td>
<td>Biopsies: (gliomas, grade I (n = 1), grade II (n = 1)m and GBM (n = 3), together with noncancerous brain samples)</td>
<td>1DE, MALDI-TOF MS, ES-MS-MS</td>
<td>Identified (\alpha)-tubulin as one of the nitrated proteins in tumor samples, relative to those of grade I or nontumoral controls</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>Shim et al. 2006</td>
<td>Cell lines: U87 glioma cell lines expressing 3 different PTEN clones</td>
<td>2DE, MALDI-TOF-MS</td>
<td>Identified 50 differentially expressed proteins, including enolase, vimentin, cathepsin L preproprotein, SRC family associated phosphoprotein, coflin 1, glutathion S-transferase chain A, dihydropyrimidinase-like 2 protein, BANP isoform b, HSPA8, (\beta)-actin</td>
<td>Vimentin was verified further</td>
</tr>
<tr>
<td>18</td>
<td>Zhou et al. 2006</td>
<td>Cell lines: glioma cell lines U251 and A172; in vivo and in vitro tumor forming ability</td>
<td>2DE, MALDI-TOF-MS</td>
<td>Identified 18 differentially expressed proteins, including transketolase, prohibitin, PR48, PARP, CYP, CypA, Grfb, DJ-1, and cathepsin-D</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>Trog et al. 2006</td>
<td>Cell lines: U87 glioma cell line with or without the following treatments: irradiation (RT), chemotherapy (CT) with TMZ or combined radio chemotherapy (RCT)</td>
<td>2DE, nanoelectrospray MS-MS</td>
<td>Identified the following differentially expressed proteins: Vimentin (elevated after CT and combined RCT; but reduced after RT alone), TIP47 (reduced after CT, RT, and RCT), Rhoda GTPase (upregulated in CT, absent in RT and reduced in RT)</td>
<td></td>
</tr>
</tbody>
</table>

