<table>
<thead>
<tr>
<th><strong>Title</strong></th>
<th>Quantitative clinical proteomic study of autopsied human infarcted brain specimens to elucidate the deregulated pathways in ischemic stroke pathology</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author(s)</strong></td>
<td>Datta, Arnab; Akatsu, Hiroyasu; Heese, Klaus; Sze, Siu Kwan</td>
</tr>
<tr>
<td><strong>Date</strong></td>
<td>2013</td>
</tr>
<tr>
<td><strong>URL</strong></td>
<td><a href="http://hdl.handle.net/10220/18364">http://hdl.handle.net/10220/18364</a></td>
</tr>
<tr>
<td><strong>Rights</strong></td>
<td>© 2013 Elsevier B.V. This is the author created version of a work that has been peer reviewed and accepted for publication by Journal of proteomics, Elsevier. It incorporates referee’s comments but changes resulting from the publishing process, such as copyediting, structural formatting, may not be reflected in this document. The published version is available at: [<a href="http://dx.doi.org/10.1016/j.jprot.2013.08.017">http://dx.doi.org/10.1016/j.jprot.2013.08.017</a>].</td>
</tr>
</tbody>
</table>
Quantitative Clinical Proteomic Study of Autopsied Human Infarcted Brain Specimens to Elucidate the Deregulated Pathways in Ischemic Stroke Pathology

Arnab Datta\textsuperscript{1}, Hiroyasu Akatsu\textsuperscript{2}, Klaus Heese\textsuperscript{3*}, Siu Kwan Sze\textsuperscript{1*}

\textsuperscript{1}School of Biological Sciences, Nanyang Technological University, Singapore 637551, Singapore

\textsuperscript{2}Choju Medical Institute, Fukushimura Hospital, Toyohashi, Aichi 441-8124, Japan.

\textsuperscript{3}Graduate School of Biomedical Science and Engineering, Hanyang University, Seongdong-gu, Seoul 133-791, Republic of Korea

*Corresponding authors

*Correspondence:
Siu Kwan Sze, PhD, Tel.: +65 65141006, Fax: +65 67913856, E-mail: sksze@ntu.edu.sg;
Klaus Heese, PhD, Tel.: +82 2 2220 0398; Fax: +82 2 2296 5943, E-mail: klaus@hanyang.ac.kr

Keywords: clinical proteomics, ferritin, ischemic stroke, 8-plex iTRAQ, malate-aspartate shuttle
ABSTRACT

Ischemic stroke, still lacking an effective neuroprotective therapy is the third leading cause of global mortality and morbidity. Here, we have applied an 8-plex iTRAQ-based 2D-LC-MS/MS strategy to study the commonly regulated infarct proteome from three different brain regions (putamen, thalamus and the parietal lobe) of female Japanese patients. Infarcts were compared with age-, post-mortem interval- and location-matched control specimens.

The iTRAQ experiment confidently identified 1520 proteins with 0.1% false discovery rate. Bioinformatic data mining and immunohistochemical validation of pivotal perturbed proteins revealed a global failure of the cellular energy metabolism in the infarcted tissues as seen by the parallel down-regulation of proteins related to glycolysis, pyruvate dehydrogenase complex, TCA cycle and oxidative phosphorylation. The concomitant down-regulation of all participating proteins (SLC25A11, SLC25A12, GOT2 and MDH2) of malate-aspartate shuttle might be responsible for the metabolic in-coordination between the cytosol and mitochondria resulting in the failure of energy metabolism. The levels of proteins related to reactive gliosis (VIM, GFAP) and anti-inflammatory response (ANXA1, ANXA2) showed an increasing trend. The elevation of ferritin (FTL, FTH1) may indicate an iron-mediated oxidative imbalance aggravating the mitochondrial failure and neurotoxicity. The deregulated proteins could be useful as potential therapeutic targets or biomarkers for ischemic stroke.

Biological significance

Clinical proteomics of stroke has been lagging behind other areas of clinical proteomics like Alzheimer's disease or schizophrenia. Our study is the first quantitative clinical proteomics study where iTRAQ-2D-LC–MS/MS has been utilized in the area of ischemic stroke to obtain a comparative profile of human ischemic infarcts and age-, sex-, location- and post-mortem interval-matched control brain specimens. Different pathological attributes of
ischemic stroke well-known through basic and pre-clinical research such as failure of cellular energy metabolism, reactive gliosis, activation of anti-inflammatory response and aberrant iron metabolism have been observed at the bedside. Our dataset could act as a reference for similar studies done in the future using ischemic brain samples from various brain banks across the world. A meta-analysis of these studies could help to map the pathological proteome specific to ischemic stroke that will guide the scientific community to better evaluate the pros and cons of the pre-clinical models for efficacy and mechanistic studies.

Infarct being the core of injury should have the most intense regulation for several key proteins involved in the pathophysiology of ischemic stroke. Hence, a part of the up-regulated proteome could leak into the general circulation that may offer candidates of interest as potential biomarkers. In support of our proposed hypothesis, we report ferritin in the current study as one of the most elevated proteins in the infarct, which has been documented as a biomarker in the context of ischemic stroke by an independent study. Overall, our approach has the potential to identify probable therapeutic targets and biomarkers in the area of ischemic stroke.
INTRODUCTION

Ischemic stroke, still lacking an effective neuroprotective therapy, continues to be a major socioeconomic burden throughout the world. The successive clinical failures in new drug development along with a bleak epidemiological landscape have led to the emergence of several novel concepts. Increased emphasis has been given on the neurovascular unit and the interaction between its different components instead of neuron alone. Participation of peripheral organs through bidirectional communications with brain following Ischemic stroke has been highlighted [1]. Consequently, a combination therapy instead of a mono-therapy or incorporation of a multi-target drug has been suggested [2]. Therefore, a global comprehensive study of the altered system(s) is the prerequisite to understand and tackle this multifactorial disorder.

