Laboratory Investigation (2007) 87, 1113–1124

Comprehensive identification of proteins in Hodgkin lymphoma-derived Reed–Sternberg cells by LC-MS/MS

Jeremy C Wallentine[1], Ki Kwon Kim[1], Charles E Seiler III[1], Cecily P Vaughn[2], David K Crockett[2], Sheryl R Tripp[2], Kojo SJ Elenitoba-Johnson[1,2] and Megan S Lim[1,2]

Mass spectrometry-based proteomics in conjunction with liquid chromatography and bioinformatics analysis provides a highly sensitive and high-throughput approach for the identification of proteins. Hodgkin lymphoma is a form of malignant lymphoma characterized by the proliferation of Reed–Sternberg cells and background reactive lymphocytes. Comprehensive analysis of proteins expressed and released by Reed–Sternberg cells would assist in the discovery of potential biomarkers and improve our understanding of its pathogenesis. The subcellular proteome of the three cellular compartments from L428 and KMH2 Hodgkin lymphoma-derived cell lines were fractionated, and analyzed by reverse-phase liquid chromatography coupled with electrospray ionization tandem mass spectrometry. Additionally, proteins released by Hodgkin lymphoma-derived L428 cells were extracted from serum-free culture media and analyzed. Peptide spectra were analyzed using TurboSEQUEST® against the UniProt protein database (5.26.05; 188 712 entries). A subset of the identified proteins was validated by Western blot analysis, immunofluorescence microscopy and immunohistochemistry. A total of 1945 proteins were identified with 785 from the cytosolic fraction, 305 from the membrane fraction, 441 from the nuclear fraction and 414 released proteins using a minimum of two peptide identifications per protein and an error rate of $<5.0\%$. Identification of proteins from diverse functional groups reflected the functional complexity of the Reed–Sternberg proteome. Proteins with previously reported oncogenic function in other cancers and from signaling pathways implicated in Hodgkin lymphoma were identified. Selected proteins without previously demonstrated expression in Hodgkin lymphoma were validated by Western blot analysis (B-RAF, Erb-B3), immunofluorescence microscopy (Axin1, Tenascin-X, Mucin-2) and immunohistochemistry using a tissue microarray (BRAF, PIM1). This study represents the first comprehensive inventory of proteins expressed by Reed–Sternberg cells of Hodgkin lymphoma and demonstrates the utility of combining cellular subfractionation, protein precipitation, tandem mass spectrometry and bioinformatics analysis for comprehensive identification of proteins that may represent potential biomarkers of the disease.

Laboratory Investigation (2007) **87,** 1113–1124; doi:10.1038/labinvest.3700672; published online 17 September 2007

KEYWORDS: biomarker discovery; Hodgkin lymphoma; shotgun proteomics; tandem mass spectrometry; bioinformatics analysis

Each year approximately 8000 new cases of Hodgkin lymphoma (HL) are diagnosed in the United States.[1] It is one of the most common forms of malignancy affecting young adults, with an average age at diagnosis of 32 years. HL is characterized morphologically by the presence of a distinct population of neoplastic Reed–Sternberg cells (RS), that are surrounded by a non-neoplastic mixed inflammatory infiltrate.[2] Accurate tissue-based diagnosis and staging are critical to the successful treatment of HL. When accurately diagnosed, HL can be treated and cured with local treatment.[3] However, 20–30% of patients experience relapse and die of complications associated with progressive disease and/or treatment.[3–6] Therefore, early diagnosis and the ability to monitor the disease more effectively during and following treatment are important for the management of HL.

Proteins involved in cellular differentiation, activation, cell cycle control, apoptosis regulation and host response have been associated with the clinical outcome of HL.[7–12] Methods previously employed in the identification of biomarkers have included enzyme-linked immunosorbent

[1]Department of Pathology, University of Utah School of Medicine, Salt Lake City, UT, USA and [2]ARUP Institute for Clinical and Experimental Pathology, Salt Lake City, UT, USA

Correspondence: Dr MS Lim, MD, PhD, Department of Pathology, University of Michigan, M5242 Medical Science Building 1, 1301 Catherine Road, Ann Arbor, MI 48105, USA. E-mail: meganlim@med.umich.edu

Received 10 June 2007; accepted 3 August 2007

assays, immunohistochemistry of tissue microarrays, immunoradiometric assays, reverse transcriptase-polymerase chain reaction, gene and protein expression profiling and mass spectrometry-based methods.[13–15] To date, the identification of biomarkers in HL has been limited to known proteins for which there exist antibodies. Furthermore, there are no robust serum biomarkers that are currently in clinical use for the diagnosis and monitoring of patients with HL.

Liquid chromatography-tandem mass spectrometry (LC-MS/MS) represents a uniquely powerful technology that enables the identification and cataloging of large numbers of proteins in complex mixtures, and has been used for biomarker discovery.[16] In addition, simplification of complex proteomes by strategies such as subcellular fractionation and one-dimensional sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) further improves the sensitivity and ability to identify proteins of low abundance, such as cytokines and transcription factors.[17]

Mass spectrometry-based methods have been employed increasingly to aid researchers in biomarker discovery.[18] The speed, sensitivity, accuracy and the ability of LC-MS/MS to identify proteins in complex mixtures is an advantageous combination favoring its use with a wide range of samples which have included various biological fluids and established cancer-derived cell lines.[19–23] Together, LC-MS/MS and high-throughput data analysis represent an important strategy in the field of protein identification and biomarker discovery.

In this study we have utilized a subcellular proteomic approach to identify the proteins expressed by the RS cells of HL-derived cell lines in an attempt to identify potential biomarkers. This study provides the first comprehensive list of proteins expressed by RS cells of HL-derived cells.

MATERIALS AND METHODS

Cell Lines and Cell Culture

The human HL L428 and KMH2 cell lines (DSMZ, Braunschweig, Germany) were grown in RPMI 1640 (Invitrogen Corporation, Carlsbad, CA, USA) with 10% heat inactivated fetal bovine serum (Nova-Tech, Inc., Grand Island, NE, USA) and a penicillin/streptomycin/amphotericin B solution (Invitrogen Corporation). Cultures were grown and maintained at 37°C with 5.0% CO_2 in a humidified atmosphere. Cells were grown to a concentration of 0.75 million cells/ml. Cells to be used in the identification of released proteins were washed three times with serum-free RPMI 1640 media to remove all serum proteins before serum-free culture conditions.

Membrane, Cytoplasmic and Nuclear Subcellular Fractionation

L428 and KMH2 cells were pelleted at 1000 *g* at 4°C. Pellets were washed 3× in ice-cold PBS. Packed cell pellets of 2×10^8 log-phase cells were gently resuspended in residual PBS and aliquoted into four 1.5 ml microcentrifuge tubes. Cytoplasmic and nuclear proteins were extracted from two aliquots using NE-PER Nuclear and Cytoplasmic Extraction kit no. 78833 (Pierce, Rockford, IL, USA). Membrane proteins were extracted from the remaining two aliquots using Mem-PER Eukaryotic Membrane Protein Extraction Reagent kit no. 89826 (Pierce). Cytoplasmic, nuclear and membrane protein fractions were dialyzed against 4°C water for 2 days for a 10^8-fold change in diasylate. Fractions were removed from the dialysis and assayed for total protein concentration using the Peirce BCA™ Protein Assay kit no. 23255. One hundred and twenty micrograms of each fraction was then prepared for electrophoresis using PAGEprep Protein Clean-UP and Enrichment kit. Briefly, 120 μg of protein binding resin was mixed with 120 μg of the subcellular fractions in 50.0% DMSO. Samples were vortexed and incubated at room temperature for 4 min. Sample pellets were centrifuged, washed two times in 50% DMSO and then eluted into elution/loading buffer.

