The Study of Biological Systems through Proteomics: A Quantitative Proteomic Analysis of Non-small Cell Lung Cancer (NSCLC)

J Watkins, E Horlick, D Love

Citation

J Watkins, E Horlick, D Love. *The Study of Biological Systems through Proteomics: A Quantitative Proteomic Analysis of Non-small Cell Lung Cancer (NSCLC)*. The Internet Journal of Genomics and Proteomics. 2013 Volume 6 Number 2.

Abstract

In the post-genomic era proteomics has emerged as one of the most important areas of research, particularly with respect to differential proteomics, the comparative analysis of the normal versus disease states of distinct proteomes. The most well established tool for proteomic analysis is 2-dimensional gel electrophoresis (2-DE). This tool offers a simple method to visually identify changes in protein abundance. 2-DE has been a staple in the examination of proteomes, and has undergone numerous technological advances over the past several decades. Isobaric tags for relative and absolute quantitation (iTRAQ) is a non-gel based alternative to 2-DE as a tool for proteomic analysis. Over the last decade or more it has become a common alternative to 2-DE technology coupled with mass spectrometry. The global use and evolution of proteomics warrants incorporation of the recent advances and uses of the technology into the undergraduate laboratory curriculum to prepare students for the proteomic era. In data presented herein we utilize iTRAQ coupled with mass spectrometry as a novel approach to teach students about the identification, quantification, statistical analysis and road blocks associated with the ever advancing field of comparative proteomics.

INTRODUCTION

From the time when the first prokaryotic and eukaryotic genomes were released in 1995 and 1996 respectively [1,2], the field of genomics has exploded into the 21st century providing an enormous amount of information, from the complete sequencing of the human genome to global gene expression in cells, uncovering a vast, fertile landscape for bioinformatics exploration.However, despite the success of these efforts, coupling gene sequence with function remains an ongoing effort.Predictably, however, another discipline, proteomics, has been expanding alongside these technologies.Proteomics describes the methodical identification and quantification of proteins expressed in biological systems.Application of proteomics allows a researcher to obtain information on protein identity, expression level, variants, and post-translational modifications.Researchers attempted to get a head start on this type of analysis twenty years before the first genome sequences were presented [3,4].It all began with the introduction of 2-dimensional electrophoresis (2-DE) and the separation of complex mixtures for comparative expression patterns.This provided much excitement for insightful interactions in human health, agriculture, environment and biotechnology would be uncovered.

2-DE has been the most favored technology for decades, to investigate the global qualitative and quantitative proteomic changes of proteomes [5-9].2-DE gel electrophoresis offers a low cost method of protein analysis. However, it has several notable shortcomings, particularly with respect to variations in repetitive gel runs, ultimately making protein analysis a very labor-intensive and experimentally error-prone task [10]. Although modifications in methodology have helped to alleviate such issues [11] making it a reliable technology for protein profiling.Despite the global use of 2-DE in proteomic research, there is a lack of proteomics education represented in the undergraduate laboratory curriculum to prepare the future workforce.The Rochester Institute of Technology is one of the few colleges that has successfully incorporated its introduction into the science curriculum [12].There are most likely several reasons but cost, reproducibility and implementation within a semester course seem to be the most likely barriers for more widespread acceptance into undergraduate science curricula.

With advances in liquid chromatography (LC), mass spectrometry (MS) and bioinformatics, newer comparative

and quantitative studies are possible.LC-MS/MS is a technique coupling high pressure liquid chromatography and tandem mass spectrometry to identify protein mixtures [13, 14].When coupled with chemical tagging, LC-MS/MS allows for quantification of comparative protein mixtures.Two methods of chemically labeling peptides are isotope coded affinity tags (ICAT) and iTRAQ.In ICAT, proteins from two experimental populations are labeled at cysteine residues with light (example, ^{12}C isotopic label) and heavy tags (example, ^{13}C isotopic label) carrying a biotin moiety.Mass spec analysis after separation of the labeled peptides reveals peaks corresponding to the same peptide as doublets in mass spectra due to the mass difference between light and heavy isotopes.Peak intensities of the peptides relate to the relative abundance of the proteins in the control versus the experimental data set.The primary limitations to this chemical labeling method include the sensitivity to cysteine residues – proteins lacking cysteine can be missed entirely and proteins with few cysteines can be misidentified.

