Jintang He[1]
Jianhui Zhu[1]
Yashu Liu[1]
Jing Wu[1]
Song Nie[1]
Jason A. Heth[2]
Karin M. Muraszko[2]
Xing Fan[2,3]
David M. Lubman[1]

[1]Department of Surgery, University of Michigan Medical Center, Ann Arbor, MI, USA
[2]Department of Neurosurgery, University of Michigan Medical Center, Ann Arbor, MI, USA
[3]Department of Cell and Developmental Biology, University of Michigan Medical Center, Ann Arbor, MI, USA

Received October 17, 2012
Revised December 9, 2012
Accepted January 14, 2013

Research Article

Immunohistochemical staining, laser capture microdissection, and filter-aided sample preparation-assisted proteomic analysis of target cell populations within tissue samples

An important problem involves isolating subpopulations of cells defined by protein markers in clinical tissue samples for proteomic studies. We describe a method termed Immunohistochemical staining, laser capture microdissection (LCM) and filter-aided sample preparation (FASP)-Assisted Proteomic analysis of Target cell populations within tissue samples (ILFAPT). The principle of ILFAPT is that a target cell population expressing a protein of interest can be lit up by immunohistochemical staining and isolated from tissue sections using LCM for FASP and proteomic analysis. Using this method, we isolated a small population of $CD90^+$ stem-like cells from glioblastoma multiforme tissue sections and identified 674 high-confidence (false discovery rate < 0.01) proteins from 32 nL of $CD90^+$ cells by LC-MS/MS using an Orbitrap Elite mass spectrometer. We further quantified the relative abundance of proteins identified from equal volumes of LCM-captured $CD90^+$ and $CD90^-$ cells, where 109 differentially expressed proteins were identified. The major group of these differentially expressed proteins was relevant to cell adhesion and cellular movement. This ILFAPT method has demonstrated the ability to provide in-depth proteome analysis of a very small specific cell population within tissues. It can be broadly applied to the study of target cell populations within clinical specimens.

Keywords:
Immunohistochemical staining / Laser capture microdissection / Proteomics / Target cell population / Tissue sample
DOI 10.1002/elps.201200566

Additional supporting information may be found in the online version of this article at the publisher's web-site

1 Introduction

Organs and tissues are composed of heterogeneous cell populations with different functions. An in-depth proteomic study of a specific cell population within tissues will improve our understanding of its functions at the protein level, and could even reveal the mechanisms of diseases if the target cell population is from clinical specimens. Some cell populations, for example, cancer stem cell populations, are of great interest where there is increasing evidence that cancer stem cells are responsible for tumor generation and ongoing growth [1, 2], including brain tumors [3–5]. Our recent work has demonstrated that $CD90^+$ cells within glioblastoma multiforme (GBM) exhibit distinct stem cell properties [6]. It is thus essential to investigate the proteome of this $CD90^+$ cell population.

A crucial step for proteome research of a target cell population within tissue samples is to isolate this population, which is often based on specific binding to a marker of the target population. Several techniques have been developed to sort different cell populations from tissues, where

Correspondence: Professor David M. Lubman, Department of Surgery, The University of Michigan Medical Center, 1150 West Medical Center Drive, Building MSRB1, Rm A510B, Ann Arbor, MI 48109-0656, USA
E-mail: dmlubman@umich.edu
Fax: +1-734-615-2088

Abbreviations: **FACS**, fluorescent-activated cell sorting; **FASP**, filter-aided sample preparation; **FDR**, false discovery rate; **GBM**, glioblastoma multiforme; **ILFAPT**, Immunohistochemical staining, LCM and FASP-Assisted Proteomic analysis of Target cell populations within tissue samples; **LCM**, laser capture microdissection; **MACS**, magnetic-activated cell sorting; **PFA**, paraformaldehyde; **TMA**, tissue microarray

Colour Online: See the article online to view Figs. 1–6 in colour.

fluorescent-activated cell sorting (FACS) [7] and magnetic-activated cell sorting (MACS) [8] have been most frequently used. However, both FACS and MACS are only suitable to sort live cells in fresh tissues using cell surface markers, while intracellular markers are not compatible with these techniques. These limitations impede the application of FACS and MACS to the study of specific cell populations within clinical tissue samples.

Laser capture microdissection (LCM) is a powerful technology that can be used to capture cells from tissue sections [9]. Most clinical samples are routinely fixed in formalin where they are readily sectioned and can be used for LCM analysis. Immunohistochemistry-guided LCM has been developed for mRNA analysis of specific cell populations in clinical specimens [10]. Recently, LCM has also been utilized for proteomic analysis of clinical samples. For example, Patel et al. [11] performed a proteomic analysis of squamous cell carcinoma of the head and neck tissues and identified around 700 proteins from 20 000 cells procured from tissue sections. In another study on human breast cancer epithelial cells [12], around 60 000 cells were captured from each of 18 tissue samples and 1623 proteins were identified from these samples. Quantitative analysis of these proteins resulted in identification of 298 differentially expressed proteins between normal and malignant breast epithelial samples. Recently, Wisniewski et al. [13] applied a filter-aided sample preparation (FASP) method to proteomic analysis of LCM-captured human tissues and greatly improved the depth of proteome analysis. By adding an additional peptide fractionation step prior to LC-MS/MS analysis, they accomplished a depth of 3600–4400 proteins from each sample containing 175 nL of collected tissue (5–7 μg of peptides). These studies demonstrated the ability of LCM to assist in proteomic analysis of tissue samples, which typically contain multiple cell populations; however, they did not demonstrate analysis of a specific cell population within tissue samples.

Herein, we describe a method for in-depth proteomic analysis of a very small number of target cell populations within tissue samples. We have termed the method as Immunohistochemical staining, LCM and FASP-Assisted Proteomic analysis of Target cell populations within tissue samples (ILFAPT). To guide the precise capture of a target cell population by LCM, we applied immunohistochemical staining to specifically light up the target population in tissue sections. After the captured cells were processed by the FASP method, the resulting peptides were analyzed by LC-MS/MS. We have demonstrated the application of this ILFAPT method for proteomic analysis of a small population of $CD90^+$ cancer stem cells within GBM tissues, where 674 high-confidence proteins were identified from 32 nL of LCM captured cells. Further comparison of $CD90^+$ and $CD90^-$ cell populations using a spectral counting method resulted in identification of 109 differentially expressed proteins.

2 Materials and methods

2.1 Tissue samples

A fresh primary GBM biopsy was obtained from the University of Michigan University Hospitals with approval from the Internal Review Board. Tissues were fixed in 4% paraformaldehyde (PFA) overnight at 4°C, cryoprotected with 30% sucrose saturation, and cryoembedded in Tissue-Tek OCT. Tissues were sectioned in 8 μm thickness and stored at −80°C for subsequent use.