Gel Running, Digestion and Peptide Extraction

One-dimensional SDS-PAGE gels were performed in a 20×20 cm 10.0% gel running at 140 V for 3 h. After coomassie staining and imaging, the gel was destained with destaining solution, washed with deionized water and then rehydrated in ammonium bicarbonate. Lanes of interest were excised into 32 equivalent slices, chopped and dried in a SpeedVac for 40 min, and finally digested and extracted using the Invitrogen protocol. Briefly, each gel slice is destained and washed, crushed and dried, and then rehydrated in ammonium bicarbonate buffer. Each tube was then subject to digestion using Lysine-C/trypsin and Glu-C. Digestion with Lysine-C (Princeton Separations Inc., Adelphia, NJ, USA) was carried out at an enzyme to protein ratio of 1:100 at 37°C for 6 h followed by digestion with trypsin (Promega Corporation, Madison, WI, USA) at a ratio of 1:20 at 37°C overnight. A separate aliquot was digested with Glucine-C (Princeton Separations Inc.) at an enzyme to protein ratio of 1:50 at 37°C overnight. The digested peptides were then extracted from the gel pieces with 50.0% acetonitrile with 0.1% trifluoracetic acid and reduced to a final volume of $\sim$25 μl. All samples were prepared in duplicate. The protein samples collected from the cell culture media were digested using Lysine-C/trypsin and Glu-C as described above.

Collection of Culture Media and Protein Precipitation

To ensure maximal cell viability during serum-free culture conditions, cells were analyzed at multiple time points using the trypan-blue method. Media were collected at 17 h when cell viability was greater than 90.0%. Proteins were precipitated from the media with 10.0% trichloracetic acid (Sigma-Aldrich, St Louis, MO, USA). Following a 2-h precipitation at 4°C, the samples were centrifuged at 10 000 *g* at 4°C for 30 min and washed three times with ether. The precipitated proteins were resuspended in 100 μl of a 100 mM TRIS buffer, pH 9 and concentrated by vacuum centrifugation. Protein concentration was determined using the Coomassie Protein Assay (Pierce).

Table 1 A subset of proteins identified from the subcellular compartments and conditioned media of HL-derived cell lines by LC-MS/MS

UniProt[a]	Protein description	Top X_{corr} (z)	Peptides[b]	Predicted % error	Protein function[c]
Cytoplasm[d]					
Q9HCU9	Breast cancer metastasis-suppressor 1	3.29 (+2)	4	0.0	Mediator of metastasis suppression
Q7Z7G8	Cohen syndrome protein 1	2.15 (+1)	3	0.0	Unknown
Q9Y6C2	EMILIN 1 precursor	2.86 (+2)	6	0.0	Cell adhesion
Q9Y2E4	Hypothetical protein KIAA0934	3.10 (+2)	3	0.0	Unknown
Q13233	Mitogen-activated protein kinase kinase kinase 1	3.66 (+2)	4	0.0	NF-κB pathway activation
P21359	Neurofibromin	3.08 (+2)	11	0.0	GTPase activator activity
Q12769	Nuclear pore complex protein Nup160	3.69 (+2)	6	0.0	Nucleocytoplasmic transporter activity
Q9ULC3	Ras-related protein Rab-23	4.32 (+3)	4	0.0	Unknown
Q9Y3S1	Serine/threonine-protein kinase WNK2	1.95 (+1)	4	0.0	Protein serine/threonine kinase activity
P17948	Vascular endothelial growth factor receptor 1 precursor	3.10 (+2)	9	0.0	Positive regulation of cell proliferation
Q86UP0	Cadherin-24 precursor	3.76 (+2)	5	0.0	Cell–cell adhesion
Q9UPT6	c-Jun-amino-terminal kinase interacting protein 3	3.06 (+2)	3	0.0	Regulation of JNK cascade
Q96AA8	Hypothetical protein KIAA0555	2.65 (+1)	4	0.0	Unknown
P15056	B-Raf proto-oncogene serine/threonine-protein kinase	3.17 (+2)	4	2.0	Protein kinase activity
Q9Y6K8	Adenylate kinase isoenzyme 5	3.09 (+2)	3	1.0	Adenylate kinase activity
P48637	Glutathione synthetase	2.05 (+1)	4	1.0	Response to oxidative stress
Q9Y490	Talin 1	3.38 (+2)	8	1.0	Cytoskeletal anchoring
Q99708	Retinoblastoma-binding protein 8	3.10 (+2)	4	1.0	Cell cycle checkpoint
O15123	Angiopoietin-2 precursor	4.17 (+3)	3	5.0	Signal transduction
Membrane					
Q9Y618	Nuclear receptor corepressor 2	3.67 (+2)	6	0.0	Notch binding
P51610	Host cell factor C1	2.09 (+1)	7	0.0	Transcription coactivator activity
Q9H165	B-cell lymphoma/leukemia 11A	3.42 (+2)	4	0.0	Myeloid and B-cell proto-oncogene
Q9NSY1	BMP-2 inducible protein kinase	3.51 (+2)	4	0.0	Osteoblast differentiation
Q9Y6X2	Protein inhibitor of activated STAT protein 3	3.76 (+2)	2	1.0	Transcriptional coregulation in the STAT pathway
P35713	Transcription factor SOX-18	3.36 (+2)	4	1.0	Transcription activation
Q9Y3A4	Hypothetical protein CGI-96	4.50 (+3)	3	5.0	Unknown
Q92794	MYST histone acetyltransferase 3	3.40 (+2)	5	5.0	Histone acetyltransferase
Nucleus					
Q9NRY6	Phospholipid scramblase 3	3.55 (+2)	2	0.0	Phospholipid scrambling
P41220	Regulator of G-protein signaling 2	3.52 (+2)	3	0.0	Regulation of G-protein signaling pathway
O14709	Zinc finger protein 197	3.24 (+2)	5	0.0	Transcriptional regulation
P25054	Adenomatous polyposis coli protein	2.17 (+1)	3	0.0	Tumor suppressor
Q9UGN5	Poly (ADP-ribose) polymerase-2	4.06 (+3)	2	0.0	DNA repair

Table 1 Continued

UniProt[a]	Protein description	Top X_{corr} (z)	Peptides[b]	Predicted % error	Protein function[c]
Q9UNK9	Protein KIAA0759	3.21 (+2)	2	0.0	Unknown
P21675	Transcription initiation factor TFIID subunit 1	3.56 (+2)	2	2.0	Protein amino-acid autophosphorylation
P17038	Zinc finger protein 43	3.83 (+2)	4	2.0	Transcriptional regulation
Released					
Q8N4C8	MAPK/ERK kinase kinase kinase 6	3.86 (+3)	5	0.0	JNK cascade kinase activity
Q9BXA7	Serine/threonine-protein kinase 22A	2.04 (+1)	2	3.0	Protein serine/threonine kinase activity
P46531	Notch 1	3.23 (+2)	5	3.0	Notch signaling pathway
P51617	Interleukin-1 receptor-associated kinase 1	4.06 (+3)	2	0.0	NF-κB-inducing kinase activity
O95835	Serine/threonine protein kinase LATS1	3.07 (+2)	2	0.4	Protein serine/threonine kinase activity
Q13164	Mitogen-activated protein kinase 7 (ERK-5) (ERK4)	2.21 (+1)	3	0.7	MAP kinase activity
O15169	Axin 1	3.61 (+3)	4	1.0	Signal transducer activity
P46013	Antigen KI-67	3.08 (+2)	2	1.3	Cell proliferation and regulation of cell cycle
P51451	Tyrosine-protein kinase BLK (B lymphocyte kinase)	2.92 (+2)	3	3.0	Protein tyrosine kinase activity

[a]Protein accession numbers taken from the Uniprot database downloaded at http://www.pir.uniprot.org.

