Applications of Formalin Fixed Paraffin-Embedded Tissue Proteomics in the Study of Cancer

**Abstract**

Among the many challenges in identifying novel tumoral diagnostic, prognostic, and personal therapy biomarkers we highlight the challenge of dealing with a single tumor’s molecular complexity and heterogeneity. Formalin-fixed paraffin-embedded (FFPE) tissue specimens are commonly used by pathologists for diagnosing cancer; this material comprises a valuable resource for molecular studies addressing tumoral molecular profiles. Here we demonstrate the effectiveness of shotgun proteomics in assessing the cellular heterogeneity using FFPE tissues slides under three different scenarios: 1) Comparing astrocytoma grade I versus glioblastoma, 2) Assessing different morphological areas from the same histological slide of a glioblastoma and comparing the identifications with those from a fresh tissue, 3) Comparing two different histopathological profiles in the same patient’s liver: cirrhosis and cancer.

Briefly, contents of FFPE slides were scraped, washed three times with xylene and then rehydrated in a grade ethanol series [100%, 90%(v/v), 70 % (v/v)]. A total of 100 µL of 0.2% (w/v) RapiGest™ in 50mM ammonium bicarbonate was then added to each sample. Tryptic digests were separated by reversed-phase capillary liquid chromatography coupled to nano-electrospray high-resolution mass spectrometry for identification. The samples were analyzed in technical replicates by LC-MS/MS on an Orbitrap XL. Our datasets enabled elucidating key tumoral regions allowing a better comprehension of such specific tumor areas; the identifications and raw data are available at http://max.ioc.fiocruz.br/juliffppecases.

**Keywords:** Formalin fixed paraffin-embedded; Glioblastoma; Hepatic cancer; Cirrhosis; Shotgun proteomics

**Introduction**

Tumors consist of cells in different stages of transformation, resulting in both molecular and cellular heterogeneity. While cellular morphology is characterized by pathologist, it is now known that what might apparently seems as morphologically healthy tissue might already be molecularly compromised [1,2]. The application of mass spectrometry based proteomics in the analysis of formalin fixed paraffin-embedded (FFPE) tissue poses as a powerful technique to enable an in depth study of specific tumoral areas [3]. Needless to say, there is a plethora of FFPE tissue specimens available, which composes a treasure-trove for retrospective studies paving the way for mass spectrometric assessment of archived material of rare diseases [4].

To exemplify the usefulness of FFPE in molecularly assessing tumoral heterogeneity we demonstrate applications of FFPE proteomics by studying three very distinct scenarios: in the first, we compare the relationship between astrocytoma grade I and glioblastoma; in the second, different morphological regions from the same glioblastoma histological slide and a fresh GBM tissue sample are contrasted; the last compares two different histopathological areas, cirrhosis and hepatic cancer, from the same patient.

**Methods**

The sample preparation and data analysis for FFPE cancer proteomics performed for generating the datasets are described henceforth. The ethics committee review board of the University Hospital Clementino Fraga Filho reviewed and approved this work under number 060-11.

**FFPE protein extraction**

The contents of microscope slides were scraped using a syringe needle and the material transferred to an Eppendorf type tube. Then the contents were washed three times with 250 µL of xylene at 60°C. Afterwards, the contents were rehydrated with 250 µL of a grade ethanol series [100%, 90%(v/v), 70 % (v/v)], followed by centrifugation during 30 minutes at 15000 x g and air dried at room temperature until complete ethanol evaporation.

A total of 100 µL of 0.2% (w/v) RapiGest™ in 50mM ammonium bicarbonate was then added to each sample. The protein mass within the samples was estimated using the BCA protein assay Kit (Sigma-Aldrich) according to the manufacturer’s instructions. The contents from each sample were reduced with 5mM of dithiothreitol (DTT) and incubated for 90 min at 95°C, followed to 2 h of incubation at 70°C. The sample incubation is necessary to hydrolyze the cross links generated during formalin fixation of the tissue. The samples were cooled to room temperature and incubated in the dark with 15mM of iodoacetamide (IAA) for 30 minutes.