Quantitative neuroproteomics has emerged as an advanced technique in the post-genomic era for unbiased probing of perturbed pathways in complex biological systems not possible by the traditional reductionist approaches. Recently iTRAQ-based neuroproteomic strategy has been applied to improve our mechanistic understanding of axon injury and Alzheimer’s disease (AD) [3, 4]. In the area of ischemia stroke, we pioneered to employ the iTRAQ-2D-LC-MS/MS-based quantitative proteomic strategy to study an in vitro neuronal model of ischemic penumbra [5] and an in vivo rodent model of transient focal stroke [6] to delineate the molecular complexities and to propose potential therapeutic targets for further studies. However, the direct translation of hypothesis generated through pre-clinical models into clinical application is limited by the complexity and fundamental difference of the biological systems as exemplified by successive clinical failures in translating the targets. The reasons may be related to the extreme heterogeneity of human stroke, the absence of long-term environmental influence or co-morbidity or risk factors in the pre-clinical models or flaws in the clinical trial design [7]. Moreover, the majority of the animal studies used rodent that are
phylogenetically separated from human millions of years ago [8]. Hence, a wealth of bedside-
back-to-bench data using post-mortem brain specimens is needed to exclude model-specific
artifacts and to identify stroke-specific ‘target space’ encompassing different pathological
events and brain cell types. However, studying the molecular pathophysiology of ischemic
stroke using clinical samples is confronted with various difficulties and challenges that may
account for pre-analytical or ex vivo stresses [9]. Additionally, for a pathological condition
like stroke, considerable neuropathological and diagnostic expertise are needed as only
specific locations have to be targeted for isolation. Unsurprisingly, although several
methodological studies using different sub-structures of human brain have been reported for
profiling either whole proteome [10-12] or sub-proteome[13], only a few studies addressed
pathological questions related to ischemic stroke [14, 15]. The first one used a targeted
approach by combining Laser Microdissection with protein array to focus on different matrix
metalloproteinases [14]. The other one is the only reported profiling study to examine
different areas of ischemic brain where 2DE with an off-line MALDI MS/MS was applied
[15]. Using a complementary online shotgun proteomic approach, such as the multi-
dimensional protein identification technology [16] with isobaric labeling of peptides (i.e.
iTRAQ) will allow simultaneous quantification of proteins in up to 8 samples by avoiding
intra-experimental variation.

Given that the iTRAQ-2D-LC-MS/MS strategy has already been successfully applied on the
pre-clinical models of ischemic stroke in our laboratory [5, 6] here we extend to
quantitatively study three well-documented and pathologically-characterized infarcted tissues
of human brain by an 8-plex iTRAQ experiment. Ethnicity-, age-, sex-, location- and post-
mortem interval (PMI)-matched areas of the non-infarcted brains were used as control for an
appropriate comparison. The location-matched perturbed proteome obtained from the three
pairs of infarct-control specimens were validated by comparing the expression of each protein
with an independent control and infarcted sample included in the same iTRAQ experiment
during the data analysis. Mining the data set revealed a mitochondrial dysfunction that may
have caused by the uncoupling of cytosolic and mitochondrial metabolism along with an
iron-mediated oxidative imbalance.

MATERIALS AND METHODS

Reagents

Unless indicated, all reagents were purchased from Sigma-Aldrich (St. Louis, MO).

Sample Collection

Autopsy and Sampling of Brain Tissues

Post-mortem brain samples of 7 subjects (infarct, n=3; control, n=4) were obtained from the
brain bank of the Choju Medical Institute of Fukushima hospital (Toyohashi, Aichi, Japan).
The protocols utilized were approved by the local ethics committee of the Fukushima
hospital. The scientific use of human material was conducted in accordance with the
Declaration of Helsinki, and informed consent was obtained from the guardians of the
patients [17]. Institutional Review Board approval from Singapore for the use of human
materials had been obtained for this study.

All subjects underwent extensive clinical and neuropathologic characterization to document
associated neurodegenerative disorders, if any. Pathological diagnosis of AD-type dementia
was carried out in accordance with the ‘Consortium to Establish a Registry for AD’
 guidelines [18]. The neurofibrillary pathology was determined by Braak staging [19].
Screening for Parkinson’s disease (PD) was performed according to the criteria described
For clinical diagnosis of vascular dementia, the criteria presented at the NINDA-AIREN International Workshop were used [21].

Brains were dissected by an experienced neuropathologist; infarcts areas were isolated with the help of neuroimaging and based on the consistency and color of the parenchyma. The control cases were isolated from the respective ipsilateral locations of subjects who never had clinically defined stroke and died due to non-neurological causes. All tissue pieces were weighed at the time of autopsy. The samples were snap frozen with liquid nitrogen before being stored at –80 ºC [20].

**Experimental Design**

The experimental design is described in Figure 1 and Supplemental Table 1. Cerebrovascular disease being a circulatory complication can affect any part of the brain unlike some other central nervous system disorders that have a brain-site specific pattern (e.g. thalamus for schizophrenia or limbic system for epilepsy). Recently, spatial pattern of differential vulnerability in ischemic stroke patients have been demonstrated using cerebral blood flow and infarction volume [22]. Thus, it is imperative to target more than one location to identify the perturbed proteome that is beyond the spatial influence. Accordingly, three functionally and structurally distinct regions (putamen, thalamus and the parietal lobe) of the brain were chosen during the sample selection. Control brain tissues were collected from the identical locations of the ipsilateral hemisphere of individuals who did not have infarcts. Infarcts of patient ID 0914 (putamen), 1006J (thalamus) and 1006H (parietal lobe) were matched by the location with control ID 0406, 0003, and 0917, respectively (biological replicate = 3, Supplemental Table 1, Supplemental Figure 1).
Ischemic stroke is mostly heterogeneous having a scattered pattern of brain lesions (archipelago pattern) rather than an uniform pattern (fried egg pattern), where penumbra nicely covers the core tissue [23]. To address this, one of the patients (ID 1006) had multiple infarcts with tissues from thalamus (slice J) and parietal lobe (slice H). Comparing them will provide an estimate about the inherent spatial heterogeneity of the infarcts at the level of protein expression. In addition, two additional samples acted as a random internal control (ID 9825, parietal occipital lobe) and infarct (ID 1008, infarct from cerebellum cortex) that was used to validate the trends, obtained from the three independent pairs of location-matched samples. In total, eight specimens of brain tissue (n=4 per group) were used for the 8-plex iTRAQ experiment (Figure 1). The groups were ethnicity- (Japanese), age- (control (mean ± SD), 86 ± 4.7 y; infarct, 82.5 ± 0.1 y) and gender- (sex, female) matched. The median PMI was 3.5 h (range 2 - 36 h) (Supplemental Table 1).
The LC-MS/MS analysis was performed thrice (technical replicate= 3) to measure the analytical variation of the MS instrument [5, 6].