[b]Number of unique peptides identified by LC-MS/MS.

[c]Protein functions were obtained through searches performed at http://www.expasy.org using the Uniprot accession number.

[d]Protein identifications from the subcellular compartments were obtained from the cell lysate of L428 and KMH2 HL-derived cell lines. Protein identifications from the released fraction were obtained from the conditioned media of L428 HL-derived cells.

Protein Digestion

The protein samples collected from the cell culture media were digested using Lysine-C/trypsin and Glu-C. Digestion with Lysine-C (Princeton Separations Inc.) was carried out at an enzyme to protein ratio of 1:100 at 37°C for 6 h followed by digestion with trypsin (Promega Corporation) at a ratio of 1:20 at 37°C overnight. A separate aliquot was digested with Glucine-C (Princeton Separations Inc.) at an enzyme to protein ratio of 1:50 at 37°C overnight.

Tandem Mass Spectrometry

Digested samples were analyzed in triplicate by the LCQ Deca XP ion trap mass spectrometer (ThermoElectron Corporation, San Jose, CA, USA). Fifteen microliters of sample was injected by the autosampler into a reverse-phase column (75 μm ID fused with silica packed with 10 cm of 5 μm C18 particles). The samples were eluted through the column using a 3-h acetonitrile gradient (0–60% solution B in 2 h; solution A: 5.0% acetonitrile, 0.4% acetic acid and 0.005% heptafluorobutyric acid (HFBA); solution B: 95% acetonitrile, 0.4% acetic acid and 0.005% HFBA) followed by electrospray ionization. Mass spectrometry scans were performed from 400 to 2000 m/z followed by MS/MS scans of the three most abundant peptides in each MS scan. Dynamic exclusion was set to a repeat count of 2 with an exclusion duration of 3 min.

Protein Database Searching

The acquired MS spectra for the samples were searched using TurboSEQUEST® (BioWorks 3.1 SR1; Thermo Electron Corporation) against the UniProt database (5.26.05 download; 188 712 entries). Searches were performed with either trypsin or Glu-C specified as the enzyme, with an allowance for up to two missed cleavage sites. Peptide mass tolerance was set at 1.4. Acceptance levels for positive peptide identifications were determined using cross-correlation scores (X_{corr}) and Δ-correlation scores (ΔC_n). The minimum acceptable X_{corr} was 1.8 for $+1$ peptides, 2.5 for $+2$ peptides and 3.5 for $+3$ peptides with a $\Delta C_n \geq 0.100$.

The identified peptides were statistically validated using INTERACT™, PeptideProphet™ and ProteinProphet™.[24,25] These software tools provide an empirical statistical model which estimates the accuracy of peptides identified by SEQUEST™. ProteinProphet also allows for the analysis of peptide data obtained from multiple samples increasing the

probability of achieving positive identifications. Receiver operator curves with corresponding sensitivity and error rates were produced for each dataset and were used to determine the rate of false positives ($\leq 5.0\%$). Proteins were ranked based on the produced probability scores. Additionally, the Gene Ontology Miner (GoMiner™) software tool was used to categorize the identified proteins according to molecular function and subcellular location.[26]

Cell Lysis and Western Blot Analysis

Cell pellets were subjected to lysis with RIPA buffer with 0.1% protease inhibitor cocktail (Sigma-Aldrich). The RIPA buffer was added to the cell pellets, incubated on ice for 20 min and centrifuged at 15 000 *g* for 10 min at 4°C. The BCA Protein Assay kit was used to determine the protein concentration from the cell lysates (Pierce). The protein from each sample (50 μg per lane) was loaded and resolved by 8.0% SDS-PAGE. The proteins were transferred from the gel to a nitrocellulose membrane using the semi-dry transfer method. Nitrocellulose membranes were then equilibrated and blocked in 5% non-fat dried milk (NFDM) overnight at 4°C.

Protein from L428 media was collected as described previously under the Collection of Culture Media and Protein Precipitation. Briefly, media were collected at similar time points, subjected to precipitation by the 10.0% trichloracetic acid method, quantified by the BCA Protein Assay kit and resolved by 8.0% SDS-PAGE. Proteins were then transferred to a nitrocellulose membrane as described above.

Primary antibodies specific to our proteins of interest were selected (Table 1). Antibodies to the following proteins were used: ErbB3 and BRAF (Santa Cruz Biotechnology Inc., Santa Cruz, CA, USA). The membranes were gently rocked for 1 h in 5.0% NFDM at room temperature. Membranes were then washed three times in 1X TTBS for 10–15 min. The secondary antibody was then added in 5.0% NFDM and gently rocked for 1 h at room temperature. Membranes were then washed three times in 1X TTBS for 10–15 min with the final wash in 1X TBS. The blots were visualized with Chemiluminescence luminol reagent (Santa Cruz Biotechnology Inc.).

Immunofluorescence Analysis

Immunofluorescence microscopy (IFM) was performed to validate the expression of selected proteins. Cells were washed with serum-free media and plated in four-well plates at a density of 1.0×10^6 cells/ml. Cells were fixed with 3.7% formaldehyde, permeabilized and blocked with 0.1% Triton X. The cells were subject to immunofluorescence analysis using the following primary antibodies: CD100, Axin1, Tenascin-X and Mucin2 (Santa Cruz Biotechnology Inc.). All antibodies were diluted 1:200 in PBS with 0.1% Triton X and incubated with cells for 1 h at room temperature. The cells were washed three times with PBS, followed by incubation with the secondary antibodies at a dilution of 4 μg/ml, for 1 h in the dark. Control wells were used for each of the secondary antibodies, Alexa Fluor® 488 donkey anti-goat IgG and fluorescein goat anti-rabbit IgG (Molecular Probes, Eugene, OR, USA). The cells were visualized by confocal fluorescent microscopy; fluorescent signal thresholds were set using the control wells.