**Fresh GBM protein extraction**

A fresh GBM sample was obtained from a 71 year old man with...
Hematoxylin and eosin stained slides from astrocytoma grade I (A, B, C) and glioblastoma (D, E, F). A, D: 40x; B, E: 100x; C, F: 400x.

In image C, the characteristic microcystic areas are prominent and little pleomorphism can be noticed, in contrast with image F, where nuclear atypia and hypercellularity are easily seen.

Results and Discussion

Scenario 1: Astrocytoma grade I versus glioblastoma

The first scenario compares astrocytoma grade I and glioblastoma from microscope slides of formalin fixed paraffin-embedded tissues. To evaluate the differentially expressed proteins in astrocytoma grade I (pilocytic astrocytoma) and grade IV (glioblastoma), the contents of five 5 µm sections obtained from each case were scraped from the microscope slides and the proteins were extracted and analyzed once as described above. The microscope slide images stained with hematoxylin and eosin are shown in Figure 1. Our results describe 715 (145 by maximum parsimony [10]) and 623 (110 by maximum parsimony) proteins identified from astrocytoma grade I and glioblastoma section, respectively.

We used the Pattern Lab for proteomics for pinpointing identified proteins from astrocytoma grade I and glioblastoma. Our results showed that 52% of proteins were common to both grades, 256 proteins were exclusively identified in astrocytoma grade I and 164 in the glioblastoma. The total number of identified proteins was 879 (178 by maximum parsimony) when considering all results simultaneously. We recall that not identifying a protein by mass spectrometry means that it is either not present in the sample or is under the detection limits of the strategy at hand; as such, it can be hypothesized as “off” or as differentially expressed when comparing to a condition where it is identified [9]. We opted for this conservative approach.
strategy as in a previous report where analyses were performed in the same conditions and with very limited sample [11].

Among the proteins exclusively identified in the glioblastoma dataset, it is worth pinpointing 14-3-3 gamma, dermcidin, and defensin. Proteins of the 14-3-3 family are linked to many key cellular processes related to cancer such as apoptosis and cell cycle key points; for these reasons, they are considered as important therapeutic targets [12]. In particular, 14-3-3 gamma was previously described, among eleven other proteins, as glioblastoma biomarker candidates by the use of large scale proteomic approaches [13]. Dermcidin and defensin were both described as having antimicrobial effects; interestingly, both were only identified in the glioblastoma sample. Dermcidin and dermcidin-derived peptides have been described to play a key role in several biological functions in neuronal and cancer cells [14]; although this protein has been described as up-regulated in various cancer cell types [14], as far as we know, this is the first report claiming such for glioblastoma. Defensin has been described as a potential strategy for cancer immunotherapy as promising immunogenes [15].

Among the proteins exclusively identified in astrocytoma grade I we highlight annexin A1 and Brevican. Annexin-1 (ANXA1) belongs to a structurally related family of calcium-binding anti-inflammatory proteins that act extracellularly as an inhibitor of iocosanoid synthesis and phospholipase A2 [16]. Schittenhelm et al. proposed that the up-regulation of annexin-1 in astrocytomas is linked to a bad prognosis for tumor; moreover showed that its expression profile is highly correlated to its substrate, EGFR [17]. Brevican has been previously correlated as a marker for low-grade gliomas [18]. Figure 2 shows a carbon metabolism metabolic pathway map; as can be noticed, different sections of the pathways are up-regulated in each type of cancer.

