



Spatial mapping of protein abundances in the mouse brain by voxelation integrated with high-throughput liquid chromatography–mass spectrometry

Vladislav A. Petyuk, Wei-Jun Qian, Mark H. Chin, et al.

Genome Res. 2007 17: 328-336 originally published online January 25, 2007
Access the most recent version at doi:[10.1101/gr.5799207](https://doi.org/10.1101/gr.5799207)

References This article cites 47 articles, 10 of which can be accessed free at:
<http://genome.cshlp.org/content/17/3/328.full.html#ref-list-1>

License

Email Alerting Service Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

To subscribe to *Genome Research* go to:
<https://genome.cshlp.org/subscriptions>

Copyright © 2007, Cold Spring Harbor Laboratory Press

Methods

Spatial mapping of protein abundances in the mouse brain by voxelation integrated with high-throughput liquid chromatography–mass spectrometry

Vladislav A. Petyuk,^{1,4} Wei-Jun Qian,^{1,4} Mark H. Chin,^{2,3,4} Haixing Wang,¹ Eric A. Livesay,¹ Matthew E. Monroe,¹ Joshua N. Adkins,¹ Navdeep Jaitly,¹ David J. Anderson,¹ David G. Camp II,¹ Desmond J. Smith,² and Richard D. Smith^{1,5}

¹Biological Sciences Division and Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, Washington 99352, USA; ²Department of Molecular and Medical Pharmacology, David Geffen School of Medicine at UCLA, Los Angeles, California 90095, USA; ³Department of Human Genetics, David Geffen School of Medicine at UCLA, Los Angeles, California 90095, USA

Temporally and spatially resolved mapping of protein abundance patterns within the mammalian brain is of significant interest for understanding brain function and molecular etiologies of neurodegenerative diseases; however, such imaging efforts have been greatly challenged by complexity of the proteome, throughput and sensitivity of applied analytical methodologies, and accurate quantitation of protein abundances across the brain. Here, we describe a methodology for comprehensive spatial proteome mapping that addresses these challenges by employing voxelation integrated with automated microscale sample processing, high-throughput liquid chromatography (LC) system coupled with high-resolution Fourier transform ion cyclotron resonance (FTICR) mass spectrometer, and a “universal” stable isotope labeled reference sample approach for robust quantitation. We applied this methodology as a proof-of-concept trial for the analysis of protein distribution within a single coronal slice of a C57BL/6J mouse brain. For relative quantitation of the protein abundances across the slice, an ¹⁸O-isotopically labeled reference sample, derived from a whole control coronal slice from another mouse, was spiked into each voxel sample, and stable isotopic intensity ratios were used to obtain measures of relative protein abundances. In total, we generated maps of protein abundance patterns for 1028 proteins. The significant agreement of the protein distributions with previously reported data supports the validity of this methodology, which opens new opportunities for studying the spatial brain proteome and its dynamics during the course of disease progression and other important biological and associated health aspects in a discovery-driven fashion.

[Supplemental material is available online at www.genome.org.]

The post-genome sequencing era has brought a new discovery-driven paradigm to life sciences, enabled by technological advances that allow for high-throughput data collection and analysis. Such high-throughput technologies have been widely applied to diverse areas of biological research including one of the most challenging areas—neuroscience (Boguski and Jones 2004), the study of the human nervous system, the brain, and the biological basis of consciousness, perception, memory, and learning. A key challenge in neuroscience derives from the molecular complexity of the brain; reportedly, about one-third of the mammalian genome is exclusively dedicated to brain function (Bantle and Hahn 1976; Chikaraishi 1979). Developing an understanding of how the brain functions is further challenged by the complicated 3D structure and anatomy of the brain.

Several on-going projects are responding to these challenges. For example, the National Institutes of Health (NIH) launched the Brain Molecular Anatomy Project to develop an understanding of gene expression and function in the nervous

system. Additionally, a few independent efforts are targeted to the comprehensive characterization of spatial gene expression in the mouse brain. The Allen Brain Atlas project is aimed at characterizing the spatial expression of all mRNAs in detail by using an automated high-throughput platform for in situ hybridization (<http://www.brainatlas.org>). The GENSAT (Gene Expression Nervous System Atlas) (<http://www.ncbi.nlm.nih.gov/projects/gensat/>) project is creating a gene expression atlas of the mouse central nervous system using transgenic mice in which coding sequences of individual genes have been replaced with enhanced green fluorescent protein (EGFP) (Gong et al. 2003). Despite higher levels of automation, comprehensive discovery-based analyses of multiple mouse brain models that represent different disease states and their progression remain a daunting task.

Undoubtedly, detailed spatial mapping of protein levels in the brain would allow better understanding of biological processes. Although the application of new proteomic technologies has greatly extended the set of proteins known to exist in the brain (Fountoulakis 2004; Wang et al. 2006), few insights as to localization have been provided. To date, reported proteomic imaging of the mouse brain has been based on high-throughput MALDI-TOF (Reyzer and Caprioli 2005) mass spectrometry (MS). The great value of this approach has been demonstrated for a

⁴These authors contributed equally to this work.

⁵Corresponding author.

E-mail rds@pnl.gov; fax (509) 376-7722.

Article published online before print. Article and publication date are at <http://www.genome.org/cgi/doi/10.1101/gr.5799207>.

number of important subjects such as diagnosis and prognosis of cancer development and studying of brain pathologies (Caprioli 2005; Reyzer and Caprioli 2005; Rohner et al. 2005). In spite of good spatial resolution, this approach provides only limited protein identifications; typically 200–400 distinct mass signals are observed, which presumably correspond to proteins in the low to middle molecular weight range. Identification of such mass signals represents a non-trivial laborious task.

Conceptually, liquid chromatography–mass spectrometry (LC-MS) based proteomics can be applied for comprehensive proteome imaging of the brain by high-throughput analyses of tissue samples isolated from different brain areas using techniques such as voxelation (Liu and Smith 2003; Singh and Smith 2003); however, such spatial imaging efforts have not been realized due to a number of challenges such as high number of samples, small sample sizes, robust quantitation approaches, and throughput of LC-MS analyses. Herein, we report a novel methodology for spatial mapping of the relative protein abundance patterns in the mouse brain with comprehensive proteome coverage employing a stable isotope labeled “universal” reference sample approach for accurate quantitation and automated microscale sample processing procedure integrated with voxelation and automated high-throughput LC-MS. Voxelation has been previously coupled with microarray analysis to spatially map brain gene expression patterns in both a mouse model of Parkinson’s disease (Brown et al. 2002b) and a post-mortem human brain tissue to study the effects of Alzheimer’s disease (Brown et al. 2002a). To address the challenge of accurate quantitation across a large number of biological samples, we designed a strategy of spiking a stable isotope labeled reference sample to each voxel sample.

In the present study, we demonstrated the quantitative mapping of the relative protein abundances by analyzing a coronal section of normal mouse brain that consisted of 71 voxels at the striatum level (bregma = 0 mm). The use of voxelation followed by automated microscale sample preparation and capillary liquid chromatography–Fourier transform ion cyclotron resonance (LC-FTICR) mass spectrometry resulted in quantitative images of abundance patterns for >1000 proteins. Proteins were identified by using the accurate mass and time (AMT) tag strategy in which accurate masses and elution times of tryptic peptides measured by LC-FTICR were matched to an AMT tag reference database (i.e., a “look-up” table) that contained peptides confidently identified from separate LC-MS/MS analyses (Qian et al. 2004; Qian et al. 2005b; Zimmer et al. 2006). For quantitative measurements of relative peptide abundances, isotopic $^{16}\text{O}/^{18}\text{O}$ labeling was applied where each voxel sample was spiked with an ^{18}O -labeled reference sample derived by tryptic digestion of proteins from an entire control coronal slice.