A relatively new technology iTRAQ, labels the amino-terminus of peptides after trypsin cleavage at lysine and arginine residues and allows for simultaneous comparison of up to four specimens.With this method, proteins from two related samples can be digested separately with trypsin and the resulting peptides labeled with two amine-specific reagents with distinguishable mass tag reagents. The labeled samples are then combined and analyzed by LC-MS/MS. The mass tags can then be cleaved readily under collision induced dissociation to liberate molecular ions of 114 to 117 Da, leaving a compensatory mass on the peptide. For any one peptide, both labeled forms contribute to the y- and b-ion fragmentation pattern and the relative intensity of the cleaved tags reflects the relative abundance of the two initial peptides. The iTRAQ method increases confidence in identification and quantification by labeling multiple peptides per protein.The iTRAQ labeling technique is unique in that there are four available isobaric tags from which four comparable experiments can be combined into one proteomic analysis.

In this experiment we will outline an undergraduate protocol for the comparative analysis of total protein lysate between human normal lung and lung squamous cell carcinoma tissue utilizing iTRAQ as opposed to 2-DE technology.The aim of this paper is to provide an approach from which to teach students about the identification, quantification, statistical analysis and roadblocks associated with the ever advancing field of comparative proteomics.To limit time related to sample preparation and collection, total cell lysate samples were acquired from the company BioChain Institute, Inc. (Newark, CA) allowing for immediate protein precipitation, proteolysis, iTRAQ labeling, and peptide analysis.

MATERIALS AND METHODS

Proteolysis and iTRAQ™ reagent labeling of total cell and mitochondrial fraction proteins.

Total protein lysate samples for human adult normal lung tissue (male, age 24) and human adult lung tumor tissue (male, age 72) were purchased from BioChain Institute, Inc.100 µg samples of protein from human adult normal lung tissue and human adult lung tumor tissue were precipitated with acetone for 1 h at -20 °C and collected by centrifugation at 13,400 rpm. The pellet of protein from each sample was then resuspended in 30 µl of iTRAQ™ dissolution buffer (Applied Biosystems, Grand Island, NY) with 0.05% SDS final concentration.Each 100 µg sample of protein was reduced, alkylated, digested with trypsin and labeled with the isobaric reagents according to the manufacturer's protocol.

LC/MS/MS analysis

iTRAQ -labeled peptides were separated by OFFGEL based on pI [15] followed by nanoLC-LTQ Orbitrap analysis.

Protein identification and quantification

A triplicate experiment was performed comparing human adult normal lung tissue and human adult lung tumor tissue. Mass spectrometer data were imported into an excel file.Each individual peptide was linked to an independent mass tag ratio.The resulting peptides that identified a single protein were used to calculate a new averaged mass tag ratio for that protein (SD<±0.14).Complete data sets from independent experiments were compared to one another and a final average ratio for each protein identified in the data set was calculated.For each identified protein, the National Center for Biotechnology Information (NCBI, http://www.ncbi.nlm.nih.gov/) was utilized to manually collate information on all proteins with Gene Ontology (GO) annotations.

Statistical analysis

Each protein was characterized by a ratio of mass tag signal

intensity reflecting the relative abundance of the protein in normal lung tissue relative to cancerous lung tissue. These abundance ratios were converted to a $\log_{10}$ scale to overcome the lack of symmetry in raw protein abundance ratios and to calculate a mean, , and standard deviation, , for the log protein abundance.The observed mean of the relative protein abundance ratio was = 1 and the log of this value was 0. The standard deviation of the log relative protein abundance was = 0.1.In order to determine the biological significance of relative protein expression data, we formulated the null hypothesis H_0 that all protein abundance data fit a normal distribution characterized by the population mean = .The hypothesis of interest is whether the relative $\log_{10}$ abundance ratio of a protein X = , or if the alternative hypothesis is correct, that X ≠ .Using a two-tailed t test, we defined criteria to determine whether the abundance of a protein was significantly distinct from the normal distribution characterized by the population mean, . We consider a protein to be significantly up or down-regulated if the p-value of its abundance ratio relative to the mean was ≤ 0.05.