**Proteomics**

**Sample Preparation**

Frozen tissue pieces were homogenized with liquid nitrogen before lyses with ice-cold lyses buffer (2% sodium dodecyl sulfate (SDS); 1 M triethylammonium bicarbonate (TEAB) with a protease inhibitor cocktail (Roche; Mannheim, Germany)). The supernatant obtained following centrifugation (20 600xg, 15 min) was precipitated and washed with ice-cold acetone to remove non-protein contaminants. The protein pellets were redissolved in lyses buffer (6% SDS, 20 mM DTT, 100 mM Tris, pH 7.8). Protein quantification was performed using a 2-D Quant kit (Amersham Biosciences, Piscataway, NJ, USA). Details of the sample preparation are provided in the SI Materials and Methods.

**In-gel Tryptic Digestion and Isobaric Labeling**

The samples (200 μg of protein/ condition) were subjected to a denaturing PAGE using a 4% - 6% - 25% gel following a similar procedure as described previously [6]. The diced gel bands were extensively washed with 50 mM TEAB in 50% ACN to completely remove Tris HCl and detergent, then reduced (5 mM tris-(2-carboxyethyl) phosphine, 60 °C, 1 h) and alkylated (10 mM methyl methanethiosulfonate in isopropanol, room temperature, 45 min in dark) before being digested with 12.5 ng/μL of sequencing-grade modified trypsin (Promega, Madison, WI, USA) in 50 mM TEAB, 2% ACN for overnight at 37 °C. The peptides were extracted with 50% ACN, 5% acetic acid and vacuum centrifuged to dryness. The dried peptides were reconstituted into 0.5 M TEAB and labeled with respective isobaric tags of 8-plex iTRAQ Reagent Multi-Plex kit (Applied Biosystems, Foster City, CA, USA) as follows:
for stroke patients (Supplemental Table 1, Figure 1). These numbers (113-121) correspond to the mass of the reporter ions that appear as distinct peaks in the MS/MS spectra after fragmentation of the labeled peptides. These peak intensities are used to deduce the relative concentration of the corresponding peptide [24]. The labeling reaction was stopped by adding excess water after 2 h. The samples were combined and dried in a vacuum centrifuge.

Electrostatic Repulsion and Hydrophilic Interaction Chromatography (ERLIC)

The combined iTRAQ sample was desalted by Sep-Pak C18 SPE cartridges (Waters, Milford, MA). A modified ERLIC using volatile salt-containing buffers was adopted [25]. The dried iTRAQ-labeled peptide was reconstituted in 200 µl of Buffer A (10 mM NH₄HCO₃, 85% ACN, 0.1% formic acid (FA)) and fractionated using a PolyWAX LP column (200 × 4.6 mm; 5 µm; 300 Å) (PolyLC, Columbia, MD, USA) on a Prominence HPLC system (Shimadzu, Kyoto, Japan) in a 65 min gradient with Buffer B (30% ACN, 0.1% FA). The HPLC gradient was composed of 100% buffer A for 10 min; 0-25% buffer B for 35 min; then 25-100% buffer B for 10 min; followed by 100% buffer B for 10 min. The chromatogram was recorded at 214 nm. Eluted fractions were collected in every 1 min, and then pooled into 34 fractions depending on the peak intensities, before drying them in a vacuum centrifuge. They were stored at -20 °C till MS analysis.

Reverse Phase LC-MS/MS Analysis using QSTAR

The iTRAQ-labeled peptides were reconstituted with 0.1% FA, 3% ACN and analyzed using a HPLC system (Shimadzu) coupled with QSTAR Elite Hybrid MS (Applied Biosystems/MDS-SCIEX) as described previously with minor modifications [6]. Briefly, most of the LC parameters for a 90 min gradient including column configuration, gradient
and flow rate were kept constant except the mobile phase A composition (0.1% FA in 3% ACN) and sample injection volume (15 µl/injection). Regarding MS parameters, the precursors with a mass range of 300–1600 m/z and calculated charge of +1 to +5 were selected for the fragmentation. The selected precursor ion was dynamically excluded for 20 s with a 50 mDa mass tolerance. All other parameters were kept identical as reported previously [6]. The peak areas of the iTRAQ reporter ions reflect the relative abundance of the proteins in the samples.

**Mass Spectrometric Raw Data Analysis**

The spectral data acquisition was performed using the Analyst QS 2.0 software (Applied Biosystems). ProteinPilot Software 3.0, Revision Number: 114 732 (Applied Biosystems) was used for peak list generation, protein identification and quantification against the concatenated target-decoy Uniprot human database (total 191242 proteins) for calculating the false discovery rate (FDR). Details of the analysis strategy have been described previously [6].

**Bioinformatics Analysis**

Enrichment analysis was performed by submitting Uniprot accession numbers of the list of proteins to open-source software DAVID [26]. Open-source Gene Pattern software (version 3.3.3) was used for clustering the final list of regulated proteins by hierarchical clustering algorithm [27]. Details of the analysis are provided in the SI Materials and Methods.