To validate the data, we compared the protein abundance images with existing data that included immunohistochemical markers (e.g., DARPP-32, MBP, and CaMKIIa), as well as with two extensive databases of mRNA abundance distributions: Allen Brain Atlas (<http://www.brainatlas.org>) and GENSAT (<http://www.ncbi.nlm.nih.gov/projects/gensat/>). Good agreement was observed with the latter databases, which lends further support to the presented methodology. The resulting protein abundance patterns included examples of proteins with previously uncharacterized distributions in the mouse brain in addition to proteins with known distributions.

The present study represents a step toward characterization of spatial abundance patterns of the brain proteome and provides

the methodological basis for future studies that are extended in scope to include coverage of major protein modification states and the use of faster separations (Shen et al. 2005; Tang et al. 2005) that provide the throughput levels needed to study a number of neurodegenerative disease mouse models. This study also successfully demonstrated a methodology based on stable isotope labeling that allows accurate quantitation across 71 biological samples, which should have broad applications in the area of biomarker discovery and clinical proteomics that require quantitative analysis of relative changes between many clinical samples.

Results

Overview of the methodology

We developed an integrated methodology for spatial proteome mapping of the mouse brain that includes tissue voxelation, automated microscale sample processing, the use of an ^{18}O -labeled reference sample, high-throughput LC-FTICR analysis, and the AMT tag strategy for peptide identification and quantitation (Qian et al. 2004; Qian et al. 2005b; Zimmer et al. 2006). An extensive mouse brain peptide/protein database was developed from an earlier global characterization of the mouse brain proteome that resulted in confident identification of 7792 different proteins from >48,000 unique peptides (Wang et al. 2006), which was based on LC-MS/MS analyses of both global tryptic and cysteinyl-peptide enriched fractions. In this work, only peptides identified from the global tryptic fraction were included in the reference database used to identify peptides/proteins in the present high-throughput LC-FTICR analyses of voxel samples. The reference database contains the calculated exact monoisotopic mass and observed normalized elution time (NET) for ~38,000 confidently identified peptides (i.e., AMT tags). To create a 2D map of relative protein abundances across a bregma = 0 mm coronal slice of a control C57BL/6J mouse, we voxelated the tissue slice into 1-mm³ cubes (Liu and Smith 2003) followed by automated sample processing (ultrasonic homogenization, protein denaturation, reduction of disulfide bonds, and digestion with trypsin; Fig. 1A).

To date, the isotopic labeling strategy was applied mainly for pairwise sample comparison. To have a quantitative comparison of the observed proteins across multiple different voxels, we used a “universal” stable isotope reference sample. This labeled reference sample was prepared from a whole coronal section of another normal mouse brain by tryptic digestion and ^{18}O -labeling via post-digestion trypsin-catalyzed oxygen exchange Qian et al. 2005b) (Fig. 1B). Equal amounts of peptide sample from the individual voxels were mixed with identical amounts of the ^{18}O -labeled reference sample prior to analysis by LC-FTICR. Peptides were confidently identified by matching detected mass and elution time features against the AMT tag database (Fig. 1C) within 2 ppm and 1% for mass accuracy and LC normalized elution time (NET) tolerances, respectively.

Both ^{16}O members of isotopically paired $^{16}\text{O}/^{18}\text{O}$ features (i.e., two coeluting features with a 4.0085-Da mass difference) and unpaired features (features not having a coeluting counterpart with 4.0085-Da mass difference) detected by the LC-FTICR were matched against the AMT tag database for peptide/protein identification. A total of 1028 proteins were identified, each with a minimum of two peptides detected (Supplemental Tables 1, 2, and 3). However, not all 1028 proteins were detected in all 71

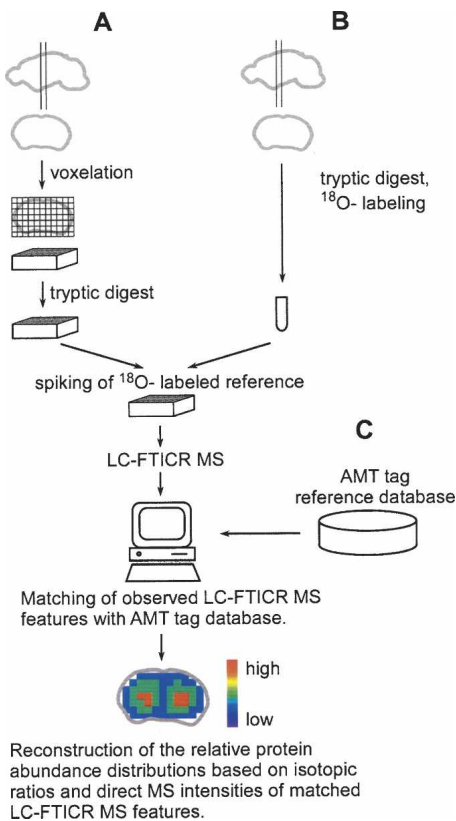


Figure 1. Strategy for spatial mapping of protein abundance patterns in the mouse brain using an example of a coronal slice. (A) Coronal section was further dissected into voxels (1-mm³ cubes) followed by tryptic digestion of each voxel sample. (B) The entire coronal slice was digested by trypsin followed by ¹⁸O-labeling and spiking into each voxel sample as a reference for robust quantitation. (C) The observed LC-FTICR features were matched against an AMT tag database containing theoretical masses and observed elution times of previously identified peptides from LC-MS/MS analyses.

voxels, which is expected for a number of reasons including the limited dynamic range of detection and the fact that some proteins may naturally have expression limited to only certain areas of the brain. Noticeably, nearly half of the identified proteins were detected in >65 out of 71 voxels (Supplemental Fig. 1). On average we identified ~600–700 proteins for a given voxel (Supplemental Fig. 2) and no single voxel exhibited the entire set of 1028 proteins. We did not observe any apparent bias in the number of protein identifications per voxel toward any brain structure or part of the brain (Supplemental Fig. 3). Also as expected, the overlap between the voxels in protein identifications is quite high and is typically 80%–90% (Supplemental Figs. 4, 5).

Isotopic ratio-based quantitative data were obtained for 838 proteins that had at least one peptide detected as an isotopically labeled pair. The relative peptide abundances in different voxels were assessed based on the ¹⁶O/¹⁸O ratios for detected peptide pairs, where natural ¹⁶O peptide abundances from different voxels were compared to the same spiked reference sample with ¹⁸O-labeled peptides.

The labeled reference sample allowed for isotopic pairing of most of the detected tryptic peptides to provide precise quantitation. Additionally, the LC-MS peak intensities of natural ¹⁶O peptides from all of the 1028 proteins were used to assess relative protein abundances across the voxels. As a result, 190 of the 1028

proteins were identified and quantified solely by ¹⁶O peptides, presumably because the ¹⁸O-labeled members of these peptides were below the detection limit. Our results showed that both quantitation approaches provide relatively good accuracy; however, the isotopic ratio-based approach is overall more precise than the label-free approach, as shown by the mean CV values of 7% and 15%, respectively (Supplemental Fig. 6). The observed agreement between the two types of independently derived isotopic ratio- and direct intensity-based patterns provided an additional level of confidence for the determined protein abundance distributions.

The confidence of peptide and protein identifications

To assess the percentage of false-positive identifications obtained with the AMT tag approach, i.e., the false discovery rate (FDR), we introduced a strategy analogous to the reversed sequence database approach used for LC-MS/MS identification of peptides (Peng et al. 2003; Qian et al. 2005a). We made a random AMT tag database of the same size as the normal database by shifting the masses of all peptides by 11 Da. In this case 11 Da was chosen as a randomly selected integer value which differs from any known common post-translational modification. Indeed the number of matches does not vary significantly as long as the mass shift value stays an integer (Supplemental Fig. 7). Features detected in all LC-FTICR datasets were matched against a merged normal and shifted database. We assumed that all matches to the shifted part of the merged database are false and should be at approximately the same level as false matches to the normal part. Indeed, the distribution of matches to the shifted part closely approximated the distribution of random false matches to the normal part (Fig. 2A). Thus, we used the ratio of the number of matches to the shifted part to the number of matches to the normal part as an estimate of FDR of our peptide and protein identifications (Fig. 2B; Table 1).

