Journal of
Proteomics & Bioinformatics

Suehara et al., J Proteomics Bioinform 2013, S3
DOI: 10.4172/jpb.S3-001

Review Article Open Access

Proteomic Technologies to Develop Biomarkers and Functional Analyses for Bone and Soft Tissue Tumors

Yoshiyuki Suehara[1,2*], Shinji Kohsaka[2], Daisuke Kubota[1], Kenta Mukaihara[1,3], Keisuke Akaike[1,4], Reiko Mineki[5], Tsutomu Fujimura[5], Kazuo Kaneko[1], Marc Ladanyi[2], Tsuyoshi Saito[4] and Tadashi Kondo[3]

[1]*Department of Orthopedic Surgery, Juntendo University School of Medicine, Tokyo, Japan*
[2]*Department of Pathology, Memorial Sloan-Ketteing Cancer Center, NY, USA*
[3]*Division of Pharmacoproteomics, National Cancer Center Research Institute, Tokyo, Japan*
[4]*Department of Human Pathology, Juntendo University School of Medicine, Tokyo, Japan*
[5]*Laboratory of Biochemical Analysis, Central Laboratory of Medical Sciences, Juntendo University School of Medicine, Tokyo, Japan*

Abstract

Proteomics suggests that global protein expression studies can provide important clues for developing biomarkers and understanding tumor biology that cannot be obtained using other approaches. Proteomic studies, such as gel-based analyses and mass spectrometry-based analyses, have provided protein expression profiles that can be used to develop novel diagnostic and therapeutic biomarkers, allowing for the molecular classification of tumors. Recently, we used proteomic approaches to develop biomarkers for bone and soft tissue tumors and identified novel biomarkers for predicting the prognosis and chemosensitivity of bone and soft tissue tumors. Although the predictive power of these biomarkers has been confirmed in large validation studies, functional analyses of the biomarkers (proteins) remain to be conducted.

In this article, we describe our proteomics methodology for identifying biomarkers and our approach to evaluating the functions of the biomarkers (proteins) and provide a few examples of our recent proteomic studies.

Keywords: Proteomics; Bone and soft tissue sarcomas; 2D-DIGE; GeLC-MS; NPM1; MYC

Introduction

Bone and soft tissue sarcomas are rare malignant tumors [1]. Patients who exhibit a poor response to chemotherapy and develop metastasis continue to have a poor prognosis. Therefore, it is critical to identify proteins associated with tumor malignancy and chemoresistance as predictive biomarkers and novel targets in patients with bone and soft tissue tumors.

The use of high-throughput screening approaches, such as array-based comparative genomic hybridization analyses and cDNA microarray technology, allows for the screening of several thousand DNA and mRNA sequences and can be used to identify genes relevant to the diagnosis and clinical features of tumors [2-14]. Comprehensive studies have identified several genes that may be involved in the development or progression of tumors, representing candidate biomarkers, and/or drug targets [2-14]. However, DNA sequencing and measurement of the mRNA expression alone cannot be used to detect posttranslational modifications of proteins, such as phosphorylation or glycosylation, or differences in protein stability, factors that play important roles in the malignant behavior of tumor cells [15-18]. Furthermore, many lines of evidence have indicated discordance between the mRNA expression and the protein expression [15-18]. Therefore, proteomic studies are critical tools for understanding the biology of tumors, as well as identifying biomarkers for various cancers. These difficulties undermine the potential advantages of global protein expression studies, an approach known as "proteomics".

Standard proteomic techniques, such as two-dimensional gel electrophoresis (2DE) and mass spectrometry (MS), have been developed over the past three decades. Since the end of the 1990s, due to the development of high-throughput platforms, proteomics has allowed the simultaneous measurement of multiple protein products and protein modifications. Recently, our studies successfully identified various candidate proteins associated with the differential diagnosis [17,19-21], prognosis [18,21-27] and prediction of the response to chemotherapy [18,23,28] in patients with bone and soft tissue tumors. We also verified the predictive power of these variables using large validation cohorts to develop clinical applications of useful biomarkers. Most of the biomarkers were successfully confirmed; however, the roles of the proteins in the tumors remain unknown and functional analyses of the biomarkers (proteins) have yet to be conducted. Therefore, we performed functional studies of these biomarkers as ongoing proteomic studies.

The following section describes (i) proteomic technologies, (ii) how proteomic approaches have been applied to identify biomarkers in bone and soft tissue tumors, and (iii) our proteomic approaches to conducting functional analyses of biomarkers (proteins), followed by a few examples of our recent proteomic studies.

Proteomic Technologies

Proteomics is the large-scale study of proteins, including their structures and functions [29-32]. Unlike studies of a single protein or pathway, proteomic methods enable the researcher to obtain a systematic overview of the profiles of the expressed proteins, which in

***Corresponding author:** Yoshiyuki Suehara, Department of Orthopedic Surgery, Juntendo University School of Medicine Hongo 2-1-1, Bunkyo-ku, Tokyo 113-8421, Japan, Tel +81-3-3813-3111; Fax: +81-3-3813-3428; E-mail: ysuehara@juntendo.ac.jp

Received September 10, 2013; **Accepted** October 22, 2013; **Published** October 25, 2013

Citation: Suehara Y, Kohsaka S, Kubota D, Mukaihara K, Akaike K, et al. (2013) Proteomic Technologies to Develop Biomarkers and Functional Analyses for Bone and Soft Tissue Tumors. J Proteomics Bioinform S3: 001. doi:10.4172/jpb.S3-001

cases involving tumors, can ultimately improve the diagnosis, prognosis and management of the patient by revealing protein interactions affecting overall tumor progression [29-32]. Technologies used in proteomics research include electrophoresis, mass spectrometric technologies, protein labeling, protein arrays, antibody-based approaches, imaging and bioinformatics technology. In particular, mass spectrometry technologies are now high-throughput, allowing for the rapid and accurate identification of thousands of proteins present within a complex tumor specimen. Therefore, various technologies are now being employed to identify tumor-specific proteins in sarcomas using proteomics technologies. In this section, we briefly describe two-dimensional difference gel electrophoresis (2D-DIGE) and GeLC-MS [33,34], as these technologies are the most frequently used methods for obtaining protein expression profiles in our proteomic studies [17-28].

2D-DIGE

We routinely employ 2D-DIGE for biomarker identification using surgical samples [17-28]. 2D-DIGE is an advanced variation of 2D-PAGE (two-dimensional polyacrylamide gel electrophoresis) that has the potential to address many of the drawbacks of classical 2D-PAGE [31,32]. 2D-DIGE is frequently applied in sarcoma proteomics, in which the overall features of the protein expression are correlated with the sarcoma phenotypes to identify the molecular background of cancer biology. 2D-DIGE generates 2,000-5,000 protein spots as quantitative proteomic data [31,32].

In 2D-DIGE, proteins are extracted from surgical samples and all protein samples are labeled with different fluorescent dyes before gel electrophoresis (Figure 1). We create a common internal control sample that includes a mixture of a small portion of all individual samples and label it with a fluorescent dye that differs from the dyes used to label the individual samples. The differently labeled internal control and individual samples are then mixed together and separated according to both the pH and molecular weight ranges using 2D-PAGE. Laser scanning can be used to obtain gel images, because all proteins are labeled with fluorescent dye before gel electrophoresis. These gel images provide data regarding protein spots as protein expression profiles. Protein spots whose intensity statistically differs between the groups examined are identified using software programs in each study [17-28]. Proteins corresponding to the spots of interest are identified using mass spectrometry.

GeLC-MS

GeLC-MS involves SDS-PAGE, followed by in-gel tryptic digestion and liquid chromatography-tandem mass spectrometry [33,34]. The technology is a powerful approach for conducting proteomic analyses, and the method directly acquires protein profiles consisting of intact proteins (not protein spots). In our GeLC-MS approaches, the technology identifies 1,500-2,000 protein expressions as semiquantitative proteomic data in one run. We usually employ GeLC-MS technology in functional analyses of bone and soft tissue sarcomas.

Using this technique, a protein sample for the analysis is separated using SDS-PAGE, and the entire gel lanes are excised and further subdivided into smaller sections (Figure 2). We usually slice each gel into 24 slices. The proteins in these gel sections are subsequently