**Western Blot (WB) Analyses**

WB was performed after SDS-PAGE by probing with primary antibodies at the indicated dilutions: anti-ACTB (beta-actin, 1:5000, mouse monoclonal; Millipore, Billerica, MA,
USA), anti-ALB (albumin, 1:5000, rabbit polyclonal; Abcam, Cambridge, UK), anti-ANXA2 (annexin II, 1:2000, mouse monoclonal; Santa Cruz Biotech, Santa Cruz, CA, USA), anti-GAPDH (glyceraldehyde-3-phosphate dehydrogenase, 1:5000, mouse monoclonal; Santa Cruz, USA), anti-ferritin (1:2000, rabbit polyclonal; Abcam), anti-HSP70 (heat shock protein 70, 1:5000, mouse monoclonal; Abcam), anti-MBP (myelin basic protein, 1:2000, mouse monoclonal; Abcam), anti-SOD1 (superoxide dismutase 1, 1:2000, rabbit polyclonal; Abcam), anti-SIRT2 (sirtuin 2, 1:1000, rabbit polyclonal; Cell Signaling, Danvers, MA), anti-SYN1 (synapsin I, 1:800, rabbit polyclonal; Cell Signaling), anti-VDAC1 (voltage dependent anion channel 1, 1:2000, mouse monoclonal; Santa Cruz), anti-VIM (vimentin, 1:4000, mouse monoclonal; Millipore) and Mitoprofile Total OXPHOS Human WB Antibody Cocktail (1:2000, mouse monoclonal, MS601; Mitoscience, Eugene, Oregon). MS601 is an optimized premixed cocktail containing five monoclonal antibodies (Complex I subunit (NDUFB8), Complex II subunit (SDHB), Complex III subunit Core 2 (UQCRC2), Complex IV subunit II, ATP synthase subunit alpha (ATP5A)) representing various enzymes of the mitochondrial electron transport chain. Twenty micrograms of protein was used for WB. Immunoreactivity was detected by using an HRP chemiluminescent substrate reagent kit (Invitrogen, Carlsbad, CA). A pooled sample was used to normalize the inter-gel variation between repeated runs for the same protein.

**Statistical Analysis**

All statistical analyses were performed using SPSS 13.0 for Windows software (SPSS Inc.). Experimental data for WB analysis were presented as mean ± SEM. One sample t-test was used to compare the significant difference of location-matched ratios from one. Statistical significance was accepted at *p < 0.05.
RESULTS

Quality Control of the iTRAQ Data Set

To minimize the false positive identification of proteins, an unused ProtScore ≥ 2 (confidence >99%) was used as the qualification criteria. With this criterion, 1520 proteins were identified with a FDR of 0.1% (Supplemental Table 2). The average number of unique peptides (having a confidence level of >95%) detected per protein was 6.96 and more than 30% of the proteins had ≥5 unique peptides that was similar to our previous findings [5, 6]. The quality of the sample preparation was determined by plotting the distribution of the proteins according to their theoretical MW and theoretical pI values as depicted in Supplemental Figure 2. The presence of proteins covering a wide MW spectrum from high (>240 kDa) to low (<10 kDa) and also a wide pI range from acidic (pI ~4) to basic (pI >12) character indicated an extensive representation of the whole proteome in the current data set.

Selecting the Final List of Perturbed Proteins

Filtering the data set using a p-value cut-off of 0.05 obtained 49, 34 and 141 proteins with significant ratios from the three pairs of location-matched samples (e.g., 117/114, 118/115, 119/116). To capture the consensus signature of the infarcts, the proteins that were significantly regulated in at least two of these three sample-pairs (i.e. at least in two locations from putamen, thalamus or parietal lobe) were included for further analysis in this study. Thus, 46 candidates were short-listed (Supplemental Table 3, Section I) from the initial list of 1520 proteins. Next, the remaining two samples, a random control from parietal occipital lobe (label 113) and a random infarct from cerebellum cortex (label 121) were introduced for internal validation of these 46 candidates (Supplemental Table 3, Section II and III). Ideally, in comparison with location-matched pairs (117/114, 118/115, and 119/116), these proteins should show no or opposite regulation in 113 (i.e., 113/114, 113/115, and 113/116) and no or similar trends in 121 (i.e., 121/114, 121/115, 121/116) to be true positives. Remarkably, 33
out of 46 (~72%) candidates were qualified through this dual validation (Supplemental Table 3, last column). Overall, this data set represents a brain-site-specific response pattern, many of which are consistent across different brain locations amid some spatially distinct signatures. Thus, the flexibility of the experimental design combined with the multiplex iTRAQ labeling technology allowed various comparisons to check the specificity and accuracy of the candidates. No cut-off based on the magnitude of the ratio was set on the final list of proteins. This list of significantly regulated proteins (33 out of 1521; 2.2% of total hits) was advanced to the next phase for rigorous validation by complementary techniques.

Post-proteomic Validation by WB Analysis

Figure 2. Post-proteomic validation of the selected proteins from the location matched control and infarcted brain tissue by WB analysis. A) The band numbers (1-6) correspond to the three pairs of location-matched samples (1→4, 2→5 and 3→6) with patient IDs as mentioned in the Figure 1, Supplemental Table 1. Band 5 and
Eight proteins (i.e. VIM, ANXA2, HSP70, ferritin, VDAC1, MBP, ATP5A and UQCRC2) out of the 33 filtered candidates and three more proteins (SYN1, ALB and SIRT2) from the initial list of 46 candidates were selected to check the reliability of the iTRAQ quantification. All eleven proteins showed consistent trends with the significantly regulated location-matched ratios from the iTRAQ data set (Figure 2). GAPDH and ACTB were included as they are commonly used as loading controls. However, our result showed a down-ward trend of both of these housekeeping proteins in the infarcted areas of the brain which is consistent with a previously published report [15]. Conversely GAPDH was found to be increased in the hippocampus of AD patients [28]. The significant increase in the SYN1 level in the thalamus (Supplemental Table 3) is due to the presence of a non-ideal control (Figure 2). Ferritin showed the most consistent increase in abundance irrespective of the location of the infarct. SOD1 was included as a marker of redox balance to estimate the oxidative stress among the two groups. A generalized decline of the proteins of oxidative phosphorylation (OxPhos) (i.e. ATP5A, UQCRC2 and complex IV subunit II) could be related to the oxidative stress as seen from the down-ward trend of SOD1 in the infarcted specimens.
Bioinformatics Analysis of Age-Sex-Location-PMI-matched Proteome of Human Infarcted Brain Tissue