We attribute the observed significant difference in the FDR for paired and unpaired features to the fact that the majority of peptides were observed as paired species, which provides excellent confidence that the features were indeed peptides and specific to the samples, whereas unpaired features could include contaminants and non-peptidic species (e.g., metabolites). This finding highlights an additional benefit beyond robust quantitation that is afforded by spiking an isotopically labeled reference sample; that is, peptide pair detection provides increased confidence in identifications.

Patterns of protein abundances and their classification

We expected the majority of protein abundance patterns in the mouse brain to be bilaterally symmetrical due to the high degree of bilateral symmetry observed in mouse brain structures. Indeed, the protein abundance patterns showed a high degree of correlation with their mirror transformations (Supplemental Fig. 8). Moreover, the relationships between voxels within each of the hemispheres showed a high degree of similarity between left and right hemispheres (Supplemental Fig. 9), which further underlines the structural bilateral symmetry of the derived abundance patterns.

While each protein showed a characteristic abundance pattern, we nevertheless selected proteins having obvious correlation of abundance pattern with major brain structures: 29 proteins being predominantly abundant in the ventromedial area collocated with diencephalon, 59 in the cortex, 17 in the stria-

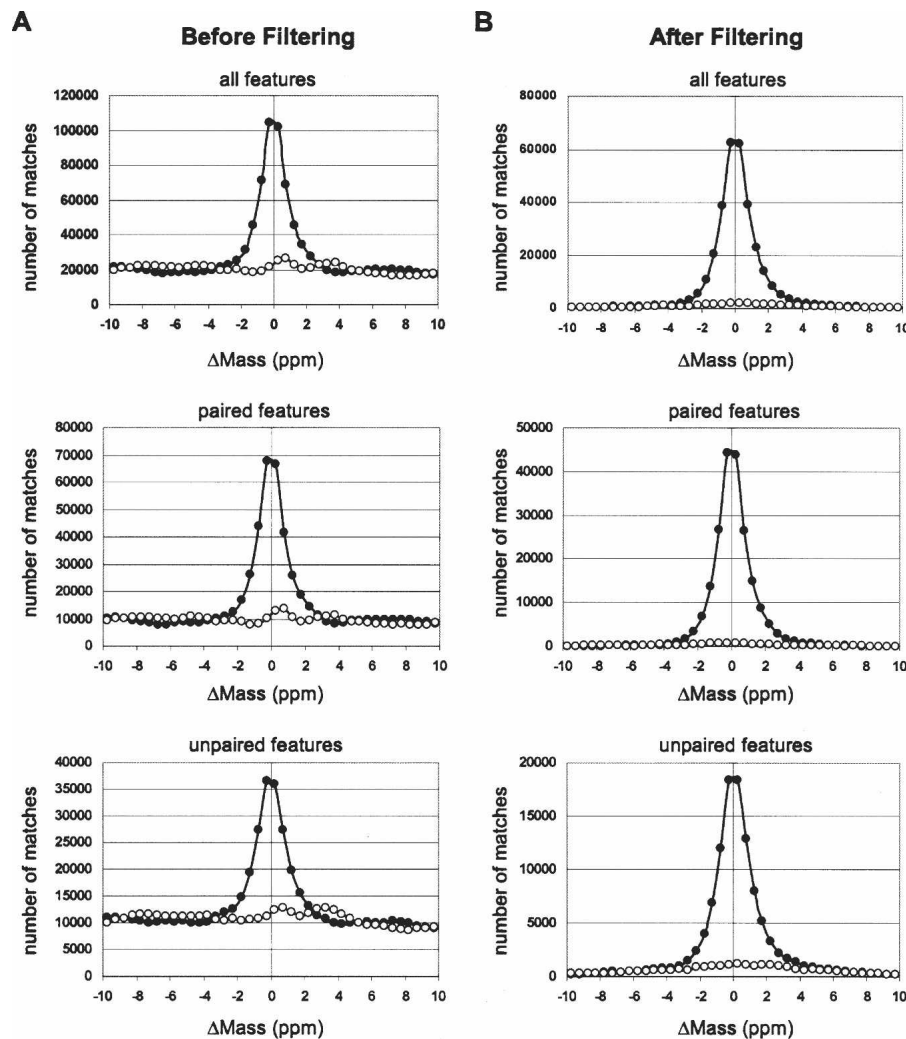


Figure 2. Mass error distribution histograms of feature-peptide matching to the merged AMT tag database, before (A) and after (B) applying filtering (step 1 in Table 1) except mass error tolerance cutoff. The number of matches combined from all 71 MS datasets. Solid circles represent matches to the normal AMT tag database. Open circles represent matches to the 11 Da shifted AMT tag database. The FDR rate after filtering (B) can be estimated as a ratio of area beneath the curve representing matches to the shifted database to the area beneath the curve representing matches to the normal database within respective mass error tolerances: ± 2.3 ppm for paired features and ± 2.0 ppm for unpaired features.

tum, and 28 in the central part of the section collocating with white matter structures like corpus callosum, fornix, and anterior commissure. In addition we observed 150 proteins without any noticeable bias in abundance across the entire coronal slice, thus representing a clearly uniform pattern.

Figure 3 shows selected examples of proteins expected to be abundant in particular areas of the brain based on mRNA abundances (Allen Brain Atlas project [<http://www.brainatlas.org>]). Syntaxin 1a (a histological molecular marker of neuronal synapses) and vesicular glutamate transporter type 1 (VGLUT1) encoded by the *Stx1a* and *Slc17a7* genes are expected to be expressed predominantly in the cortex (Honer et al. 1997; Fremeau et al. 2004). Although we observed these proteins almost across the entire slice, the greatest abundance was clearly in the cortex. The ProSAAS protein and GABA-4 transporter (encoded by genes *Pcsk1n* and *Slc6a11*), involved in neuropeptide hormone process-

ing and cleaning of the postsynaptic cleft of GABA neurotransmitter, respectively (Jursky and Nelson 1996; Feng et al. 2001), were detected predominantly in the ventromedial area of the section collocating with diencephalon. Guanine nucleotide binding protein (G-protein) gamma 7 subunit and dopamine- and cAMP-regulated phosphoprotein (DARPP-32) encoded by *Gng7* and *Ppp1r1b* genes, respectively, were predominantly detected in the striatal region, which is consistent with their known expression (Foster et al. 1987; Schwindinger et al. 2003). In fact DARPP-32 is a widely used immunohistochemical marker of medium-sized striatal neurons. Myelin-basic protein (MBP) is a major component of myelin nerve ensheathment and predominantly expressed in the white matter represented by corpus callosum, fornix, and anterior commissure and located in the central part of this coronal section. The cyclic nucleotide phosphodiesterase (CNP) encoded by *Cnp1* is another white matter oligodendrocyte-specific protein important for maintaining axoglial integrity (Rasband et al. 2005) and colocalized well with myelin basic protein. Ubiquitously expressed cytoplasmic actin (*Actg1*) and mitochondrial NADH-ubiquinone oxidoreductase (*Ndufs3*) are components of cytoskeleton and the first enzyme in oxidative phosphorylation, respectively; both proteins displayed relatively uniform patterns, as expected. We also observed proteins that displayed combinations of the above patterns, as well as other apparently unique patterns (Fig. 4A).

