Application iTRAQ Reagent for Analysis of Embryonal Carcinoma Cell Line

Qishan Lin*, Liming Yu, Jinghua Zhu*, Tiffany Devine and Paulette McCormick
UAlbany Proteomics Facility*, Center for Functional Genomics, SUNY Albany, Rensselaer, NY 12144

OVERVIEW

One of the great strengths of proteomics is its potential to aid in the generation of novel hypotheses and/or to lead to new insights into biological systems. The aim of this study was to isolate and identify novel proteins that are involved in the process of tumor initiation and progression through comparison of a tumorigenic vs. a mutant non-tumorigenic embryonal carcinoma (EC) cell line. This application described how the use of iTRAQ isobaric tagging combined with off-line multidimensional liquid chromatography have enabled the identification and differential profiling of low level proteins from a cancer cell line.

INTRODUCTION

Embryonal carcinoma (EC) cells are stem cells of tetracarcinomas-tumors that develop in the gonads of both humans and mice. These tumor cells are essentially indistinguishable, at least in terms of differentiative and neoplastic potentials, from the embryonal stem (ES) cells of early (peri-implantation) mammalian embryo. We have used retroviral insertion to create a mutant embronyonal carcinoma cell line NR1-6, which is unique in its morphological, adhesive, tumorigenic, and differential properties (Figure 1). Genetic analysis of mutant, hybrid, and revertant cell lines indicates there is only a single retroviral insertion site that we have mapped to the proximal portion of the mouse X chromosome. Mouse mammary gland and brain expressed sequence tags (ESTs) have greater than 97% homology to a region within the > 18 kb of the insertion site flanking genomic DNA that we have sequenced. Utilizing these ESTs, we have identified two transcripts expressed in parental, but not in mutant cells. These are 3.8 and 2.3 kb. The predominant transcript, ~2.3 kb, contains two exons, the second of which is disrupted by insertion. Expression analysis indicates that these transcripts are widely expressed both temporally and spatially. However, although expression of these transcripts is highly reproducible and has been confirmed by multiple assays, neither we, nor various different sequencing facilities, have been able to find an open reading frame (ORF) within them. Thus it is likely that they represent regulatory RNAs. We have conducted large-scale differential display (over 80 primer pairs and 90,000 bands) and micro-array analyses (1076 genes of ESTs) to identify downstream targets of these transcripts or RNAs with little success. That is to say, that by this technology only these two transcripts, and one for a bHLH gene called MyoR, are differentially expressed. We have tested the MyoR gene in transient transfections and have found that its expression, or lack thereof, does no regulate the mutant phenotypes. Thus, we have been unable to identify the relevant altered downstream transcripts. It is possible that the amplification steps of the PCR based differential display and microarray techniques are damping out differences. To identify proteins that may play a role for those phenotype changes, we employed proteomic approach for comparative study the protein profiles from those two cell lines.

EXPERIMENTAL

Protein preparation and iTRAQ isobaric labeling. EC Cells (6×10^6) were cultured in Dulbecco's modified Eagle's medium containing 10% fetal bovine serum. The cells were washed with cold PBS 3 times and scrapped in 400 µl sample buffer (100 mM Tris-HCl, 8M Urea, 0.05% SDS, 5 mM TBP, pH 8.3). The sample mixtures were probe-sonicated on ice for 5×20 seconds followed by iodoacetimide alkylation for 1hr. Protein concentration was measured by MicroBCA by using a Pierce protein assay kit before the reaction was quenched by the addition of DTT. Equal aliquots (500 µg) from each lysate were then tryptic digested overnight at 37oC. The tryptic peptide mixture from each sample was desalted with a C18 cartridge followed by the iTRAQ labeling as described in the iTRAQ protocol (Figure 2).