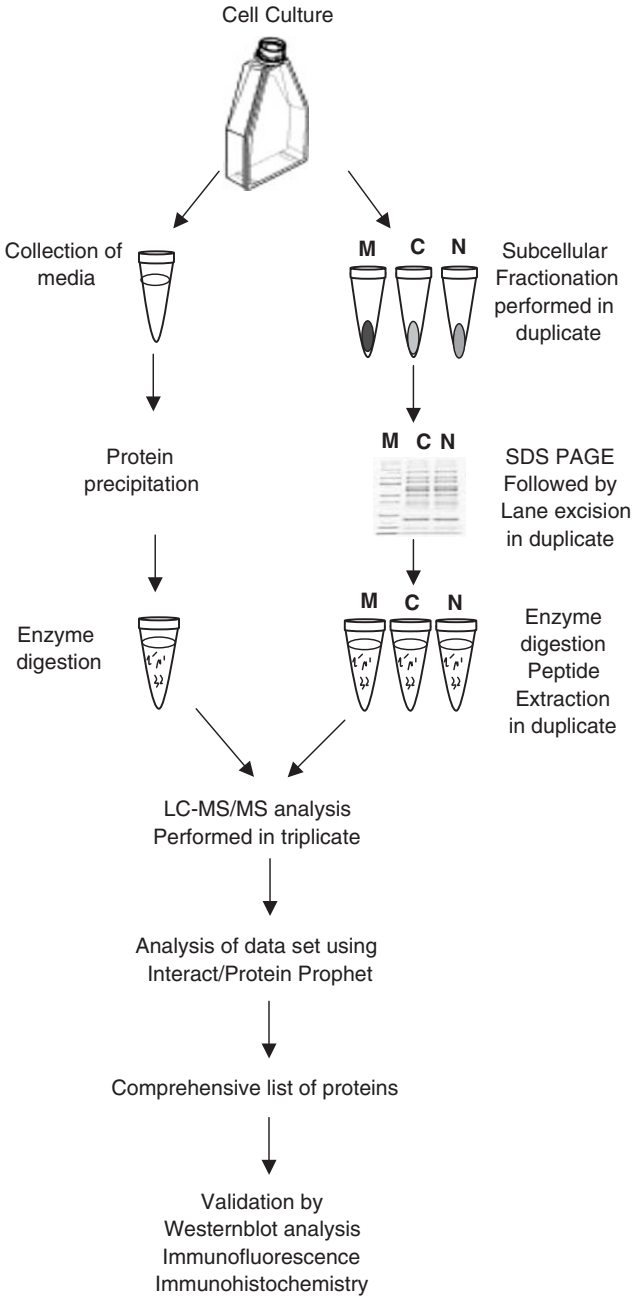
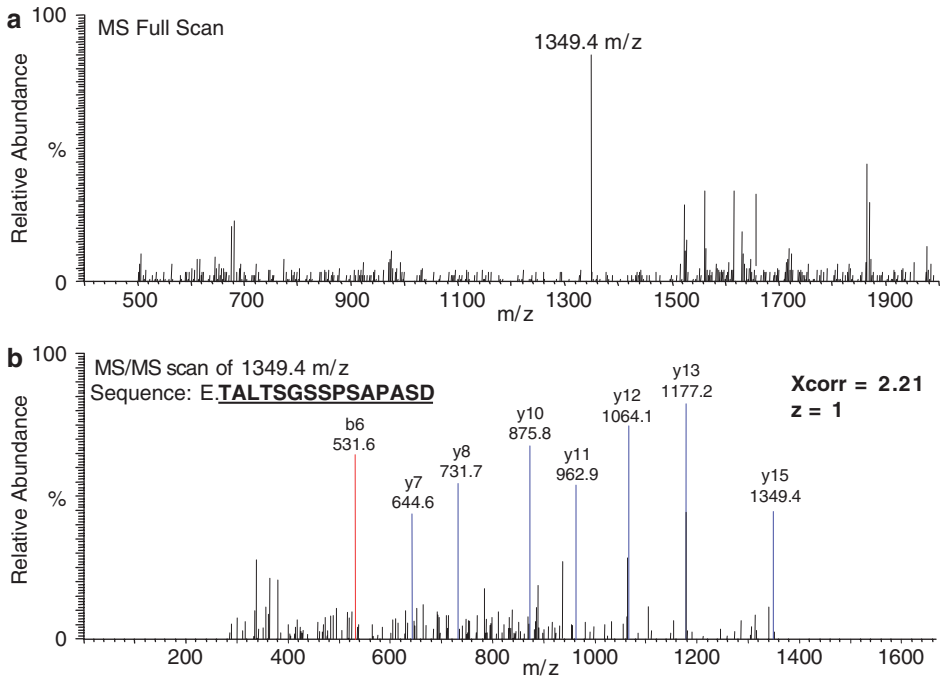


Figure 1 Experimental design used for the identification of proteins from the Hodgkin lymphoma-derived L428 and KMH2 cell lines by LC-MS/MS.

Construction of Tissue Microarrays and Immunohistochemical Studies

Tumor specimens from HL and reactive tissues including chronic tonsillitis and lymphadenitis were obtained from the surgical pathology files of the Department of Pathology, University of Utah School of Medicine, Salt Lake City, UT,



AA		B ions	Y ions	
T	1	102.11	1349.39	15
A	2	173.19	1248.28	14
L	3	286.35	1177.21	13
T	4	387.45	1064.05	12
S	5	474.53	962.94	11
G	6	531.58	875.86	10
S	7	618.66	818.81	9
S	8	705.74	731.73	8
P	9	802.85	644.65	7
S	10	889.93	547.54	6
A	11	961.01	460.46	5
P	12	1058.13	389.38	4
A	13	1129.21	292.26	3
S	14	1216.29	221.19	2
D	15	1331.37	134.11	1

c **Protein:** BRAF_HUMAN(P15056) B-Raf proto-oncogene serine/threonine-protein kinase

MAALSGGGGGGAEPGQALFNGDMEPEAGAGAGAAASSAADPAIPEEVWNIKQMIKLTQEHIEALLDKFGGEHNPPSIYLEAYEEYTSKLDAL
QQREQQLLESLGNGTDFSVSSSASMDTVTSSSSSSLSVLPSSLSVFQNPTDVARSNPKSPQKPIVRVFLPNKQRTVVPARCGVTVRDSLKKAL
MMRGLIPECCAVYRIQDGEKKPIGWDTDISWLTGEELHVEVLENVPLTTHNFVRKTFFTLAFCDFCRKLLFQGFRCQTCGYKFHQRCSTEVPL
MCVNYDQLDLLFVSKFFEHHPIPQEEASLAE**TALTSGSSPSAPASD**SIGPQILTSPSPSKSIPIPQPFRPADEDHRNQFGQRDRSSSAPNVHI
NTIEPVNIDDLIRDQGFRGDGGSTTGLSATPPASLPGSLTNVKALQKSPGPQRERKSSSSSEDRNRMKTLGRRDSSDDWEIPDGQITVGQRI
GSGSFGTVYKGKWHGDVAVKMLNVTAPTPQQLQAFKNEVGVLRKTRHVNILLFMGYSTKPQLAIVTQWCEGSSLYHHLHIIETKFEMIKLIDI
ARQTAQGMDYLHAKSIIHRDLKSNNIFLHEDLTVKIGDFGLATVKSRWSGSHQFEQLSGSILWMAPEVIRMQDKNPYSFQSDVYAFGIVLYE
LMTGQLPYSNINNRDQIIFMVGRGYLSPDLSKVRSNCPKAMKRLMAECLKKKRDERPLFPQILASIELLARSLPKIHRSASEPSLNRAGFQTED
FSLYACASPKTPIQAGGYGAFPVH

Figure 2 Tandem mass spectrometry sequencing of tryptic peptides identifying the protein BRAF_HUMAN (P15056). (**a**) Data-dependent MS full scans are followed by (**b**) MS/MS sequencing scans of top peptide candidates that lead to the identification of the (**c**) 15 amino-acid peptide.

USA, from the period 1995 to 1998. This study was approved by the Institutional Review Board (IRB no. 11188) of the University of Utah. Tissues were formalin-fixed and paraffin-embedded for histological diagnosis and immunohistochemical (IHC) study. Cases were categorized according to the World Health Organization classification of lymphoid neoplasms[27] and reviewed by two hematopathologists (MSL and KSJE-J). Tissue microarrays were manually constructed using the 2-mm needle with the tissue microarray work station (Beecham Instruments, Hackensack, NJ, USA). Five micrometer-thick serial sections were mounted on glass slides coated with 2.0% aminopropyltrioxysilane in acetone. Sections were dewaxed in xylene and rehydrated in graded ethanols. Endogenous peroxidase activity was blocked by immersion in 0.3% methanolic peroxide for 15 min. Antigen–antibody reactions were visualized with diaminobenzi-

dine as the chromogen. Normal mouse serum containing mixed immunoglobulins at a concentration approximating that of the primary antibodies (PIM1-1:60, B-RAF-1:40) was used as a negative control. Sections were counterstained with hematoxylin. A normal tonsil was used as a positive control for all antibodies.