Comparison of mRNA and protein abundance patterns

Our confidence in these protein abundance patterns is further reinforced by their correspondence to known patterns from the GENSAT and Allen Brain Atlas projects (Fig. 3). Despite the significant methodological differences, we found a substantial agreement between mRNA and protein abundance patterns. That is, a signifi-

Table 1. Estimated FDRs (%) for peptide and protein identifications

	Paired	Unpaired	Overall
Step 1, ^a peptides	2.7	8.6	4.5
Step 2, ^b peptides	1.9	6.4	3.2
Step 2, ^b proteins			2.6

^aStep 1 data filtering criteria are based on mass and elution time differences between observed features and matched peptides from the reference database.

^bA Step 2 filtering criterion is based on the number of peptide hits per protein. Peptides from proteins having only one hit were discarded.

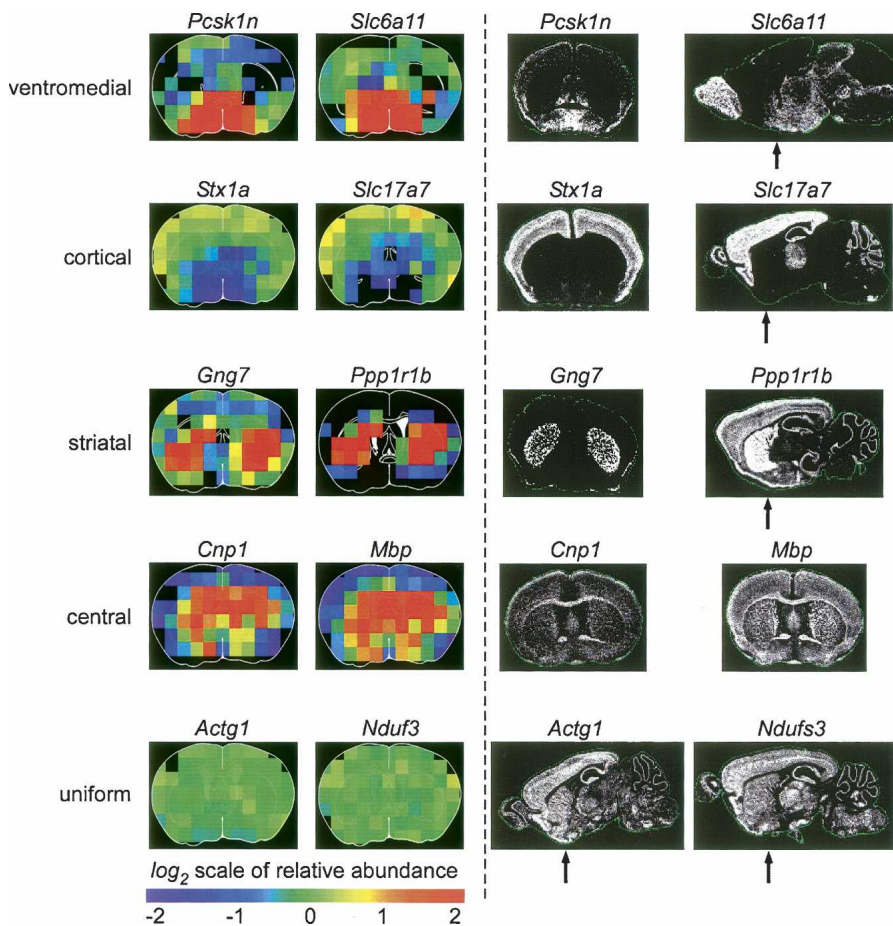


Figure 3. Comparison of protein and mRNA distribution for selected genes with characteristic abundance patterns: ventromedial, cortical, striatal, central, and uniform. The determined protein abundance patterns are on the left side. Relative protein abundance values are encoded by color. Black regions inside the contour of the coronal slice indicate that the protein was not detected for the given voxels. On the right side, the mRNA abundance data provided under license by The Allen Institute for Brain Science. All rights reserved. Sagittal sections were used if the coronal section data were not available. The arrows on the sagittal sections indicate the corresponding position of bregma 0 mm coronal section.

cant portion of the proteins having increased abundance in a certain region also have increased expression levels of the corresponding mRNA in that particular region relative to the rest of the coronal slice. For example, 68% of the genes encoding protein products predominantly abundant in the ventromedial area, 82% in the cortex, 100% in the striatum, and 52% in the central area also revealed relatively increased mRNA abundances in the corresponding areas (Supplemental Table 1). We observed less overall agreement with mRNA abundances in the central cluster area due to the striking and reproducible discrepancy among the abundance distributions of the histone mRNAs and their protein products (genes: *H1f0*, *Hist1h1c*, *Hist1h2bb*, *H2afx*, and *Hist2h3c2*) (Fig. 4B). On the other side, we observed 100% correlation of protein with mRNA abundances for all glia-specific genes abundant in the central area such as *Mbp*, *Cnp1*, *Mog*, *Mobp*, *Plp1*, and *Gfap*.

Overall, the agreement observed between mRNA and protein spatial abundance patterns was significant, but somewhat surprising given the previous reports that have highlighted the disagreement of protein abundances measured by MS with mRNA expressions derived from microarray studies (Griffin et al.

2002; Washburn et al. 2003). However, discrepancies between mRNA and protein abundance patterns are not unexpected and may be attributed to factors such as mRNA stability, differential mRNA translation efficiencies, and differential protein stability in different parts of the brain. Additional factors may be related to protein post-translational modifications (PTMs). The extent of the PTMs may affect the levels of the unmodified peptides used for protein quantitation; for example, histone proteins are likely candidates for having PTMs that vary among some types of brain tissues and cells.

Observation of previously uncharacterized genes

Of interest is the fraction of genes in our proteomic database for which spatial abundance patterns have not yet been characterized at the protein level or at the mRNA level. For example, we found the protein corresponding to the gene *Atp2b2* encoding Ca^{2+} transporting ATPase to be highly abundant in the cortex (Fig. 4C). It was reported that mutation of the gene *Atp2b2* causes slower growth, balance problems, and hearing loss in mice (<http://www.informatics.jax.org/>). Mice having homozygous mutations of the gene *Pde1b* encoding Ca^{2+} calmodulin-dependent phosphodiesterase are reportedly hyperactive with increased exploratory behavior and learning deficits (<http://www.informatics.jax.org/>). This gene was found to be predominantly expressed in the striatum. Additionally, gene *Ass1*, encoding argininosuccinate synthetase 1, involved

in production of nitric oxide by recycling citrulline (Husson et al. 2003; <http://www.informatics.jax.org/>), was found only in the ventromedial area. Another example of a protein with newly characterized distribution in the brain is Sirtuin 2 deacetylase, encoded by *Sirt2* gene, a homolog of yeast *Sir2* and mammalian *Sirt1* genes, which regulate chromatin silencing and cellular stress response and presumably are involved in organismal life span regulation (Bordone and Guarente 2005). This protein was found to be predominantly abundant in the central part of the coronal section.