Continued
<table>
<thead>
<tr>
<th>#</th>
<th>Manuscript, Year</th>
<th>Source</th>
<th>Proteomic Method</th>
<th>Biomarkers Identified</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>Billecke et al. 2006</td>
<td>Cell line: LNZ308 glioma cells; untreated vs. response to chemotherapeutic agents (cisplatin, BBR3464, BBR3610)</td>
<td>PF2D, MALDI MS/MS</td>
<td>Identified TIM as a protein differentially regulated upon BBR3610 treatment</td>
<td>TIM was further verified by WB</td>
</tr>
<tr>
<td>21</td>
<td>Khwaja et al. 2006a</td>
<td>Cell line: LNZ308 with tetracycline inducible or repressible wt-p53 expression</td>
<td>2DE, cleavable ICAT</td>
<td>Identified 111 secreted proteins: 39 and 21 showed enhanced and inhibited secretion, respectively, in response to wt-p53 expression</td>
<td>Galectin-1, galectin-3, beta-2-microglobulin, SPARC, fibroblast growth factor-4, and transforming growth factor beta were further verified by WB</td>
</tr>
<tr>
<td>22</td>
<td>Khwaja et al. 2006b</td>
<td>CSF (n = 32): CNS neoplasms (n = 10), CNS inflammatory disease samples (n = 12), control cohorts (n = 10)</td>
<td>1DE, SELDI-TOF-MS, MALDI-TOF/TOF, MALDI-TOF</td>
<td>Identified carbonic anhydrase as a prognostic marker of neoplastic disease</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>Khwaja et al. 2006c</td>
<td>CSF (n = 60): A grade II vs. GBM</td>
<td>2DE, cICAT, MALDI-MS/MS</td>
<td>Identified attractin as an upregulated marker in GBM</td>
<td>Attractin was further verified by functional assays and migration studies</td>
</tr>
<tr>
<td>24</td>
<td>Li et al. 2006</td>
<td>Biopsies (n = 14): A, grades II, III, GBM (n = 10) vs. normal brain controls (n = 4)</td>
<td>2DE, nanoLC-MS/MS</td>
<td>Identified the following upregulated proteins in GBM: Ki-67, N-CoR, and IRS-2</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>Jiang et al. 2006</td>
<td>Biopsies (n = 82): A (n = 8), OD (n = 7), OA (n = 3), AA (n = 10), AOD (n = 11), AO (n = 6), and GBM (n = 37)</td>
<td>Reverse phase protein array</td>
<td>Identified 18 proteins differentially expressed in GBM, including IkappaB, EGFRpTyr845, AKTpThr308, PI3K, IGFBP2/5, MAMP, VEGF, pRB, Bcl-2, and c-Abl</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>Goplen et al. 2006</td>
<td>Animal models: GBM xenografts (n = 2) in nude rats; angiogenic vs. nonangiogenic tumors</td>
<td>2DE, MALDI-TOF MS</td>
<td>Identified PDI A6 to be upregulated in nonangiogenic phenotype compared with the angiogenic one</td>
<td>PDI A6 was further verified</td>
</tr>
<tr>
<td>27</td>
<td>Wibom et al. 2006</td>
<td>Animal models: BT4C rat glioma model in BDIX rats (n = 24); response to radiotherapy</td>
<td>SELDI-TOF-MS</td>
<td>Identified 77 different peaks when comparing treated vs. untreated</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>Ngo et al. 2007</td>
<td>Cell lines: U251 (1p+/+), A172 (1p+/-) glioma cell lines, AOD</td>
<td>2D DIGE, MALDI MS/MS</td>
<td>Identified 19 differentially expressed 1p-encoded proteins, including stathmin, a-enolase, and DJ-1</td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>Bouamrani et al. 2007</td>
<td>Biopsies: Brain cancer tissue and mouse tissue samples</td>
<td>SELDI-TOF MS</td>
<td>Identified 3 potential markers that discriminate GBMs from ODs</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>Khalil 2007</td>
<td>Biopsies (n = 50): grades I, II, III, and IV, together with normal brain tissue</td>
<td>2D DIGE, MALDI-TOF MS</td>
<td>Identified 91 differentially expressed proteins in GBM, including peroxiredoxin 4, SH3 domain binding glutamic acid-rich-like protein 3, aldolase C fructose-biphosphatase, creatine kinase, leucine aminopeptidase, neurofilament triplet L, stathmin, vacuolar ATP synthase subunit E, and Alb protein</td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>Okamoto et al. 2007</td>
<td>Biopsies: OD and AOD (n = 9), without (n = 5) and with (n = 4) 1pLOH</td>
<td>2D, LC-MS</td>
<td>Identified a panel of 19 protein markers that distinguish ODs with and without 1pLOH, as an indication for sensitivity to chemoresistance</td>
<td></td>
</tr>
<tr>
<td>#</td>
<td>Manuscript, Year</td>
<td>Source</td>
<td>Proteomic Method</td>
<td>Biomarkers Identified</td>
<td>Notes</td>
</tr>
<tr>
<td>----</td>
<td>-----------------</td>
<td>--------</td>
<td>------------------</td>
<td>-----------------------</td>
<td>-------</td>
</tr>
<tr>
<td>32</td>
<td>Puchades et al. 2007</td>
<td>Cell lines: U87 cell line with and without wt p53; treatment with SN38</td>
<td>2DE, MALDI-TOF</td>
<td>Identified galactokinase 1, GRP78, caspase 14, galectin-1, differentially expressed in TP53 cells treated with SN38</td>
<td>Galectin-1 was further verified by WB and IHC</td>
</tr>
<tr>
<td>33</td>
<td>Mustafa et al. 2007</td>
<td>Biopsies (n = 20): GBM (n = 10) and normal brain samples (n = 10); comparing blood vessels</td>
<td>MALDI-FTMS, nanoLC-MALDI-TOF/TOF</td>
<td>Identified 16 differentially detected peptides in glioma vessels for acidic calponin 3, fibronectin, and collagen 2</td>
<td></td>
</tr>
<tr>
<td>34</td>
<td>Guo et al. 2007</td>
<td>Biopsies: snap-frozen and FFPE GBM tissues</td>
<td>Integrated CIEF/nanoRPLC coupled with nanoESI-LTQ-MS/MS</td>
<td>Identified 2733 nonredundant SwissProt protein entries from FFPE GBM tissues. 83% of identified proteins overlapped with those obtained from snap-frozen tissues of the same patient</td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>Khwaja et al. 2007</td>
<td>CSF (n = 73): A, grades II, III, and GBM, other primary brain tumors, metastatic and infectious control samples</td>
<td>2DE, MALDI-MS/MS, cICAT, 2D-LC-MS</td>
<td>Identified a panel of 103 tumor-specific markers, including upregulated SPARCL1, FGF14, VEGF-B, tau, b2M, b-defensin 6/7, and attractin</td>
<td>SPARCL1, FGF14, VEGF-B, tau, b2M, b-defensin 6/7, and attractin were further verified by WB</td>
</tr>
<tr>
<td>36</td>
<td>Zhang et al. 2007</td>
<td>Serum (n = 140): A, grade I and II (n = 30), AA, and GBM (n = 43), normal controls (n = 56)</td>
<td>SELDI-TOF-MS chip</td>
<td>Detected 7 putative differentially regulated biomarkers</td>