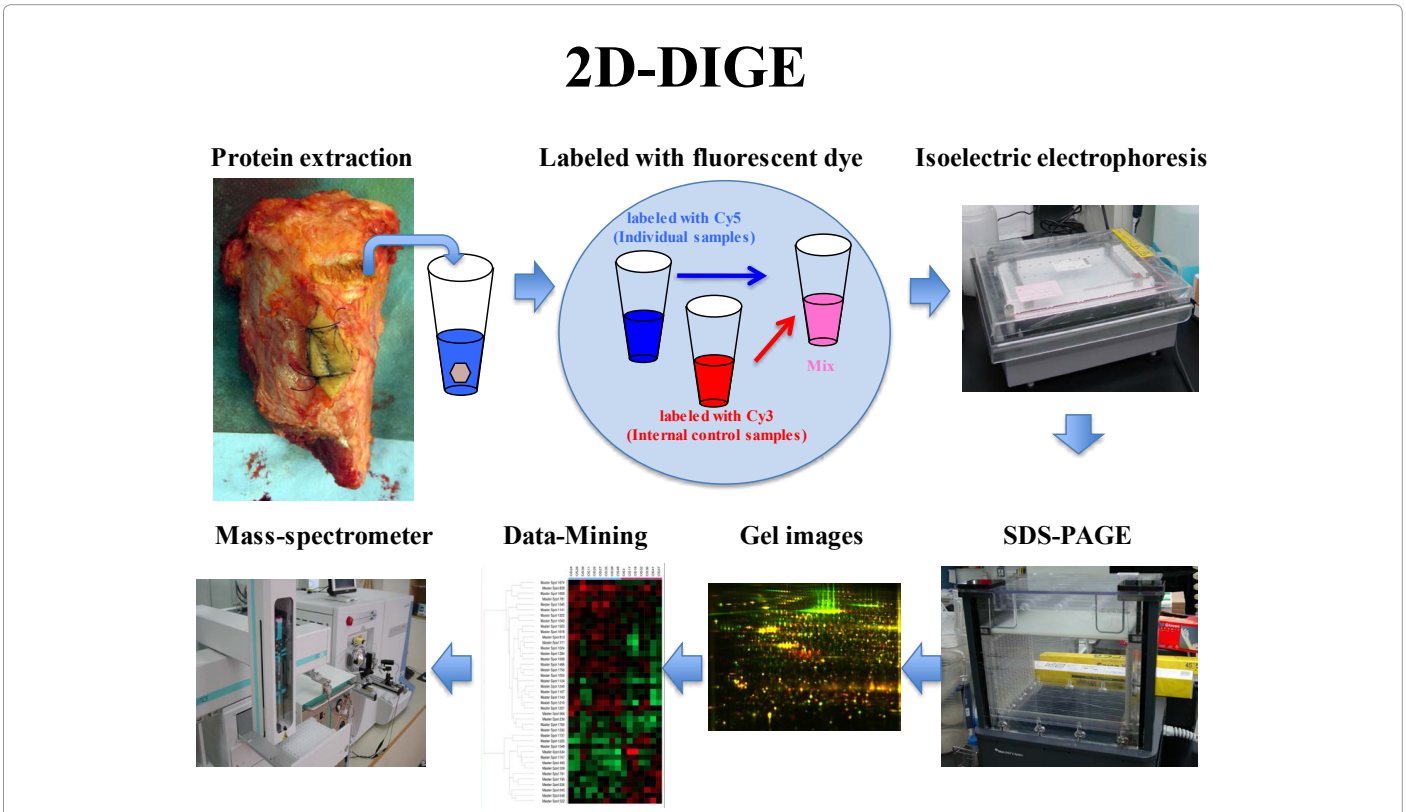

Figure 1: 2D-DIGE: Proteins extracted from surgical samples. All protein samples are labeled with different fluorescent dyes. The internal control sample, a mixture of a small portion of all individual samples is labeled by Cy3, and the individual samples are labeled by Cy5. The differently labeled samples are then mixed together. The samples are separated according to both the pH and molecular weight ranges. Gel images are then acquired using laser scanning. Finally, interest protein spots selected using data mining are identified in the intact proteins using a mass spectrometer.

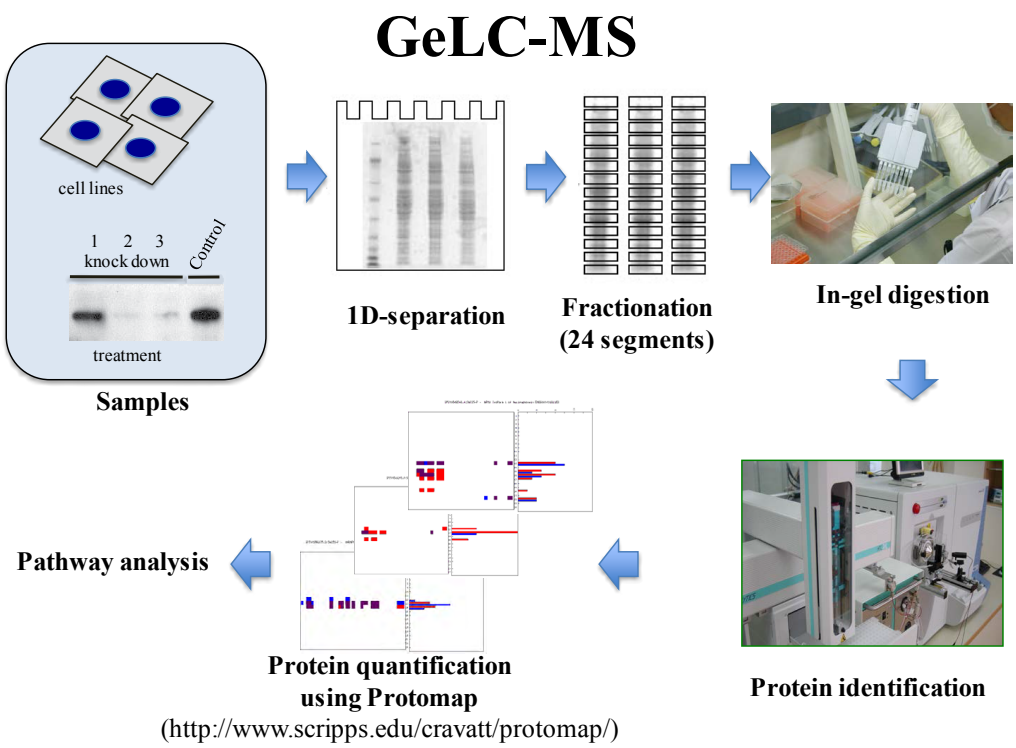

Figure 2: GeLC-MS: Proteins are extracted from cell lines, including normal cell lines and treated cell lines (for example, those treated with siRNA). All protein samples are separated using SDS-PAGE. The gels are sliced into 24 gels in each lane. The cut gels are digested using in-gel trypsinization. The acquired peptides are analyzed using LC-MS to identify proteins. The results obtained from all slices are combined to generate semiquantified data using the Protomap method [35].

digested within the gel using trypsin. In addition, the generated peptides are analyzed using an LC-MS experiment to acquire information regarding peptide sequence coverage, and the spectral count values in order to identify proteins present in a particular sample of each study. The database search results for all slices of a biological sample are combined, yielding global protein identification and semiquantification for each sample using the Protomap method [35]. The Protomap method provides a rich set of protein data that reveal global changes in the volume, size, topography and abundance of proteins in complex biological samples.

A comparison of the 2D-DIGE and GeLC-MS methods used for our proteomic studies

With respect to the comparison between the 2D-DIGE and GeLC-MS methods, there are two important differences: "quantification" and "protein identification". The 2D-DIGE can provide accurate quantification of protein spots, but the method cannot demonstrate the protein identity directly. Therefore, the protein spots need to be assessed by an additional process to identify the protein names. On the other hand, the GeLC-MS can provide all of protein names directly based on the profiles. However, the GeLC-MS cannot provide accurate quantification because it is only semi-quantitative. Our studies include the discovery of biomarkers and a functional analysis of the findings from the discovery studies. In the discovery study, we usually employ 2D-DIGE to identify novel biomarkers, because we need to obtain the exact expression profiles. In functional studies, we need to know the identity of the most abundant proteins that are related to the protein expression dynamics, including upregulation, downregulation and no change. Therefore, we usually use GeLC-MS for the functional analyses.

Identification of Biomarkers and Therapeutic Targets in Soft Tissue Tumors

Identifying predictive biomarkers and drug targets for tumors is the most important goal of global protein and gene expression studies. Current gene expression profiling technologies have been used to identify upregulated or downregulated genes with prognostic value that can be used to predict the prognosis or chemosensitivity of soft tissue sarcomas [3,4,11-14].

In order to identify useful biomarkers using global protein expression studies, we conduct high-integrity and reliable studies consisting of three sets (Figure 3): (1) a discovery set that attempts to identify candidate biomarkers from the global protein expression profiles of the tissue samples (in our studies, we usually use 2D-DIGE for these analyses); (2) a confirmation set that is used to confirm the protein expression differences identified in the discovery set using other proteomic tools (in our studies, we usually use a Western blot analysis); (3) a validation set that is used to verify the predictive power of a biomarker on a large scale using numerous samples in order to develop biomarkers for clinical application (in our studies, we usually use immunohistochemistry and Western blot analyses).

With respect to the number of samples included in the discovery set, we usually employ 10 to 20 samples (example 10 vs 10, 7 vs 8, 5 vs 5, and so on) to develop the novel biomarkers. Using a large number of samples may generate abundant protein profiles, and then these results may provide a large amount of information that can be used to choose candidate novel biomarkers. However, we believe that it is critical for the discovery analyses to eliminate noise from samples, even