Gene ontology and KEGG pathway analysis

Unlike microarray measurements, where an expression level value can be obtained for almost every single gene, in LC-MS-based proteomics the number of detected proteins is limited by the dynamic range of detection. Most likely, the 1028 detected proteins are among the more abundant proteins of whole brain proteome. To provide an overview of the biological processes and pathways accessed by the presented approach, we applied GO (Beissbarth and Speed 2004; <http://gostat.wehi.edu.au>) and KEGG (Kanehisa and Goto 2000; Al-Shahrour et al. 2006; [332 Genome Research
www.genome.org](http://</p>
</div>
<div data-bbox=)

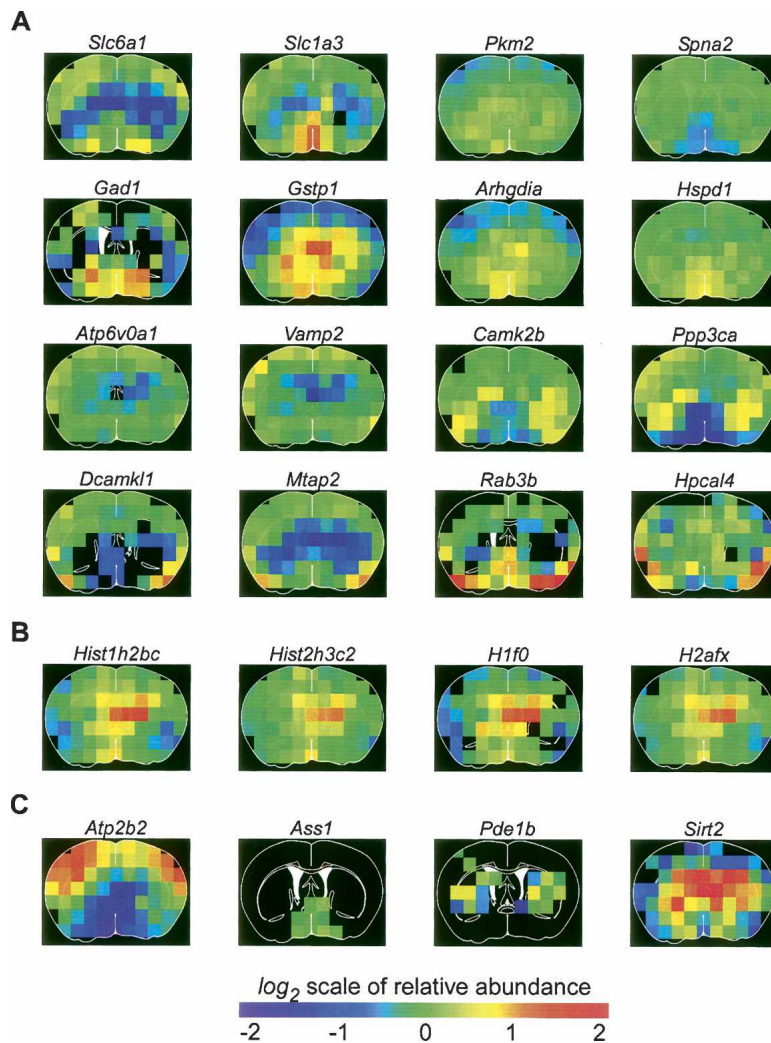


Figure 4. (A) Examples of proteins having complex abundance patterns. (B) Histone proteins classified with predominant abundance in the center have no correlation with mRNA abundance data. (C) Selected examples of protein abundance patterns of genes with previously uncharacterized expression patterns in the mouse brain. Relative protein abundance values are encoded by color. Black regions inside the contour of the coronal slice indicate that the protein was not detected for the given voxels.

babelomics.bioinfo.cipf.es) classification to analyze the entire list of 1028 proteins. As expected we found a significant amount of high abundant proteins, e.g., mitochondrial proteins (GO:0005739), proteins involved into the generation of precursor metabolites and energy (GO:0006091), cytoskeletal proteins (GO:0005856), and other structural and housekeeping proteins. However, the detected protein set is noticeably under-represented by presumably low abundant proteins such as transcription factors (GO:0006350) and transmembrane receptors (GO:0004888). The top three KEGG pathways detected are oxidative phosphorylation, regulation of actin cytoskeleton, and tight junction with 62, 38, and 31 observed proteins, respectively. In addition to the most abundant housekeeping proteins, the present approach also allowed the detection of proteins involved into signaling pathways e.g., calcium, Wnt, and MAPK signaling pathways with 27, 21, and 20 proteins, respectively. Importantly, we also detected a number of proteins involved in etiologies of a number of neurodegenerative disorders, including Alzheimer's and Parkinson's diseases (genes *ApoE*, *Ide*, *Mapt*,

Park7, *Snca*, *Stx1a*, *Th*, *Ube2l3*, *Ube1x*, *Uchl1*; Supplemental Fig. 10).

Discussion

The array of discovery-driven approaches previously applied to unravel brain function on the molecular level utilize a range of methods and technologies, from 2D PAGE and microarray to automated in situ hybridization and confocal microscopy (Gong et al. 2003; Baranzini 2004; Boguski and Jones 2004; Yang et al. 2005). However, the complexity of the brain and limitations of these approaches dictate the need for higher throughput and higher coverage discovery-driven approaches. In response to this need, we designed a spatial proteome mapping methodology that utilizes tissue voxelation and high-throughput quantitative proteomic measurements and demonstrated its applicability by analyzing a mouse brain coronal section. We mapped the relative abundance pattern for a total of 1028 proteins, with each protein identified by two or more unique peptides. Robust relative quantitation was achieved by spiking a "comprehensive" internal standard of ^{18}O -labeled tryptic peptides prepared from an entire coronal slice from the control mouse brain to ensure virtually every detected peptide had a labeled partner. However, due to the limited dynamic range of detection, some proteins were identified solely on the basis of detected ^{16}O peptides. Nevertheless, 838 (82%) of the identified proteins were quantified by using isotopically paired peptides, and an additional 190 proteins were quantified by using ^{16}O peptide intensities. These 190 proteins

are presumably relatively low-abundant proteins with localized expression of the corresponding genes (e.g., *Impact* and *Th*), as they are the likely candidates to fall below the detection limit in the pooled ^{18}O reference sample. In addition to dynamic range constraints for $^{16}\text{O}/^{18}\text{O}$ pair identification, other potential reason for not identifying those 190 proteins with paired peptides may include failure of the $^{16}\text{O}/^{18}\text{O}$ "pair finding" algorithm.

Compared with other technologies for analyzing the protein distribution across the brain, such as antibody hybridization or MALDI-TOF-based imaging methods (Reyzer and Caprioli 2005; Rohner et al. 2005) that provide impressive spatial resolution, the methodology presented here provides a much greater "depth" of proteome coverage in terms of protein identifications. This proteome coverage could be further increased with improved data analysis algorithms. For example, about half of the observed $^{16}\text{O}/^{18}\text{O}$ paired mass and NET features, which are likely to be peptidic species, did not confidently match any peptides from the reference AMT tag database. While some of these peptidic species may be chemically modified as a result of sample preparation and

handling steps (e.g., methionine oxidation), other modified species may reveal biologically important post-translation modification states. In either case, identification of such chemical modifications would increase proteome coverage. We are currently pursuing development and application of approaches for identifying such modifications in LC-FTICR datasets, as well as application of data-directed analyses (Masselon et al. 2005) for identifying presently unidentifiable features that display interesting patterns. In addition to the 1028 proteins identified with two or more peptides, we observed an additional 696 proteins identified with only one peptide (not included in this report). Given the acceptable FDR of peptide identification (Table 1), these proteins are likely to contain useful and biologically important information on brain protein distributions.

We estimate that completion of the Allen Brain Atlas project will require analyses of ~60,000 mice (~30,000 genes, sagittal and coronal sections per gene). Even greater efforts are required for the GENSAT project given the need to create transgenic EGFP-expressing mice for every single gene. Despite advantages and benefits of the visualization approaches used in these projects, it is unrealistic to widely apply them in a discovery-driven fashion for studying models of neurodegenerative diseases and their progression. This realization highlights another attribute of the present approach; that is, since the protein abundance measurements are effectively multiplexed, ultimately only one mouse is required for protein abundance mapping, thus making it possible to perform comparative studies having statistically reasonable sizes of mouse populations.