<td>Prognostic value was further verified in a different cohort of GBM (n = 72)</td>
</tr>
<tr>
<td>37</td>
<td>Petrik et al. 2008</td>
<td>Serum (n = 200): A, grades II, III, GBM, and controls</td>
<td>SELDI-TOF-MS, CM10 chip</td>
<td>Identified B-chain of a2-Heremans-Schmid glycoprotein as a downregulated marker with increasing tumor grade</td>
<td>Peroxiredoxin 6 and rho GDP were verified further</td>
</tr>
<tr>
<td>38</td>
<td>Park et al. 2008</td>
<td>Biopsies (n = 1): OD and AOD (from the same patient)</td>
<td>2DE, MALDI-TOF MS</td>
<td>Identified 23 differentially expressed spots, including overexpression of peroxiredoxin 6 and underexpression of rho GDP, as predictive factors for the malignant transformation of OD</td>
<td></td>
</tr>
<tr>
<td>39</td>
<td>Li et al. 2008</td>
<td>Biopsies (n = 8): GBM (n = 4) and normal brain cortex tissue (n = 4)</td>
<td>Capillary IEF-nRP-LC-MS/MS</td>
<td>Identified the following upregulated proteins in GBM: SMCS, BSG9, prothymosin alpha, WHSC1, and Ki-67</td>
<td>WHSC1 was verified further</td>
</tr>
<tr>
<td>40</td>
<td>Bian et al. 2008</td>
<td>Cell lines: CHG-5 A cell line, sHG-44 AA cell line, and U87 glioma cell line</td>
<td>2DE, MALDI-TOF</td>
<td>Identified 10 differentially expressed proteins, including PAG-A, beta galactoside binding lectin, ASF-3, Eif-5A, coflin-1, GST-pi, glyceraldehyde-3-phosphate dehydrogenase, a-enolase, and cyclophilin</td>
<td></td>
</tr>
<tr>
<td>41</td>
<td>Seyfried et al. 2008</td>
<td>Cell lines: C8-D30 astrocyte cell line compared with A cell line, CT-2A</td>
<td>LC-MS/MS LTQ-linear ion trap</td>
<td>Identified 25 differentially regulated proteins in CT-2A, including upregulated CSGP4 and IFIT1M/13</td>
<td></td>
</tr>
<tr>
<td>42</td>
<td>Fischer et al. 2008</td>
<td>Plasma: GBM (n = 10)</td>
<td>2D DIGE</td>
<td>Identified anti-thrombin and fibrinogen as prognostic markers for perillyl alcohol treatment of GBM</td>
<td></td>
</tr>
<tr>
<td>43</td>
<td>Zhang et al. 2009</td>
<td>Cell lines: U251 astrocytoma cell lines (treated or untreated with RA inhibitor)</td>
<td>2DE, MALDI-TOF MS</td>
<td>Identified 36 altered proteins, including PHGDH</td>
<td>PHGDH was further verified</td>
</tr>
<tr>
<td>#</td>
<td>Manuscript, Year</td>
<td>Source</td>
<td>Proteomic Method</td>
<td>Biomarkers Identified</td>
<td>Notes</td>
</tr>
<tr>
<td>----</td>
<td>-----------------</td>
<td>--------</td>
<td>------------------</td>
<td>----------------------</td>
<td>-------</td>
</tr>
<tr>
<td>44</td>
<td>Hill et al. 2009</td>
<td>Cell lines: U87 serum-free conditioned medium (n = 3)</td>
<td>NanoLC-MS, LC-MS/MS</td>
<td>Identified 35 differentially expressed proteins, including upregulated cathepsin-L, FSTL1, NBL1, TFP12, and downregulated tenasin C and IGF2R</td>
<td>Identified proteins were further verified by WB and DNA microarrays</td>
</tr>
<tr>
<td>45</td>
<td>Park et al. 2009a</td>
<td>Biopsies: gliomas, grades II and III, GBM (n = 3)</td>
<td>2DE, MALDI-TOF MS</td>
<td>Identified UCHL1 upregulated in grade II compared with grades III/IV; as well as TTR downregulated in grade II</td>
<td></td>
</tr>
<tr>
<td>46</td>
<td>An et al. 2009</td>
<td>Biopsies and cell lines: GBM (n = 10) and normal cell lines U87, HEK293T, F3 neural stem cells</td>
<td>2DE, MALDI-TOF</td>
<td>Identified 16 upregulated proteins in GBM, including AnxA2, TIMP-1, COL11A1, Bax, CD74, TNFRSF8, and SPTLC2</td>
<td></td>
</tr>
<tr>
<td>47</td>
<td>Park et al. 2009b</td>
<td>Biopsies: long-term survival GBM (LTSGBL) samples</td>
<td>2DE, MALDI-TOF MS</td>
<td>Identified the absence of overexpression of MnSOD in LTSGBL patients as a prognostic marker for GBM patients</td>
<td>MnSOD was further verified</td>
</tr>
<tr>
<td>48</td>
<td>Schuhmann et al. 2009</td>
<td>CSF (n = 24): GBM vs. normal controls</td>
<td>RP-LC MALDI-MS, nanoESI-qTOF-MS/MS</td>
<td>Identified 4 proteins upregulated in GBM, i.e. C-terminal peptides of α-1-antichymotrypsin, osteopontin, TTR, and N-terminal albumin peptide</td>
<td>Proteins were verified by TMAs</td>
</tr>
<tr>
<td>49</td>
<td>Rajcevic et al. 2009</td>
<td>Animal models: GBM xenografts in nude rats (n = 4); nonangiogenic vs. angiogenic phenotype</td>
<td>iTRAQ, 2D-LC, MALDI-TOF/TOF</td>
<td>Identified a panel of 60 upregulated proteins in angiogenic GBM; a panel of 6 upregulated proteins in nonangiogenic GBM</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>Grzendowski et al. 2009</td>
<td>Biopsies (n = 10): OD (n = 5) and AOD (n = 5)</td>
<td>2D DIGE, MALDI-MS</td>
<td>Identified 8 differentially regulated proteins, including glial fibrillary acidic protein and creatine kinase B-type (CKB)</td>
<td>CKB was further verified by WB</td>
</tr>
<tr>
<td>51</td>
<td>Koncarevic et al. 2009</td>
<td>Cell lines: NCH89, NCH82; response to platinum based drugs</td>
<td>2DE, MALDI-TOF/TOF</td>
<td>Identified differentially regulated proteins in response to drug treatment, including TrxR, GR, p53_S15Phos, AnxA1, vimentin, α-enolase, SOD, and EIF-5A</td>
<td></td>
</tr>
<tr>
<td>52</td>
<td>Melchior et al. 2009</td>
<td>Biopsies: GBM (n = 1)</td>
<td>2D-LC, MALDI-TOF/TIF, IP-RP HPLC</td>
<td>Demonstrated how bottom-up and semi-top-bottom approach can be complementary; identified the following proteins: calnexin, neurofascin, transmembrane protein 65, gamma-glutamyltransferase 5, mitochondrial import receptor subunit TOM22 homolog, and adipocyte plasma membrane-associated protein</td>
<td></td>
</tr>
<tr>
<td>53</td>
<td>Ohnishi et al. 2009</td>
<td>CSF: A, AA and GBM</td>
<td>2DE</td>
<td>Identified gelsolin as a prognostic marker whose expression decreased with histological grade</td>
<td>Gelsolin was further verified by IHC</td>
</tr>
<tr>
<td>54</td>
<td>Fischer et al. 2010</td>
<td>Cell lines: untreated GBM, A172 cells or exposed to perillyl alcohol for different time-points</td>
<td>2D-LC-MS</td>
<td>Identified an extensive list of 4000 proteins, with some of the time-point-specific proteins, i.e. PICALM isoform 2 and AP2A1 isoform B for the 24 h treatment time point</td>
<td></td>
</tr>
</tbody>
</table>
makes CSF a valuable source of biomarkers for brain malignancies.