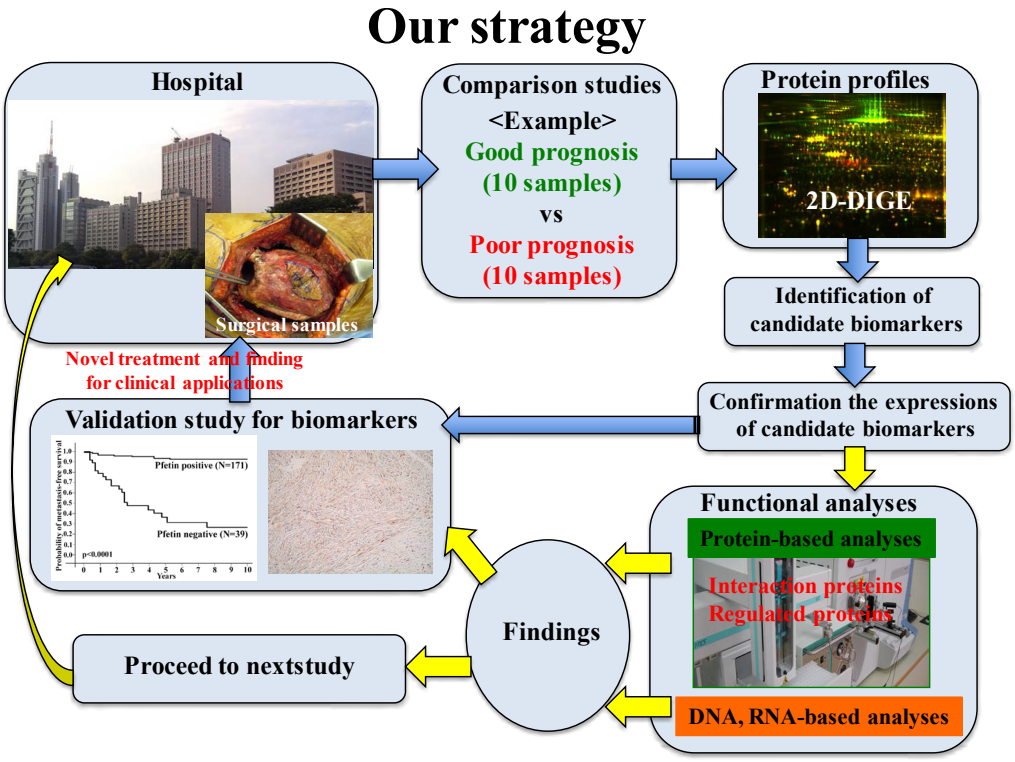

Figure 3: Our strategy for conducting proteomic studies using bone and soft tissue sarcomas is herein described. To develop biomarkers (blue arrows), we usually employ a three-step process: (i) 2D-DIGE-based target identification, (ii) confirmation, and (iii) validation. For the functional analyses (yellow arrows), we employ protein-based analyses (proteomic technologies) and DNA- and RNA-based analyses. In this article, we described the protein-based analyses used for the functional studies of the identified biomarkers (proteins).

To develop biomarkers (blue arrows), surgical samples are collected from patients with bone and soft tissue tumors. We organize both the clinical samples and information to establish efficient strategies. Protein expression profiles are generated using 2D-DIGE and analyzed using data mining to identify biomarker candidates. The protein expression levels of the candidates are confirmed using Western blotting analyses, and/or immunohistochemistry. The diagnostic value of the biomarker candidates is verified using additional large variation cohorts. Finally, the validated biomarkers are subjected to novel clinical applications.

In the functional analyses (yellow arrows), we focus on both the interaction proteins and regulated proteins associated with the biomarker proteins as proteomic approaches. The novel findings generated by the functional analyses are verified in validation studies, and/or are used in subsequent studies. Finally, we hope that the novel findings will provide beneficial effects to patients.

if the sample set will be small. A noisy sample can easily obstruct the identification of novel findings, and provides incorrect results. In our experience, sample sets of 10 to 20 are able to identify novel biomarkers in the bone and soft tissue tumors successfully. Therefore, we believe our strategies regarding the samples are acceptable for sarcoma research.

In this section, we introduce pertinent proteomic studies that have been previously used to identify prognostic biomarkers for GISTs, synovial sarcomas and Ewing's sarcomas, and chemosensitivity biomarkers for osteosarcomas.

GISTs

Gastrointestinal stromal tumors (GISTs) are the most common mesenchymal tumors of the gastrointestinal tract and are characterized by the expression of the kit oncogene. The tyrosine kinase inhibitor, imatinib, has been proven to be highly effective in treating these tumors [36,37].

In order to identify protein expression profiles that correlate with the prognosis of GISTs, we conducted a quantitative expression study of the intact proteins in GIST samples [24]. We compared the protein expression profiles between a poor prognosis group (eight cases) and a good prognosis group (nine samples). These comparisons identified 43 protein spots with different intensities in the two types of samples. Eight of the 43 protein spots corresponded to pfetin and had higher intensity in the good prognosis group. We confirmed the expression of pfetin using Western blot analyses.

As validation studies, we verified the expression of pfetin in 210 GIST cases using immunohistochemistry. These studies revealed 5-year metastasis-free survival rates of 93.9% and 36.2% for the patients with pfetin-positive and pfetin-negative tumors, respectively ($P<0.0001$) [24]. Univariate and multivariate analyses demonstrated the pfetin expression to be an independent prognostic factor in patients with GISTs. These results demonstrate that the pfetin expression can be used to correctly distinguish poor prognosis cases from good prognosis cases and suggest that pfetin is a useful biomarker that may contribute to the development of novel therapeutic strategies for treating GIST patients.

Synovial sarcoma

Synovial sarcomas are malignant mesenchymal tumors that are primarily characterized by the presence of a chromosomal translocation, t(X;18)(p11.2;11.2), representing the fusion of the SYT gene with *SSX1*, *SSX2* or *SSX4* [1].

In our study, we used a proteomic approach to develop prognostic biomarkers for synovial sarcomas using 2D-DIGE [22]. We used 13 surgical samples (obtained from eight synovial sarcoma patients with a good prognosis and five synovial sarcoma patients with a poor prognosis), and identified 20 protein spots whose intensity statistically differed between the two groups. Mass spectrometric protein identification demonstrated that these 20 spots corresponded to 17 distinct gene products. Three of the 20 spots corresponded to secernin-1 and had higher intensity in the good prognosis group.

With respect to validation studies, the prognostic performance of secernin-1 was also examined immunohistochemically in 45 synovial sarcoma patients. The 5-year survival rates were 77.6% and 21.8% for the patients with secernin-1-positive and -negative primary tumors, respectively ($p<0.01$). We concluded that secernin-1 may be used as a biomarker to predict overall and metastasis-free survival in synovial sarcoma patients.

Ewing's sarcoma

Ewing's sarcomas are malignant neoplasms of the bone and soft tissue. Ewing's sarcomas are genetically characterized by the presence of EWS-FLI1 or another related gene fusion, and recent studies suggest that Ewing's sarcomas may arise from the malignant transformation of mesenchymal, and/or neural crest stem cells [1].

Kikuta et al. [27] reported that the protein expression level of nucleophosmin (NPM1) is correlated with the prognosis of Ewing's sarcoma.That study investigated the global protein expression profiles of Ewing's sarcomas using 2D-DIGE and found statistically significant differences in the NPM1 protein expression levels between Ewing's sarcoma patients with a poor prognosis and those with a good prognosis.

Furthermore, the prognostic performance of nucleophosmin was evaluated immunohistochemically in an additional 34 Ewing's sarcoma cases. A univariate analysis revealed that the expression of NPM1 was significantly correlated with the overall survival ($P<0.01$). Additionally, in 29 of the 34 patients with localized disease at diagnosis, the univariate analysis demonstrated that NPM1 positivity was also a strong negative predictor of the overall survival ($P<0.01$). These results suggest that the expression of NPM1 defines a more aggressive subset of Ewing's sarcoma patients and is a candidate prognostic marker for Ewing's sarcoma.

Osteosarcoma

Osteosarcoma is the most common primary malignant bone tumor. It most frequently occurs in the second decade of life, with 60% of patients being under 25 years of age [1]. The response to preoperative chemotherapy provides critical information regarding the patient, and chemosensitive patients are divided into two groups based on the pathological percentage of necrosis [1].

To identify novel biomarkers of the chemosensitivity of osteosarcoma, we employed a proteomic approach (2D-DIGE) [18,23,38]. We generated protein profiles of 12 biopsy samples, including six poor chemosensitivity osteosarcomas and six good chemosensitivity osteosarcomas, according to the Huvos grading system. We compared the expression profiles between the two groups and found 55 spots that corresponded to 38 distinct proteins, including peroxiredoxin 2 (PRDX2). The protein expression of PRDX2 exhibited higher intensity in the poor responder group.

In order to validate the predictive value for chemosensitivity, we conducted a validation study using a Western blot analysis of additional osteosarcoma samples. The validation study also demonstrated that the poor responders had higher PRDX2 expression levels than the good responders. We concluded that PRDX2 is a candidate marker for chemosensitivity in osteosarcoma patients.

Functional Analyses of Biomarkers

We previously reported that our proteomic approaches successfully identified various novel biomarkers for predicting the prognosis and chemosensitivity of bone and soft tissue tumors [17-28]. However, the predictive power of these biomarkers must be confirmed in large validation studies and functional analyses of the biomarkers (proteins) remain to be conducted. Therefore, we continue to research functional analyses of our identified biomarkers using proteomic technologies to identify their functions and roles in tumors. We usually focus on interaction proteins and regulated proteins (Figure 3). Hence, in this section, we describe (i) the identification of interaction proteins, and (ii) the identification of regulated proteins, as well as (iii) demonstrate our the results of our functional analyses of NPM1 in Ewing's sarcoma using these proteomic technologies.

Identification of interaction proteins

Protein–protein interaction (PPI) networks provide valuable information regarding the understanding of cellular functions and biological processes [39-42]. With the tremendous increase in human protein interaction data, a network approach is used to understand the molecular mechanisms of disease, particularly with regard to cancer phenomena [39-42]. In the setting of cancer, PPI data provide insight into the distinct topological features of cancer genes, cancer classification and cancer-related subnetworks [39-42]. PPI data form signaling nodes and hubs that transmit pathophysiological cues along molecular networks that also provide integrated biological outputs, thereby promoting tumorigenesis and tumor progression, invasion and/or metastasis [39-42]. Therefore, analyses of PPIs are critical for understanding biological processes and developing effective strategies for cancer treatment. In our studies, we focus on the PPIs of the biomarkers identified in our proteomic studies in order to understand the functions of these protein biomarkers.