Mitochondrion was found to be the most enriched organelle in the cellular component category of the gene ontology analysis as 50% of the annotated proteins were of mitochondrial origin or associated with mitochondria. Notably, only 13 proteins of the mitochondrial proteome are being encoded by the mitochondrial genome and the rest must be imported from the cytosol [29]. Most of the mitochondrial proteins are made as larger precursors in the cytoplasm with a transit peptide at the N-terminal and are post-translationally incorporated into mitochondria. This was reflected by the ‘SP_PIR_KEYWORDS’ analysis, where the top two qualified and significantly enriched keywords was ‘mitochondrion’ and ‘transit’ having 15 and 11 (out of 33) proteins respectively (Supplemental Table 3).

Figure 3. Dendrogram from the three location-matched pairs (as indicated by the locations) along with the location-unmatched ratios of random control and random infarct that was used for the internal validation. Log-
transformed ratios (i.e. infarct/control) of each protein (row) were presented for all conditions (column) Pearson correlation was applied for measurement of row and column distance. Globally normalized view was presented here. The color scale of the heat map ranges from saturated blue (value, -3.8) to saturated red (value, 4.1) in the natural logarithmic scale. The proteins were mainly clustered into two parts as shown by A and B. The columns were clustered into two sections, where the control (iTRAQ label: 113) was separated from location-matched and -unmatched infarct samples (C). The protein names and accession numbers were taken from the Uniprot protein database. The gene symbols are provided within brackets along with the protein name, wherever available.

In the category of molecular function, calcium-dependent phospholipid binding, lipase or phospholipase inhibitor activity (ANXA2, ANXA1, and ANXA5) and eukaryotic cell surface binding (ATP5B, ATP5A1, and ANXA5) were featured as significantly enriched. The enrichment analysis with the KEGG pathway analysis tool showed TCA cycle (MDH2, PDHB, DLAT, and CS) and OxPhos (ATP5B, ATP5A1, ATP6V1H, UQCRC2, ATP6V0A1) as the two significantly enriched pathways which are consistent with the enrichment of mitochondrial proteins.

The hierarchical clustering analysis showed a clear separation of the random control specimen (ID 9825, label: 113) from the infarct samples (Figure 3C) thus justifying the utility of multiple comparisons to exclude the noise (proteins unrelated to the pathology of interest). Overall, the proteins were grouped in two main clusters (A: 19 proteins, 57.6% and B: 14 proteins, 42.4%) consisting of predominantly down-regulated (i.e. A) and up-regulated proteins (i.e. B) from the infarcts. Intriguingly, all mitochondria associated proteins were grouped in the cluster A (Figure 3) and were related to energy metabolism (ATP synthesis). The enriched pathways (e.g. TCA cycle and OxPhos) were also part of this cluster. Concordant down-regulation of all mitochondrial proteins in the infarcts irrespective of their location when compared to the control samples may be a consequence of loss of neurons and electrophysiological activities in the infarct regions. This may also indicate reduced energy
requirements by the non-functional cells and an increased degradation or a decreased synthesis of the proteins in the infarct. Notably, the generalized failure of the mitochondrial machinery is also observed in chronic neurological disorders like AD and PD [30, 31]. The enriched functional category of calcium-dependent phospholipid binding proteins comprising of the annexin family (ANXA1, ANXA2 and ANXA5) were concurrently up-regulated with GFAP and VIM. This response is also partly observed in our previous studies involving the pre-clinical models of cerebral ischemia [5, 6].

Despite being parts of the same individual’s brain (ID 1006), the overall regulation pattern of infarcts from thalamus (label: 118) and parietal lobe (label: 119) showed a distinct spatial pattern of differential expression (e.g. SYN1, SIRT2) in comparison to the respective controls (Figure 2, 3). This could be related to the variability in the tissue composition, site-specific differences in susceptibility [22] or severity of the lesion or co-morbidity. Finally, guided by the above trends, the selected list of commonly regulated proteins were manually classified and curated to understand their regulation in context of the pathology of ischemic stroke (Table 1).

**DISCUSSION**

Proteomic characterization of post-mortem samples has been used widely in the recent past to generate valuable bedside data from subjects of different neurodegenerative or psychiatric disorders (e.g. Alzheimer’s, Parkinson’s disease, and Schizophrenia) [28, 32]. In contrast, the area of clinical proteomics of ischemic stroke is lagging behind as few proteomic studies used human brain tissue [14, 15]. Here, we functionally analyzed the deregulated consensus proteome of human brain infarcts obtained by comparing three pairs of infarcts and matched controls from various brain locations.
Down-regulation of the Malate-Aspartate Shuttle (MAS) and PDC

Our data indicates for the first time the deregulation of the MAS as seen by the concomitant down-regulation of all participating proteins (SLC25A11, SLC25A12, GOT2 and MDH2) in the human brain mitochondria of infarcted specimens (Table 1, Figure 4). The crucial role of the MAS has been described in the heart development during fetal and perinatal growth, in the physiology of pancreatic β-cells by influencing glucose-induced activation of mitochondrial metabolism and insulin secretion apart from involvement in pathological brain conditions like autism [33].

**Figure 4.** Schematic diagram showing the complex interplay of glycolysis, TCA cycle, MAS and OxPhos. Blue arrows designate glycolysis. Green and brown arrows are for TCA cycle and MAS respectively. The product of glycolysis, pyruvate is generated in the cytosol and enters the mitochondria and is converted to acetyl Co-A by enzymes of PDC. Mitochondrial MDH (MDH2) is an enzyme used by both TCA cycle and MAS. OGC and AGC1 are two transporters present on the mitochondrial membrane. The proteins in red color (OGC, AGC1, AATM, MDH, PDC and CS) are quantified by the iTRAQ experiment with acceptable levels of confidence (p-
value < 0.05). All of them were down-regulated as well as the proteins of OxPhos. Aspartate is generated in the neuronal mitochondria and transported through MAS to the cytosol. 1,3- BPG, 1,3-bisphosphoglycerate; LDH, lactate dehydrogenase; IMM, inner mitochondrial membrane; OMM, outer mitochondrial membrane.