![Figure 2: Metabolic pathway map of the carbon metabolism generated using KEGG mapper. Green links are associated to up-regulated proteins in astrocytoma grade I, red links to glioblastoma, and blue links indicate no significant change in regulation.](image-url)
Scenario 2: Different morphological regions from the same glioblastoma histological slide and comparison with fresh GBM tissue

The second dataset compares FFPE tissue of two distinct areas of the same glioblastoma tumor and GBM from fresh tissue; each condition had three technical replicate analyses. Figure 3 shows a picture of the histological slide used for obtaining the material for this analysis. Glioblastomas characteristically exhibit a wide range of morphological aspects, from small cell areas to giant or highly pleomorphic areas, although considered to be monoclonal in origin. Irrespectively, the clinical outcome is said to be unrelated to this variation. The goal was to verify if a proteomic assessment of FFPE could pinpoint differences in the proteomic profiles of different regions of the same tumor that presented a significantly different morphology as in accordance to a pathologist’s judgment. Two regions of interest were selectively scraped from the slide and the proteins were extracted according to our methods section.

Our proteomic analysis revealed 1463 proteins; 439 were identified in region A (Figure 4) 1058 in region B (Figure 4) and 705 in the fresh tissue sample. We expected to identify more proteins in the fresh tissue as in according to previous works [19]; yet, our results ended up being different. We believe this was due to the fresh tissue sample originating from a different patient than that of the FFPE tissue; as previously reported, there is a tremendous variation between the proteome of different individuals [20]. As an aggravator, according to Fowler et al, “even for matched fresh tissue specimens, overlap in the identified proteins can be as low as 50% due to variations in sample preparation and under sampling of the resulting peptides” [21].

Of the 1463 proteins, 368 were identified only in the fresh tissue and another 731 were not identified in the fresh tissue. There was a significantly greater overlap in the protein identifications originating from the two different FFPE areas (92% overlap) than from FFPE versus fresh tissue (46% overlap). Details on IDs are provided at the end of this manuscript and in the supplementary table. We hypothesize that more proteins were identified in region B because more material was obtained during the protein extraction. As can be noted in Figure 3, region B is approximately 4 x larger than region A. Our results pinpoint 18 proteins exclusively identified in region A when comparing with region B. We consider this a positive result, as even though significantly fewer proteins were identified in this area, we were still able to find some “uniqueness” in it; this raises our confidence that these proteins shed important biological aspects of the disease. Such result demonstrates the great potential of this methodology to reveal the cellular heterogeneity of the tumor as reflected in the hematoxylin and eosin stains (Figure 4). Among the unique proteins identified in region A we point out cold-inducible RNA binding protein and THO complex subunit 4. The former has been described as decreased in breast cancer cells [22] and the latter to have a key role in anti-cancer mechanisms [23]. That said, we emphasize that specific proteins correlate to particular histological sections and therefore could ultimately be used for molecularly defining distinct areas, as tumors are very heterogeneous.

Scenario 3: Different areas, cirrhosis and hepatic cancer, from the same patient

The third dataset compares the proteomic profiles of different pathologies in the liver of the same patient obtained in the same surgical procedure; three technical replicate analyses from each condition were performed. In this case, two different histopathological areas, cirrhosis and cancer, were identified by three independent pathologists. This makes the material at hand extremely valuable as it is hard to be obtained when belonging to the same patient. That said, this case demonstrates the usefulness of FFPE proteomics to enable retrospective studies from archival material. For the case at hand, FFPE proteomics enabled identifying proteins that could be related to the progression between these two distinct pathologies helping to elucidate the relationship between these diseases. The microscope slides were assessed by a pathologist (Figure 5); areas of necrosis were marked and excluded from our proteomic analysis.

A total of 277 proteins (69 by maximum parsimony) were identified in both cases; 115 for cirrhosis and 267 for hepatic cancer; 10 exclusively identified in cirrhosis and 162 in cancer. Among the exclusive proteins in cancer, again, the protein family of 14-3-3 poses as interesting candidate for progression biomarker to hepatic cancer. These consist a family, containing 7 isoforms of highly conserved...
and largely expressed eukaryotic acidic proteins (25-33kDa) that dynamically regulates the activity of target proteins in various signaling pathways and control diverse physiological and pathological processes (e.g., apoptosis and cell cycle regulation), including prion diseases [12,24,25]. 14-3-3 is integrated into the core of regulatory pathways that are key to the normal growth and development, and therefore posing this protein as progression biomarker candidate as a disturbance in these regulatory pathways can lead to cancer.