Identification of regulated proteins

The protein profiles regulated by biomarker proteins provide critical information for understanding the functions of the biomarker proteins [43]. These protein lists have the potential to offer important clues for understanding tumor biology and may include candidates for biomarkers and therapeutic targets. In our studies, we routinely use proteomic approaches to identify proteins regulated by the biomarker proteins using a transfection system. The cell lines are treated by either introducing genes encoding the biomarker proteins into the cells without an expression of the proteins (gain-of-protein effect), or removing the biomarker protein expression from the cell lines constantly expressing the proteins using RNAi (loss-of-protein effect). These analyses can be used to identify candidates for regulatory proteins of the biomarker proteins. This approach can also be used to provide critical information for understanding the functions and roles of these biomarkers in tumors.

Functional analyses of NPM1 in Ewing's sarcoma

We previously reported NPM1 to be a predictive biomarker for the

prognosis of Ewing's sarcoma in patients identified using proteomic [27]. NPM1 is a ubiquitously expressed protein belonging to the nucleoplasmin family of nuclear chaperones, and a highly conserved nucleocytoplasmic shuttling protein that shows restricted nucleolar localization [44-48]. NPM1 is frequently translocated or mutated in hematological malignancies, and mutations of the NPM1 gene leading to aberrant cytoplasmic dislocation of nucleophosmin (NPMc+) occur in approximately one-third of acute myeloid leukemia patients, who exhibit distinct biological and clinical features [44-48]. Although one article revealed a list of interaction proteins with NPM1 in Ewing's sarcoma, the functions of NPM1 in Ewing's sarcoma still remain unknown [49]. Therefore, we used proteomic approaches that consisted of the identification of both interaction proteins and regulated proteins associated with NPM1 proteins.

In the PPI analyses, we performed immunoprecipitation (IP) assays using two Ewing's sarcoma cell lines (SKES1 and CHP100) and NPM1 antibodies to identify the expression profiles of interaction proteins physiologically associated with NPM1 (Figure 4). Proteins extracted from Ewing's sarcoma cell lines were immunoprecipitated using either NPM1 antibodies or IgG antibodies (control). The IP samples were separated using SDS-PAGE and the gel images were compared between the NPM1 IP samples and the control samples. We found 20 bands with significantly different densities between the two groups (Table 1). The bands were treated with in-gel digestion, and the proteins were identified using MS spectrometry (Table 1). The proteins interacting with NPM1 are shown in Table 1.

To identify protein expression profiles regulated by NPM1, we employed siRNA knockdown and GeLC-MS in four Ewing's sarcoma cell lines (A673, TC71, SKES1 and CHP100), using NPM1 siRNA (Figure 5). The cell lines were transfected with either NPM1 siRNA or control siRNA and harvested after 72 hours. Proteins extracted from the cell lines were analyzed using GeLC-MS. We compared the acquired proteomic profiles between the control group and the siRNA group to calculate the semiquantitative expressions. The comparisons identified approximately 1,500 proteins that exhibited upregulation, downregulation or no changes in each of the four cell lines (Figure 5 and Table 2). We analyzed the four profiles to identify commonly regulated proteins in the four cell lines and found 36 upregulated and 18 downregulated commonly regulated proteins (Figure 5 and Table 3). The regulated proteins are shown in Table 3.

In order to further understand the biological processes and networks and determine whether the proteins were direct or indirect proteins, we routinely employed network analyses using the Ingenuity Pathways Analysis (IPA) system (Ingenuity Systems, Inc, CA, USA) (Figure 6). In this study, we performed network analyses using each PPI profile (Table 1) and regulated protein profile (Table 3 and Figure 7). In both independent analyses using each set of data, the network analyses identified the MYC pathway as playing a critical functional role as an upstream regulator of NPM1 in Ewing's sarcoma (Table 4 and Figure 7). Additionally, in order to confirm the relationships between MYC and NPM1, we conducted siRNA assays of the Ewing's sarcoma cell lines using MYC siRNA and verified the protein expressions of both MYC and NPM1 in the cells using Western blotting. The results revealed that

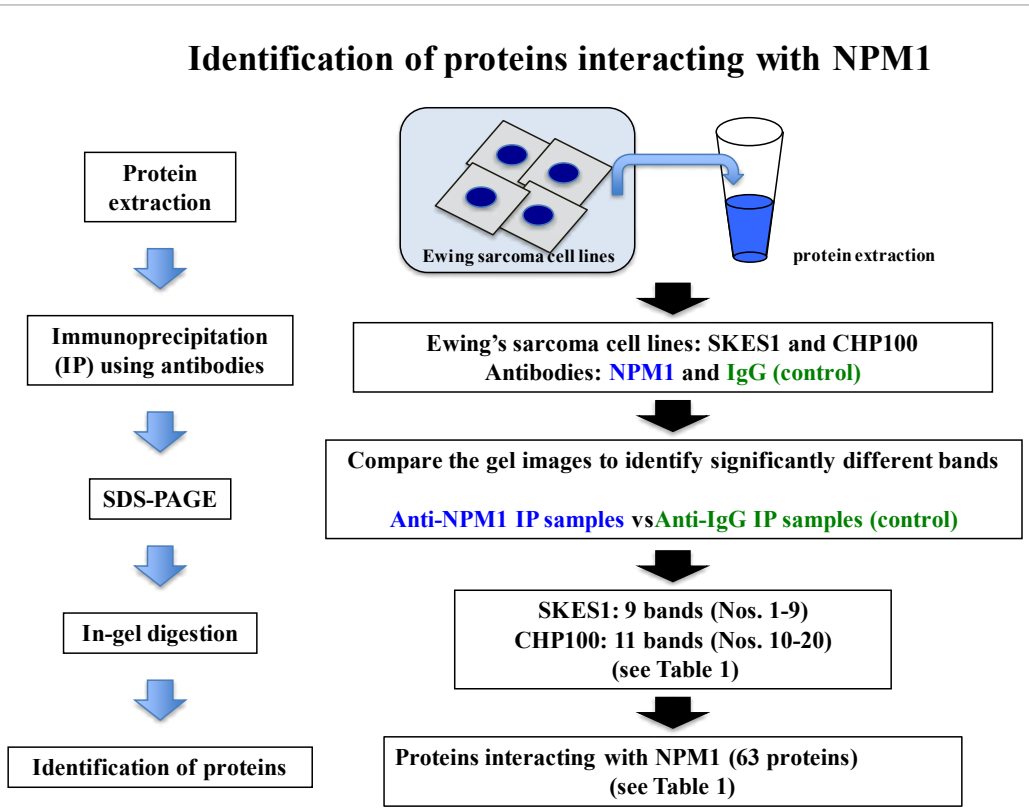

Figure 4: Identification of interaction proteins associated with NPM1: Immunoprecipitation (IP) was performed using two Ewing's cell lines (SKES1 and CHP100) and antibodies (NPM1 and IgG (Santa Cruz, TX)). The IP samples were separated using SDS-PAGE, and the gel images were compared between the NPM1 samples and IgG samples in each cell line. In this study, we identified 20 bands with significantly different densities in the two cell lines. We then identified the proteins included in each band using a mass spectrometer. The identified proteins are listed in Table 1.