2.2 Immunohistochemical staining for $CD90^+$ cells

The PFA-fixed GBM tissue sections were thawed for 20–30 s and rinsed with PBS. Endogenous peroxidase activity was blocked using 6% H_2O_2 in 80% methanol for 10 min. The tissue sections were then blocked with 2% normal goat serum and incubated with a rabbit anti-human CD90 monoclonal antibody (1:100 dilution, Abcam, Cambridge, MA, USA) for 2 h, washed with PBS three times (10 min each), and incubated with a goat anti-rabbit biotinylated antibody (Vector Laboratories, Burlingame, CA, USA) for 1 h. The slides were washed again with PBS three times (10 min each), and immunodetection was performed using the VECTASTAIN Elite ABC system and DAB substrates (Vector Laboratories), where the expression of antigen CD90 was visualized as a brown color. Nuclei were lightly counterstained with hematoxylin (blue). The slides were dehydrated in 70%, 95%, and 100% ethanol for 30 s each, rinsed in xylene, air-dried, and then placed in a desiccator overnight to ensure complete drying.

2.3 LCM

$CD90^+$ cells were captured from completely dried tissue sections using the Veritas Laser Capture Microdissection System (Arcturus, Molecular Devices, CA, USA). $CD90^+$ areas were selected under a 20× objective and microdissected using CapSure LCM macro caps. LCM was performed on six tissue sections and ~4 mm^2 of $CD90^+$ cells were collected. Equal areas of $CD90^-$ cells were also collected, which served as a control against $CD90^+$ cells.

2.4 Protein extraction and digestion by FASP

Proteins in microdissected cells were extracted and digested using the FASP protocol [13, 14]. A FFPE-FASP™ Protein Digestion Kit (Protein Discovery, San Diego, CA, USA) was used in this work. Briefly, microdissected cells were lysed with 40 μL of protein extraction buffer (0.1 M Tris-HCl, pH 8.0, 0.1 M DTT, 4% SDS) with agitation at 99°C for 1 h. The samples were then centrifuged at 15 000 × *g* for 10 min and the clarified lysates were mixed with 200 μL of 8 M urea in 0.1 M

Tris-HCl, pH 8.5 in spin filters. The samples were centrifuged at 14 000 × *g* for 30 min. The step was performed twice. Then 100 μL of 0.05 M iodoacetamide in 8 M urea was added to the filters and the samples were incubated in the dark for 20 min, followed by two washes with 100 μL of 8 M urea and two washes with 100 μL of 50 mM ammonium bicarbonate. Finally, trypsin was added in 75 μL of 50 mM ammonium bicarbonate to each filter. Protein digestion was performed at 37°C overnight. The released peptides were collected by centrifugation at 14 000 × *g* for 15 min and evaporated to dryness using a SpeedVac concentrator (Thermo Savant, Milford, MA, USA).

2.5 LC-MS/MS

The peptide mixtures were analyzed by LC-MS/MS using an Orbitrap Elite mass spectrometer (Thermo Fisher Scientific). Chromatographic separation of peptides was performed on a Proxeon EASY nLC II System (Thermo) equipped with an in-house packed 25 cm column (inner diameter 75 μm, Magic C18 AQ 100 Angstrom, 5 μm). Peptides were separated with a gradient of ACN/water containing 0.3% formic acid at a flow rate of 400 nL/min. A 130 min linear gradient from 2 to 35% ACN was used. MS spectra were acquired in data-dependent mode. The resolution of full-scan MS was 120 000 (m/z 400.0–1800.0). The 20 most intense ions in the full mass scan were selected for MS/MS analysis by CID in the linear ion trap. The normalized collision energy was set at 35% for MS/MS. MS/MS data were obtained for up to three ions of charge 2+, 3+, or 4+ detected in the survey scan. Dynamic exclusion was defined by a list size of 500 features and exclusion duration of 60 s, and expiration was disabled to decrease the resequencing of isotope clusters.

2.6 Data analysis

All MS/MS spectra were searched against the UniProtKB/Swiss-Prot (reviewed, human proteins) protein knowledgebase release 2012_05. The search was performed using the SEQUEST algorithm incorporated in the Proteome Discoverer software (version 1.1.0.263). The search parameters were as follows: (i) Fixed modification: carbamidomethyl of cysteine; (ii) variable modification: oxidation of methionine; (iii) maximum missed cleavage sites: 1; (iv) precursor mass tolerance: 10 ppm; (v) fragment mass tolerance: 0.8 Da. False discovery rate (FDR) was set to be < 0.01. The Xcorr values for all the charge states (+1, +2, and +3) were automatically adjusted to achieve the predetermined FDR value. If multiple proteins share the same peptide sequences, they will be reported as a protein group.

We measured the relative abundance between proteins identified from $CD90^+$ and $CD90^-$ cells using the spectral counting method [15]. The missing values were assigned a spectral count of 1 to avoid dividing by zero. The spectral count of each individual protein was normalized by the sum of the spectral counts per LC-MS/MS run [16]. Spectral count fold-change was calculated as the ratio of the average spectral count for the target protein. A fold-change ratio of >3.0 was considered as a significant difference between the two samples.

2.7 Bioinformatic analysis

Gene ontology (GO) annotation for cellular component and biological process of the identified proteins were performed using the Database for Annotation, Visualization and Integrated Discovery (DAVID) bioinformatics resources [17, 18]. Protein–protein interaction of the differentially expressed proteins was analyzed using the STRING database [19]. The molecular function and biological networks of the differentially expressed proteins were analyzed using the Ingenuity Pathway Analysis software (Ingenuity Systems, Mountain View, CA, USA).

2.8 Double immunofluorescence analysis of FN1 and CD90

A tissue microarray (TMA) of formalin-fixed paraffin-embedded brain tumor and normal tissues was purchased from US Biomax (Catalog No. T174), which contains one each of astrocytoma (grade 2), anaplastic astrocytoma (AA, grade 3), glioblastoma (GBM, grade 4) and oligodendroglioma, and two of normal tissues. The tissue samples originated from different donors and each sample had four replicates.