RESULTS AND DISCUSSION

Proteomic analysis of total cell lysates

Total cell protein was prepared from human adult normal lung tissue (male age 24) and human adult lung tumor tissue (male age 72) for quantitative LC-MS/MS analysis of isobaric tagged peptides as described in Materials and Methods. We collected quantitative data on multiple peptides for 91 proteins.All results obtained are representative of experiments performed in triplicate.In approximately 80% of the identifications, multiple qualifying peptides were observed in each experiment (Table 1).

Table 1

Accession	Gene	Protein Name	No. Pep.	Ratio Pep.	Ratio	STD	P-value	No. Exp
Antibody								
IPI00384938	IGHG1	Putative uncharacterized protein DKFZp686N02209	10	4	0.9	0.13	0.38	3
IPI00930124	IGHG1	Putative uncharacterized protein DKFZp686C11235	10	4	0.9	0.13	0.38	3
IPI00399007	IGHG2	Putative uncharacterized protein DKFZp686I04196	7	3	0.9	0.16	0.63	3
IPI00550731	IGKC	Putative uncharacterized protein	3	2	1.3	0.11	0.03	3
IPI00382577	Kappa1	light chain variable region (Fragment)	2	2	1.3	0.07	0.02	3
IPI00807428	Unknown	Putative uncharacterized protein	4	4	1.0	0.13	0.87	3
IPI00386524	IGHA1	CDNA FLJ25298 fis	3	2	1.0	0.13	0.69	3
IPI00154742	IGLV2-14	LOC100293277	4	3	1.0	0.14	0.69	3
IPI00829956	Ig-like	Rheumatoid factor C6 light chain (Fragment)	2	2	1.2	0.02	0.07	3
Anti-inflammatory/immuno-modulatory								
IPI00006705	SCGB1A1	Uteroglobin	1	1	0.6	0.00	0.00	3
Calcium Ion Binding								
IPI00418169	ANXA2	Isoform 2 of Annexin A2	7	3	0.3	0.12	0.00	3
IPI00013895	S100A11	Protein S100-A11	3	1	1.6	0.00	0.00	3
IPI00329801	ANXA5	Annexin A5	4	2	0.8	0.15	0.20	3
IPI00024095	ANXA3	Annexin A3	1	1	1.2	0.00	0.07	3
IPI00025084	CAPNS1	Calpain small subunit 1	1	1	0.7	0.00	0.04	3
IPI00218918	ANXA1	Annexin A1	5	3	1.0	0.13	0.81	3
IPI00027463	S100A6	Protein S100-A6	3	2	1.3	0.37	0.47	3
Cell Cycle/Cell Differentiation								
IPI00550020	PTMS	Parathymosin	1	1	0.8	0.00	0.07	3
IPI00217465	HIST1H1C	Histone H1.2	3	3	0.7	0.00	0.01	3
Cytoskeletal								
IPI00021439	ACTB	Actin, cytoplasmic 1	12	8	0.7	0.11	0.02	3
IPI00218343	TUBA1C	Tubulin alpha-1C chain	5	3	1.1	0.10	0.42	3
IPI00180675	TUBA1A	Tubulin alpha-1A chain	5	3	0.9	0.26	0.38	3
IPI00216138	TAGLN	Transgelin	4	3	1.5	0.21	0.00	3