Since the protein abundance patterns were generated from a single mouse brain slice, the biological conclusions may suffer due to the lack of statistical confidence. However, there are several lines of evidences supporting the quality of the data and reproducibility of the current approach. First, the consistency of the protein abundance values among adjacent voxels for the majority of the proteins (Figs. 3, 4; Supplemental Table 3) indicates good reproducibility of the quantitation approach. Second, the observed bilateral symmetry of the protein abundance patterns (Supplemental Figs. 8, 9) and high agreement of abundance patterns of individual peptides derived from the same protein (Supplemental Fig. 6). Finally, the agreement of the presented patterns with gene expression patterns obtained by other orthogonal methods (Foster et al. 1987; Jursky and Nelson 1996; Honer et al. 1997; Feng et al. 2001; Gong et al. 2003; Schwindinger et al. 2003; Freneau et al. 2004; Rasband et al. 2005; <http://www.brainatlas.org>; <http://www.ncbi.nlm.nih.gov/projects/genSAT/>) (Fig. 3) further supports the confidence in the results.

The spatial resolution of the presented method is currently limited by sample preparation techniques rather than by LC-FTICR measurement sensitivity. The sample preparation procedure we used, based on 2,2,2-trifluoroethanol as a denaturing agent (Wang et al. 2005), is a very simplified one compared to conventional protocols for bottom-up proteomic sample preparation. Such a simplification enabled us to handle a large number of samples with small amount of material without critical sample losses and to make it easily compatible with automated liquid handling robotics. Microwave-assisted methods for enzymatic and chemical digestion of proteins (Hua et al. 2006; Sun et al. 2006) lend some additional potential for further acceleration and simplification of sample preparation procedures, and will be considered in the future studies. However, sample handling with conventional liquid handling techniques results in some degree of sample losses, e.g., due to nonspecific binding of the analyte to

tube and pipette tip plastic surfaces. Although it was not significant for the 1-mm³ voxel size, the percentage of sample loss may be substantial as the sample size decreases and may even result in complete loss of a sample. As a result, improvements in spatial resolution require some conceptually fundamental changes in sample preparation techniques, which may be forthcoming within a few years in the form of nanotechnology and microfluidic devices known as “lab-on-a-chip” (Freire and Wheeler 2006).

With the present methodology, the size of a voxel (i.e., 1 mm³) should make it possible to map the protein abundance patterns to the major structures of a mouse brain, such as the cortex, cerebellum, hippocampus, thalamus, hypothalamus, striatum, and some others. Furthermore, the voxel-to-structure assignment should be even more accurate for physically larger brains. For example, one cubic millimeter resolution is routinely used in MRI studies of human brains. Despite the limited resolution, the presented spatial mapping methodology affords a primary approach for genome-scale screening of protein candidates for follow-up single-cell resolution studies. Importantly, the protein abundance patterns generated with this methodology may be informative even without single-cell resolution data, as these patterns can be developed into quantitative traits and correlated with genetic variation, neuropathological and neurobehavioral phenotypes (Gillette et al. 2005). Additionally, the extensive identification and abundance pattern data generated from the presented approach can facilitate identification of higher resolution images of unidentified features from MALDI-TOF-based imaging by correlating the derived protein abundance patterns with high-resolution images of mass features. In such a tandem approach the identification of the observed mass features should be significantly easier as the candidate protein list could be narrowed down from the entire proteome to a list of a few protein candidates with the most cross-correlating patterns.

Given the successful proof-of-concept trial, our next step is to apply the presented methodology that combines voxelation with high-throughput quantitative LC-FTICR-based proteomic analysis to build 3D patterns of protein abundances for the study of normal and neurodegenerative disease mouse model brains. The ability of quantifying relative protein abundance from spatially localized regions should provide additional increase in dynamic range for discovering disease-related proteins compared to the whole brain sample analysis. In general, this methodology can be applied for comprehensive proteomic imaging of other organs and tissues and we expect it to be one of the major discovery-driven proteomic methodologies in the emerging field of functional neurogenomics.

Methods

Sample preparation

C57BL/6J male mouse was euthanized at 8 wk of age. The coronal section bregma 0 was dissected into 1-mm³ voxels as described previously (Liu and Smith 2003). All further liquid handling steps with the exception of ultrasonic disintegration were performed in 96-well format on a Biomek FX robotic station (Beckman Coulter). Samples were resuspended in 50% 2,2,2-trifluoroethanol (Sigma-Aldrich) in 50 mM NH₄HCO₃ (pH 7.8) with 5 mM tributylphosphine (Sigma-Aldrich) and homogenized in a 5510 Branson ultrasonic water bath (Branson Ultrasonics), followed by incubation for 2 h at 60°C to reduce disulfide bonds (Wang et al. 2005). For tryptic digestion, samples were diluted fivefold with 50 mM NH₄HCO₃, supplied with 1 mM CaCl₂ and 3 μg of trypsin per sample, and incubated for 7 h at 37°C with gentle shaking. After

lyophilization, samples were redissolved in 25 mM NH_4HCO_3 , boiled for 5 min, and then quickly chilled in an ice bath to inactivate trypsin. Equal amounts of 0.875 μg from each sample were mixed with equal amounts of an ^{18}O -labeled internal standard prepared from the entire coronal brain slice of another mouse. Preparation of the internal standard involved the same protocol as described above to derive tryptic peptides from the entire coronal slice followed by ^{18}O -labeling described elsewhere (Qian et al. 2005b).

Capillary LC-FTICR

The 1.75- μg aliquots were analyzed on a previously described 11 Tesla FTICR mass spectrometer (Gorshkov et al. 1998) that was interfaced with a 150 μm i.d. \times 65 cm long LC column packed with 3- μm Jupiter C_{18} particles (Phenomenex). The mobile phase solvents consisted of (A) 0.2% acetic acid and 0.05% TFA in water and (B) 0.1% TFA in 90% acetonitrile. An exponential gradient was used for the separation, which started with 100% A and gradually increased to 60% B over 100 min. To maximize the utilization of the instrument time we used an automated dual column LC system, which enabled us to run 71 voxel samples in a one-week time interval in a nonstop fashion.

Peptide and protein identification

LC-FTICR datasets were analyzed by in-house software VIPER (<http://ncrr.pnl.gov/software>; Zimmer et al. 2006) that detected features in mass-elution time space and assigned them to peptides in AMT tag database. Typically an LC-FTICR run results in a dataset with 12,000–16,000 features, of which 2500–4500 form $^{16}\text{O}/^{18}\text{O}$ isotopic pairs that coelute with a 4.0085-Da mass difference. The AMT tag reference database contained mass and NET values for pre-filtered set of ~38,000 the most confidently identified tryptic peptides from the previously done LC-MS/MS analyses of the entire mouse brain (Wang et al. 2006). The estimated values of standard deviations from LC-FTICR analyses for mass (σ_{mass}) and NET (σ_{NET}) error distributions were 1.4 ppm and 0.75%, respectively. The probabilistic distance D between the observed LC-MS feature and a matching peptide from the AMT tag database was calculated based on mass (ΔMass) and elution time (ΔNET) assignment errors and σ_{mass} and σ_{NET} standard deviations (Norbeck et al. 2005).

$$D = \sqrt{\left(\frac{\Delta\text{Mass}}{\sigma_{\text{mass}}}\right)^2 + \left(\frac{\Delta\text{NET}}{\sigma_{\text{NET}}}\right)^2} \quad (1)$$

After initial matching of features with relatively wide error tolerances (20 ppm mass and 10% NET) we performed initial filtering (step 1) based on mass and NET measurement errors. We employed two different stringency criteria for paired and unpaired features to remove false peptide identifications, maintaining relative FDRs. That is, we retained only those matches that had $D < 1.645$ for paired features and $D < 1.44$ for unpaired features, thereby retaining 90% and 85%, respectively, of true matches (assuming normality of mass and NET measurement error distributions). Probabilistic distances $D = 1.645$ and $D = 1.44$ define ellipses with a and b parameters corresponding to tolerances of ± 2.3 ppm mass, $\pm 1.2\%$ NET, and ± 2.0 ppm mass, $\pm 1.1\%$ NET, respectively. Further, we used only those peptide matches that had a SLiC (spatially localized confidence) score (Norbeck et al. 2005) > 0.3 for paired features and only the top matches, i.e., peptides with the highest SLiC score matching the same feature, for unpaired features.

$$\text{SLiC}_{ij} = \frac{e^{(-D_{ij}^2/2)}}{\sum_{k=1}^N e^{(-D_{ik}^2/2)}} \quad (2)$$

The resulting list of peptides was analyzed, using ProteinProphet software (Nesvizhskii et al. 2003), to derive the minimal set of proteins that had the maximal likelihood of observation. As an additional filtering criterion (step 2), we considered only those protein identifications that were identified by at least two peptides. To avoid ambiguity, only those peptides that matched a single protein from the list of confident identifications were selected for further quantitative analysis; peptides matching multiple proteins were discarded. Peptides with a nontryptic C terminus also were not considered for quantitation, as such peptides from voxel samples would be indistinguishable from peptides that originated from the spiked ^{18}O -labeled reference sample.