RESULTS

In an attempt to comprehensively identify the proteins expressed by HL-derived RS cells, a subcellular proteomic-based approach was performed. The experimental design is outlined in Figure 1. Briefly, HL-derived cell lines were grown, pelleted, lysed, fractionated, separated by one-dimensional SDS-PAGE, subjected to enzymatic digestion and analyzed by LC-MS/MS. In addition, proteins released into serum-free culture were collected and analyzed by LC-MS/MS. Tandem mass spectrometry spectra were analyzed using the TurboSEQUEST software against the UniProt non-redundant human protein database. The data were then combined using INTERACT/PeptideProphet/ProteinProphet. Analysis of the RPMI media alone, using the same criteria, yielded a negligible number of proteins ($n = 7$), of which none were found in the L428 serum-free media.

In total, 67 662 spectra were generated from the LC-MS/MS analysis of all L428 and KMH2 samples. These spectra represented 16 732 peptides when searched against the UniProt database using criteria described in Materials and methods. From these peptides we identified a total of 1945 proteins (785 from the cytosolic fraction, 305 from the membrane fraction, 441 from the nuclear fraction and 414 released proteins) using a minimum of two peptide identifications per protein, and an error rate cutoff of $<5.0\%$ (INTERACT probability score ≥ 0.80). Figure 2 depicts the experimental progression in the LC-MS/MS sequencing of Glu-C peptides identifying one of these proteins, BRAF. A significant number of common proteins were identified in each of the three L428 and KMH2 fractions and is depicted graphically in Figure 3. Table 1 provides a list of the selected proteins identified from each of the three subcellular fractions and a list of released proteins identified from the serum-free media of L428 cells. Notably, many proteins previously reported to be expressed by HL-RS cells, such as CD30, CD21, CD22, CD45 and CD79a, were identified. Among the proteins identified are several components of known signaling pathways such as NF-κB, Notch and Ras/Raf/MAPK/ERK, thought to have important roles in the pathogenesis of HL. A selected list of these proteins and their corresponding pathways are provided in Table 2.

The GoMiner software tool was employed to group the proteins according to cellular function as shown in Figure 4. The proteins identified in the released, membrane, cytoplasmic and nuclear fractions were of diverse function. In each of the four fractions the majority were involved in binding, catalytic and signaling activities (Figure 4).

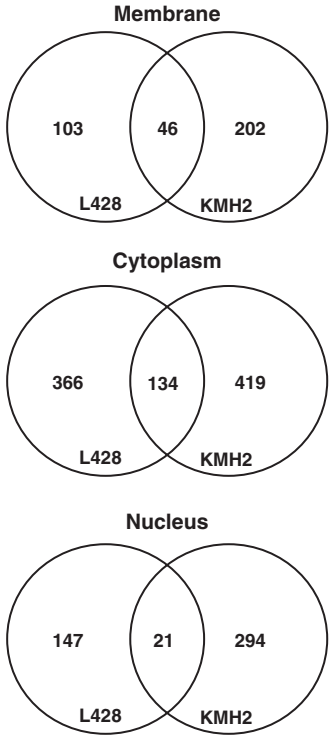


Figure 3 LC-MS/MS analysis of L428- and KMH2 HL-derived cell lines reveals proteomic similarities. The number of proteins identified in each of the L428 and KMH2 fractions is depicted in the circles with the overlapping area representing the number of common proteins identified between the two cell lines.

Seven proteins were selected for validation by Western blot (WB), IFM and IHC analysis, and are listed in Table 3. Proteins for validation were selected based on criteria including error rate $<5.0\%$, cellular location, availability of commercial antibodies, novelty and potential use as biomarkers of disease.

Figure 5a demonstrates two WBs in which an 84 kDa band corresponding to the full-length BRAF protein and a 68 kDa band consistent with the proposed 68 kDa extracellular domain of ErbB-3 were expressed in the lysate of L428 cells. Figure 5b demonstrates the expression of Axin1, Tenascin-X, Mucin2 and CD100 by IFM of L428 cells. Immunofluorescent staining of Axin1 exhibits a moderately intense cytoplasmic distribution of the protein. The expression of Tenascin-X confirmed by IFM exhibits a diffuse cytoplasmic staining pattern. The expression of Mucin2 confirmed by IFM exhibits a cytoplasmic distribution of the protein. Finally, the expression of CD100 confirmed by IFM exhibits moderately intense areas of staining corresponding to a localized membranous distribution of the protein. IHC analysis using a tissue microarray confirmed the expression of BRAF and PIM1 by HL-RS cells (Figure 5c).

DISCUSSION

The work described herein represents the first unbiased characterization of proteins expressed by L428 and KMH2

Table 2 Identification of proteins with known expression in Hodgkin lymphoma by LC-MS/MS

UniProt[a]	Protein description	Top X_{corr} (z)	Peptides[b]	Predicted % error	Protein function[c]
NF-κB pathway					
O75791	GRB2-related adaptor protein 2 (P38)	3.12 (+2)	3	0.0	Ras protein signal transduction
P51617	Interleukin-1 receptor-associated kinase 1	3.62 (+3)	3	0.0	NF-B-inducing kinase activity
Q9BXA7	Serine/threonine-protein kinase 22A	2.07 (+1)	2	3.0	Protein serine/threonine kinase activity
Q9UBC1	NF-B inhibitor-like protein 1	2.89 (+2)	2	45.0	Negative regulator of NF-B activation
P28908	CD30	2.16 (+1)	3	45.0	Regulates gene expression through activation of NF-B.
Ras/Raf/MAPK/ERK pathway					
Q8N4C8	MAPK/ERK kinase kinase kinase 6	3.86 (+3)	5	0.0	Protein serine/threonine kinase activity
Q13164	ERK-5	2.71 (+2)	3	1.0	Involved in signal transduction through MAP kinase activity
P28482	ERK-2	3.98 (+2)	2	1.0	Induction of apoptosis through MAP kinase activity
P15056	B-Raf proto-oncogene serine/threonine-protein kinase	3.17 (+2)	4	2.0	Transduction of anti-apoptotic mitogenic signals from membrane to nucleus
P53778	ERK-6	2.18 (+1)	2	14.0	Plays a role in myoblast differentiation through MAP kinase activity
Q12851	MAPK/ERK kinase kinase kinase 2	2.95 (+2)	3	20.0	Protein serine/threonine kinase activity
P45985	MAP kinase kinase 4	3.18 (+2)	4	20.0	Activates MAPK8 and MAPK9 as well as MAPK14 (p38)
Q15078	p35	2.98 (+2)	2	26.0	Protein serine/threonine kinase activator activity
P45984	c-Jun N-terminal kinase 2	3.77 (+3)	3	26.0	Protein binding and Jun kinase activity
Q99683	MAPK/ERK kinase kinase 5	1.93 (+1)	2	31.0	Phosphorylates and activates MAP2K4 and MAP2K6
O95819	MAPK/ERK kinase kinase kinase 4	1.97 (+1)	7	28.0	Serine/threonine kinase that acts in response to cytokines
Q9Y6R4	MAPK/ERK kinase kinase 4	2.96 (+2)	3	38.0	Activation of MAPKK activity
P36507	MAPK/ERK kinase 2	4.13 (+3)	4	38.0	Activates the ERK1 and ERK2 MAP kinases
P45983	c-Jun N-terminal kinase 1	2.12 (+1)	3	45.0	Jun kinase activity
Notch pathway					
Q04721	Notch 2	2.98 (+2)	4	2.0	Receptor for membrane-bound ligands Jagged1, Jagged2 and Delta1 to regulate cell-fate determination
P46531	Notch 1	3.23 (+2)	5	3.0	Receptor for membrane-bound ligands Jagged1, Jagged2 and Delta1 to regulate cell-fate determination
P78504	Jagged1	4.15 (+3)	4	7.0	Ligand for multiple Notch receptors
Q9Y219	Jagged2	3.09 (+2)	2	7.0	Ligand for multiple Notch receptors
Q9HCC6	Transcription factor HES-4	3.01 (+2)	4	26.0	Transcriptional repressor induced upon activation of notch
Q9Y543	Transcription factor HES-2	4.13 (+3)	2	31.0	Transcriptional repressor induced upon activation of notch
P23769	Endothelial transcription factor GATA-2	3.93 (+3)	3	38.0	Early transcription factor important in the mediation of notch signaling

[a]Protein accession numbers taken from the Uniprot database downloaded at http://www.pir.uniprot.org.
[b]Number of unique peptides identified by LC-MS/MS.
[c]Protein functions were obtained through searches performed at http://www.expasy.org using the Uniprot accession number.