The TMA was dewaxed in xylene for 10 min twice and rehydrated through a series of alcohol solutions (100% ethanol twice, 90% ethanol, 70% ethanol, 5 min each) to water. Antigen retrieval was achieved by boiling the array in citrate buffer at pH 6.0 (Invitrogen) for 15 min. The TMA was incubated with 2% normal goat serum in PBS for 1 h at room temperature to block nonspecific binding. To achieve immunofluorescence staining, mouse anti-FN1 (LifeSpan Biosciences, Seattle, WA, USA) and rabbit anti-CD90 (Abcam) were incubated with the TMA overnight at 4°C at a dilution of 1:100. Then DyLight 488 anti-rabbit IgG (H+L) and DyLight 549 anti-mouse IgG (H+L) (Vector laboratories) were diluted (1:200) and incubated with the TMA for 1 h at room temperature. The nuclei visualization was explored by DAPI counterstaining. Between each step, there were three washes with PBST for 5 min each. Finally, the TMA was dehydrated in alcohol (70% ethanol, 90% ethanol, 100% ethanol twice, 5 min each), and coverslipped using a CC/Mount™ permanent mounting medium (Sigma).

3 Results

3.1 ILFAPT

Our recent work has revealed that CD90 is a marker for GBM-derived stem-like neurospheres and the $CD90^+$ cell population has distinct stem cell properties including self-renewal

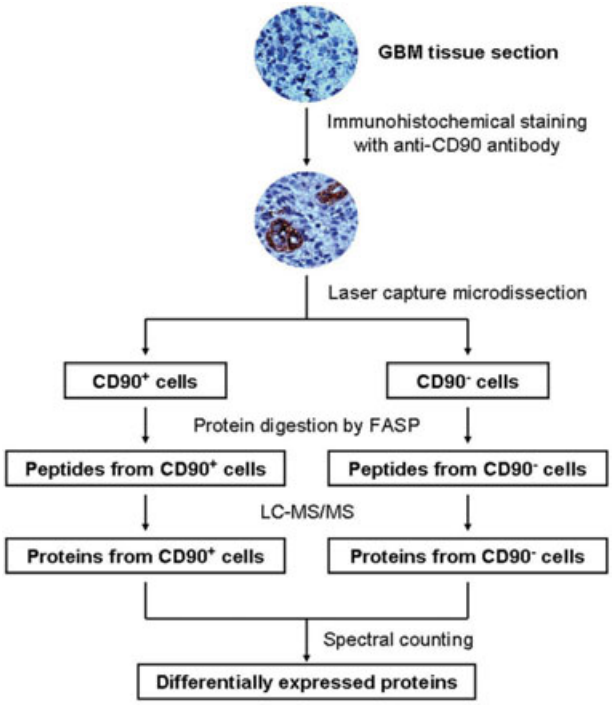

Figure 1. Workflow of the ILFAPT strategy. In this work, the proteome of $CD90^+$ and $CD90^-$ cells within GBM tissues was compared using the ILFAPT strategy. GBM tissue sections were immunohistochemically stained using an anti-human CD90 antibody, followed by isolation of $CD90^+$ and $CD90^-$ cell populations via laser capture microdissection. After tryptic digestion of proteins in each cell population using an FASP approach, the proteins were analyzed by LC-MS/MS and spectral counting, which allowed identification of differentially expressed proteins between $CD90^+$ and $CD90^-$ cell populations.

and multilineage differentiation [6, 20]. It is thus essential to investigate the protein expression of $CD90^+$ cells within GBM tissue samples and to investigate the differences at the proteome level between $CD90^+$ and $CD90^-$ cells, which do not exhibit stem-like properties. To achieve this goal, we have used an ILFAPT method. The principle of ILFAPT is that a target cell population expressing a protein of interest can be lit up by immunohistochemical staining and isolated from tissue sections using LCM for further FASP-based proteomic analysis.

In this work, the $CD90^+$ population within GBM tissues is our target cell population. The workflow of ILFAPT is shown in Fig. 1. Briefly, GBM tissue sections were stained using an anti-human CD90 antibody, and $CD90^+$ cells were isolated by LCM. $CD90^-$ cells were also collected which serve as a control against $CD90^+$ cells. Proteins from each cell population were extracted and digested using a FASP method and further analyzed by LC-MS/MS and spectral counting. The ILFAPT strategy is not only suitable to study the proteome of a target cell population within tissues, such as $CD90^+$ cells in GBM tissues in this work, but also allows us to identify differentially expressed proteins between different cell populations within tissues (e.g., $CD90^+$ versus $CD90^-$ cells).

3.2 Isolation of a $CD90^+$ cell population from GBM tissue sections

PFA-fixed GBM tissue sections were immunohistochemically stained with an anti-CD90 antibody and counterstained with hematoxylin. As shown in Fig. 2A, $CD90^+$ cells form vascular-like structures (*brown*), which are surrounded by $CD90^-$ cells (*blue*). $CD90^+$ cells were captured from GBM tissues using the LCM technology. After LCM, $CD90^+$ cells were efficiently separated from $CD90^-$ cells, where $CD90^-$ cells remained on the tissue slide (Fig. 2B) while most of $CD90^+$ cells were captured to the LCM cap (Fig. 2C).

$CD90^+$ cells are a small population within GBM tissues. In the GBM specimen used in this work, the percentage of $CD90^+$ cells was only ~2%, where ~0.7 mm^2 of $CD90^+$ cells were obtained from each ~36 mm^2 tissue section. To collect an adequate number of cells for the subsequent proteomic analysis, we combined $CD90^+$ cells captured from six tissue sections (~4 mm^2). Equal areas of $CD90^-$ cells were also captured from GBM tissues, which serve as a control against $CD90^+$ cells for quantitative proteomic analysis.

3.3 Identification of proteins from $CD90^+$ and $CD90^-$ cell populations

A FASP method was used to extract and digest proteins from 32 nL (4 mm^2 × 8 μm) of each of $CD90^+$ and $CD90^-$ cell populations. We estimate that 32 nL of LCM-captured GBM

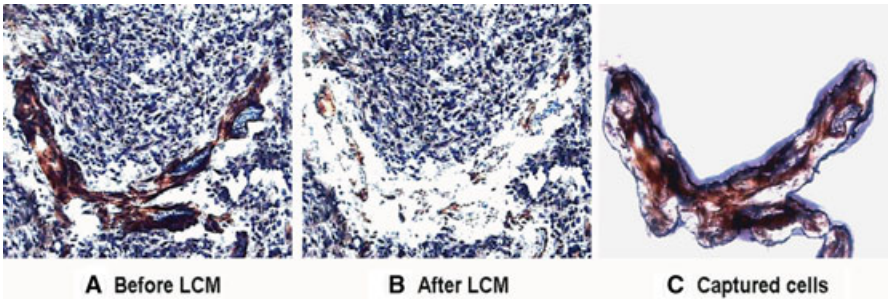

Figure 2. Isolation of $CD90^+$ cells from GBM tissues by LCM. A frozen GBM tissue section was immunohistochemically stained using an anti-CD90 antibody. (A) Before LCM, $CD90^+$ cells were clearly observed within the tissue section. (B) After LCM, most of $CD90^+$ cells were removed while $CD90^-$ cells remained on the tissue slide. (C) $CD90^+$ cells were captured to the LCM cap.