Since 2001, CSF has been used as a key biospecimen for proteomic profiling in 6 manuscripts, resulting in the identification of candidate diagnostic protein and peptide markers in gliomas.°2°-°2° Such studies are expected to be facilitated in the near future because the protein composition of normal human CSF has recently been established.°2°

One of the initial studies by Khwaja et al, utilized surface-enhanced laser desorption/ionization (SELDI) on 32 CSF samples, including neoplastic and inflammatory CNS disease groups, to identify carbonyl anhydrase as a candidate diagnostic biomarker of neoplastic disease.°2° Carbonyl anhydrase is upregulated by hypoxia; thus, it came as no surprise that it may be increased in GBM, which displays extensive areas of micronecroses surrounded by hypoxic pseudopalisading cells.°2° Follow-up studies from the same group using 2-dimensional gel electrophoresis (2DE) and isopeptide-coded affinity tag (ICAT) methodologies on 73 CSF and cyst fluid samples subsequently identified 103 astrocytoma grade-specific markers and proposed the first candidate biomarker panel that could distinguish different grades of astrocytoma from normal brain or other primary or metastatic brain tumors. One selected marker, attractin, was studied further using functional assays and was shown to be a key mediator of glioma cell migration.°2°°2° These initial studies provided proof-of-principle that global proteomic approaches can be used for the discovery of biomarkers and new disease mediators in the CSF of brain tumor patients.

Tissue. Direct analysis of tumor tissue represents an attractive approach for glioma-specific protein biomarker discovery, as it is able to circumvent the highly dynamic and complex protein composition of biofluids.°10° Of course, not all markers identified by this approach will be detected in bodily fluids, as they may not be secreted proteins, their diffusion might be limited upon tumor cell death, or they may be eliminated through phagocytosis by macrophages/microglia.

Several proteomic profiling studies in tissues have been reported to date, including analysis of patient biopsies, glioma cell lines, and animal models. However, the invasive nature of biopsy, the inadequate amount of tissue to obtain statistically significant results, and the inherent sampling errors that can occur will limit the use of tissue markers as a clinical diagnostic. To circumvent these limitations, other methods (see below) provide means of sampling and allow proteomics on increasingly smaller amounts of tissue. On the whole, tissue-derived biomarkers continue to offer important uses for the characterization of resected tumor specimens and the understanding of their underlying biology.

Patient Biopsies, Cell Lines, and Animal Models. The largest percentage of glioma biomarker discovery

Table 1. Continued

<table>
<thead>
<tr>
<th>#</th>
<th>Manuscript, Year</th>
<th>Source</th>
<th>Proteomic Method</th>
<th>Biomarkers Identified</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>55</td>
<td>Grzendowski et al., 2010°11°</td>
<td>Biopsies: OA (n = 9) with (n = 4) or without 1p/19q deletions (n = 5)</td>
<td>2D DIGE, 2DE, MALDI-TOF-TOF MS</td>
<td>Identified 22 differentially expressed proteins in tumors with or without combined deletion of 1p and 19q, including vimentin, villin 2, annexin A1, and glial fibrillary acidic protein</td>
<td>Vimentin, villin 2, annexin A1, and glial fibrillary acidic protein were further verified by WB and IHC</td>
</tr>
<tr>
<td>56</td>
<td>Rostomily et al., 2010°12°</td>
<td>Biopsies: OD (n = 10) with (n = 5) and without 1p/19q deletions (n = 5)</td>
<td>Microcapillary LC-MS, ICAT</td>
<td>Identified 163 differentially expressed proteins when comparing tumors with and without 1p/19q deletion, including BCAN</td>
<td>BCAN was further verified by WB</td>
</tr>
<tr>
<td>57</td>
<td>Geiger et al., 2010°13°</td>
<td>Biopsies and cell lines: A cell line 1321N1, AA cell line CCF-STTG1, and GBM cell lines U87MG, U118MG, and U373MG</td>
<td>Super-SILAC, MS analysis using LTQ-Orbitrap Velos</td>
<td>Identified 5183 proteins and quantified 4318 of them when comparing a lysate of astrocytoma protoplasmaticum mixed with Super-SILAC mix composed of 5 brain tumor cell lines</td>
<td>The paper provided a new super-SILAC methodology that allows for a more accurate quantification of proteins in brain tumor tissues</td>
</tr>
<tr>
<td>58</td>
<td>Fischer et al., 2011°14°</td>
<td>Cell lines: GBM A172 cell line or A172r (perillyl alcohol-treatment resistant) cell line</td>
<td>2D-LC-MS</td>
<td>Identified 57 differentially expressed proteins when comparing A172 and A172r cell lines, including an elevated expression of HSP70 in A172r</td>
<td>HSP70 was further verified by WB</td>
</tr>
</tbody>
</table>