Gel band No	Cell line name	MW in the gel image(KDa)[1)]	Name	Protein Name[2)]	Molecular Weight[2)]	Mascot Score[2)]
1	SKES1	54	TBA1A_HUMAN	Tubulin alpha-1A chain	Mass: 50956	Score: 71
1	SKES1	54	SERA_HUMAN	D-3-phosphoglycerate dehydrogenase	Mass: 57538	Score: 45
2	SKES1	52	VIME_HUMAN	Vimentin	Mass: 53690	Score: 206
2	SKES1	52	TBB2C_HUMAN	Tubulin beta-2C chain	Mass: 50367	Score: 128
2	SKES1	52	TBA1A_HUMAN	Tubulin alpha-1A chain	Mass: 50956	Score: 117
2	SKES1	52	GFAP_HUMAN	Glial fibrillary acidic protein	Mass: 49921	Score: 95
2	SKES1	52	ATPA_HUMAN	ATP synthase subunit alpha, mitochondrial	Mass: 59856	Score: 63
3	SKES1	49	TBB5_HUMAN	Tubulin beta chain	Mass: 50207	Score: 372
3	SKES1	49	ATPA_HUMAN	ATP synthase subunit alpha, mitochondrial	Mass: 59856	Score: 164
3	SKES1	49	ATPB_HUMAN	ATP synthase subunit beta, mitochondrial	Mass: 56525	Score: 143
4	SKES1	36	NPM_HUMAN	Nucleophosmin	Mass: 32768	Score: 134
4	SKES1	36	ROA2_HUMAN	Heterogeneous nuclear ribonucleoproteins A2/B1	Mass: 37478	Score: 99
4	SKES1	36	PCBP2_HUMAN	Poly(rC)-binding protein 2	Mass: 39053	Score: 73
4	SKES1	36	RA1L3_HUMAN	Putative heterogeneous nuclear ribonucleoprotein A1-like protein 3	Mass: 34415	Score: 48
4	SKES1	36	ROA3_HUMAN	Heterogeneous nuclear ribonucleoprotein A3	Mass: 39855	Score: 44
5	SKES1	35	G3P_HUMAN	Glyceraldehyde-3-phosphate dehydrogenase	Mass: 36244	Score: 190
5	SKES1	35	ROA2_HUMAN	Heterogeneous nuclear ribonucleoproteins A2/B1	Mass: 37478	Score: 184
5	SKES1	35	CAZA1_HUMAN	F-actin-capping protein subunit alpha-1	Mass: 33115	Score: 135
5	SKES1	35	HNRH3_HUMAN	Heterogeneous nuclear ribonucleoprotein H3	Mass: 36974	Score: 75
5	SKES1	35	PCBP2_HUMAN	Poly(rC)-binding protein 2	Mass: 39053	Score: 57
6	SKES1	33	ROA1_HUMAN	Heterogeneous nuclear ribonucleoprotein A1	Mass: 38964	Score: 547
6	SKES1	33	PHB2_HUMAN	Prohibitin-2	Mass: 33276	Score: 279
6	SKES1	33	LDHA_HUMAN	L-lactate dehydrogenase A chain	Mass: 37021	Score: 93
6	SKES1	33	ROA0_HUMAN	Heterogeneous nuclear ribonucleoprotein A0	Mass: 31035	Score: 80
6	SKES1	33	ROA2_HUMAN	Heterogeneous nuclear ribonucleoproteins A2/B1	Mass: 37478	Score: 77
6	SKES1	33	VDAC2_HUMAN	Voltage-dependent anion-selective channel protein 2	Mass: 32186	Score: 71
7	SKES1	30	EFHD2_HUMAN	EF-hand domain-containing protein D2	Mass: 26823	Score: 116
7	SKES1	30	RS3_HUMAN	40S ribosomal protein S3	Mass: 26885	Score: 105
7	SKES1	30	RL8_HUMAN	60S ribosomal protein L8	Mass: 28291	Score: 66
7	SKES1	30	CAPZB_HUMAN	F-actin-capping protein subunit beta	Mass: 31686	Score: 50
7	SKES1	30	SFR2B_HUMAN	Splicing factor, arginine/serine-rich 2B	Mass: 32410	Score: 46
7	SKES1	30	RS2_HUMAN	40S ribosomal protein S2	Mass: 31660	Score: 42
7	SKES1	30	RFA2_HUMAN	Replication protein A 32 kDa subunit	Mass: 29371	Score: 39
8	SKES1	26	TPIS_HUMAN	Triosephosphate isomerase	Mass: 27008	Score: 74
8	SKES1	26	BAP31_HUMAN	B-cell receptor-associated protein 31	Mass: 28045	Score: 72
8	SKES1	26	RAB21_HUMAN	Ras-related protein Rab-21	Mass: 24830	Score: 62
8	SKES1	26	RALA_HUMAN	Ras-related protein Ral-A	Mass: 23765	Score: 57
8	SKES1	26	SNP23_HUMAN	Synaptosomal-associated protein 23	Mass: 23766	Score: 38
8	SKES1	26	RL19_HUMAN	60S ribosomal protein L19	Mass: 23593	Score: 35
9	SKES1	16	GAPR1_HUMAN	Golgi-associated plant pathogenesis-related protein 1	Mass: 17350	Score: 178
9	SKES1	16	H4_HUMAN	Histone H4	Mass: 11360	Score: 176
9	SKES1	16	H2B1C_HUMAN	Histone H2B type 1-C/E/F/G/I	Mass: 13811	Score: 174
9	SKES1	16	H2B1B_HUMAN	Histone H2B type 1-B	Mass: 13942	Score: 160
9	SKES1	16	PPIA_HUMAN	Peptidyl-prolyl cis-trans isomerase A	Mass: 18285	Score: 107
9	SKES1	16	H31T_HUMAN	Histone H3.1t	Mass: 15641	Score: 84
9	SKES1	16	DCD_HUMAN	Dermcidin	Mass: 11419	Score: 80
9	SKES1	16	RL31_HUMAN	60S ribosomal protein L31	Mass: 14454	Score: 63
9	SKES1	16	RLA2_HUMAN	60S acidic ribosomal protein P2	Mass: 11658	Score: 53
9	SKES1	16	H2A1A_HUMAN	Histone H2A type 1-A	Mass: 14225	Score: 53
9	SKES1	16	MYL6_HUMAN	Myosin light polypeptide 6	Mass: 17132	Score: 47
9	SKES1	16	RL35_HUMAN	60S ribosomal protein L35	Mass: 14543	Score: 42
10	CHP100	68	PLAK_HUMAN	Junction plakoglobin	Mass: 82572	Score: 419
10	CHP100	68	KPRP_HUMAN	Keratinocyte proline-rich protein	Mass: 67929	Score: 47
11	CHP100	54	TBA1A_HUMAN	Tubulin alpha-1A chain	Mass: 50956	Score: 120
11	CHP100	54	HNRPK_HUMAN	Heterogeneous nuclear ribonucleoprotein K	Mass: 51300	Score: 87
11	CHP100	54	SPB12_HUMAN	Serpin B12	Mass: 46744	Score: 60
12	CHP100	53	VIME_HUMAN	Vimentin	Mass: 53690	Score: 264
12	CHP100	53	TBA1A_HUMAN	Tubulin alpha-1A chain	Mass: 50956	Score: 248

12	CHP100	53	TBB5_HUMAN	Tubulin beta chain	Mass: 50207	Score: 63
12	CHP100	53	RBBP4_HUMAN	Histone-binding protein RBBP4	Mass: 47981	Score: 58
12	CHP100	53	ATPA_HUMAN	ATP synthase subunit alpha, mitochondrial	Mass: 59856	Score: 54
12	CHP100	53	GFAP_HUMAN	Glial fibrillary acidic protein	Mass: 49921	Score: 38
13	CHP100	49	ATPB_HUMAN	ATP synthase subunit beta, mitochondrial	Mass: 56525	Score: 364
13	CHP100	49	ATPA_HUMAN	ATP synthase subunit alpha, mitochondrial	Mass: 59856	Score: 196
13	CHP100	49	HNRH1_HUMAN	Heterogeneous nuclear ribonucleoprotein H	Mass: 49554	Score: 95
14	CHP100	37	NPM_HUMAN	Nucleophosmin	Mass: 32768	Score: 106
14	CHP100	37	PCBP2_HUMAN	Poly(rC)-binding protein 2	Mass: 39053	Score: 46
14	CHP100	37	ARGI1_HUMAN	Arginase-1	Mass: 34926	Score: 37
15	CHP100	36	NPM_HUMAN	Nucleophosmin	Mass: 32768	Score: 68
15	CHP100	36	ROA2_HUMAN	Heterogeneous nuclear ribonucleoproteins A2/B1	Mass: 37478	Score: 54
15	CHP100	36	ROA3_HUMAN	Heterogeneous nuclear ribonucleoprotein A3	Mass: 39855	Score: 40
16	CHP100	33	ROA1_HUMAN	Heterogeneous nuclear ribonucleoprotein A1	Mass: 38964	Score: 268
16	CHP100	33	PHB2_HUMAN	Prohibitin-2	Mass: 33276	Score: 227
16	CHP100	33	ROA0_HUMAN	Heterogeneous nuclear ribonucleoprotein A0	Mass: 31035	Score: 112
16	CHP100	33	LDHA_HUMAN	L-lactate dehydrogenase A chain	Mass: 37021	Score: 67
16	CHP100	33	ROA2_HUMAN	Heterogeneous nuclear ribonucleoproteins A2/B1	Mass: 37478	Score: 42
17	CHP100	30	TPM3_HUMAN	Tropomyosin alpha-3 chain	Mass: 32870	Score: 115
17	CHP100	30	VDAC1_HUMAN	Voltage-dependent anion-selective channel protein 1	Mass: 30896	Score: 38
17	CHP100	30	VDAC3_HUMAN	Voltage-dependent anion-selective channel protein 3	Mass: 31066	Score: 38
17	CHP100	30	MTCH2_HUMAN	Mitochondrial carrier homolog 2	Mass: 34090	Score: 36
18	CHP100	29	RS3_HUMAN	40S ribosomal protein S3	Mass: 26885	Score: 92
18	CHP100	29	EFHD2_HUMAN	EF-hand domain-containing protein D2	Mass: 26823	Score: 71
18	CHP100	29	CAPZB_HUMAN	F-actin-capping protein subunit beta	Mass: 31686	Score: 57
18	CHP100	29	RFA2_HUMAN	Replication protein A 32 kDa subunit	Mass: 29371	Score: 39
19	CHP100	27	PHB_HUMAN	Prohibitin	Mass: 29857	Score: 173
19	CHP100	27	ADT3_HUMAN	ADP/ATP translocase 3	Mass: 33129	Score: 164
19	CHP100	27	ADT2_HUMAN	ADP/ATP translocase 2	Mass: 33158	Score: 140
19	CHP100	27	RL7_HUMAN	60S ribosomal protein L7	Mass: 29278	Score: 60
19	CHP100	27	1433B_HUMAN	14-3-3 protein beta/alpha	Mass: 28207	Score: 37
19	CHP100	27	1433S_HUMAN	14-3-3 protein sigma	Mass: 27899	Score: 37
20	CHP100	26	CHCH3_HUMAN	Coiled-coil-helix-coiled-coil-helix domain-containing protein 3	Mass: 26491	Score: 101
20	CHP100	26	SNP23_HUMAN	Synaptosomal-associated protein 23	Mass: 23766	Score: 71
20	CHP100	26	TPIS_HUMAN	Triosephosphate isomerase	Mass: 27008	Score: 52
20	CHP100	26	RL19_HUMAN	60S ribosomal protein L19	Mass: 23593	Score: 43
20	CHP100	26	RALA_HUMAN	Ras-related protein Ral-A	Mass: 23765	Score: 41
20	CHP100	26	BAP31_HUMAN	B-cell receptor-associated protein 31	Mass: 28045	Score: 39

1) MW: Molecular Weight
2) Mascot score for the identified proteins based on the peptide ions score ($p < 0.05$) (http://www.matrixscience.com)

Table 1 : Protein list of interaction proteins associated with NPM1.

silencing MYC in parallel inhibited the NPM1 expression, indicating that MYC is an upstream regulator of NPM1 in Ewing's sarcoma . We believe that the findings obtained in the functional analyses will contribute to improving understanding of the relationship between NPM1 and malignant behavior in Ewing's sarcoma and lead to the development of novel therapeutic strategies.