IPI00003269	ACTBL2	Beta-actin-like protein 2	6	2	0.7	0.12	0.02	3
IPI00739539	POTEF	POTE ankyrin domain family member F	6	2	0.8	0.06	0.15	3
IPI00216691	PFN1	Profilin-1	2	2	1.0	0.30	0.69	3
IPI00219365	MSN	Moesin	3	1	0.9	0.00	0.42	3
IPI00012011	CFL1	Cofilin-1	2	2	0.8	0.09	0.15	3
IPI00007752	TUBB2C	Tubulin beta-2C chain	4	1	1.6	0.00	0.00	3
IPI00180240	TMSL3	Thymosin beta-4-like protein 3	3	1	0.9	0.00	0.34	3
IPI00027341	CAP0	Macrophage-capping protein	1	1	1.0	0.00	0.94	3
IPI00886911	LOC646048	similar to actin alpha 1 skeletal muscle protein	1	1	0.5	0.00	0.00	3
IPI00643920	TKT	TKT cDNA FLJ54957, highly similar to Transketolase	2	1	0.7	0.00	0.01	3
IPI00298994	TLN1	Talin-1	2	1	0.4	0.00	0.00	3
IPI00291175	VCL	Isoform 1 of Vinculin	2	1	1.2	0.00	0.08	3
IPI00418471	VIM	Vimentin	2	1	0.4	0.00	0.00	3
Hemoglobin binding								
IPI00478493	HP	haptoglobin isoform 2 preproprotein	7	4	1.3	0.63	0.02	3
Metabolism								
IPI00465248	ENO1	Isoform alpha-enolase of Alpha-enolase	6	2	0.9	0.40	0.47	3
IPI00220644	PKM2	Isoform M1 of Pyruvate kinase isozymes M1/M2	6	3	2.1	2.16	0.00	3
IPI00465343	ADH1C	Alcohol dehydrogenase 1C	4	1	0.9	0.00	0.52	3
IPI00169383	PGK1	Phosphoglycerate kinase 1	5	1	1.0	0.21	0.81	3
IPI00219018	GAPDH	Glyceraldehyde-3-phosphate dehydrogenase	3	1	1.1	0.00	0.47	3
IPI00465439	ALDOA	Fructose-bisphosphate aldolase A	2	1	3.2	0.00	0.00	3
IPI00465028	RCTPI1	TPI1 triosephosphate isomerase 1 isoform 2	4	2	1.0	0.08	0.75	3
IPI00290462	CBR3	Carbonyl reductase [NADPH] 3	1	1	0.9	0.00	0.63	3
IPI00218570	PGAM2	Phosphoglycerate mutase 2	1	1	1.2	0.00	0.20	3
IPI00006663	ALDH2	Aldehyde dehydrogenase, mitochondrial	1	1	0.9	0.00	0.23	3
Mitochondrial ATP Synthetase								
IPI00303476	ATP5B	ATP synthase subunit beta, mitochondrial	3	2	1.1	0.58	0.58	3
Mitochondrial Oxidative Stress								
IPI00218733	SOD1	Superoxide dismutase [Cu-Zn]	1	1	0.8	0.00	0.11	3
Neurogenesis								
IPI00021263	YWHAZ	14-3-3 protein zeta/delta	5	4	1.4	0.90	0.00	3
IPI00000816	YWHAE	14-3-3 protein epsilon	5	2	1.2	0.12	0.18	3
IPI00013890	SFN	Isoform 1 of 14-3-3 protein sigma	5	2	1.2	0.12	0.18	3
IPI00257508	DPYSL2	Dihydropyrimidinase-related protein 2	2	1	0.7	0.00	0.04	3
IPI00216318	YWHAB	Isoform Long of 14-3-3 protein beta/alpha	4	1	1.1	0.00	0.47	3