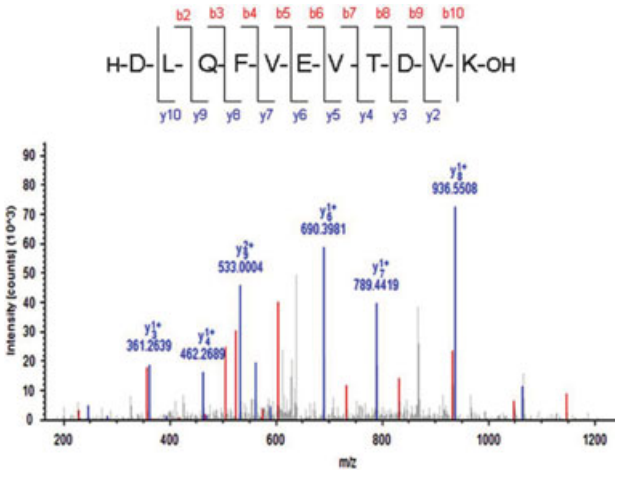

Figure 3. A representative MS/MS spectrum of a peptide. The sequence of the peptide was identified as DLQFVEVTDVK, which was from fibronectin.

tissue is equivalent to ~5000 cells, although the number may not be accurate because different cells vary in size. According to the relationship between the volume of LCM-captured tissue and the peptide yield measured in a previous study [13], 32 nL of tissue yields around 1.1 μg of peptides. The resulting peptides were analyzed by three replicates of LC-MS/MS using an Orbitrap Elite mass spectrometer. When the FDR was set to be <0.01, the LC-MS/MS analyses resulted in identification of 674 proteins from $CD90^+$ cells and 1034 proteins from $CD90^-$ cells. The detailed information of these proteins is listed in Supporting Information Table S1. As an example, Fig. 3 shows a representative MS/MS spectrum of a peptide from fibronectin.

GO cellular component annotation (Fig. 4A) showed that these proteins were primarily located to the cytosol, nucleus, and cytoskeleton. A major difference in subcellular location between the two cell populations was that the percentage of extracellular proteins was much lower for $CD90^+$ cells compared with $CD90^-$ cells. The GO biological process annotation (Fig. 4B) revealed that 0.7% of identified proteins were relevant to cell adhesion for $CD90^+$ cells, compared with 5.7% for $CD90^-$ cells. This was paralleled by the corresponding decrease of extracellular proteins in $CD90^+$ cells where a large proportion of cell adhesion-related proteins are located in the extracellular matrix.

We evaluated the reproducibility of three replicates of LC-MS/MS analyses using R^2 values from linear regression for spectral counts of identified proteins. The R^2 values for the measurements of both $CD90^+$ (Fig. 5A and 5B) and $CD90^-$ (Fig. 5C and 5D) cells were ~0.99, which indicates that the LC-MS/MS analyses were highly reproducible in this work. We further placed all the identified proteins from three replicates of LC-MS/MS into an unsupervised hierarchical clustering (Fig. 5E). Consequently, the clustering algorithm grouped the analyses of each cell population together. This suggests that $CD90^+$ cells can be separated from $CD90^-$ cells based on the spectral counts of identified proteins.

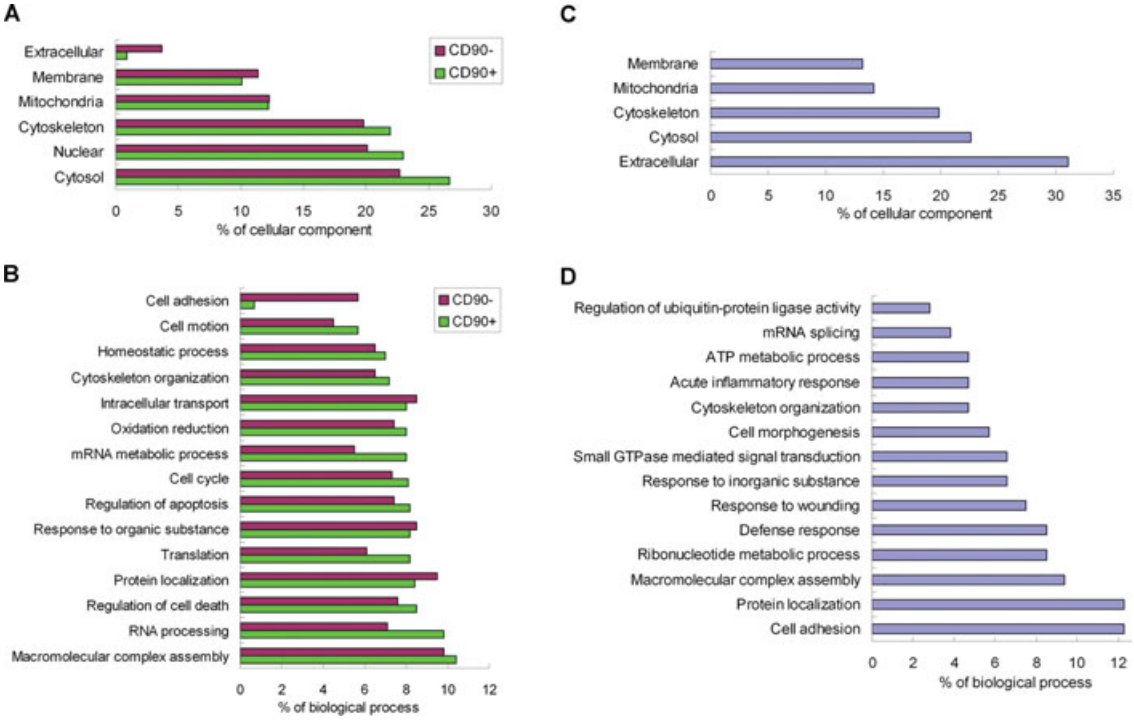

Figure 4. Gene ontology annotation of identified proteins. Comparison of subcellular location (A) and biological process (B) annotations of proteins identified from $CD90^+$ and $CD90^-$ cells. Subcellular location (C) and biological process (D) annotations of differentially expressed proteins. Gene ontology cellular component and biological process terms were derived using the DAVID bioinformatics database.