SLC25A12 is a brain-specific aspartate/glutamate carrier (AGC1 or aralar) expressed predominantly in neurons. It is an electrogenic Ca\(^{2+}\)-dependent and unidirectional carrier that transports a protonated glutamate into the mitochondria in exchange of an aspartate anion, driven by the mitochondrial membrane electrochemical gradient (Figure 4). AGC1 requires the presence of mitochondrial aspartate aminotransferase (GOT2 or AATM) to generate aspartate from oxaloacetate (OAA) that is coupled with the generation of α-ketoglutarate (α-KG) from glutamic acid (protonated glutamate). Glutamate dehydrogenase, although can generate α-KG using glutamate, doesn’t support AGC1 because it cannot supplement aspartate. SLC25A11 (oxoglutarate carrier, OGC), which localizes to the inner mitochondrial membrane (IMM) like AGC1, mediates the electroneutral exchange of α-KG for malate. This malate is converted to OAA by mitochondrial malate dehydrogenase (MDH2) to provide the substrate for AATM. In this way, AGC1 and AATM work in concert with OGC to complete the MAS that drives the reduction equivalents (NADH) to the polarized mitochondria in an irreversible manner under physiological condition to sustain cytosolic glycolysis, the mitochondrial TCA cycle and OxPhos (Figure 4)[34]. Hence, cytosolic and mitochondrial metabolism will fail to synchronize due to aconcomitant down-regulation of these two trans-mitochondrial carriers (e.g. AGC1 and OGC) in the infarct.

Further, translocation of protoporphyrin IX and coproporphyrin III through the down-regulated OGC constitutes an important mechanism for the synthesis of heme within the mitochondrial matrix, where it acts as a precursor for several heme proteins including the cytochromes (Figure 4)[35]. This is consistent with the significant down-regulation of UQCRC2 (mitochondrial Cytochrome b-c1 complex subunit) in our data set (Figure 2).
Mitochondria being the primary site for oxygen consumption are liable to oxidative stress due to the generation of reactive oxygen species. Oxidative stress can inactivate large number of enzymes including dehydrogenases (e.g. NDUFS1) and transport ATPases due to the oxidation of the critical sulfhydryl groups that must be maintained in the reduced form for proper functioning. Glutathione (GSH), a low molecular weight thiol supplies the vital sulfhydryl group and is indispensable for upholding of the redox status of mitochondria. Depletion of intracellular GSH has been associated with increased infarction in the MCAO model of ischemic stroke [36, 37]. Studies have established OGC as one of the major transporters of GSH from the cytosol into the mitochondria (Figure 4) [38]. Thus, decreased expression of NDUFS1 (an enzyme having NADH dehydrogenase activity) in our data set may be related to the alteration of the mitochondrial GSH level that corroborates with the down-regulation of OGC.

Our data set also reveals two components (i.e. PDHB and DLAT) of the mitochondrial PDC, both of which were down-regulated (Table 1, Figure 4). PDC catalyzes the oxidative decarboxylation of pyruvate to form acetyl Co-A and NADH using one molecule of NAD⁺, linking glycolysis to the TCA cycle and fatty acid synthesis. A tight regulation of PDC activity is crucial under stress and its inactivation has been suggested as an adaptive response under conditions of glucose deprivation[39]. PDC deficiency due to mutations of PDHB or DLAT has been associated with lactic acidosis and neurological dysfunctions like ataxia and hypotonia [40]. This indicates that the presence of similar pathological (i.e. lactic acidosis) or neurological (i.e. ataxia) features in stroke patients may partially be related to PDC deficiency [41].

**Increased Level of Ferritin – Iron-mediated Neurotoxicity**
Ferritin is the major iron storage protein in healthy human brain and can sequester >4000 iron atoms. The heavy ferritin chain (FTH1) has antioxidant functions as well and can convert the highly toxic ferrous (Fe\(^{+2}\)) molecule to the less reactive ferric (Fe\(^{+3}\)) form. A pathological increase of brain iron and ferritin following stroke has been documented in several animal and clinical studies [42].

The localization of increased ferritin (light and heavy subunits) in the infarcts could be intracellular (i.e. brain parenchyma) or extracellular that may come from vascular leakage or released from brain cells or accumulated as a consequence of tissue disruption. An increased presence of intracellular free iron may lead to an adaptive increase in ferritin levels in the infarcts. Accordingly, it may have a protective influence against oxidative stress by trapping the pro-oxidant iron that could play a deleterious role in the mitochondrial failure as discussed above. Being infarcts, the presence of microglia or resident macrophages are possible in the infarcted specimens. Macrophages can actively secrete ferritin. An elevation of extracellular ferritin can be neurotoxic and cause microglia and macrophage activation [43]. Recently, VIM is found to control microglia activation and neurotoxicity during cerebral ischemia [44]. Thus, significant increase in VIM level (Table 1, Figure 2) in the infarcts may contribute to an indirect increase in the level of ferritin through microglial involvement. Conversely, extracellular ferritin can also stimulate the formation of new oligodendrocytes during the recovery phase following brain injury [43]. Oligodendrocytes are the primary iron-containing cells in the CNS while astrocytes generally have low ferritin and iron levels. The increased presence of oligodendrocyte-specific MBP in the infarcts from putamen and parietal lobe may be related to this phenomenon (Table 1, Figure 2). Thus, whether this increase in iron storage protein is cytoprotective or cytotoxic is not clear from our study. Immunohistochemical and imaging studies on post-mortem and live patients respectively could answer some of the raised issues.
Tissue-based Proteomic Profiling – An Alternative Approach to Identify Circulatory Biomarkers