Abbreviations: A, astrocytoma (grade II); AA, anaplastic astrocytoma (grade III); GBM, glioblastoma multiforme (grade IV); OD, oligodendroglioma; AOD, anaplastic oligodendroglioma; OA, oligoastrocytoma; AOA, anaplastic oligoastrocytoma. CREB, cyclic adenosine monophosphate–responsive element binding protein 1; GSTp, glutathione S–transferase pi; GTP, guanosine triphosphate; IGFBP, insulin-like growth factor-binding protein; LTSGBL, long-term survival glioblastoma; tTG, tissue transglutaminase; TTR, transthyretin; WB, western blot.
In sum, the aggregate of these studies supports a role for biomarkers, such as attractin and galectin-1, that were identified in 2 or more studies and have been verified in patient-derived tissues by immunohistochemistry (IHC). Until now, however, there has been minimal overlap between the tissue-identified biomarkers and those found in biofluids, with VEGF being one of the few biomarkers identified using both biological sources. Finally, any one of these markers has yet to be translated into the clinic.

**Experimental Design, Sample Selection, and Handling**

A crucial step for obtaining meaningful results from any proteomic study is a well-developed experimental design and sample selection. The following key points, among others, need careful attention:

1. The test set has to be large enough to offer statistical significance.
2. Variation in individual samples needs to be addressed.
3. Adequate amounts of specimens have to be available for multiple repetitions of proteomic analysis (minimal amount for a single analysis of tissue: \( \approx 1 \text{ mm}^2 \) tissue area; of proximal fluids: \( \approx 0.05 \text{ mL} \)).
4. Tissue samples need to be well defined in terms of pathology/morphology/histology.
5. Bodily fluids need to be handled and stored uniformly throughout to minimize protein degradation and proteolysis.
6. CSF samples have to be consistent in terms of the site (lumbar, ventricular, or subdural) and timing (pre-, intra- or postop) of the procurement.
7. Patients’ safety may not be jeopardized; that is, lumbar punctures should be avoided as a choice for CSF procurement in patients with increased intracranial pressure, as this may cause tissue shifts and vascular obstruction resulting in brain infarction and rapid death.
8. Samples have to be examined for contamination and degradation.
9. Optimal experimental conditions for each sample type need to be established, including the need for processing prior to proteomic analysis to enrich protein concentration, the removal of interfering proteins, the use of suitable buffer conditions, minimal lysis times, and high-yield protein isolation techniques.

As an example, a “good” CSF sample is one that is well characterized, supported by clinical data, and contains a reliable diagnosis. It was obtained immediately after procurement, is clear of whole blood and of any other cell content contamination by centrifugation, and has been rapidly stored below \(-80^\circ\text{C}\) to conserve protein integrity against degradation. Prior to proteomic analysis,
A CSF sample may need to be further processed by desalting, prefractionation, dialysis, ultrafiltration,\textsuperscript{63} depletion of high-abundance proteins (e.g., albumin and immunoglobulin), and protein precipitation (with acetone and/or trichloroacetic acid).\textsuperscript{61} These are carried out to enrich and thereby improve identification and quantification of protein biomarkers, which may otherwise be present in scarce amounts or masked by abundant proteins, and to eliminate salts (>150 mmol/L), which may interfere with electrophoretic analysis.

Methodologies for Data Acquisition and Processing

Two major and complementary methodologies are widely utilized for downstream quantitative protein and peptide analysis in glioma proteomics research (Fig. 1). Although gel-based protein profiling predominated in the early stages of proteomics, a mass spectrometry (MS)–based approach is emerging as the mainstream technology in proteomics today because it offers high levels of resolution, facilitating subsequent analyses and improving confidence for accurately determining differential protein expression profiles.

Gel-based technique. In 2DE-based proteomics, proteins in a complex sample are separated on a polyacrylamide gel medium using electrokinetic methods (most typically electrophoresis) in 2 dimensions, one based on isoelectric point (pI) and the other on molecular weight, thus providing a unique 2D profile of the sample’s proteome.\textsuperscript{12,64} Two-dimensional liquid-phase electrophoresis is similar to 2DE except for the isoelectric focusing (IEF) that is performed in the liquid phase.\textsuperscript{65} After electrophoretic separation, the spots of interest are identified and quantified with the help of MS after in-gel digestion of stained or fluorescently tagged protein spots.

The gel-based technique was used for the analysis of neoplastic CSF from different grades of astrocytoma, where CSF from neoplastic CSF was compared with the CSF of nonneoplastic and inflammatory patients.\textsuperscript{24} Another recent study used a combination of 2DE gel and analysis by matrix-assisted laser desorption/ionization (MALDI)–time-of-flight mass spectrometer (TOF) to identify and characterize 16 proteins elevated in glioma tissues.\textsuperscript{44}

While gel-based technologies have the advantage of providing a detailed analysis of post-translational modification, several limitations of the methodology hinder unambiguous and robust protein spot identification/quantification: (1) proteins underrepresented in the mix may be masked or completely undetected, as they often comigrate with other proteins into the same spot, and (2) proteins with extreme size, pI, or hydrophobic properties are often missed as well. Nonetheless, 2DE technology is a simple approach that can be practiced in any wet lab. With continuous improvement, it is still widely used in neuroscience laboratories. Several strategies, including more sensitive staining methods, alternative separation techniques, larger gels, sample fractionation, limited range IEF, and pH gradient (or IPG) strips, have improved 2DE reproducibility, resolution, and thus overall analysis immensely (reviewed in\textsuperscript{66,67}). In 2005, the 2D difference gel electrophoresis (DIGE) technology emerged, demonstrating improved reproducibility by avoiding gel-to-gel variation, which permitted high-resolution separation of modified proteins and allowed hundreds of proteins to be quantified simultaneously.\textsuperscript{68} As an example, using 2D DIGE as an approach of choice, Ngo et al identified 19 unique proteins,\textsuperscript{50} while Khalil et al reported numerous candidate biomarkers elevated in GBM.\textsuperscript{34}