Off-line strong cation exchange chromatography. The labeled samples were then pooled and acidified to a total volume of 2.0 ml followed by injection into a Agilent 1100 HPLC system with a Zorbax 300-SCX column (4.6 ID ×250 mm). Solvent A was 5mM K2HPO4 + 25% acetonitrile, pH 3.0 and B was 350 mM KCl +A. Peptides were eluted from the column with a 40 min mobile phase B gradient. A total of 26 fractions were collected. The samples were dried by a speed-vac prior to LC-MS/MS analysis.

On-line nano-LC ESI QqTOF MS Analysis. The nanobore LC system was from Dionex and interfaced to a QSTAR XL QqTOF mass spectrometer with a NanoSpray ion source. The Picofrit column (Magic C18, 100 A pore 75 µm ID ×150 mm) was packed in house. Solvent A was 3% CH3CN+ 0.1% formic acid+0.01%TFA, and B was 98% CH3CN+0.1% formic acid +0.01% TFA. Peptides mixture (reconstituted in 200 µl of 5% formic acid) was injected and eluted from the column with a 110 min mobile phase B gradient (5 -5 %B in 5 min, 5 -18%B in 10 min, 18-30% B in 65min, 30 -60%B in 10 min, 60-90 % B in 10 min, and 90-90% in 5 min) at a flow rate of 250 nl/min. The mass spectrometer was operated in an information dependent acquisition mode whereby following the interrogation of MS data (m/z 350-1600) using a 1 second survey scan, ions were selected for MS/MS analysis based on their intensity (>20 cpm) and charge state (+2, +3, and +4). Total of 3 product ion scans (2,3,3 s each) were set from each survey scan. Rolling collision energies were chosen automatically based on the m/z and charge-state of the selected precursor ions. The IDA Extensions II script was set to one repetition before dynamic exclusion. Identification and quantitation was performed using PRO QUANT software (AB, Forst City, CA) using IPI mouse database (v3.14), with an MS and MS/MS mass tolerance of 0.15 Da. Protein identification with confidence scores of >90% were considered.

Figure 1. Morphology of NR1-0 and NR1-6

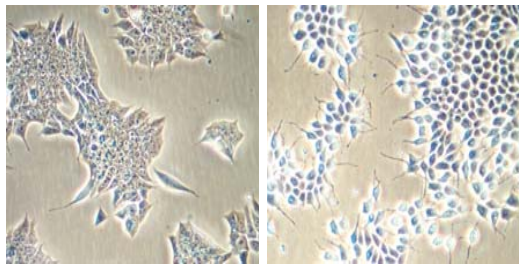
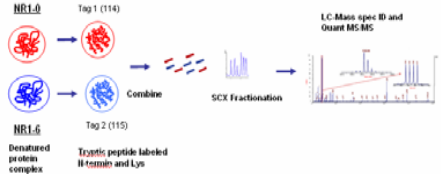


Figure 2. Workflow for the identification of changes in the proteome in EC line, using iTRAQ isobaric tagging



RESULTS

Figure 3. Fractionation of pooled iTRAQ tagged peptides through SCX. Fraction 0.8 ml are collected

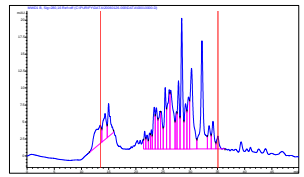


Figure 4. Representative MS/MS spectrum of labeled GVVDSEDLPLNISR obtained from LC-MS/MS duplex analysis of protein mixture. The expansion region shows the "reporter" ion distribution obtained upon peptide fragmentation

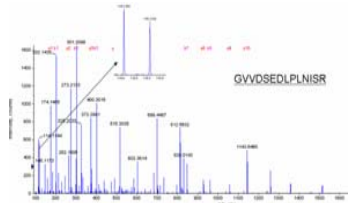


Table 1. Summary of significant protein expression change in NR1-0 *vs.* NR1-6 comparative study