The infarct being the core of injury is the area of tissue having the maximum magnitude of change for most of the proteins during an ischemic stroke event. Hence, one of the primary objectives behind identifying the consensus proteome from the infarcts of various locations was to look for up-regulated candidates that had the possibility of leaking into the circulation during the acute phase of the injury through a disrupted blood brain barrier. Accordingly, proteins with consistent over-abundance (e.g. ferritin, histone H4, ANXA1) in the proteomic data set was searched systematically by combining with keywords such as ‘stroke biomarker’, ‘cerebrovascular disorder biomarker’. Consistent with our hypothesis, ferritin has been reported as a circulatory biomarker of ischemic stroke by an independent study [45]. Hence, ANXA1 and histone H4 that were not reported previously as biomarkers in the context of cerebrovascular disease could be new candidates of potential interest for future tissue or blood-based biomarker studies. The increased presence of a nuclear protein (i.e. histone H4), despite the general down-regulation of energy metabolism, may point towards underlying epigenetic changes ongoing in the infarcts.

LIMITATIONS

These patients are a surviving population following a non-fatal attack of ischemic stroke, thus representing cases of chronic ischemia. Hence, the duration between the onset of stroke symptoms and death is variable and not controlled. The patency of the arteries in the infarcted region was also not known at the time of death [15]. In addition, the small sample size and advanced age group (average age > 80 y) of the subjects could limit the general applicability of the data generated from this study. Higher PMI for some of the subjects (Supplemental
Table 1) may influence the pattern or magnitude of protein expression. The quality of post-mortem human brain based on mRNA stability and pH has been found to be at least at par with the surgical biopsy tissue having a freezing delay of maximum 1 h [46]. This indicates that the variability in PMI may not be a critically limiting factor especially when proteins are analyzed. Typical PMI has been reported to range between 4-36 h among studies utilizing autopsied human brain [47]. Nonetheless, most of these limitations are common to similar studies done on human autopsied tissues due to logistics, technical or availability related constraints [15, 47]. Hence, further studies using larger cohort of patients from various age-groups are recommended to confirm the perturbed pathways/proteins found in this small cohort of female Japanese patients.

CONCLUDING REMARKS
This study reports the deregulated infarct proteome perturbed irrespective of the location using an iTRAQ-2D-LC-MS/MS based quantitative proteomic profiling approach. The failure of mitochondrial energy metabolism is apparent through the down-regulation of the MAS/PDC-complex, enzymes of glycolysis and OxPhos that is most likely caused the death of brain cells and loss of electrophysiological activities of the neurons in the infarct. The significant elevation of ferritin may be related to the aberrations in iron metabolism that could explain the free radical-mediated toxicity and subsequent mitochondrial failure. Collectively, we provide ample evidences about the dynamic changes happening even in the chronic infarcts of aged human brain. Nevertheless, the identification of an opposite regulation of candidate proteins within and between individuals exposes the inherent complexity of the infarct pathophysiology. In conclusion, our study could stimulate further clinical proteomic studies involving post-mortem stroke specimens apart from the follow-up pre-clinical studies on targeted proteins.
ASSOCIATED CONTENT

Supporting Information

Table 1. Detailed Clinical and Neuropathological Information of the Studied Subjects

Table 2. Complete information of the full list of the qualified proteins (Unused prot score >2) obtained from the bias and background corrected iTRAQ data set.

Table 3. Quantitative Information of the Significantly Regulated Proteins for the Three Location-matched Pairs along with the Internal Validation Using One Random Control and One Random Infarcted Specimen

Table 4. SP_PIR_KEYWORDS Analysis of the Final List of Selected Proteins by DAVID

Figure 1. Representative photographs of the autopsied brain sections of five Japanese subjects with their respective IDs.

Figure 2. Distribution of the identified proteins from the bias- and background-corrected iTRAQ data set (unused prot score >2) according to their molecular weight (Mol weight, kDa) and theoretical isoelectric point (pI).

ACKNOWLEDGEMENT

We thank Mr. T. Kanesaka, Mr. N. Ogawa, Mr. Y. Tanian and Ms. C. Taniguchi (all at the Choju Medical Institute, Fukushima Hospital) for technical assistance, patient care, sampling and tissue data acquisition. This research is supported by the Singapore National Research Foundation under its NMRC-CBRG and administered by the Singapore Ministry of Health’s National Medical Research Council (NMRC/CBRG/0004/2012).

CONFLICT OF INTEREST STATEMENT
The authors have declared no conflict of interest.
REFERENCES


### Table 1. Final List of Selected Regulated Proteins from the Location-matched Infarcted Human Brain Samples along with Quantitative Expression Data from the Random Control and Infarct Samples