The first 2DE map of normal patient CSF was reported in 1985 and comprised a mere 70 proteins.\textsuperscript{69,70} The number of peptide and protein products in CSF profiles rapidly grew to over 700 as technologies improved.\textsuperscript{21} Using 2DE combined with MALDI-TOF-MS, nearly 1,000 protein spots from a single CSF sample can now be analyzed.\textsuperscript{71}

MS-based technique. Several MS-based approaches have been used for multidimensional protein/peptide separation and analysis. In contrast to gel-based
quantifications followed by MS identification, these approaches rely on the identification and quantification of digested peptides obtained from unfraccionated or partially fractionated protein mixtures. A commonly used shotgun 2D liquid chromatography (LC)-MS/MS technology separates digested peptides (following a protease treatment of proteins) on a strong cation exchange chromatography coupled to an automated reverse phase LC-MS/MS platform for the detection of peptides to achieve identification and quantification of thousands of proteins with femtomolar or even subfemtomolar sensitivity. Because the analysis is performed on tryptic peptides, as opposed to proteins, this technology is also known as a bottom-up approach.

Another common analytic procedure combines LC-MS/MS with prior separation of intact proteins on a sodium dodecyl sulfate gel. In this approach, separated proteins are cut into pieces after electrophoresis, in-gel digested with a protease, and subjected to LC-MS/MS analysis. Finally, a combined IEF- and MS-based approach utilizes the separation power of IEF of digested fragments, which are then extracted for LC-MS/MS analysis.

After acquisition of the MS data, protein identification is achieved mainly through a database search to match MS/MS spectra with peptide/protein sequences by computer software (ie, SEQUEST, Mascot, etc.), whereas relative protein amounts can be achieved by label-free or stable isotope labeling approaches. Label-free proteomics use spectral counting (SC) or extracted ion current (XIC) as a quantitative index. SC represents the total numbers of MS/MS spectra identifying a protein of interest, and it increases proportionally to protein abundance. XIC of the peptides is the peak intensity of corresponding ions in MS survey scans.

In contrast to label-free quantification, heavy labeled proteins/peptides are used as internal standards for accurate quantification by numerous isotope labeling approaches, such as ICAT, isobaric tags for relative and absolute quantification of peptides (iTRAQ), and stable isotope labeling of amino acids in cell culture (SILAC). The first 2 approaches label proteins being compared with the respective structurally identical but isotopically distinct (of different molecular weight) tags on either thiol- or amine-containing residues, respectively, thus allowing comparison of peptides identical in sequence but differing in mass. SILAC could be the approach of choice if one utilizes metabolic labeling of cultured cells and even animals in that the amino acids of the stable isotope are added to the cell culture medium and are metabolically incorporated into the proteins, thereby serving as internal standard controls for subsequent quantification. In all 3 methods, contribution of the internal standard allows accurate and reliable quantification of proteins of interest via comparison of relative abundances of labeled peptides.

Many of the MS-based approaches were used to analyze brain tumor samples. As an example, in a study examining CSF samples from astrocytoma patients, complementary 2DE and ICAT analyses were employed to identify differentially expressed protein biomarker panels for each grade of astrocytoma. Another group used a combination of iTRAQ and LC-MALDI technologies to quantify 1460 proteins in membrane-enriched fractions of serially transplanted glioma xenografts in rats, generating molecular signatures of nonangiogenic versus angiogenic GBM. The limitations of these techniques, however, are potential inefficient chemical labeling, sample loss due to additional purification steps, and chemical side reactions.

### Challenges

Proteomic studies have already generated a sizable number of candidate diagnostic and prognostic markers in human brain cancer. Before adapting discovered biomarkers to the clinic, it is necessary to address some of the significant challenges that still remain in the existing methodologies and technologies hindering an in-depth, nonbiased profiling of the human glioma proteome. These include reproducible detection, intertechnology reproducibility, lack of standards for proper study design and biospecimen collection/handling, low sample numbers that result in inadequate statistical power of protein aberrations, and overall nonagreement when cancer gene expression profiles are translated into functional differences at the proteomic level. A recent initiative of the United States National Cancer Institute aims at overcoming some of these limitations by developing a network of Proteome Characterization Centers that use data and selected biospecimens from cancer genomics programs to identify changes at the protein level that may reflect changes in cancer genomes. These centers are expected to systematically define the functional cancer proteome, to discover and verify well-credentialed protein biomarkers, and to drive the improvement/development of proteomic technologies.

### Biomarker Verification

Verification of biomarker panels is a natural and mandatory next step after the discovery phase in the biomarker development pipeline. The proteomic techniques used in biomarker discovery vary significantly, and often little overlap is found among identified biomarkers in different studies. In addition, low patient numbers and the absence of standardized sample processing protocols may lead to the discovery of false-positive proteins. Furthermore, many hundreds of potential markers are identified, and a prioritization process for their verification needs to be implemented. Studies combining the information gained from genetic, transcriptomic, and proteomic analyses of the same set of samples need to be performed for optimal selection of biomarkers.