Conclusion

Our proteomic studies of soft tissue sarcomas identified various candidate biomarkers relevant to the prognosis and chemosensitivity of tumors [17-28]. These proteomic studies successfully verified the value of the biomarkers in validation sets using immunohistochemistry. We believe that these proteins are potentially useful biomarkers for various clinical applications. However, although we identified useful biomarkers in our proteomic studies, the functions of the biomarker proteins in tumors remain unknown. Therefore, we conducted functional studies in order to identify the roles and functions of these proteins in the tumors. In particular, we employed proteomic technologies as a tool for conducting functional studies, which revealed novel findings. These results indicate that our proteomic approaches used to perform functional analyses are efficient. Therefore, we should continue these studies in order to further understand these functions. Proteomic analyses are more directly linked to aberrant tumor phenotypes; therefore, there are limitations in our approaches to revealing all processes of molecular biology. In fact, in comparison to cDNA microarray analyses (50,000 probe sets), the sensitivity of the current 2D-DIGE analysis (5,000 spots) remains unsatisfactory. Therefore, these technologies, including CGH arrays, cDNA microarrays, whole genome sequences and proteomic techniques should be used in combination to overcome their individual disadvantages. We believe that hybrid comprehensive studies consisting of genomics, transcriptomics and proteomics will provide important, novel clues for understanding the biology of tumors and identifying biomarkers and therapeutic targets.

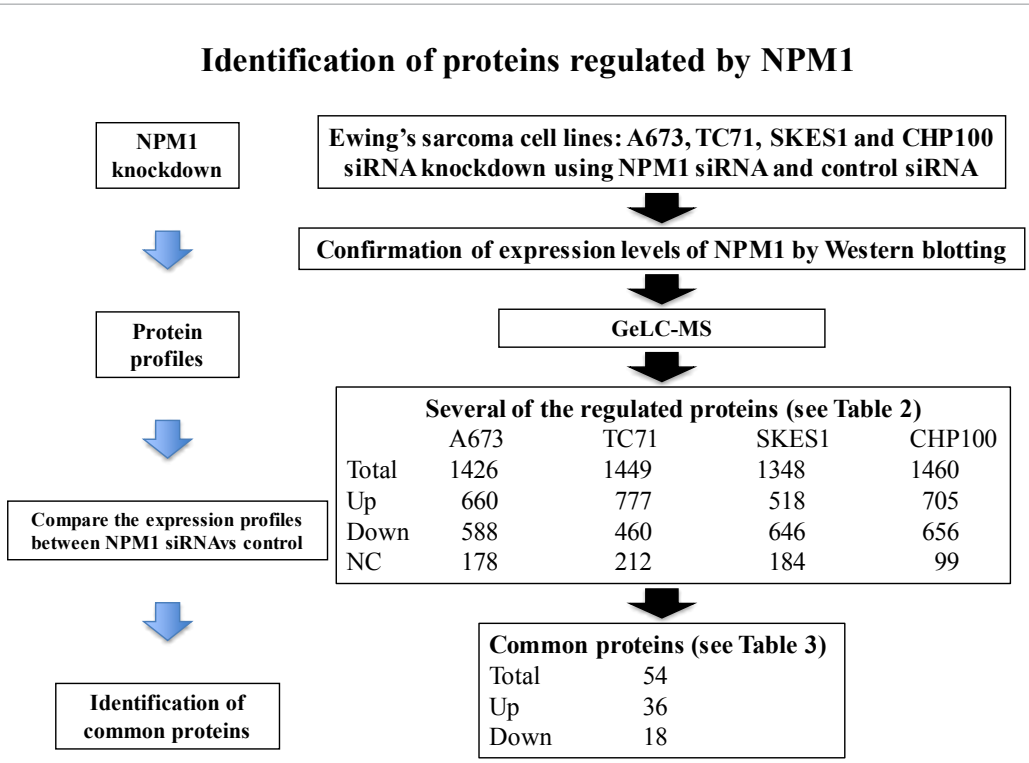

Figure 5: Identification of proteins regulated by NPM1: In order to identify proteins regulated by the NPM1 expression, we performed siRNA knockdown and GeLC-MS analyses in four Ewing's sarcoma cell lines (A673, TC71, SKES1 and CHP100). The four Ewing's sarcoma cell lines were treated with either NPM1 siRNA (SASI_Hs01_00214118; SIGMA-ALDRICH) or negative control siRNA (SIGMA-ALDRICH). A Western blot analysis confirmed that the cells treated with NPM1 siRNA exhibited a significant decrease in the NPM1 expression compared to the controls. These protein samples were then analyzed using GeLC-MS to obtain their protein profiles, and the acquired data were calculated as semiquantitative expressions (control vs siRNA). Approximately 1,500 proteins were identified in each cell line (Table 2). Finally, 36 upregulated proteins and 18 downregulated proteins were identified as common proteins in the four cell lines (Table 3).

	Cell line name			
	A673	**TC71**	**SKES1**	**CHP100**
Downregulation	588	460	646	656
Upregulation	660	777	518	705
No change	178	212	184	99
Total	1426	1449	1348	1460

Table 2 : Number of regulated proteins.

Accession number	Description	Up or Down regulation
IPI00549248	NPM1 Isoform 1 of Nucleophosmin	Down regulation
IPI00646304	PPIB peptidylprolyl isomerase B precursor	Down regulation
IPI00742682	TPR nuclear pore complex-associated protein	Down regulation
IPI00221226	ANXA6 Annexin A6	Down regulation
IPI00418313	ILF3 Isoform 4 of Interleukin enhancer-binding factor 3	Down regulation
IPI00003918	RPL4 60S ribosomal protein L4	Down regulation
IPI00329745	LRPPRC 159 kDa protein	Down regulation
IPI00218236	PPP1CB Serine/threonine-protein phosphatase PP1-beta catalytic subunit	Down regulation
IPI00647337	PPP1CB Serine/threonine-protein phosphatase PP1-beta catalytic subunit	Down regulation
IPI00301263	CAD CAD protein	Down regulation
IPI00217966	LDHA Isoform 1 of L-lactate dehydrogenase A chain	Down regulation
IPI00296053	FH Isoform Mitochondrial of Fumarate hydratase, mitochondrial precursor	Down regulation
IPI00293867	DDT D-dopachrome decarboxylase	Down regulation
IPI00376798	RPL11 Isoform 1 of 60S ribosomal protein L11	Down regulation
IPI00298547	PARK7 Protein DJ-1	Down regulation
IPI00480032	LOC653156 similar to ribosomal protein L21 isoform 2	Down regulation
IPI00472864	LOC285053 Uncharacterized protein	Down regulation
IPI00794221	DBN1 76 kDa protein	Down regulation

IPI00004534	PFAS Phosphoribosylformylglycinamidine synthase	Up regulation
IPI00010896	DDAH2;CLIC1 Chloride intracellular channel protein 1	Up regulation
IPI00746205	PSME2 proteasome activator subunit 2	Up regulation
IPI00784131	AARS Uncharacterized protein AARS	Up regulation
IPI00103994	LARS Leucyl-tRNA synthetase, cytoplasmic	Up regulation
IPI00034049	UPF1 Isoform 1 of Regulator of nonsense transcripts 1	Up regulation
IPI00029997	PGLS 6-phosphogluconolactonase	Up regulation
IPI00016862	GSR Isoform Mitochondrial of Glutathione reductase, mitochondrial precursor	Up regulation
IPI00140420	SND1 Staphylococcal nuclease domain-containing protein 1	Up regulation
IPI00030781	STAT1 Isoform Alpha of Signal transducer and activator of transcription 1-alpha/beta	Up regulation
IPI00011603	PSMD3 26S proteasome non-ATPase regulatory subunit 3	Up regulation
IPI00009904	PDIA4 Protein disulfide-isomerase A4 precursor	Up regulation
IPI00001636	ATXN10 Ataxin-10	Up regulation
IPI00305092	WIBG Isoform 1 of Protein wibg homolog	Up regulation
IPI00021766	RTN4 Isoform 1 of Reticulon-4	Up regulation
IPI00009342	IQGAP1 Ras GTPase-activating-like protein IQGAP1	Up regulation
IPI00022462	TFRC Transferrin receptor protein 1	Up regulation
IPI00607818	MYH14 Isoform 2 of Myosin-14	Up regulation
IPI00307155	ROCK2 Rho-associated protein kinase 2	Up regulation
IPI00013290	HDGF2 hepatoma-derived growth factor-related protein 2 isoform 1	Up regulation
IPI00375144	ARS2 Uncharacterized protein	Up regulation
IPI00018350	MCM5 DNA replication licensing factor MCM5	Up regulation
IPI00477313	HNRNPC Isoform C2 of Heterogeneous nuclear ribonucleoproteins C1/C2	Up regulation
IPI00295386	CBR1 Carbonyl reductase [NADPH] 1	Up regulation
IPI00295098	SRPRB Signal recognition particle receptor subunit beta	Up regulation
IPI00021370	HIP2 Isoform 1 of Ubiquitin-conjugating enzyme E2-25 kDa	Up regulation
IPI00640817	AK1 Adenylate kinase 1	Up regulation
IPI00001757	RBM8A Isoform 1 of RNA-binding protein 8A	Up regulation
IPI00339269	HSPA6 Heat shock 70 kDa protein 6	Up regulation
IPI00184330	MCM2 DNA replication licensing factor MCM2	Up regulation
IPI00645431	BAT3 HLA-B associated transcript 3	Up regulation
IPI00007401	IPO8 Importin-8	Up regulation
IPI00604707	DLAT Dihydrolipoamide S-acetyltransferase	Up regulation
IPI00828150	SUGT1 Isoform 1 of Suppressor of G2 allele of SKP1 homolog	Up regulation
IPI00718888	PRPS2 Isoform 2 of Ribose-phosphate pyrophosphokinase II	Up regulation
IPI00016077	GBAS Protein NipSnap2	Up regulation
IPI00021570	EDF1 Isoform 1 of Endothelial differentiation-related factor 1	Up regulation

Table 3: List of proteins regulated by NPM1 suppression.