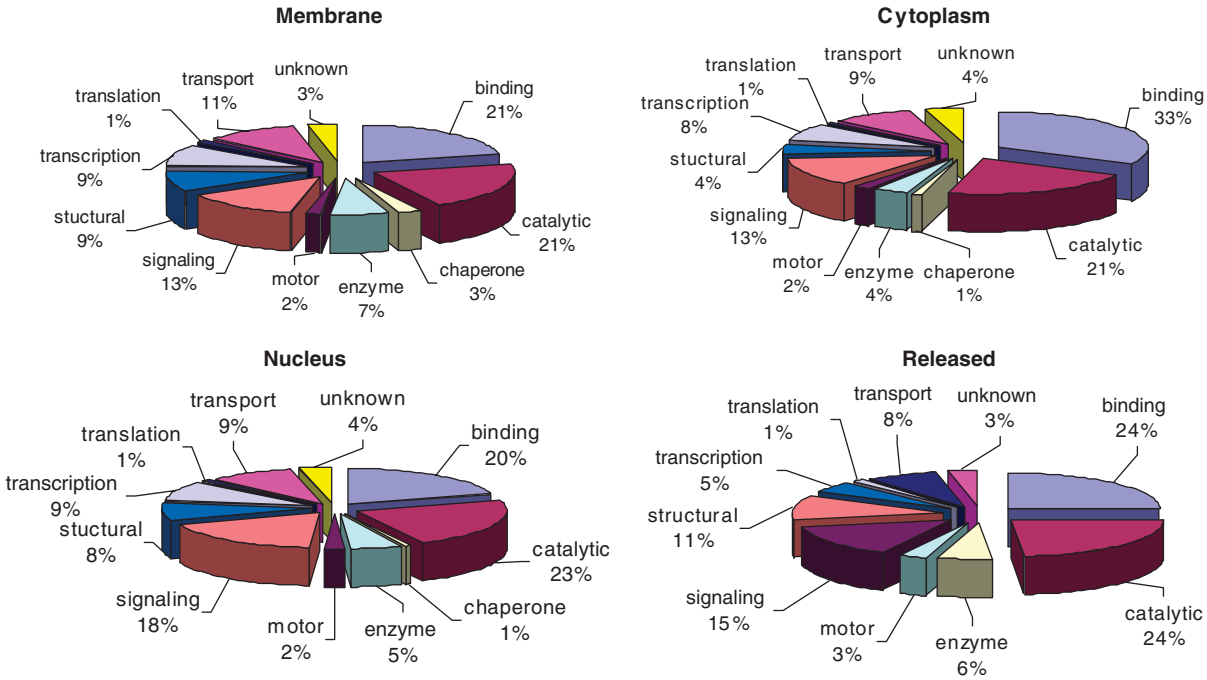


Figure 4 Functional categories of proteins expressed from the released, membrane, cytoplasmic and nuclear fractions. Categorization of protein identifications from the L428 and KMH2 fractions according to function was performed with the assistance of GoMiner and reveals diverse functional categories.

Table 3 Selected proteins validated by WB analysis, IHC and IFM

UniProt[a]	Protein	Observed fraction(s)[b]	Antibody source[c]
P15056	BRAF	KMH2 cytoplasm L428 cytoplasm, released	Santa Cruz, C-19 (WB) Abgent, AP7810a (IHC)
P11309	PIM1	KMH2 nucleus	Abgent, 7932a (IHC)
P21860	ErbB3	L428 released	Santa Cruz, C-17 (WB)
O15169	Axin1	L428 cytoplasm, released	Santa Cruz, sc-8568 (IFM)
Q02817	Mucin2	KMH2 cytoplasm L428 released	Santa Cruz, P18 (IFM)
P22105	Tenascin-X	KMH2 nucleus L428 cytoplasm, released	Santa Cruz, R-20 (IFM)
Q92854	CD100	L428 released	Santa Cruz, C-19 (IFM)

IFM, immunofluorescence microscopy; IHC, immunohistochemical; WB, Western blot.

[a]Protein accession numbers obtained from the Uniprot database downloaded at http://www.pir.uniprot.org.

[b]Cell line and fraction from which the listed protein was identified by LC-MS/MS.

[c]Antibodies were obtained from the following commercial venders: Santa Cruz Biotechnology Inc., www.scbt.com and ABGENT www.abgent.com.

RS-derived HL cell lines using a subcellular LC-MS/MS-based approach. The cell lines utilized in this study were derived from HL patients in the late stages of disease and represent the neoplastic B-lineage RS-cells of HL. In this study we identified 1945 proteins with high statistical confidence, using two or more peptide identifications and a maximum error rate cutoff of $<5.0\%$. By employing duplicate subcellular fractionation techniques, two distinct enzymes for digestion, and analyzing each duplicate sample in triplicate by LC-MS/MS, we improved our ability to identify many low abundance proteins, which may have otherwise been missed in our analysis.

Many of the proteins identified in our data have previously been demonstrated to be expressed in HL providing independent validation of our approach. In addition, our data contain many proteins that have not been previously demonstrated to be expressed in HL, representing possible novel participants in the pathogenesis of HL. Proteins of particular interest are those involved or associated with cell signaling and proliferation. We found multiple proteins within our list directly involved within the NF-κB, Notch and Ras/Raf/MAPK/ERK signaling pathways (Table 2). Of note are multiple components of the Notch signaling cascade (Notch1-2, Jagged1-2, GATA-2,5 and HES-4,2). Notch signaling regulates self-renewal and cell lineage decisions in many different tissues.[28] Aberrant Notch activation has been shown to promote neoplastic transformation in many cell types, including T-cell acute lymphoblastic leukemia, anaplastic large-cell lymphoma, pancreatic cancer and neuro-