Accession	Protein Name	Ratio 114:115	P Val 114:115
IPI00330804.3	Heat shock protein HSP 90-alpha	1.6695	0.0001
IPI00480560.1	PREDICTED: similar to heat shock protein 8	1.2069	0.0031
IPI00407130.3	NOD-derived CD11c +ve dendritic cells cDNA, RIKEN full-length enriched library, clone:F	1.3099	0.0028
IPI00121788.1	Peroxiredoxin 1	1.2952	0
IPI00462072.2	Alpha-enolase	1.3816	0.0024
IPI00319992.1	78 kDa glucose-regulated protein precursor	1.1986	0.012
IPI00133222.4	60S ribosomal protein L6	1.2705	0.0111
IPI00133903.1	Stress-70 protein, mitochondrial precursor	1.4661	0.0052
IPI00230108.5	CRL-1722 L5178Y-R cDNA, RIKEN full-length enriched library, clone:I730028I05 product:gl	1.3341	0.001
IPI00460069.2	Elongation factor 2	1.5011	0.0001
IPI00225961.4	D-3-phosphoglycerate dehydrogenase	1.391	0.0014
IPI00469481.2	ATP synthase beta chain, mitochondrial precursor	1.2548	0.0028
IPI00459493.4	T-complex protein 1, alpha subunit B	1.252	0.006
IPI00218841.3	Elongation factor 1-gamma	1.1691	0
IPI00223415.3	Activated spleen cDNA, RIKEN full-length enriched library, clone:F830225J05 product:asp	1.5162	0.0015
IPI00554989.1	Peptidyl-prolyl cis-trans isomerase A	1.3723	0.0041
IPI00856269.1	Splice Isoform 2 of 14-3-3 protein theta	1.5128	0.0187
IPI00123862.3	C-1-tetrahydrofolate synthase, cytoplasmic	1.3933	0.0185