A: Interaction proteins associated with NPM1			
Upstream Regulator	p-value of overlap	Molecule Type	Target molecules in dataset
MYC	5.78E-06	transcription regulator	CAPZB,LDHA,PHB,PHB2,PPIA,RBBP4,VDAC2
MYCN	1.59E-03	transcription regulator	LDHA,PHB,RBBP4
ALX3	2.02E-03	transcription regulator	GFAP
E2F1	4.68E-03	transcription regulator	HNRNPK,PHB,RBBP4
OLIG2	5.06E-03	transcription regulator	GFAP
MYCBP	7.07E-03	transcription regulator	LDHA
Pdx1	8.08E-03	transcription regulator	GFAP
HNF4A	1.07E-02	transcription regulator	MYL6,PHB,PHB2,RBBP4,VDAC1,VDAC2
PURA	1.41E-02	transcription regulator	GFAP
KCNIP3	1.51E-02	transcription regulator	GFAP
NFIX	1.81E-02	transcription regulator	GFAP
SUPT16H	2.11E-02	transcription regulator	HNRNPK
NR2E1	2.21E-02	ligand-dependent nuclear receptor	GFAP
HDAC4	2.60E-02	transcription regulator	LDHA
Nuclear factor 1	2.90E-02	group	GFAP
HIF1A	3.50E-02	transcription regulator	LDHA,PPIA
NRF1	4.37E-02	transcription regulator	VDAC1
E2F6	4.46E-02	transcription regulator	RBBP4

B: Proteins regulated by NPM1			
Upstream Regulator	p-value of overlap	Molecule Type	Target molecules in dataset
MYCN	3.28E-05	transcription regulator	CAD,LDHA,NPM1,PDIA4,RPL11,RPL4
MYC	1.39E-04	transcription regulator	ANXA6,CAD,DBN1,GSR,LDHA,MCM5,NPM1,ROCK2,TFRC
MYCBP	1.42E-04	transcription regulator	CAD,LDHA
NFE2L2	4.15E-04	transcription regulator	CBR1 (includes EG:100360507),GSR,PDIA4,PPIB,PSMD3,UBE2K
TP53	9.16E-04	transcription regulator	AK1,ANXA6,GSR,LDHA,MCM2,MCM5,NPM1,PARK7,PSMD3,STAT1
Meg3	1.05E-02	transcription regulator	IQGAP1
E2F2	1.13E-02	transcription regulator	MCM2,MCM5
RBL1	1.27E-02	transcription regulator	MCM2,MCM5
XBP1	1.28E-02	transcription regulator	PDIA4,PPIB,SRPRB
GTF2H4	1.31E-02	transcription regulator	CAD
MYCL1	1.31E-02	transcription regulator	CAD
CDKN2A	1.49E-02	transcription regulator	AK1,MCM5,NPM1
E2F3	1.51E-02	transcription regulator	MCM2,MCM5
TBX2	1.67E-02	transcription regulator	MCM2,MCM5
MAX	1.80E-02	transcription regulator	CAD,NPM1
TLE1	1.83E-02	transcription regulator	ROCK2
ERG	2.45E-02	transcription regulator	DBN1,ROCK2
KDM5A	2.86E-02	transcription regulator	MCM2
CCNT1	2.86E-02	transcription regulator	CAD
ZNF148	3.12E-02	transcription regulator	STAT1
E2f	3.40E-02	group	MCM2,MCM5
HTT	3.57E-02	transcription regulator	CBR1 (includes EG:100360507),GSR,LDHA,PSMD3,TFRC
MXI1	3.63E-02	transcription regulator	LARS
HR	3.88E-02	transcription regulator	HNRNPC
GTF2I	3.88E-02	transcription regulator	PDIA4
Cyclin E	4.13E-02	group	ROCK2
HIF1A	4.19E-02	transcription regulator	LDHA,NPM1,TFRC
SP100	4.39E-02	transcription regulator	HSPA6
NR1D1	4.39E-02	ligand-dependent nuclear receptor	STAT1
IRF7	4.42E-02	transcription regulator	PSME2,STAT1
FOXO3	4.42E-02	transcription regulator	CAD,LARS
KLF2	4.73E-02	transcription regulator	STAT1,TFRC
BRCA1	4.73E-02	transcription regulator	AK1,STAT1

Table 4 : Upstream regulators.

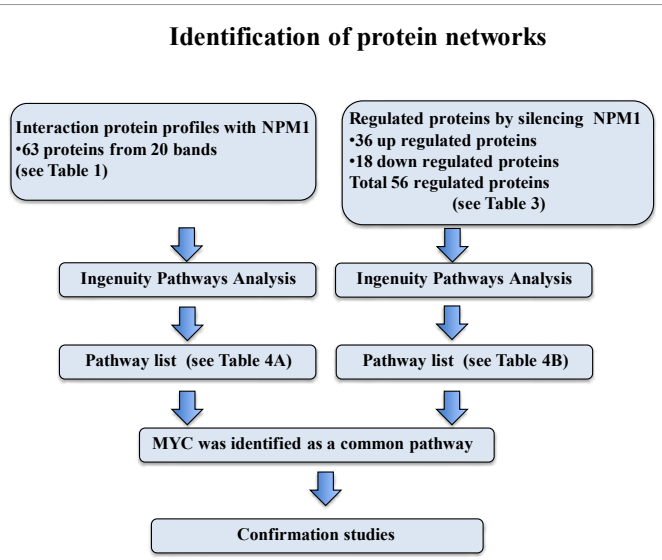

Figure 6: Identification of protein networks: To identify networks and upstream proteins, we routinely employed the Ingenuity Pathways Analysis (IPA) system. We analyzed these networks using either the interaction protein profiles (Table 1), or regulated protein profiles independently (Table 3). The results of these analyses are demonstrated in Table 4. We found that both pathway lists included the MYC pathway as an upstream protein. We conducted a confirmation study and successfully confirmed MYC to be an upstream regulator of NPM1 in Ewing's sarcoma (data not shown).

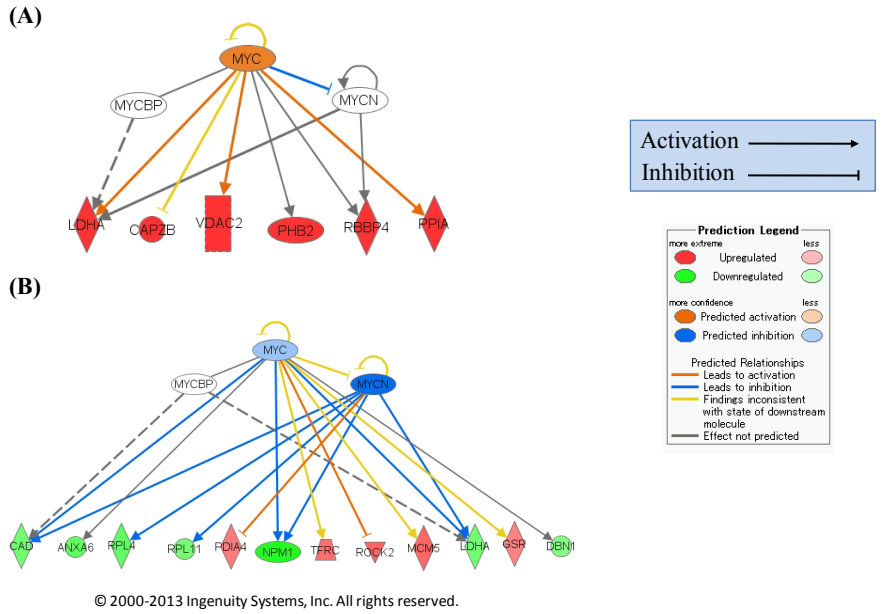

Figure 7: (A) The Ingenuity Pathways Analysis (IPA) system demonstrated pathways based on the interaction proteins with NPM1 (Table 1 and Table 4A). (B) The IPA system demonstrated pathways based on the proteins regulated by NPM1 (Table 3 and Table 4B).

Acknowledgments

This work was supported by a grant from the Japan Society for the Promotion of Science (JSPS) and science Grants-in- Aid for Young Scientists B (No- 25861342 to YS) and Scientific Research (No. 23590434 to TS).

Conflicts of Interest

The corresponding author declares that there are no conflicts of interest.