Table 1 Cont'd

Identification of 91 proteins and the changes in their abundance: The iTRAQ® ratio of protein abundance in Normal v. Cancerous Lung is shown with the standard deviation, S.D. of peptide abundance. Columns on the right show the number of peptides observed by mass spectrometry analysis in repetitive experiments and the number of experiments in which the protein was detected.

Protein Binding								
IPI00021812	AHNAK	Neuroblast differentiation-associated protein AHNAK	17	4	0.9	0.17	0.63	3
IPI00479743	POTEE	Isoform 1 of POTE ankyrin domain family member E	7	3	0.8	0.09	0.07	3
IPI00478003	A2M	Alpha-2-macroglobulin	3	2	1.2	0.02	0.23	3
IPI00003817	ARHGDIB	Rho GDP-dissociation inhibitor 2	1	1	0.7	0.00	0.03	3
IPI00555610	AHNAK-like	AHNAK, Putative uncharacterized protein AHNAK	12	2	0.9	0.04	0.30	3
Protein Folding/Chaperone								
IPI00304925	HSPA1A	Heat shock 70 kDa protein 1A/1B	6	3	1.0	0.44	1.00	3
IPI00382470	HSP90AA1	Isoform 2 of Heat shock protein HSP 90-alpha	5	4	0.9	0.20	0.30	3
IPI00003865	HSPA8	Isoform 1 of Heat shock cognate 71 kDa protein	5	3	0.8	0.12	0.09	3
IPI00414676	HSP90AB1	Heat shock protein HSP 90-beta	5	3	1.0	0.11	0.69	3
IPI00003362	HSPA5	HSPA5 protein	4	2	0.8	0.16	0.05	3
IPI00643152	HSPA1L	cDNA FLJ56386	4	2	0.8	0.16	0.05	3
IPI00419585	PPIA	Peptidyl-prolyl cis-trans isomerase A	4	1	0.4	0.00	0.00	3
IPI00025512	HSPB1	Heat shock protein beta-1	3	1	2.1	0.00	0.00	3
IPI00339269	HSPA6	Heat shock 70 kDa protein 6	3	1	0.9	0.00	0.27	3
IPI00935692	LOC100290509	LOC100290509 hypothetical protein XP_002348229	2	1	0.4	0.00	0.00	3
IPI00027230	HSP90B1	Endoplasmin	2	1	0.9	0.00	0.30	3
IPI00555565	HSP90AB4P	Putative heat shock protein HSP 90-beta 4	2	1	0.9	0.00	0.38	3
IPI00012889	SFTPA1	Pulmonary surfactant-associated protein A1	1	1	0.4	0.00	0.00	3
Protein inhibitor								
IPI00553177	SERPINA1	Isoform 1 of Alpha-1-antitrypsin	8	3	1.3	0.54	0.03	3
Proteolysis								
IPI00011229	CTSD	Cathepsin D	5	3	0.6	0.05	0.00	3
Response to Stress/Oxidative Stress								
IPI00027350	PRDX2	Peroxiredoxin-2	3	3	1.0	0.13	0.94	3
IPI00000874	PRDX1	Peroxiredoxin-1	2	2	0.9	0.28	0.58	3
IPI00465436	CAT	Catalase	1	1	0.7	0.00	0.02	3
Translation/Protein Biosynthesis								
IPI00179330	RPS27A	ribosomal protein S27a precursor	3	2	0.6	0.25	0.00	3
IPI00025447	EEF1A1	Elongation factor 1-alpha	3	2	0.8	0.09	0.06	3
Transport								
IPI00745872	ALB	Isoform 1 of Serum albumin	38	25	1.1	0.27	0.47	3
IPI00654755	HBB	Hemoglobin subunit beta	10	9	1.2	0.06	0.18	3
IPI00473011	HBD	Hemoglobin subunit delta	8	7	1.2	0.09	0.15	3
IPI00410714	HBA2	Hemoglobin subunit alpha	7	5	1.1	0.15	0.38	3
IPI00829896	HBD	Hemoglobin Lepore-Baltimore (Fragment)	7	6	1.2	0.09	0.15	3
IPI00930351	HBD (Fragment)	Hbbm fused globin protein (Fragment)	5	4	1.2	0.11	0.11	3
IPI00022463	TF	Serotransferrin	7	3	1.2	0.13	0.08	3
IPI00215983	CA1	Carbonic anhydrase 1	3	2	1.1	0.12	0.30	3
IPI00021854	APOA2	Apolipoprotein A-II	3	2	1.1	0.38	0.27	3
IPI00021841	APOA1	Apolipoprotein A-I	5	1	1.5	0.00	0.00	3
IPI00010796	P4HB	Protein disulfide-isomerase	1	1	0.9	0.00	0.38	3
IPI00216298	TXN	Thioredoxin	2	2	1.2	0.24	0.20	3