Accession	Protein Name	Ratio 114:115	P Val 114:115
IPI00126010.5	Laminin receptor 1	1.2737	0.0455
IPI00224505.6	60S ribosomal protein L13	1.3274	0.0165
IPI00139795.2	60S acidic ribosomal protein P2	1.2442	0.0191
IPI00317740.3	Guanine nucleotide-binding protein beta subunit 2-like 1	1.2723	0.0164
IPI00468688.4	Threonyl-tRNA synthetase, cytoplasmic	1.3757	0.0094
IPI00230602.5	12 days embryo eyeball cDNA, RIKEN full-length enriched library, clone:D230033A18 produ	1.5162	0.0003
IPI00130959.4	Nuclear autoantigenic sperm protein	1.4362	0.0003
IPI00111412.1	60S ribosomal protein L4	1.4166	0.0466
IPI00323806.4	Adult male epididymis cDNA, RIKEN full-length enriched library, clone:9230011J09 produc	1.3397	0.0363
IPI00620156.1	18 kDa protein	1.2448	0.0174
IPI00457704.1	PREDICTED: similar to ribosomal protein L27a	1.4185	0.0159
IPI00115620.1	Phosphoserine aminotransferase	1.3355	0.0354
IPI00224739.1	2 days neonate thymus thymic cells cDNA, RIKEN full-length enriched library, clone:E430	1.1359	0.0153
IPI00132950.1	40S ribosomal protein S21	1.2797	0.0047
IPI00652705.1	12 days embryo spinal ganglion cDNA, RIKEN full-length enriched library, clone:D130054D	1.2059	0.0001
IPI00626390.1	40S ribosomal protein S16	1.1553	0.0104
IPI00118875.3	Eukaryotic translation elongation factor 1 delta, isoform a	1.2085	0.0258
IPI00554850.1	40S ribosomal protein S12	1.4259	0.0214
IPI00660168.1	PREDICTED: similar to 60S ribosomal protein L18	1.4196	0.0288
IPI00459570.1	2 days neonate thymus thymic cells cDNA, RIKEN full-length enriched library, clone:C920	1.2028	0.0004
IPI00114269.1	Glutamate dehydrogenase 1, mitochondrial precursor	1.3495	0.0332
IPI00323819.3	40S ribosomal protein S20	1.3355	0.0002
IPI00122815.3	17 days embryo kidney cDNA, RIKEN full-length enriched library, clone:I920160L24 produc	1.4336	0.0044
IPI00223371.1	Splice Isoform 2 of RNA-binding region containing protein 2	1.3238	0.0367
IPI00347814.2	Hypothetical protein MGC103372	1.1643	0.0483
IPI00312637.4	51 kDa protein	1.7597	0.0216
IPI00355563.2	Fascin	1.5758	0.0371
IPI00116308.1	Hsc70-interacting protein	1.2856	0.0276
IPI00128522.1	Splice Isoform A of Heat-shock protein beta-1	2.2099	0.0102
IPI00123494.3	26S proteasome non-ATPase regulatory subunit 2	1.0988	0.0333
IPI00323353.3	**Uridine phosphorylase 1**	**0.6718**	**0.0036**
IPI00670422.1	PREDICTED: similar to NHP2-like protein 1 (High mobility group-like nuclear protein 2 h	1.1848	0.0149
IPI00606062.2	PREDICTED: similar to eukaryotic translation initiation factor 5A	1.2819	0.0048
IPI00621625.1	13 kDa protein	1.3055	0.0135
IPI00453819.2	Leucyl-tRNA synthetase, cytoplasmic	1.6116	0.005
IPI00119224.1	Small nuclear ribonucleoprotein Sm D3	1.3924	0.0441
IPI00116279.3	T-complex protein 1, epsilon subunit	1.3084	0.0006
IPI00115679.1	Splice Isoform 2 of Neutral alpha-glucosidase AB precursor	1.4276	0.0396
IPI00620417.2	PREDICTED: similar to acidic nuclear phosphoprotein 32 family, member B isoform 2	1.4041	0.0463
IPI00128230.1	Metastasis-associated protein MTA2	2.832	0.0251
IPI00462934.2	Activated spleen cDNA, RIKEN full-length enriched library, clone:F830029A18 product:KH-	1.1979	0.0024
IPI00331173.7	**Adult male thymus cDNA, RIKEN full-length enriched library, clone:5830469E03 product:la**	**0.8143**	**0.0242**
IPI00124444.1	Splice Isoform 1 of 6-phosphofructokinase type C	1.4156	0.0298
IPI00122223.1	DNA topoisomerase 2-alpha	1.1268	0.0432
IPI00153986.2	CRL-1722 L5178Y-R cDNA, RIKEN full-length enriched library, clone:I730071A03 product:Si	1.1672	0.0101
IPI00474805.1	12 kDa protein	1.4799	0.0221
IPI00459898.3	Separin	1.6185	0.0138
IPI00126943.1	**Serine/threonine-protein kinase 38**	**0.332**	**0.0471**
IPI00355960.1	PREDICTED: similar to melanoma antigen, family B, 1	1.6837	0.0239
IPI00137409.3	Transketolase	1.2068	0.0499
IPI00130988.1	7-dehydrocholesterol reductase	1.6028	0.0255
IPI00113911.2	T-cell ecto-ADP-ribosyltransferase 1 precursor	1.6646	0.018

CONCLUSIONS

- iTRAQ isobaric labeling coupled with off-line multidimensional chromatography has been developed and applied for analysis protein changes in a cancer cell line
- Total of 26 SCX fractions were used for the LC-MS/MS analysis resulted in 10498 MS/MS spectra. Total of 3638 spectra were identified with a confidence of 99%
- Over 337 proteins including integral membrane proteins,hydrophobic proteins and proteins with extreme pIs have been identified in an EC mutant cell line.
- Total of 67 up-regulated proteins and 3 down-regulated proteins with statistic significant differences were identified in NR1-0 cell line
- Future work will use other independent approaches such as western, immuno-histochemical and activity assays to confirm and verify the observed changes

ACKNOWLEDGEMENTS

The UAlbany Proteomics Facility is supported by the SUNY Research Foundation. The QSTAR XL mass spectrometer used in this work was funded through the National Institutes of Health under Grant 1S10 RR01782-01A1