### Statistical Significance of Identified Biomarker Panels

Prior to embarking on confirmation, an essential step in moving from biomarker discovery to experimental
verification and then clinical validation is a rigorous statistical examination of the identified biomarker datasets to reduce often surfacing false positives and false negatives. While sufficient statistical power has to be itself ensured by a large number of biospecimens, large panels of proteins identified mean little if they are not analyzed properly. A natural and effective method to remove false positives is to introduce replicates or repeat the analysis. A complementary strategy is to calculate the false discovery rate (FDR) or the family-wise error rate (FWER), which represents the percent of false predictions or the probability of making false discoveries when analyzing multiple pairwise tests, respectively.83 Another method to evaluate and reduce experimental errors in proteomic analysis involves a 2-step process of pairwise comparison. In the first step, control and diseased groups are divided into 2 pools. When 2 individual samples from 2 pools of the same group are compared, the changes in protein abundance represent a summary of biological and technical variances in the group, which after a statistical transformation, can be translated into a normal distribution curve, representing an acceptable range of variance. While the false positive rate of accepted peptide matches ranges from 0% to 5% in the literature presently, a stricter FDR needs to be employed to control for the false positive problem. In the second step, the data on the identified markers are integrated with information from the literature gathered independently to weed out unlikely candidates and guide subsequent hypothesis-driven validation.

Finally, multiple publicly available online software tools aid with statistical analysis and validation of large datasets produced with proteomic biomarker profiling. These, in combination with the Proteomics Standard Initiative, launched in 2002 by the Human Proteome Organization (HUPO) to define community standards for data representation in proteomics for facilitated data comparison, exchange, and verification, will not only improve the quality of proteomics data but are likely to provide advanced integrated capabilities for dataset analysis in the future.84

Current Technologies

A number of classical biochemistry methodologies, as well as multiplexed quantitative assays and strategies for robustly and reproducibly verifying discovered biomarkers in a wide range of sample matrices, have been developed already and are continuously being improved.

Antibody-driven techniques. Besides being confirmed through one another, gel-based and MS-based analyses are further verified by exploiting antibodies raised against the putative biomarkers. Some of the current antibody-driven technologies for biomarker verification include western blot, dot blot, ELISA, protein chip, and peptide, protein, and tissue microarrays.

Western and Dot Blotting, ELISA, and Immunohistochemistry. Classical biochemistry techniques, including western and dot blots, ELISA, and IHC, are analytical methodologies used to detect target proteins in a given sample by utilizing specific antibodies. These methods have been routinely used for verification of putative protein biomarkers identified during the discovery phase of proteomics. For example, attractin was identified as a novel marker for malignant astrocytoma (WHO grades III and IV) in the CSF.24 To verify attractin’s importance in gliomagenesis, its expression was examined by western blotting and IHC in an independent set of 100 brain tumor samples. These verification studies confirmed that attractin is elevated in most malignant astrocytomas and further demonstrated that it is produced and secreted from tumor cells and not stroma.22 However, largely relying on the availability of antibodies specific to the target proteins, and using “one-at-a-time” approaches (unable to analyze large numbers of proteins simultaneously), these techniques, while useful, presently give way to other more robust and multiplex methodologies.

Peptide, Protein, and Tissue Arraying. Protein arrays, specifically antibody and aptamer arrays, represent one of the high-throughput techniques enabling simultaneous detection and monitoring of multiple proteins. In an antibody array, hundreds of antibodies directed at biomarkers of interest are immobilized on a solid surface and act as bait molecules that capture the solution phase analyte molecules from individual biospecimen protein mixes. One of the main advantages of the technology over other proteomic approaches is that the identities of the measured proteins are known at the time of experimental design or can be readily characterized, facilitating subsequent biological interpretation. Ideally, such arrays allow for efficient biomarker validation and may potentially permit putative biomarker panels to be translated into clinical applications in an affordable and convenient manner in the future.

Because many antibody-binding activities are directed toward peptides, a platform coated with hundreds of immobilized peptides offers an unprecedented opportunity for identifying or verifying interactions of proteins. A layered peptide array is one such platform that allows for simultaneous semiquantitative multiplex measurements of proteins, peptides, and/or antibodies in a variety of patient specimens.85

As an alternative proteomic technology, tissue microarrays (TMAs) play a key role in accelerating validation of identified protein biomarkers, with the potential for translating basic research findings into clinical applications. This technology allows in situ histopathologic analysis to take place simultaneously with in situ hybridization and/or IHC analysis of hundreds of tissue samples on a single glass slide array. TMA has become a widely utilized tool for all types of tissue-based research, especially allowing detection of tumor-associated proteins in high-throughput screens of tumor tissue.86,87 As an example, 378 paraffin-
embedded brain tumors, including tissue samples of human gliomas of WHO grades I–IV were incorporated into a TMA, demonstrating that 6 protein biomarkers, including nestin, neurofilament protein-MH, and neuronal class III beta tubulin clone TUJ1 were differentially expressed in high-grade astrocytomas (WHO grades III and IV) in comparison with low-grade astrocytomas (WHO grades I and II).