Approximately 33% of the 91 proteins identified were differentially expressed at the p=0.05 confidence level, with 12 up-regulated and 18 down-regulated (Figure 1). This result indicates the cancerous lung entails a re-organization of cell structure and function. The proteins listed in Table 1 were grouped according to gene ontology classifications to reveal a number of categories that showed significant disparities in the numbers of up- and down-regulated proteins, as shown in Table 2. The most consistent patterns of alteration in protein abundance reflect decreases within the tissue for proteins involved in anti-inflammatory/immuno-modulatory response, cell cycle/differentiation, proteolysis, and translation/protein biosynthesis.On the other side there was an increase in proteins of hemoglobin binding, metabolism, and protein inhibition.Although this data shows some alterations in the proteome of the cancerous state versus the normal lung tissue, it is clear that the protein coverage is in sufficient to make any conclusive statements.However this analysis bring up key drawbacks for class discussion on how to expand and improve the protein coverage in GO categories to discover possible biomarkers for disease or central pathways involved in the onset of disease.

Figure 1

Changes in abundance of total cell proteins during neuronal differentiation. The distribution of the number of proteins with log iTRAQ® ratios in the indicated ranges is shown.

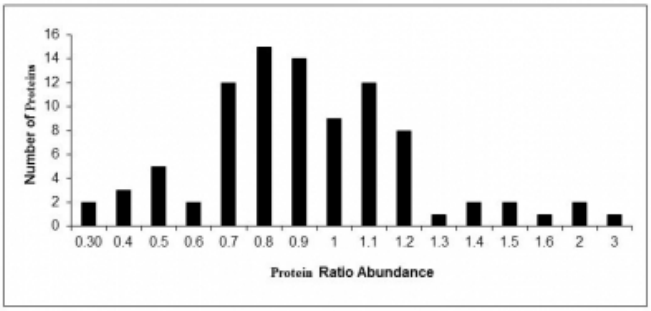

Table 2

Quantitative changes in protein expression are represented based on gene ontology. Proteins identified by LC-MS/MS were classified according to their GO annotations. The numbers of proteins in each class that were increased, decreased or unchanged (NC) in abundance are indicated along with the average fold-change for proteins in each class.

GO	Up	NC	Dn	Avg ratio
Antibody	2	7	0	1.04
Anti-inflammatory/immuno-modulatory	0	0	1	0.60
Calcium Ion Binding	1	4	2	1.01
Cell Cycle/Cell Diifferentiation	1	1	0	0.73
Cytoskeletal	2	9	6	0.88
Hemoglobin binding	1	0	0	1.30
Metabolism	2	8	0	1.32
Mitochondrial ATP Synthetase	0	1	0	1.10
Mitochondrial Oxidative Stress	0	1	0	0.80
Neurogenesis	1	3	1	1.11
Protein Binding	0	4	1	0.89
Protein Folding/Chaperone	1	7	5	0.86
Protein inhibitor	1	0	0	1.30
Proteolysis	0	0	1	0.60
Response to Stress/Oxidative Stress	0	2	1	0.88
Translation/Protein Biosynthesis	0	0	2	0.67
Transport	1	11	0	1.16

The 2-DE and iTRAQ proteome anlaysis approached have their pluses and minuses in the anlaysis of proteome changes of cells/tissues.The iTRAQ technology avoids the hurdles that typically affect protein analysis, such as poor solubility of membrane-bound proteins, limited loading capacities (as seen with pH strips), limitations on proteins of 10-120