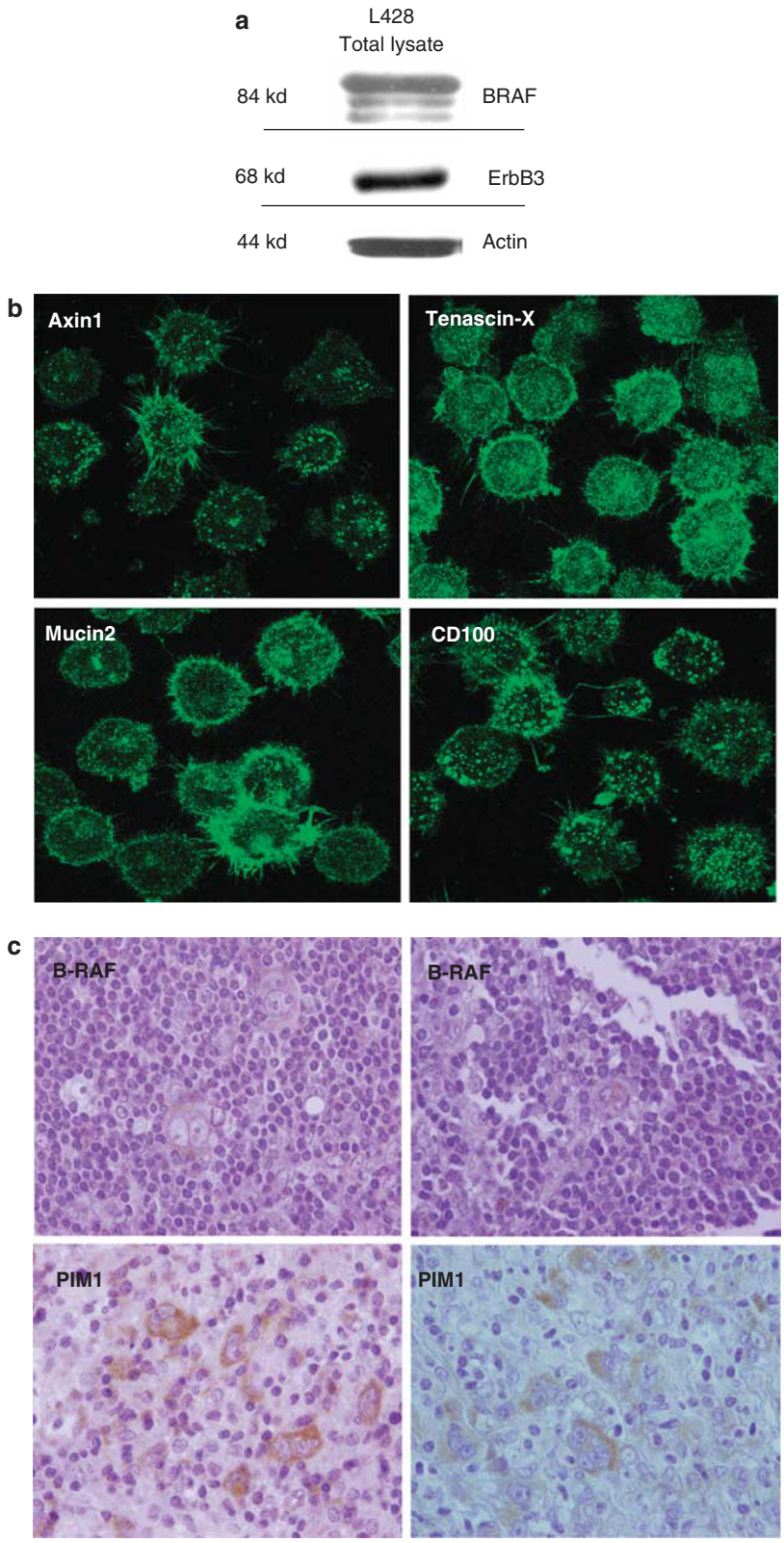


Figure 5 (**a**–**c**) Validation of select proteins identified by LC-MS/MS by WB analysis, IFM and IHC. (**a**) Protein validation of the 84 kDa BRAF full-length protein and the proposed 68-kDa extracellular product of ErbB-3 is demonstrated by WB analysis. Cell lysate was obtained from the HL-derived L428 cell line, resolved by SDS-PAGE and transferred to a nitrocellulose membrane for chemiluminescence and visualization. (**b**) HL-derived L428 cells were plated, fixed, permeabilized, blocked and subject to immunofluorescence analysis using the following antibodies; Axin1, Tenascin-X, Mucin2 and CD100. (**c**) Protein validation by IHC using a tissue microarray demonstrates the specific cytoplasmic expression of BRAF and PIM1 in HL-derived RS cells.

blastoma.[29,30] Recently its role in the pathogenesis of HL has been demonstrated.[31] NF-κB and Ras/Raf/MAPK/ERK are signaling pathways also implicated in the pathogenesis of HL.[2,32]

To validate the MS-based protein identification strategy, we used a combination of WB analysis, IFM and IHC analysis for selected candidate proteins (Figure 5a–c). Our studies revealed the expression of BRAF, PIM1 and ErbB-3 in the cell lysate of the HL-derived L428 cells. Expression of these proteins in RS cells has not been reported previously. BRAF is an 84 kDa serine/threonine protein kinase within the Ras/Raf/MAPK/ERK signaling cascade that is involved in the control of cell growth and behavior. Active mutations have been shown to play a role in a wide variety of human cancer development through increased activation of downstream kinases.[33,34] PIM1 is a serine/threonine kinase involved in several biological functions including cell survival, proliferation and differentiation.[35] PIM1 has been shown to be involved in several hematopoietic cancers, and has recently been identified as a target of aberrant somatic hypermutation in diffuse large B-cell lymphomas.[35,36] Epstein–Barr virus (EBV) has been shown to induce the upregulation of PIM kinases which is thought to contribute to the ability of EBV to immortalize B cells and predispose them to malignant transformation.[37] ErbB-3 is a 148 kDa type I membrane receptor tyrosine kinase protein. Overexpression of ErbB-3 has been demonstrated in breast cancer and its expression demonstrated in multiple myeloma.[38–40] The expression of these proteins is significant in that they suggest the activity of other important signaling pathways in the pathogenesis of HL.

In addition, the expressions of Axin1, Tenascin-X, Mucin-2 and CD100 have not been demonstrated previously in the RS cells of HL. These proteins have important roles in a myriad of cell activities including cell survival, signaling, and the mediation of cell–cell and cell–matrix interactions. Axin1 is a 95 kDa cytoplasmic protein involved in signal transduction within the Wnt signaling pathway.[41] Defective Wnt signaling has been shown to promote tumourigenesis and tumor progression with axin acting as the primary limiting factor in the pathway.[41] Tenascin-X is a member of the tenascin family of extracellular matrix proteins which function in modulating cell adhesion, migration and growth, and may play a role in supporting the growth of tumors.[42,43] Mucin2 is a 540 kDa secreted protein that is overexpressed in intraductal papillary mucinous tumors of the pancreas and bile duct cystadenocarcinomas of the liver.[44] Lastly, CD100 is a 96 kDa type I membrane/secreted protein involved in immunomodulation, B-cell activation and survival, and monocyte migration.[45] These findings again suggest the activity of other proteins and signaling pathways that may be acting independently and/or in concert with previously established pathways in the development of HL.

In conclusion, the present study demonstrates the first unbiased characterization of expressed and released proteins by HL-derived cells utilizing a subcellular proteomic methodology. Its validity is demonstrated by the identification of multiple proteins with previously demonstrated expression in HL and the validation of several novel proteins. In addition, our dataset contains an abundant number of novel proteins, representing possible participants in the pathogenesis of HL. These proteins serve as a working inventory of candidate proteins for the identification of possible biomarkers to be used in the diagnosis and monitoring of HL.