Multiple Reaction Monitoring/Selective Reaction Monitoring. Two MS-based techniques, selective reaction monitoring (SRM-MS) and multiple reaction monitoring (MRM-MS), are the main fully validated technologies used for biomarker verification presently. Experiments are accomplished by first digesting a selected protein with trypsin, followed by a selection of a few resulting peptides for MS/MS fragmentation. If the fragmentation of a single peptide is measured, the process is deemed SRM. If multiple peptides (either from a single protein or many proteins) are followed, the process is called MRM. Inclusion of isotope-labeled synthetic peptides as internal standards in the mixture results in a pair of isotopically different peptides in the final plot, the signal ratio between which provides an accurate measurement of the levels of selected peptides/proteins in the biological samples tested. Simplicity, sensitivity, reproducibility, better dynamic range, less bias toward abundant proteins, and lack of an antibody requirement for detection in these techniques make SRM/MRM ideal for sensitive and specific quantification of proteins of interest. Finally, various iterations of this technology (ie, immune capture, protein digestion strategies, and use of intact labeled proteins, as opposed to peptides, as standards) can be utilized to improve detection and quantification of peptides/proteins of interest, especially those that are post-translationally modified.

As an example, the MRM strategy was recently used to quantify 32 core postsynaptic density proteins in cortex and cerebellum using computationally selected peptides as internal standards, with a success rate at quantification of 97% (31/32). In sharp contrast, the same samples analyzed by the ICAT method using one-dimensional gel–based LC/LC-MS/MS allowed for quantification of only 60% (19/32) of proteins.

Emerging Technologies

MALDI-MS. MALDI imaging is an advanced MS-based technique that enables label-free molecular imaging of peptides and proteins directly from biological tissue sections. Unique with respect to other techniques, MALDI-MS requires neither the use of antibodies nor the use of tagged or labeled probes. In fact, multiple images of thousands of ions are obtained by simply irradiating the sample and monitoring the m/z, the mass-to-charge ratio, of generated ions from the tissue. While the resolution of the technique is 30–50 μm, detectable mass range is limited from an m/z of 400 to 30,000. However, the technique is being actively improved in terms of its sensitivity of detection and reproducibility.

Metabolic Profiling. A metabolic profile refers to a measurement of endogenous metabolite quantities in a test biospecimen at any given time. Complementary to genomics and proteomics, metabolomics is currently an active area of research. The high-throughput assessment of metabolites is achieved either via nuclear MR, such as solid-state high-resolution magic angle spinning nuclear MR (HRMAS-NMR; also known as MR spectroscopy, or via MS-based approaches. Generated findings can then be combined with other techniques to garner valuable information in protein target selection and verification.

Bottlenecks

While a few well-established verification methodologies are already in place and some promising approaches are well on their way, the challenge of accurately characterizing and quantifying putative protein biomarkers demands technological advances beyond the classical methodologies of protein biochemistry. In addition, high costs, lengthy production times and most importantly lack of high specificity antibodies complicate and slow the accurate detection, quantification and verification of putative protein biomarkers. Since 2003, the Human Protein Atlas (www.proteinatlas.org), part of HUPO, has actively tried to compensate for the antibody-related drawbacks in proteomics research by producing specific and validated antibodies for human target proteins. Likewise, the National Cancer Institute’s Clinical Proteomic Technology Assessment for Cancer program has both produced a large array of candidate peptides and demonstrated the feasibility of using MS-based assays as clinical cancer markers.

With this rapidly growing antibody and peptide collection, the challenges for biomarker validation should be reduced in the near future.

Biomarker Translation into Clinics

The fundamental goal of discovery proteomics is to identify novel biomarkers that can be clinically validated for diagnostic, prognostic, predictive, and therapeutic response assessment. In addition, it has the potential to reveal important players in the disease process for further research and possible therapeutic targeting. With the development of validated clinical protein panels readily measurable from proximal fluids and/or tissue, one can envision measuring tumor responses not only by clinical outcomes, but through biological measures of entire sets of nodes along defined pathogenic signaling pathways. However, of the literally hundreds of brain tumor biomarkers identified through the use of proteomic techniques in the past years, none have yet attained broad application for use in clinical
prognosis, therapeutic target selection or molecular classification. The key reason for the lack of this translation is still the absence of reliable follow-up verification and clinical validation techniques post biomarker discovery. Specifically, there is still a pressing need for maturation of existing and novel validation technologies, including the development of multidisciplinary approaches for the systematic assessment of analytical variability, breadth and quality of acquired data. Finally, integration of proteomic studies with the complementary genetic datasets, as recently described by Verhaak et al., is necessary for establishing comprehensive profiles of distinct molecular subtypes of gliomas in the future. Provided that these hurdles are addressed, the initiation of the translational phase for discovered biomarkers may happen very soon.

Conclusions

The application of proteomics in neuro-oncology is still in its developing stage, resulting in few studies with clinical applications. However, there is little doubt that the proteomic approach has the potential to identify novel diagnostic, prognostic, and therapeutic biomarkers for human gliomas. In fact, in the near future, improved proteomic profiling is anticipated to bring about a merger of biology, engineering, and informatics, with a profound impact on glioma research and treatment.

Excellent sources for disease biomarker discovery and a number of well-established technologies for the robust discovery of protein biomarkers in neuro-oncology are already available. Optimization of experimental design and validation in independent cohorts, improved multiplex proteomic methodologies and bioinformatics tools, and integration with genetic and metabolomic profiling technologies promise to play critical roles in the post proteomics era of cancer diagnosis and treatment. Collectively, these are anticipated to eliminate false-positive targets and bridge the current gap between discovery, validation, and translation of biomarkers into clinically relevant glioma targets in the near future, as well as produce novel insights into the pathogenesis of gliomas.

Acknowledgments

We thank Dr. Jeffrey J. Olson for advice on the collection of biofluids from patients with brain tumors.

Conflict of interest statement. None declared.

Funding

This work was supported by a Brain Tumor Funders Collaborative grant to EGVM and by an American Cancer Society grant (RSG-09-181) to J.P.

References