kilodaltons (kD) with neutral to acidic isoelectric points,difficulty in resolving very basic proteins, and intra-experimental reproducibility.Both the iTRAQ and 2-DE methods of protein analysis are limited with respect to the dynamic range of analysis.As with 2-DE, iTRAQ only identifies highly abundant proteins from complex mixtures, leaving low abundant proteins such as signaling proteins challenging to identify.To overcome this drawback, fractionation is a strategy that may be used to simplify the protein mixture, and allow for detection at the lower levels.The fractionation is carried out prior to LC-MS/MS by liquid chromatography steps, such as strong cation exchange chromatography, creating unique population pools of protein to introduce for mass spectrometry analysis.This additional step provides larger coverage of the proteome especially with proteins in low concentration.Although the dynamic range can be improved by fractionation another contributing factor that decreasesthe ability to acquire low abundant proteins is the presence of highly abundant proteins like human serum albumin (HSA) and immunoglobulin gamma (IgG), which together make up approximately seventy five percent of total serum protein content and inhibit the analysis of low abundance proteins.It can be seen in the data provided herin that a fractionation step may have provided a more in-depth look at proteins involved in GO catagories to solidify key changes affiliated with cancerous lung tissue as seen with the identification of serum albumin with the largest number of peptides at 38.In comparison each of the remaining proteins were identified by a range of 1-17 peptides at the most.To increase identifications in association with fractionation, the use of affinity-based depletion strategies are used to selectively remove HSA and IgG, to facilitate a more comprehensive analysis of proteins from high to low abundance .

This experiment provides a step by step method for the assessment of a comparative sample using iTRAQ technology. This data shows a level of changes taking place and the ability of mass spectrometry to identify and quantitate proteins.The experiment is simplified by utilizing total cell lysates already prepared by a commercial company.This dramatically decreases sample source acquisition, preparation and collection time.By acquiring the comparative samples students can concentrate on the necessary chemistry required to reduce, alkylate, and digest proteins with trypsin, followed by isobaric labeling of the amine terminus of peptides.Discussion on the proper formalities needed for the collection of patient samples, or the growth and maintenance of tissue culture cells can be discussed in detail as a part of laboratory introductions and discussion sections of lab reports.

The cost of iTRAQ analysis can be equivalent to 2-DE provided that access to LCMS/MS equipment is available.However the costs for LCMS/MS are going down and can be negotiated with willing institutions.Data turnaround is relatively quick at about two to five days depending on the load of the proteomic center service.The format of the data returned allows for easy import into excel files to organize GO categories, perform quantitative analysis, statistical analysis, create charts and graphs of representative data.This entire protocol can be performed in just 6 two and a half hour lab periods (without any complication in data acquisition from the proteomics facility) with students working diligently in groups composed of two to four students with tight instructor supervision.A simplified schedule of events can be seen in Table 3.

Table 3

Time table of laboratory activities by class period.

Lab Session	Procedure	Lecture/Course work	Class Time
1	Protein precipitation, reduce, alkylate & trypsin digest of samples (Note: The trypsin digest will take 12-16 hours. Upon completion the instructor will need to store samples in a freezer.)	Review of protein handling procedures	2.5 hours
2	iTRAQ labeling (Note: Send samples for proteomic analysis)	Discussion on sample sources & labeling chemistry	2.5 hours
3	Assessment of previously published proteomic data on cancer	Discussion on comparative proteomics approaches [16,17] Discussion on proteomics, biomarkers & medicine [17]	2.5 hours
4	Organize acquired data	Overview of organizing data Introductions to NCBI Demo on organization of proteomic data Homework: Groups are to begin organization of data	2.5 hours
5	Continue organization and interpretations of data	Overview of presentation requirements Homework: Groups are to finish organization of data	2.5hours
6	Group Presentations	Identify sources of error, discuss conclusions and future directions	2.5 hours

Acknowledgments

We would like to thank Cristian Ruse and Cexiong Fu of the Cold Spring Harbor Laboratory proteomics shared resource.