We used isotopic intensity ratios of ^{16}O species from the voxels to ^{18}O species from the spiked reference sample, as well as directly measured ^{16}O intensities as measures of peptide relative abundance. To remove any systematic trends which arise from the sample handling or other technical issues other than from differences between the protein abundances from the peptides intensity measures, we performed peptide intensity normalization between the LC-MS runs similar to between microarray normalization procedure (Park et al. 2003; Callister et al. 2006). We used a simple global normalization procedure, where the peptides abundance measures as $^{16}\text{O}/^{18}\text{O}$ intensity ratio values or ^{16}O intensities P_{ij} ($j = 1, \dots, J$) for a given LC-MS run ($i = 1, \dots, I$) were scaled by a global factor F_i . For computation of the factor F_i we used only peptides, which observed across all of the LC-MS runs. The F_i factor was computed as a median of peptide abundance ratios to the median value of peptide abundances across all the runs: $F_i = \text{median}_j(P_{ij}/M_j)$, where M_j is the median abundance value for a peptide j across all the runs $M_j = \text{median}_i(P_{ij})$. Then the normalized peptide abundance values P_{ij}^* were computed as $P_{ij}^* = P_{ij}/F_i$. An arbitrary abundance value of a protein was calculated as an average of the normalized peptide isotopic $^{16}\text{O}/^{18}\text{O}$ ratios or peptide ^{16}O abundances used to identify that particular protein. For the purpose of quantitation, to avoid interference between protein isoforms, we used only peptides with unique protein assignments. Finally, a protein relative abundance pattern was generated by using \log_2 of the ratio of protein arbitrary abundance value in each voxel to the averaged value across all voxels.

Acknowledgments

We thank The Allen Institute for Brain Science for permission to use their data in this publication. Portions of the research were supported by the NIH National Center for Research Resources (RR18522 to R.D.S.) and NIH grants R01 DA015802 and R01 NS050148 to D.J.S. Proteomic analyses were performed in the Environmental Molecular Sciences Laboratory, a US Department of Energy (DOE) national scientific user facility located at the Pacific Northwest National Laboratory (PNNL) in Richland, Washington. PNNL is a multi-program national laboratory operated by Battelle Memorial Institute for the DOE under Contract DE-AC05-76RL01830.

References

- Al-Shahrour, F., Minguez, P., Tarraga, J., Montaner, D., Alloza, E., Vaquerizas, J.M., Conde, L., Blaschke, C., Vera, J., and Dopazo, J. 2006. BABELOMICS: A systems biology perspective in the functional annotation of genome-scale experiments. *Nucleic Acids Res.* **34**: W472–W476.
- Bantle, J.A. and Hahn, W.E. 1976. Complexity and characterization of polyadenylated RNA in the mouse brain. *Cell* **8**: 139–150.
- Baranzini, S.E. 2004. Gene expression profiling in neurological disorders: Toward a systems-level understanding of the brain. *Neuromolecular Med.* **6**: 31–51.