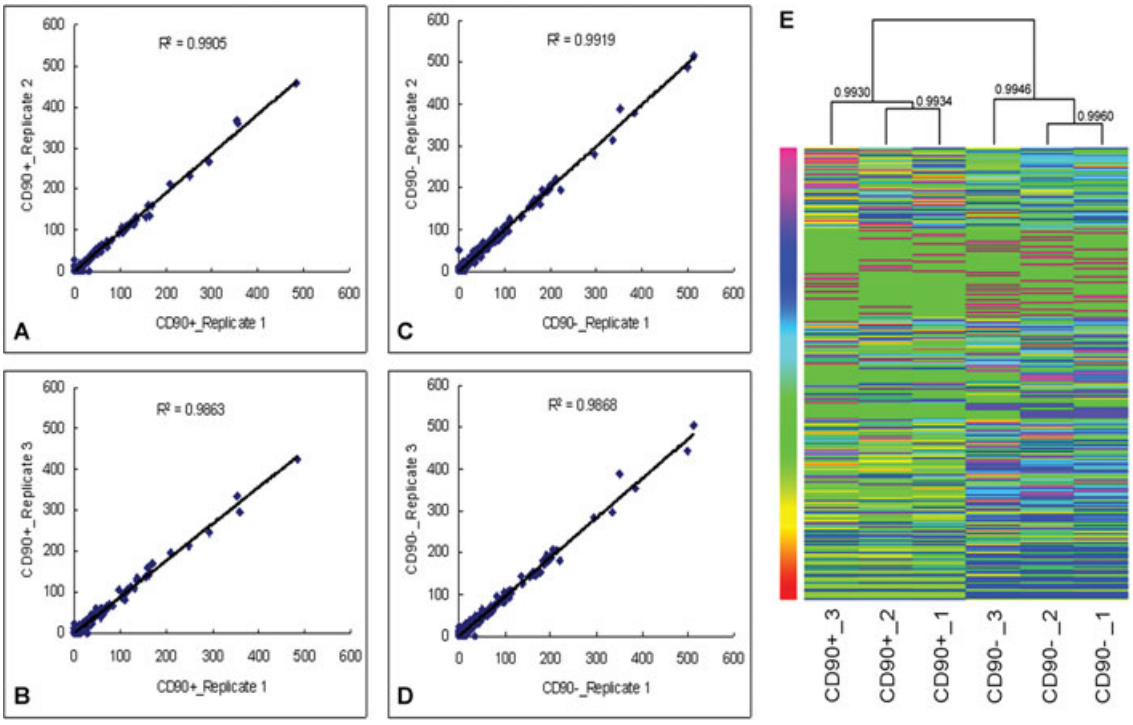

Figure 5. Reproducibility and hierarchical clustering of triplicate LC-MS/MS experiments. Both reproducibility and hierarchical clustering assays were based on the spectral count information of identified proteins from each LC-MS/MS analysis. (A and B) Correlation between LC-MS/MS analyses of $CD90^+$ cells. (C and D) Correlation between LC-MS/MS analyses of $CD90^-$ cells. (E) Hierarchical clustering of triplicate LC-MS/MS analyses of $CD90^+$ and $CD90^-$ cells, where each column indicates an individual LC-MS/MS analysis of each cell population, and each row indicates an identified protein. The number of spectral count increased from red to purple. The Pearson correlation coefficients were indicated at each branching point.

3.4 Differentially expressed proteins between $CD90^+$ and $CD90^-$ cell populations

We applied a spectral counting method to quantify the relative protein abundance between $CD90^+$ and $CD90^-$ cells isolated from GBM clinical samples. To minimize the variation among different replicates of LC-MS/MS analyses, we normalized the spectral count of each individual protein by the sum of the spectral counts per LC-MS/MS analysis. When the cut-off criteria were set with a fold-change ratio of >3.0 and a p value of <0.05, a total of 109 differentially expressed proteins were identified (Supporting Information Table S2). GO annotation revealed that the major subcellular locations of these proteins were extracellular (31.1%), cytosol (22.6%), and cytoskeleton (19.8%) (Fig. 4C). These differentially expressed proteins were involved in diverse biological processes, such as cell adhesion (12.3%), protein localization (12.3%), macromolecular complex assembly (9.4%), ribonucleotide metabolic process (8.5%), defense response (8.5%), and so on.

We analyzed protein–protein interactions among the differentially expressed proteins using the STRING database, which integrates interaction data from genomic context, high-throughput experiments, coexpression, and previous knowledge. When the confidence score was set to be higher than 0.7, a high-confidence protein–protein interaction network was generated (Fig. 6A). This network contains 58 high-confidence interactions for 39 differentially expressed proteins. The detailed information of these proteins is listed in Table 1. Note that among the 39 proteins, 19 are extracellular proteins and 8 are cytoskeletal proteins. Interestingly, the extracellular protein fibronectin (FN1) is imported as a central node to generate this network by STRING, suggesting that FN1 may play an important role in this network.

We further analyzed the function and potential pathways of the differentially expressed proteins using the Ingenuity Pathway Analysis software and identified a high-scoring (score 35) biological network (Fig. 6B). Twenty-one proteins were involved in this network, which is relevant to cell–cell signaling, cell adhesion, and cellular movement. All 21 proteins were downregulated in $CD90^+$ cells. Consistent with the STRING analysis, most of these proteins were extracellular and cytoskeletal proteins (Table 2).

3.5 Double immunofluorescence analysis of tissue microarray with FN1 and CD90

To further validate the expression of fibronectin (FN1) in $CD90^+$ cells of GBM tissues, double immunofluorescence staining of FN1 and CD90 was explored on a brain TMA. The staining was consistent with our previous publication