References

1. R. D. Fleischmann, M. D. Adams, O. White, R. A.

Clayton, E. F. Kirkness, A. R. Kerlavage, C. J. Bult, J. F. Tomb, B. A. Dougherty, J. M. Merrick; et al. (1995) Whole-genome random sequencing and assembly of Haemophilus influenzae Rd, Science 269, 496–512.
2. A. Goffeau, B. G. Barrell, H. Bussey, R. W. Davis, B. Dujon, H. Feldmann, F. Galibert, J. D. Hoheisel, C. Jacq, M. Johnston, E. J. Louis, H. W. Mewes, Y. Murakami, P. Philippsen, H. Tettelin, S. G. Oliver (1996) Life with 6000 genes, Science 274, 546–547.
3. J. Klose (1975) Protein mapping by combined isoelectric focusing and electrophoresis of mouse tissues. A novel approach to testing for induced point mutations in mammals, Humangenetik 26, 231– 243.
4. P. H. O'Farrell (1975) High resolution two-dimensional electrophoresis of proteins, J. Biol. Chem. 250, 4007–4021.
5. Ray, W. J. and D. I. Gottlieb (1996). "Regulation of Protein Abundance in Pluripotent Cells Undergoing Commitment to the Neural Lineage." Journal of Cellular Physiology(168): 264-275.
6. Wang, D. and L. Gao (2005). "Proteomic analysis of neural differentiation of mouse embryonic stem cells." Proteomics 5(17): 4414-26.
7. Maurer, M. H., R. E. Feldmann, Jr., C. D. Futterer, J. Butlin and W. Kuschinsky (2004). "Comprehensive proteome expression profiling of undifferentiated versus differentiated neural stem cells from adult rat hippocampus." Neurochem Res 29(6): 1129-44.
8. Hoffrogge, R., S. Mikkat, C. Scharf, S. Beyer, H. Christoph, J. Pahnke, E. Mix, M. Berth, A. Uhrmacher, I. Z. Zubrzycki, E. Miljan, U. Volker and A. Rolfs (2006). "2-DE proteome analysis of a proliferating and differentiating human neuronal stem cell line (ReNcell VM)." Proteomics 6(6): 1833-47.
9. An, J., Q. Yuan, C. Wang, L. Liu, K. Tang, H. Y. Tian, N. H. Jing and F. K. Zhao (2005). "Differential display of proteins involved in the neural differentiation of mouse embryonic carcinoma P19 cells by comparative proteomic analysis." Proteomics 5(6): 1656-68
10. Righetti, P. G., A. Castagna, F. Antonucci, C. Piubelli, D. Cecconi, N. Campostrini, P. Antonioli, H. Astner and M. Hamdan (2004). "Critical survey of quantitative proteomics in two-dimensional electrophoretic approaches." J Chromatogr A 1051(1-2): 3-17.
11. Tannu NS, Hemby SE. (2006). Two-dimensional fluorescence difference gel electrophoresis for comparative proteomics profiling. Nat Protoc. 1(4):1732-42.
12. Kim TD, Craig PA (2010). Introducing proteomics in the undergraduate curriculum: A simple 2D gel electrophoresis exercise with serum proteins. Biochem Mol Biol Educ. Jan;38(1):29-34. doi: 10.1002/bmb.20353.
13. Shi, Y., R. Xiang, C. Horvath and J. A. Wilkins (2004). "The role of liquid chromatography in proteomics." J Chromatogr A 1053(1-2): 27-36.
14. Issaq, H. J., K. C. Chan, G. M. Janini, T. P. Conrads and T. D. Veenstra (2005). "Multidimensional separation of peptides for effective proteomic analysis." J Chromatogr B Analyt Technol Biomed Life Sci 817(1): 35-47.
15. Hörth P, Miller CA, Preckel T, Wenz C. (2006). Efficient fractionation and improved protein identification by peptide OFFGEL electrophoresis. Mol Cell Proteomics. Oct;5(10):1968-74. Epub 2006 Jul 14.
16. Bantscheff M, Lemeer S, Savitski MM, Kuster B. (2012). Quantitative mass spectrometry in proteomics: critical review. Anal Bioanal Chem. Sep;404(4):939-65. doi: 10.1007/s00216-012-6203-4.
17. Honda K, Ono M, Shitashige M, Masuda M, Kamita M, Miura N, Yamada T. (2012). Proteomic Approaches to the Discovery of Cancer Biomarkers for Early Detection and Personalized Medicine. Jpn. J. Clin. Oncol. doi: 10.1093/jjco/hys200

Author Information

Jermel Watkins, Ph.D.
Cold Spring Harbor Laboratory
Cold Spring Harbor, New York
jwatkins@cshl.edu

Elliot Horlick, Research - Mentee
Cold Spring Harbor Laboratory
Cold Spring Harbor, New York

Damon Love, Ph.D.
Weill Cornell Medical College
New York, NY