Final Considerations

A general trend in proteomics is comparing protein profiles from cancer biopsies versus those from healthy tissues ultimately entailing the discussion on how many samples are required to pinpoint candidate biomarkers. An emerging, and complementary strategy, is to dig deeper into an individual’s proteome by molecularly dissecting it in association with morphological and spatial data as performed in this work. We advocate that each cancer should be considered as a disease of its own; needless to say, patients respond differently to treatment. As far as we know, the idea of associating large-scale protein identification to spatial data was pioneered by Richard Smith’s group with two seminal works that sectioned a mammalian brain and produce images of identified protein distributions along the brain [26,27]. Our group’s efforts are also aligned with “spatial proteomics”; we recently developed software automatically generating shotgun images and exemplified its usefulness by delineating a gastric cancer tumoral region based on Mud PIT data [28]. The literature is full of examples showing applications of FFPE for identifying candidate biomarkers through various strategies using FFPR, including Single Reaction Monitoring [29], and laser micro-dissection [30]. Clearly, there is a drive for shotgun proteomics to becoming more and more spatially centric, ultimately aiming at a single-cell’s proteome [31]. In our view, studies that combine FFPE with shotgun proteomics are in a sweet-spot in the compromise between large-scale protein identification and spatial resolution; this strategy is under continuous perfection by the scientific community [32].

Here, we provide key datasets demonstrating the great potential of FFPE proteomics for identification of tumor biomarkers candidates (dataset 1), comparing distinct proteomic profiles from different regions of the same tumor versus a fresh tissue sample (dataset 2), and two distinct histopathological areas in the liver of the same patient which could possibly describe a molecular linkage between progression from cirrhosis to cancer (dataset 3). These datasets are unique as few proteomic datasets disclose shotgun identifications of tumoral specific region delineated by microscopy and with the corresponding histopathological assessment. Supplementary Table I discriminates all identifications from each biological condition of each case followed by its respective protein length (in amino acids), number of unique peptides identified, molecular weight, sequence
count, spectrum count, protein coverage, number of technical replicates identified in, and the protein’s description.

Figure 6 summarizes the results obtained herein using area proportional Venn diagrams. All data, including raw files and SEPro identifications are available at http://max.ioc.fiocruz.br/julii/fpsecases/. The data is organized in three separate folders enumerated as in according to this article.

Author Contributions

JSGF, PFA, KMSG, and BA were responsible for sample preparation and helped in experimental design. PCC and MGCC wrote the manuscript and participated in discussion. PCC did the mass spectrometry data analysis. MGCC, NHSC, and VLNP are medical doctors and performed sample analysis procedures.

Acknowledgement

The authors thank CNPq, FAPERJ, and Fundação do Câncer for financial support. The authors thank the Program for Technological Development in Tools for Health-PDTIS-FIOCRUZ (PDTIS) for use of the mass spectrometry facility RPT02H PDTIS / Carlos Chagas Institute - Fiocruz Paraná and the mass spectrometry core facility from Fiocruz - Rio. The authors acknowledge Dr. Michel Batista from Instituto - Fiocruz Parana and the mass spectrometry core facility of the mass spectrometry facility RPT02H PDTIS / Carlos Chagas Institute - Fiocruz – Rio. The authors thank CNPq, FAPERJ, and the financial support. The authors thank the Program for Technological Development in Tools for Health-PDTIS-FIOCRUZ (PDTIS) for use of the mass spectrometry facility RPT02H PDTIS / Carlos Chagas Institute - Fiocruz Paraná and the mass spectrometry core facility from Fiocruz - Rio. The authors acknowledge Dr. Michel Batista from Instituto - Fiocruz Parana and the mass spectrometry core facility of the mass spectrometry facility RPT02H PDTIS / Carlos Chagas Institute - Fiocruz – Rio. For aiding in the mass spectrometry procedures. The authors declare no conflict of interests.

References