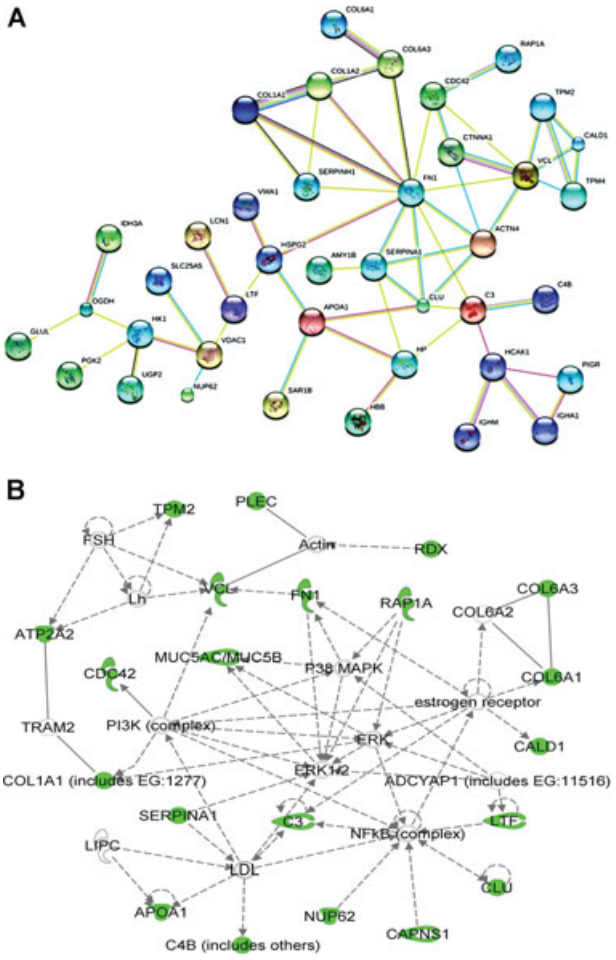

Figure 6. Protein interaction network generated from the differentially expressed proteins. (A) High-confidence protein–protein interaction network derived from the STRING database. Each node represents a protein, and each edge represents an interaction. (B) Biological network generated by Ingenuity Pathway Analysis. The functions of the network include cell–cell signaling, cell adhesion, and cellular movement. Solid arrows represent known direct interactions, and dotted arrows represent indirect interactions. Green represents the differentially expressed proteins identified in this work; White indicates proteins that were not identified in this study but related to the network.

[6], where CD90 was not detected in normal brain and astrocytoma grade 2 samples (data not shown), whereas strong staining of CD90 was observed in anaplastic astrocytoma grade 3 and GBM grade 4 samples, as shown in Supporting Information Fig. S1. The expression of fibronectin (FN1) was also highly elevated in anaplastic astrocytoma grade 3 and GBM samples compared to normal and astrocytoma grade 2 samples. Interestingly, as shown in Supporting Information Fig. S1, FN1 was coexpressed with CD90 in grade 3 tumor; however, the localization of FN1 and CD90 was separated in GBM tissue, indicating that FN1 expression in CD90+ cells was dramatically decreased. The significant downregulation of FN1 in $CD90^+$ GBM cells was consistent with the result observed in our MS analysis.

4 Discussion

The challenge for investigating the proteome of a target cell population expressing a protein of interest within clinical tissue samples is to specifically isolate the target cells from the tissues. FACS and MACS have been extensively used to sort different cell populations (especially cancer stem cells) from fresh tissues. However, the application of FACS and MACS is limited by the sample source because they are only suitable for sorting live cells from fresh tissues. Neither of these methods can be used to sort cells from the vast majority of clinical tissue samples, which are fixed with formalin or similar reagents. In this study, we developed an ILFAPT method to isolate target cell populations from fixed tissue sections for proteomic analysis. The method is based on the hypothesis that a target cell population within tissues can be specifically lit up by immunohistochemical staining with an antibody against a marker for this population and captured from tissue sections via the LCM technology. Compared with FACS and MACS, another advantage of ILFAPT is that antibodies against both cell surface and intracellular markers of target cell populations are compatible with this method.

We have applied the ILFAPT method to proteomic analysis of a $CD90^+$ cell population within GBM tissues, which has been shown to be a stem-like cell population in our recent work [6]. We found that the method allowed efficient capture of $CD90^+$ cells and effective separation of $CD90^+$ and $CD90^-$ cell populations (Fig. 2). Analysis of ~1.1 μg of peptides from 32 nL of PFA-fixed $CD90^+$ cells over three LC-MS/MS runs resulted in identification of 674 high-confidence (FDR < 0.01) proteins. Considering the extremely small amount of sample for each MS analysis, the number is reasonable compared with other studies using similar numbers of cells [21, 22].

We also compared the protein expression between $CD90^+$ and $CD90^-$ cells, and more proteins were identified from $CD90^-$ cells. A possible reason is that $CD90^-$ cells contain multiple cell populations, which makes the protein components of $CD90^-$ cells more complex than those of $CD90^+$ cells. Also, the shapes and sizes of different cell populations within tissues are highly heterogeneous. The label-free quantitation resulted in identification of 109 differentially expressed proteins, where 95 were downregulated in the $CD90^+$ cells. It is possible that the CD90 antibody and the corresponding secondary antibody remaining within the $CD90^+$ cells after LCM could lead to partial ionization suppression of peptide ions in the $CD90^+$ cell group. However, considering the low abundance of CD90 within these cells, the amount of antibodies should be very low and have limited effect to suppress ionization of peptide ions. The "housekeeping" protein beta-actin was not differentially expressed between $CD90^+$ and $CD90^-$ cells, where the $CD90^+/CD90^-$ fold-change ratio was 0.97 (data not shown).

The major groups of differentially expressed proteins are extracellular and cytoskeletal proteins associated with cell adhesion, cell signaling, and cellular movement (Figs. 4 and 6). The downregulation of this group of proteins in $CD90^+$