1. Baris D, Zahm SH. Epidemiology of lymphomas. Curr Opin Oncol 2000;12:383–394.
2. Lee IS, Kim SH, Song HG, *et al.* The molecular basis for the generation of Hodgkin and Reed–Sternberg cells in Hodgkin's lymphoma. Int J Hematol 2003;77:330–335.
3. Jacobs P. Hodgkin's disease and the malignant lymphomas. Dis Mon 1993;39:213–297.
4. Hasenclever D, Diehl V. A prognostic score for advanced Hodgkin's disease. International Prognostic Factors Project on Advanced Hodgkin's Disease. N Engl J Med 1998;339:1506–1514.
5. Pfreundschuh MG, Rueffer U, Lathan B, *et al.* Dexa-BEAM in patients with Hodgkin's disease refractory to multidrug chemotherapy regimens: a trial of the German Hodgkin's Disease Study Group. J Clin Oncol 1994;12:580–586.
6. Sureda A, Arranz R, Iriondo A, *et al.* Autologous stem-cell transplantation for Hodgkin's disease: results and prognostic factors in 494 patients from the Grupo Espanol de Linfomas/Transplante Autologo de Medula Osea Spanish Cooperative Group. J Clin Oncol 2001;19:1395–1404.
7. Clodi K, Asgari Z, Younes M, *et al.* Expression of CD40 ligand (CD154) in B and T lymphocytes of Hodgkin disease: potential therapeutic significance. Cancer 2002;94:1–5.
8. Fischer M, Bijman M, Molin D, *et al.* Increased serum levels of interleukin-9 correlate to negative prognostic factors in Hodgkin's lymphoma. Leukemia 2003;17:2513–2516.
9. Giles FJ, Vose JM, Do KA, *et al.* Circulating CD20 and CD52 in patients with non-Hodgkin's lymphoma or Hodgkin's disease. Br J Haematol 2003;123:850–857.
10. Montalban C, Garcia JF, Abraira V, *et al.* Influence of biologic markers on the outcome of Hodgkin's lymphoma: a study by the Spanish Hodgkin's Lymphoma Study Group. J Clin Oncol 2004;22:1664–1673.
11. Newcom SR, Ansari AA, Gu L. Interleukin-4 is an autocrine growth factor secreted by the L-428 Reed–Sternberg cell. Blood 1992;79: 191–197.
12. Newcom SR, Kadin ME, Ansari AA, *et al.* L-428 nodular sclerosing Hodgkin's cell secretes a unique transforming growth factor-beta active at physiologic pH. J Clin Invest 1988;82:1915–1921.
13. Blagosklonny MV, Neckers LM. The role of Bcl-2 protein and autocrine growth factors in a human follicular lymphoma-derived B cell line. Eur Cytokine Netw 1995;6:21–27.
14. Jimeno A, Hidalgo M. Molecular biomarkers: their increasing role in the diagnosis, characterization, and therapy guidance in pancreatic cancer. Mol Cancer Ther 2006;5:787–796.
15. Salles G, Bienvenu J, Bastion Y, *et al.* Elevated circulating levels of TNFalpha and its p55 soluble receptor are associated with an adverse prognosis in lymphoma patients. Br J Haematol 1996;93:352–359.
16. Lim MS, Elenitoba-Johnson KS. Mass spectrometry-based proteomic studies of human anaplastic large cell lymphoma. Mol Cell Proteomics 2006;5:1787–1798.
17. Jiang XS, Zhou H, Zhang L, *et al.* A high-throughput approach for subcellular proteome: identification of rat liver proteins using subcellular fractionation coupled with two-dimensional liquid chromatography tandem mass spectrometry and bioinformatic analysis. Mol Cell Proteomics 2004;3:441–455.
18. Lim MS, Elenitoba-Johnson KS. Proteomics in pathology research. Lab Invest 2004;84:1227–1244.
19. Theodorescu D, Wittke S, Ross MM, *et al.* Discovery and validation of new protein biomarkers for urothelial cancer: a prospective analysis. Lancet Oncol 2006;7:230–240.

20. Lin YW, Lin CY, Lai HC, *et al.* Plasma proteomic pattern as biomarkers for ovarian cancer. Int J Gynecol Cancer 2006;16(Suppl 1):139–146.
21. Li J, Zhao J, Yu X, *et al.* Identification of biomarkers for breast cancer in nipple aspiration and ductal lavage fluid. Clin Cancer Res 2005;11:8312–8320.
22. Fujii K, Kondo T, Yokoo H, *et al.* Protein expression pattern distinguishes different lymphoid neoplasms. Proteomics 2005;5: 4274–4286.
23. Wu SL, Hancock WS, Goodrich GG, *et al.* An approach to the proteomic analysis of a breast cancer cell line (SKBR-3). Proteomics 2003;3: 1037–1046.
24. Keller A, Nesvizhskii AI, Kolker E, *et al.* Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. Anal Chem 2002;74:5383–5392.
25. Nesvizhskii AI, Keller A, Kolker E, *et al.* A statistical model for identifying proteins by tandem mass spectrometry. Anal Chem 2003;75: 4646–4658.
26. Zeeberg BR, Feng W, Wang G, *et al.* GoMiner: a resource for biological interpretation of genomic and proteomic data. Genome Biol 2003;4:R28.
27. Jaffe ES, Harris NL, Stein H, *et al.* Pathology and Genetics of Tumours of Haematopoietic and Lymphoid Tissues 2001. Lyon: IARC Press.
28. Harper JA, Yuan JS, Tan JB, *et al.* Notch signaling in development and disease. Clin Genet 2003;64:461–472.
29. Weng AP, Aster JC. Multiple niches for Notch in cancer: context is everything. Curr Opin Genet Dev 2004;14:48–54.
30. Pahlman S, Stockhausen MT, Fredlund E, *et al.* Notch signaling in neuroblastoma. Semin Cancer Biol 2004;14:365–373.
31. Jundt F, Anagnostopoulos I, Forster R, *et al.* Activated Notch1 signaling promotes tumor cell proliferation and survival in Hodgkin and anaplastic large cell lymphoma. Blood 2002;99:3398–3403.
32. Knoops L, Renauld JC. IL-9 and its receptor: from signal transduction to tumorigenesis. Growth Factors 2004;22:207–215.
33. Lee JW, Yoo NJ, Soung YH, *et al.* BRAF mutations in non-Hodgkin's lymphoma. Br J Cancer 2003;89:1958–1960.
34. Smalley KS. A pivotal role for ERK in the oncogenic behaviour of malignant melanoma? Int J Cancer 2003;104:527–532.
35. Kumar A, Mandiyan V, Suzuki Y, *et al.* Crystal structures of proto-oncogene kinase Pim1: a target of aberrant somatic hypermutations in diffuse large cell lymphoma. J Mol Biol 2005;348:183–193.
36. Montesinos-Rongen M, Van Roost D, Schaller C, *et al.* Primary diffuse large B-cell lymphomas of the central nervous system are targeted by aberrant somatic hypermutation. Blood 2004;103:1869–1875.
37. Rainio EM, Ahlfors H, Carter KL, *et al.* Pim kinases are upregulated during Epstein–Barr virus infection and enhance EBNA2 activity. Virology 2005;333:201–206.
38. Bieche I, Onody P, Tozlu S, *et al.* Prognostic value of ERBB family mRNA expression in breast carcinomas. Int J Cancer 2003;106:758–765.
39. Quinn CM, Ostrowski JL, Lane SA, *et al.* c-erbB-3 protein expression in human breast cancer: comparison with other tumour variables and survival. Histopathology 1994;25:247–252.
40. Walters DK, French JD, Arendt BK, *et al.* Atypical expression of ErbB3 in myeloma cells: cross-talk between ErbB3 and the interferon-alpha signaling complex. Oncogene 2003;22:3598–3607.
41. Salahshor S, Woodgett JR. The links between axin and carcinogenesis. J Clin Pathol 2005;58:225–236.
42. Chiquet-Ehrismann R, Chiquet M. Tenascins: regulation and putative functions during pathological stress. J Pathol 2003;200:488–499.
43. Chiquet-Ehrismann R. Tenascins, a growing family of extracellular matrix proteins. Experientia 1995;51:853–862.
44. Yonezawa S, Sato E. Expression of mucin antigens in human cancers and its relationship with malignancy potential. Pathol Int 1997;47: 813–830.
45. Elhabazi A, Marie-Cardine A, Chabbert-de Ponnat I, *et al.* Structure and function of the immune semaphorin CD100/SEMA4D. Crit Rev Immunol 2003;23:65–81.