- Beissbarth, T. and Speed, T.P. 2004. Gostat: Find statistically overrepresented Gene Ontologies within a group of genes. *Bioinformatics* **20**: 1464–1465.
- Boguski, M.S. and Jones, A.R. 2004. Neurogenomics: At the intersection of neurobiology and genome sciences. *Nat. Neurosci.* **7**: 429–433.
- Bordone, L. and Guarente, L. 2005. Calorie restriction, SIRT1 and metabolism: Understanding longevity. *Nat. Rev. Mol. Cell Biol.* **6**: 298–305.
- Brown, V.M., Ossadtchi, A., Khan, A.H., Cherry, S.R., Leahy, R.M., and Smith, D.J. 2002a. High-throughput imaging of brain gene expression. *Genome Res.* **12**: 244–254.
- Brown, V.M., Ossadtchi, A., Khan, A.H., Yee, S., Lacan, G., Melega, W.P., Cherry, S.R., Leahy, R.M., and Smith, D.J. 2002b. Multiplex three-dimensional brain gene expression mapping in a mouse model of Parkinson's disease. *Genome Res.* **12**: 868–884.
- Callister, S.J., Barry, R.C., Adkins, J.N., Johnson, E.T., Qian, W., Webb-Robertson, B.M., Smith, R.D., and Lipton, M.S. 2006. Normalization approaches for removing systematic biases associated with mass spectrometry and label-free proteomics. *J. Proteome Res.* **5**: 277–286.
- Caprioli, R.M. 2005. Deciphering protein molecular signatures in cancer tissues to aid in diagnosis, prognosis, and therapy. *Cancer Res.* **65**: 10642–10645.
- Chikaraishi, D.M. 1979. Complexity of cytoplasmic polyadenylated and nonpolyadenylated rat brain ribonucleic acids. *Biochemistry* **18**: 3249–3256.
- Feng, Y., Reznik, S.E., and Fricker, L.D. 2001. Distribution of proSAAS-derived peptides in rat neuroendocrine tissues. *Neuroscience* **105**: 469–478.
- Foster, G.A., Schultzberg, M., Hokfelt, T., Goldstein, M., Hemmings Jr., H.C., Ouimet, C.C., Walaas, S.I., and Greengard, P. 1987. Development of a dopamine- and cyclic adenosine 3':5'-monophosphate-regulated phosphoprotein (DARPP-32) in the prenatal rat central nervous system, and its relationship to the arrival of presumptive dopaminergic innervation. *J. Neurosci.* **7**: 1994–2018.
- Fountoulakis, M. 2004. Application of proteomics technologies in the investigation of the brain. *Mass Spectrom. Rev.* **23**: 231–258.
- Freire, S.L.S. and Wheeler, A.R. 2006. Proteome-on-a-chip: Mirage, or on the horizon? *Lab Chip* **6**: 1415–1423.
- Freneau Jr., R.T., Kam, K., Qureshi, T., Johnson, J., Copenhagen, D.R., Storm-Mathisen, J., Chaudhry, F.A., Nicoll, R.A., and Edwards, R.H. 2004. Vesicular glutamate transporters 1 and 2 target to functionally distinct synaptic release sites. *Science* **304**: 1815–1819.
- Gillette, M.A., Mani, D.R., and Carr, S.A. 2005. Place of pattern in proteomic biomarker discovery. *J. Proteome Res.* **4**: 1143–1154.
- Gong, S., Zheng, C., Doughty, M.L., Losos, K., Didkovsky, N., Schambra, U.B., Nowak, N.J., Joyner, A., Leblanc, G., Hatten, M.E., et al. 2003. A gene expression atlas of the central nervous system based on bacterial artificial chromosomes. *Nature* **425**: 917–925.
- Gorshkov, M.V., Pasa Tolic, L., Udseth, H.R., Anderson, G.A., Huang, B.M., Bruce, J.E., Prior, D.C., Hofstadler, S.A., Tang, L., Chen, L.Z., et al. 1998. Electrospray ionization-Fourier transform ion cyclotron resonance mass spectrometry at 11.5 tesla: Instrument design and initial results. *J. Am. Soc. Mass Spectrom.* **9**: 692–700.
- Griffin, T.J., Gygi, S.P., Iderk, T., Rist, B., Eng, J., Hood, L., and Aebersold, R. 2002. Complementary profiling of gene expression at the transcriptome and proteome levels in *Saccharomyces cerevisiae*. *Mol. Cell. Proteomics* **1**: 323–333.
- Honer, W.G., Falkai, P., Young, C., Wang, T., Xie, J., Bonner, J., Hu, L., Boulianne, G.L., Luo, Z., and Trimble, W.S. 1997. Cingulate cortex synaptic terminal proteins and neural cell adhesion molecule in schizophrenia. *Neuroscience* **78**: 99–110.
- Hua, L., Low, T.Y., and Sze, S.K. 2006. Microwave-assisted specific chemical digestion for rapid protein identification. *Proteomics* **6**: 586–591.
- Husson, A., Brasse-Lagnel, C., Fairand, A., Renouf, S., and Lavoinne, A. 2003. Argininosuccinate synthetase from the urea cycle to the citrulline-NO cycle. *Eur. J. Biochem.* **270**: 1887–1899.
- Jursky, F. and Nelson, N. 1996. Developmental expression of GABA transporters GAT1 and GAT4 suggests involvement in brain maturation. *J. Neurochem.* **67**: 857–867.
- Kanehisa, M. and Goto, S. 2000. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**: 27–30.
- Liu, D. and Smith, D.J. 2003. Voxelation and gene expression tomography for the acquisition of 3-D gene expression maps in the brain. *Methods* **31**: 317–325.
- Masselon, C., Pasa-Tolic, L., Tolic, N., Anderson, G.A., Bogdanov, B., Vilkov, A.N., Shen, Y., Zhao, R., Qian, W.J., Lipton, M.S., et al. 2005. Targeted comparative proteomics by liquid chromatography-tandem Fourier ion cyclotron resonance mass spectrometry. *Anal. Chem.* **77**: 400–406.
- Nesvizhskii, A.I., Keller, A., Kolker, E., and Aebersold, R. 2003. A statistical model for identifying proteins by tandem mass spectrometry. *Anal. Chem.* **75**: 4646–4658.
- Norbeck, A.D., Monroe, M.E., Adkins, J.N., Anderson, K.K., Daly, D.S., and Smith, R.D. 2005. The utility of accurate mass and LC elution time information in the analysis of complex proteomes. *J. Am. Soc. Mass Spectrom.* **16**: 1239–1249.
- Park, T., Yi, S.G., Kang, S.H., Lee, S., Lee, Y.S., and Simon, R. 2003. Evaluation of normalization methods for microarray data. *BMC Bioinformatics* **4**: 33.
- Peng, J., Elias, J.E., Thoreen, C.C., Licklider, L.J., and Gygi, S.P. 2003. Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS) for large-scale protein analysis: The yeast proteome. *J. Proteome Res.* **2**: 43–50.
- Qian, W.J., Camp II, D.G., and Smith, R.D. 2004. High-throughput proteomics using Fourier transform ion cyclotron resonance mass spectrometry. *Expert Rev. Proteomics* **1**: 87–95.
- Qian, W.J., Liu, T., Monroe, M.E., Strittmatter, E.F., Jacobs, J.M., Kangas, L.J., Petritis, K., Camp II, D.G., and Smith, R.D. 2005a. Probability-based evaluation of peptide and protein identifications from tandem mass spectrometry and SEQUEST analysis: The human proteome. *J. Proteome Res.* **4**: 53–62.
- Qian, W.J., Monroe, M.E., Liu, T., Jacobs, J.M., Anderson, G.A., Shen, Y., Moore, R.J., Anderson, D.J., Zhang, R., Calvano, S.E., et al. 2005b. Quantitative proteome analysis of human plasma following in vivo lipopolysaccharide administration using ¹⁶O/¹⁸O labeling and the accurate mass and time tag approach. *Mol. Cell. Proteomics* **4**: 700–709.
- Rasband, M.N., Tayler, J., Kaga, Y., Yang, Y., Lappe-Siefke, C., Nave, K.A., and Bansal, R. 2005. CNP is required for maintenance of axon-glia interactions at nodes of Ranvier in the CNS. *Glia* **50**: 86–90.
- Reyzer, M.L. and Caprioli, R.M. 2005. MALDI mass spectrometry for direct tissue analysis: A new tool for biomarker discovery. *J. Proteome Res.* **4**: 1138–1142.
- Rohrer, T.C., Staab, D., and Stoekli, M. 2005. MALDI mass spectrometric imaging of biological tissue sections. *Mech. Ageing Dev.* **126**: 177–185.
- Schwindinger, W.F., Betz, K.S., Giger, K.E., Sabol, A., Bronson, S.K., and Robishaw, J.D. 2003. Loss of G protein γ_7 alters behavior and reduces striatal α_{Olf} level and cAMP production. *J. Biol. Chem.* **278**: 6575–6579.
- Shen, Y., Strittmatter, E.F., Zhang, R., Metz, T.O., Moore, R.J., Li, F., Udseth, H.R., Smith, R.D., Unger, K.K., Kumar, D., et al. 2005. Making broad proteome protein measurements in 1–5 min using high-speed RPLC separations and high-accuracy mass measurements. *Anal. Chem.* **77**: 7763–7773.
- Singh, R.P. and Smith, D.J. 2003. Genome scale mapping of brain gene expression. *Biol. Psychiatry* **53**: 1069–1074.
- Sun, W., Gao, S., Wang, L., Chen, Y., Wu, S., Wang, X., Zheng, D., and Gao, Y. 2006. Microwave-assisted protein preparation and enzymatic digestion in proteomics. *Mol. Cell. Proteomics* **5**: 769–776.
- Tang, K., Li, F., Shvartsburg, A.A., Strittmatter, E.F., and Smith, R.D. 2005. Two-dimensional gas-phase separations coupled to mass spectrometry for analysis of complex mixtures. *Anal. Chem.* **77**: 6381–6388.
- Wang, H., Qian, W.J., Mottaz, H.M., Clauss, T.R., Anderson, D.J., Moore, R.J., Camp II, D.G., Khan, A.H., Sforza, D.M., Pallavicini, M., et al. 2005. Development and evaluation of a micro- and nanoscale proteomic sample preparation method. *J. Proteome Res.* **4**: 2397–2403.
- Wang, H., Qian, W., Chin, M.H., Petyuk, V.A., Barry, R.C., Liu, T., Gritsenko, M.A., Mottaz, H.M., Moore, R.J., Camp II, D.G., et al. 2006. Characterization of the mouse brain proteome using global proteomic analysis complemented with cysteinyl-peptide enrichment. *J. Proteome Res.* **5**: 361–369.
- Washburn, M.P., Koller, A., Oshiro, G., Ulaszek, R.R., Plouffe, D., Deciu, C., Winzeler, E., and Yates III, J.R. 2003. Protein pathway and complex clustering of correlated mRNA and protein expression analyses in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci.* **100**: 3107–3112.
- Yang, J.W., Juranville, J.F., Hoger, H., Fountoulakis, M., and Lubec, G. 2005. Molecular diversity of rat brain proteins as revealed by proteomic analysis. *Mol. Divers.* **9**: 385–396.
- Zimmer, S.D., Monroe, M.E., Qian, W., and Smith, R.D. 2006. Advances in proteomics data analysis and display using an accurate mass and time tag approach. *Mass Spectrom. Rev.* **25**: 450–482.

Received July 26, 2006; accepted in revised form November 22, 2006.