<table>
<thead>
<tr>
<th>N</th>
<th>Unused</th>
<th>%Cov(95)</th>
<th>Accession</th>
<th>Protein name</th>
<th>Peptides (95%)</th>
<th>Location-matched pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>117:114 (Putamen)</td>
</tr>
<tr>
<td><strong>Energy Metabolism</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Glycolysis</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>51.7</td>
<td>32.8</td>
<td>Q59FD4</td>
<td>Hexokinase 1 isoform HKI variant (Fragment)</td>
<td>43</td>
<td>0.43</td>
</tr>
<tr>
<td><strong>Pyruvate dehydrogenase complex (PDC)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>196</td>
<td>15.2</td>
<td>17.3</td>
<td>Q69Y15</td>
<td>Dihydrolipoamide S-acetyltransferase (DLAT)</td>
<td>9</td>
<td>0.64</td>
</tr>
<tr>
<td>229</td>
<td>13.3</td>
<td>27.6</td>
<td>P11177</td>
<td>Pyruvate dehydrogenase E1 component subunit beta, mitochondrial (PDHB)</td>
<td>10</td>
<td>0.66</td>
</tr>
<tr>
<td><strong>Tricarboxylic acid cycle (TCA) cycle</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>36</td>
<td>39.8</td>
<td>42.2</td>
<td>O75746</td>
<td>Calcium-binding mitochondrial carrier protein Aralar1 (SLC25A12)</td>
<td>21</td>
<td>0.49</td>
</tr>
<tr>
<td>45</td>
<td>33.9</td>
<td>59.2</td>
<td>Q69HZ0</td>
<td>Malate dehydrogenase (MDH2)</td>
<td>40</td>
<td>0.52</td>
</tr>
<tr>
<td>48</td>
<td>33.4</td>
<td>43.3</td>
<td>P00505</td>
<td>Aspartate aminotransferase, mitochondrial (GOT2)</td>
<td>28</td>
<td>0.35</td>
</tr>
<tr>
<td>86</td>
<td>24.7</td>
<td>24.2</td>
<td>A8K132</td>
<td>cDNA FLJ75476, highly similar to Homo sapiens glutaminase (GLS), mRNA</td>
<td>13</td>
<td>0.64</td>
</tr>
<tr>
<td>118</td>
<td>20.6</td>
<td>41.4</td>
<td>Q6BH0</td>
<td>SLC25A11 protein (SLC25A11)</td>
<td>13</td>
<td>0.53</td>
</tr>
<tr>
<td>105</td>
<td>22.2</td>
<td>32.6</td>
<td>O75390</td>
<td>Citrate synthase, mitochondrial (CS)</td>
<td>15</td>
<td>0.49</td>
</tr>
<tr>
<td><strong>Oxidative phosphorylation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>58.7</td>
<td>62.2</td>
<td>P06576</td>
<td>ATP synthase subunit beta, mitochondrial (ATP5B)</td>
<td>71</td>
<td>0.51</td>
</tr>
<tr>
<td>31</td>
<td>43.6</td>
<td>42.7</td>
<td>P25705</td>
<td>ATP synthase subunit alpha, mitochondrial (ATP5A1)</td>
<td>54</td>
<td>0.44</td>
</tr>
<tr>
<td>61</td>
<td>29.1</td>
<td>31.9</td>
<td>B4DJA0</td>
<td>NADH-ubiquinone oxidoreductase 75 kDa subunit (NDUFS1)</td>
<td>16</td>
<td>0.63</td>
</tr>
<tr>
<td>101</td>
<td>22.6</td>
<td>38.6</td>
<td>P22695</td>
<td>Cytochrome b-c1 complex subunit 2, mitochondrial (UQCRCC2)</td>
<td>17</td>
<td>0.65</td>
</tr>
<tr>
<td><strong>Reactive Gliosis</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>97.1</td>
<td>67.6</td>
<td>P14136</td>
<td>Glial fibrillary acidic protein (GFAP)</td>
<td>146</td>
<td>1.06</td>
</tr>
<tr>
<td>30</td>
<td>44.5</td>
<td>41.6</td>
<td>P08670</td>
<td>Vimentin (VIM)</td>
<td>32</td>
<td>1.09</td>
</tr>
<tr>
<td><strong>Anti-inflammatory Proteins</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>128</td>
<td>19.9</td>
<td>40.7</td>
<td>Q5TZZ9</td>
<td>ANXA1 protein (ANXA1)</td>
<td>12</td>
<td>2.31</td>
</tr>
<tr>
<td>145</td>
<td>18.2</td>
<td>32.4</td>
<td>P07355</td>
<td>Annexin A2 (ANXA2)</td>
<td>10</td>
<td>1.16</td>
</tr>
<tr>
<td>88</td>
<td>24.3</td>
<td>51.9</td>
<td>P08758</td>
<td>Annexin A5 (ANXA5)</td>
<td>19</td>
<td>1.45</td>
</tr>
<tr>
<td><strong>Chaperonic Proteins</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>136</td>
<td>18.8</td>
<td>18.6</td>
<td>P38646</td>
<td>Heat shock 70 kDa protein 9 (mt-HSP70)</td>
<td>13</td>
<td>0.55</td>
</tr>
<tr>
<td>185</td>
<td>16.0</td>
<td>25.8</td>
<td>P54652</td>
<td>Heat shock-related 70 kDa protein 2 (HSPA2)</td>
<td>20</td>
<td>1.94</td>
</tr>
<tr>
<td>Iron Storage Proteins</td>
<td>Others</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------------------------------</td>
<td>--------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ferritin (FTL)</td>
<td>Myelin basic protein (MBP)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ferritin (Fragment) (FTH1)</td>
<td>V-type proton ATPase subunit a isoform 1 (ATP6V0A1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carboxyl reductase [NADPH] 1 (CBR1)</td>
<td>Neurofascin (NFASC)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Myelin basic protein (MBP)</td>
<td>4-aminobutyrate aminotransferase, mitochondrial (ABAT)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V-type proton ATPase subunit H (ATP6V1H)</td>
<td>Histone H4 (HIST1H4H)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V-type proton ATPase subunit H (ATP6V1H)</td>
<td>Histone H4 (HIST1H4H)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carboxyl reductase [NADPH] 1 (CBR1)</td>
<td>Neurofascin (NFASC)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Myelin basic protein (MBP)</td>
<td>V-type proton ATPase subunit H (ATP6V1H)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>voltage-dependent anion-selective channel protein 2 (VDAC2)</td>
<td>ATPase subunit H (ATP6V1H)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phosphoserine aminotransferase</td>
<td>Phosphoserine aminotransferase</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Selenium-binding protein 1 (SELENBP1)</td>
<td>Phosphoserine aminotransferase</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**The final list of proteins along with the quantitative values that was used for the cluster analysis. The p-value column is not shown. The ratios with bigger font (in bold) have significant p-value (<0.05). The proteins have been classified based on their respective primary function. The ‘energy metabolism’ group is sub-divided into three groups. The proteins of energy metabolism were reduced, whereas proteins related to reactive gliosis and anti-inflammatory response were elevated. Proteins participating in iron storage were also increased. Chaperonic and other proteins exhibited a mixed trend.**