Table 1. Differentially expressed proteins involved in the STRING network

Accession	Gene	Protein name	Fold ($CD90^+/CD90^-$)	Location	*p*-value
P04745	AMY1B	Alpha-amylase 1	28.099	Extracellular	5.1E-06
P37198	NUP62	Nuclear pore glycoprotein p62 (Nup62)	0.322	Nuclear	0.00082
P68871	HBB	Hemoglobin subunit beta	0.322	Extracellular	7.1E-05
P01834	HCAK1	Ig kappa chain C region	0.318	Extracellular	0.00405
Q9Y6B6	SAR1B	GTP-binding protein SAR1b	0.304	Golgi	0.00307
P67936	TPM4	Tropomyosin alpha-4 chain	0.288	Cytoskeleton	0.01538
P98160	HSPG2	Basement membrane-specific heparan sulfate proteoglycan core protein	0.288	Extracellular	0.01075
Q05682	CALD1	Caldesmon	0.272	Cytoskeleton	8.3E-05
P15104	GLUL	Glutamine synthetase	0.271	Cytosol	0.00035
Q6PCB0	VWA1	von Willebrand factor A domain-containing protein 1	0.259	Extracellular	0.02378
P62834	RAP1A	Ras-related protein Rap-1A	0.258	Membrane	4.8E-05
P01871	IGHM	Ig mu chain C region	0.256	Extracellular	0.00089
P0C0L4	C4B	Complement C4-A	0.246	Extracellular	0.00057
P31025	LCN1	Lipocalin-1	0.233	Extracellular	0.04474
P60953	CDC42	Cell division control protein 42 homolog	0.232	Cytoskeleton	0.00072
Q16851	UGP2	UTP–glucose-1-phosphate uridylyltransferase	0.227	Cytosol	0.00326
P07205	PGK2	Phosphoglycerate kinase 2	0.206	Cytosol	0.00069
P19367	HK1	Hexokinase-1	0.206	Mitochondrion	0.00244
P01009	SERPINA1	Alpha-1-antitrypsin (Serpin A1)	0.191	Extracellular	0.01601
P00738	HP	Haptoglobin	0.172	Extracellular	2.8E-07
P02452	COL1A1	Collagen alpha-1	0.172	Extracellular	0.00017
P01024	C3	Complement C3	0.169	Extracellular	0.00207
Q02218	OGDH	2-oxoglutarate dehydrogenase	0.157	Mitochondrion	0.00051
P12111	COL6A3	Collagen alpha-3(VI) chain	0.153	Extracellular	0.02814
P10909	CLU	Clusterin	0.140	Extracellular	0.00292
P35221	CTNNA1	Catenin alpha-1	0.139	Cytoskeleton	0.00091
P18206	VCL	Vinculin	0.139	Cytoskeleton	0.00047
P01876	IGHA1	Ig alpha-1 chain C region	0.135	Cytoskeleton	0.00047
P21796	VDAC1	Voltage-dependent anion-selective channel protein 1	0.133	Membrane	0.00391
P07951	TPM2	Tropomyosin beta chain	0.129	Cytoskeleton	0.00065
P01833	PIGR	Polymeric immunoglobulin receptor	0.098	Membrane	0.0011
P08123	COL1A2	Collagen alpha-2(I) chain	0.092	Extracellular	0.00166
P02751	FN1	Fibronectin	0.071	Extracellular	0.00047
P02647	APOA1	Apolipoprotein A-I	0.071	Extracellular	7.8E-05
P50213	IDH3A	Isocitrate dehydrogenase [NAD] subunit alpha	0.067	Mitochondrion	8.6E-05
P12109	COL6A1	Collagen alpha-1(VI) chain	0.066	Extracellular	4.3E-05
P50454	SERPINH1	Serpin H1	0.058	ER	1.3E-05
O43707	ACTN4	Alpha-actinin-4	0.047	Cytoskeleton	6.7E-05
P02788	LTF	Lactotransferrin	0.030	Extracellular	0.0006

cells may play a critical role in maintaining the undifferentiated status of $CD90^+$ cancer stem cells. The results are similar to the findings of a previous study where cell migration–related genes are downregulated in $CD15^+$ brain tumor stem cells [23]. As shown in Fig. 6A, the extracellular protein fibronectin (FN1) seemed to be a key protein in this network. Immunohistochemical staining of a TMA revealed that FN1 was coexpressed with CD90 in grade 3 tumor, however, the localization of FN1 and CD90 was separated in GBM tissue, indicating that FN1 expression in $CD90^+$ GBM cells was dramatically decreased. Fibronectin is involved in cell adhesion and cell motility, and the fibronectin polymer superfibronectin inhibits tumor growth and angiogenesis [24, 25]. We have previously demonstrated that $CD90^+$ cells have stem cell properties and are colocalized with tumor vasculature within GBM tissues [6]. The downregulation of fibronectin in $CD90^+$ cells may play an important role in promoting the growth of these cells and the formation of GBM vasculature.

The major advantage of ILFAPT is that it allows specific and deep proteome analysis of target cell populations within clinical specimens. The method combines the specific staining of the antibody-based immunohistochemistry technology, the precise capture of target cells by LCM, and the efficient extraction and digestion of minute amounts of protein by FASP. In addition, ILFAPT overcomes the limited availability of fresh tissue samples and is able to accommodate the majority of clinical specimens. Therefore, ILFAPT could be an important tool for proteome analysis of target cell populations of clinical specimens.

Antibodies and blocking serum used in ILFAPT for efficient isolation of target cell populations from clinical specimens do not affect protein identification in the subsequent

Table 2. Differentially expressed proteins in the Ingenuity Pathway Analysis (IPA) network which are relevant to cell–cell signaling, cell adhesion, and cellular movement

Accession	Gene	Protein name	Fold ($CD90^+/CD90^-$)	Location	*p*-value
P37198	NUP62	Nuclear pore glycoprotein p62 (Nup62)	0.322	Nuclear	0.00082
Q05682	CALD1	Caldesmon	0.272	Cytoskeleton	8.3E-05
P62834	RAP1A	Ras-related protein Rap-1A	0.258	Membrane	4.8E-05
P0C0L4	C4B	Complement C4-A	0.246	Extracellular	0.00057
Q15149	PLEC	Plectin	0.240	Cytoskeleton	0.02222
P60953	CDC42	Cell division control protein 42 homolog	0.232	Cytoskeleton	0.00072
Q9HC84	MUC5B	Mucin-5B (MUC-5B) (Cervical mucin)	0.226	Extracellular	0.03887
P16615	ATP2A2	Sarcoplasmic/endoplasmic reticulum calcium ATPase 2	0.220	ER	0.00093
P01009	SERPINA1	Alpha-1-antitrypsin (Serpin A1)	0.191	Extracellular	0.01601
P04632	CAPNS1	Calpain small subunit 1	0.172	Membrane	0.00012
P02452	COL1A1	Collagen alpha-1	0.172	Extracellular	0.00017
P01024	C3	Complement C3	0.169	Extracellular	0.00207
P12111	COL6A3	Collagen alpha-3(VI) chain	0.153	Extracellular	0.02814
P10909	CLU	Clusterin	0.140	Extracellular	0.00292
P18206	VCL	Vinculin	0.139	Cytoskeleton	0.00047
P07951	TPM2	Tropomyosin beta chain	0.129	Cytoskeleton	0.00065
P35241	RDX	Radixin	0.076	Cytoskeleton	0.00058
P02751	FN1	Fibronectin	0.071	Extracellular	0.00047
P02647	APOA1	Apolipoprotein A-I	0.071	Extracellular	7.8E-05
P12109	COL6A1	Collagen alpha-1(VI) chain	0.066	Extracellular	4.3E-05
P02788	LTF	Lactotransferrin	0.030	Extracellular	0.0006

LC-MS/MS analysis. The antibodies or blocking serum are from a variety of animal species rather than human, for example, the anti-CD90 antibody and blocking serum applied in this study were from rabbit and goat, respectively. However, the tissue samples were from human and in the MS/MS data analysis, all the spectra were searched against a human protein database. Therefore, the rabbit or goat proteins should be excluded from the proteins identified. Moreover, three washes with PBS after incubation with primary and secondary antibodies remove most animal proteins from the slides. Furthermore, Mouledous et al. had demonstrated that the immunostaining process had minimal effect on protein identification by comparing 2DE results obtained from immunostained LCM brain tissue sections against the extracellular signal-regulated kinases (ERK 1 and 2, as these proteins are highly and ubiquitously expressed throughout the brain) to those obtained from unstained, dissected samples [26].

A potential issue for application of ILFAPT to quantitative analysis of very small amounts of sample is that the total protein amount of each sample is based on the volume of each cell population where the protein concentration is too low to be determined using traditional methods. However, the density of different cell populations within tissues may be different. This will cause bias in quantitation of proteins between different samples. In order to reduce the bias, we normalized the spectral count of each protein by the sum of the spectral counts of each MS analysis, and set the fold-change threshold to a relatively high value of 3. Development of ultrasensitive protein concentration determination methods will greatly improve the performance of ILFAPT on quantitative proteomic studies.

In summary, we have described an ILFAPT method, which provided efficient isolation of target cell populations from fixed clinical specimens and in-depth proteome analysis of a very small number of these cell populations. The ILFAPT method is particularly suitable for proteome research of cancer stem cell populations expressing unique markers. Comparative analysis of different cancer stem cell populations using ILFAPT could identify novel biomarkers and reveal the molecular and functional differences between these populations.

We thank Dr. Philip Gafken of the Fred Hutchinson Cancer Research Center for running the LC-MS/MS analyses on the Orbitrap Elite mass spectrometer. This work was funded under the National Institutes of Health Grant No. R01 49500 (D.M.L.) and the National Cancer Institutes Grant No. R01CA148621 (X.F.). We would like to acknowledge grant support to Dr. Fan from Accelerate Brain Cancer Cure Project Award, American Brain Tumor Association Translational Grant, and Voices Against Brain Cancer Research Grant.

The authors have declared no conflict of interest.

5 References

[1] Reya, T., Morrison, S. J., Clarke, M. F., Weissman, I. L., *Nature* 2001, *414*, 105–111.

[2] Visvader, J. E., Lindeman, G. J., *Nat. Rev. Cancer* 2008, *8*, 755–768.

[3] Fan, X., Matsui, W., Khaki, L., Stearns, D., Chun, J., Li, Y. M., Eberhart, C. G., *Cancer Res.* 2006, *66*, 7445–7452.

[4] Fan, X., Eberhart, C. G., *J. Clin. Oncol.* 2008, *26*, 2821–2827.

[5] Fan, X., Khaki, L., Zhu, T. S., Soules, M. E., Talsma, C. E., Gul, N., Koh, C., Zhang, J., Li, Y. M., Maciaczyk, J., Nikkhah, G., Dimeco, F., Piccirillo, S., Vescovi, A. L., Eberhart, C. G., *Stem Cells* 2010, *28*, 5–16.

[6] He, J., Liu, Y., Zhu, T., Zhu, J., Dimeco, F., Vescovi, A. L., Heth, J. A., Muraszko, K. M., Fan, X., Lubman, D. M., *Mol. Cell Proteomics* 2012, *11*, M111.010744.

[7] Herzenberg, L. A., Parks, D., Sahaf, B., Perez, O., Roederer, M., *Clin. Chem.* 2002, *48*, 1819–1827.

[8] Miltenyi, S., Muller, W., Weichel, W., Radbruch, A., *Cytometry* 1990, *11*, 231–238.

[9] Emmert-Buck, M. R., Bonner, R. F., Smith, P. D., Chuaqui, R. F., Zhuang, Z., Goldstein, S. R., Weiss, R. A., Liotta, L. A., *Science* 1996, *274*, 998–1001.

[10] Buckanovich, R. J., Sasaroli, D., O'Brien-Jenkins, A., Botbyl, J., Conejo-Garcia, J. R., Benencia, F., Liotta, L. A., Gimotty, P. A., Coukos, G., *Cancer Biol. Ther.* 2006, *5*, 635–642.

[11] Patel, V., Hood, B. L., Molinolo, A. A., Lee, N. H., Conrads, T. P., Braisted, J. C., Krizman, D. B., Veenstra, T. D., Gutkind, J. S., *Clin. Cancer Res.* 2008, *14*, 1002–1014.

[12] Cha, S., Imielinski, M. B., Rejtar, T., Richardson, E. A., Thakur, D., Sgroi, D. C., Karger, B. L., *Mol. Cell Proteomics* 2010, *9*, 2529–2544.

[13] Wisniewski, J. R., Ostasiewicz, P., Mann, M., *J. Proteome Res.* 2011, *10*, 3040–3049.

[14] Wisniewski, J. R., Zougman, A., Nagaraj, N., Mann, M., *Nat. Methods* 2009, *6*, 359–362.

[15] Liu, H., Sadygov, R. G., Yates, J. R., 3rd, *Anal. Chem.* 2004, *76*, 4193–4201.

[16] Dong, M. Q., Venable, J. D., Au, N., Xu, T., Park, S. K., Cociorva, D., Johnson, J. R., Dillin, A., Yates, J. R., 3rd, *Science* 2007, *317*, 660–663.

[17] Huang da, W., Sherman, B. T., Lempicki, R. A., *Nat. Protoc.* 2009, *4*, 44–57.

[18] Huang da, W., Sherman, B. T., Tan, Q., Kir, J., Liu, D., Bryant, D., Guo, Y., Stephens, R., Baseler, M. W., Lane, H. C., Lempicki, R. A., *Nucleic Acids Res.* 2007, *35*, W169–W175.

[19] Jensen, L. J., Kuhn, M., Stark, M., Chaffron, S., Creevey, C., Muller, J., Doerks, T., Julien, P., Roth, A., Simonovic, M., Bork, P., von Mering, C., *Nucleic Acids Res.* 2009, *37*, D412–D416.

[20] He, J., Liu, Y., Xie, X., Zhu, T., Soules, M., DiMeco, F., Vescovi, A. L., Fan, X., Lubman, D. M., *J. Proteome Res.* 2010, *9*, 2565–2572.

[21] Wang, N., Xu, M., Wang, P., Li, L., *Anal. Chem.* 2010, *82*, 2262–2271.

[22] Umar, A., Luider, T. M., Foekens, J. A., Pasa-Tolic, L., *Proteomics* 2007, *7*, 323–329.

[23] Read, T. A., Fogarty, M. P., Markant, S. L., McLendon, R. E., Wei, Z., Ellison, D. W., Febbo, P. G., Wechsler-Reya, R. J., *Cancer Cell* 2009, *15*, 135–147.

[24] Yi, M., Ruoslahti, E., *Proc. Natl. Acad. Sci. U S A* 2001, *98*, 620–624.

[25] Morla, A., Zhang, Z., Ruoslahti, E., *Nature* 1994, *367*, 193–196.

[26] Mouledous, L., Hunt, S., Harcourt, R., Harry, J. L., Williams, K. L., Gutstein, H. B., *Electrophoresis* 2003, *24*, 296–302.