References

1. Fletcher CDM, Unni KK, Mertens F (2002) Pathology and genetics of tumours of soft tissue and bone. World Health Organization Classification of Tumours, IARC Press, Lyon, France.
2. Allander SV, Illei PB, Chen Y, Antonescu CR, Bittner M, et al. (2002) Expression profiling of synovial sarcoma by cDNA microarrays: association of ERBB2, IGFBP2, and ELF3 with epithelial differentiation. Am J Pathol 161: 1587-1595.
3. Nakagawa Y, Numoto K, Yoshida A, Kunisada T, Ohata H, et al. (2006) Chromosomal and genetic imbalances in synovial sarcoma detected by conventional and microarray comparative genomic hybridization. J Cancer Res Clin Oncol 132: 444-450.
4. Lee YF, John M, Falconer A, Edwards S, Clark J, et al. (2004) A gene expression signature associated with metastatic outcome in human leiomyosarcomas. Cancer Res 64: 7201-7204.
5. Nielsen TO, West RB, Linn SC, Alter O, Knowling MA, et al. (2002) Molecular characterisation of soft tissue tumours: a gene expression study. Lancet 359: 1301-1307.
6. Nagayama S, Katagiri T, Tsunoda T, Hosaka T, Nakashima Y, et al. (2002) Genome-wide analysis of gene expression in synovial sarcomas using a cDNA microarray. Cancer Res 62: 5859-5866.
7. Nakayama R, Mitani S, Nakagawa T, Hasegawa T, Kawai A, et al. (2010) Gene expression profiling of synovial sarcoma: Distinct signature of poorly differentiated type. Am J Surg Pathol 34: 1599-1607.
8. Antonescu CR (2008) Molecular profiling in the diagnosis and treatment of high grade sarcomas. Ultrastruct Pathol 32: 37-42.
9. Tschoep K, Kohlmann A, Schlemmer M, Haferlach T, Issels RD (2007) Gene expression profiling in sarcomas. Crit Rev Oncol Hematol 63: 111-124.
10. Nielsen TO (2006) Microarray analysis of sarcomas. Adv Anat Pathol 13: 166-173.
11. Francis P, Namløs HM, Müller C, Edén P, Fernebro J, et al. (2007) Diagnostic and prognostic gene expression signatures in 177 soft tissue sarcomas: Hypoxia-induced transcription profile signifies metastatic potential. BMC Genomics 8: 73.
12. Ochi K, Daigo Y, Katagiri T, Nagayama S, Tsunoda T, et al. (2004) Prediction of response to neoadjuvant chemotherapy for osteosarcoma by gene-expression profiles. Int J Oncol 24: 647-655.
13. Mintz MB, Sowers R, Brown KM, Hilmer SC, Mazza B, et al. (2005) An expression signature classifies chemotherapy-resistant pediatric osteosarcoma. Cancer Res 65: 1748-1754.
14. Barretina J, Taylor BS, Banerji S, Ramos AH, Lagos-Quintana M, et al. (2010) Subtype-specific genomic alterations define new targets for soft-tissue sarcoma therapy. Nat Genet 42: 715-721.
15. Gygi SP, Rochon Y, Franza BR, Aebersold R (1999) Correlation between protein and mRNA abundance in yeast. Mol Cell Biol 19: 1720-1730.
16. Chen G, Gharib TG, Huang CC, Taylor JM, Misek DE, et al. (2002) Discordant protein and mRNA expression in lung adenocarcinomas. Mol Cell Proteomics 1: 304-313.
17. Suehara Y, Kondo T, Fujii K, Hasegawa T, Kawai A, et al. (2006) Proteomic signatures corresponding to histological classification and grading of soft-tissue sarcomas. Proteomics 6: 4402-4409.
18. Kawai A, Kondo T, Suehara Y, Kikuta K, Hirohashi S (2008) Global protein-expression analysis of bone and soft tissue sarcomas. Clin Orthop Relat Res 466: 2099-2106.
19. Suehara Y, Kikuta K, Nakayama R, Tochigi N, Seki K, et al. (2009) GST-P1 as a histological biomarker of synovial sarcoma revealed by proteomics. Proteomics Clin Appl 3: 623-634.
20. Suehara Y, Kikuta K, Nakayama R, Fujii K, Ichikawa H, et al. (2009) Anatomic site-specific proteomic signatures of gastrointestinal stromal tumors. Proteomics Clin Appl 3: 584-596.
21. Suehara Y (2011) Proteomic analysis of soft tissue sarcoma. Int J Clin Oncol 16: 92-100.
22. Suehara Y, Tochigi N, Kubota D, Kikuta K, Nakayama R, et al. (2011) Secernin-1 as a novel prognostic biomarker candidate of synovial sarcoma revealed by proteomics. J Proteomics 74: 829-842.
23. Suehara Y, Kubota D, Kikuta K, Kaneko K, Kawai A, et al. (2012) Discovery

of biomarkers for osteosarcoma by proteomics approaches. Sarcoma 2012: 425636.

24. Suehara Y, Kondo T, Seki K, Shibata T, Fujii K, et al. (2008) Pfetin as a prognostic biomarker of gastrointestinal stromal tumors revealed by proteomics. Clin Cancer Res 14: 1707-1717.

25. Kubota D, Yoshida A, Tsuda H, Suehara Y, Okubo T, et al. (2013) Gene expression network analysis of ETV1 reveals KCTD10 as a novel prognostic biomarker in gastrointestinal stromal tumor. PLoS One 8: e73896.

26. Kikuta K, Kubota D, Saito T, Orita H, Yoshida A, et al. (2012) Clinical proteomics identified ATP-dependent RNA helicase DDX39 as a novel biomarker to predict poor prognosis of patients with gastrointestinal stromal tumor. J Proteomics 75: 1089-1098.

27. Kikuta K, Tochigi N, Shimoda T, Yabe H, Morioka H, et al. (2009) Nucleophosmin as a candidate prognostic biomarker of Ewing's sarcoma revealed by proteomics. Clin Cancer Res 15: 2885-2894.

28. Kikuta K, Tochigi N, Saito S, Shimoda T, Morioka H, et al. (2010) Peroxiredoxin 2 as a chemotherapy responsiveness biomarker candidate in osteosarcoma revealed by proteomics. Proteomics Clin Appl 4: 560-567.

29. Conrad DH, Goyette J, Thomas PS (2008) Proteomics as a method for early detection of cancer: A review of proteomics, exhaled breath condensate, and lung cancer screening. J Gen Intern Med 23: 78-84.

30. Alessandro R, Fontana S, Kohn E, De Leo G (2005) Proteomic strategies and their application in cancer research. Tumori 91: 447-455.

31. Kondo T, Hirohashi S (2009) Application of 2D-DIGE in cancer proteomics toward personalized medicine. Methods Mol Biol 577: 135-154.

32. Kondo T, Hirohashi S (2006) Application of highly sensitive fluorescent dyes (CyDye DIGE Fluor saturation dyes) to laser microdissection and two-dimensional difference gel electrophoresis (2D-DIGE) for cancer proteomics. Nat Protoc 1: 2940-2956.

33. Piersma SR, Warmoes MO, de Wit M, de Reus I, Knol JC, et al. (2013) Whole gel processing procedure for GeLC-MS/MS based proteomics. Proteome Sci 11: 17.

34. Paulo JA, Lee LS, Wu B, Repas K, Mortele KJ, et al. (2010) Identification of pancreas-specific proteins in endoscopically (endoscopic pancreatic function test) collected pancreatic fluid with liquid chromatography--Tandem mass spectrometry. Pancreas 39: 889-896.

35. Dix MM, Simon GM, Cravatt BF (2008) Global mapping of the topography and magnitude of proteolytic events in apoptosis. Cell 134: 679-691.

36. Joensuu H, Roberts PJ, Sarlomo-Rikala M, Andersson LC, Tervahartiala P, et al. (2001) Effect of the tyrosine kinase inhibitor STI571 in a patient with a metastatic gastrointestinal stromal tumor. N Engl J Med 344: 1052-1056.

37. van der Zwan SM, DeMatteo RP (2005) Gastrointestinal stromal tumor: 5 years later. Cancer 104: 1781-1788.

38. Kikuta K, Tochigi N, Saito S, Shimoda T, Morioka H, et al. (2010) Peroxiredoxin 2 as a chemotherapy responsiveness biomarker candidate in osteosarcoma revealed by proteomics. Proteomics Clin Appl 4: 560-567.

39. Kann MG (2007) Protein interactions and disease: Computational approaches to uncover the etiology of diseases. Brief Bioinform 8: 333-346.

40. Kar G, Gursoy A, Keskin O (2009) Human cancer protein-protein interaction network: A structural perspective. PLoS Comput Biol 5: e1000601.

41. Ivanov AA, Khuri FR, Fu H (2013) Targeting protein-protein interactions as an anticancer strategy. Trends Pharmacol Sci 34: 393-400.

42. Wells JA, McClendon CL (2007) Reaching for high-hanging fruit in drug discovery at protein-protein interfaces. Nature 450: 1001-1009.

43. Kobos R, Nagai M, Tsuda M, Merl MY, Saito T, et al. (2013) Combining integrated genomics and functional genomics to dissect the biology of a cancer-associated, aberrant transcription factor, the ASPSCR1-TFE3 fusion oncoprotein. J Pathol 229: 743-754.

44. Falini B, Bolli N, Liso A, Martelli MP, Mannucci R, et al. (2009) Altered nucleophosmin transport in acute myeloid leukaemia with mutated NPM1: molecular basis and clinical implications. Leukemia 23: 1731-1743.

45. Bolli N, De Marco MF, Martelli MP, Bigerna B, Pucciarini A, et al. (2009) A dose-dependent tug of war involving the NPM1 leukaemic mutant, nucleophosmin, and ARF. Leukemia 23: 501-509.

46. Garzon R, Garofalo M, Martelli MP, Briesewitz R, Wang L, et al. (2008) Distinctive microRNA signature of acute myeloid leukemia bearing cytoplasmic mutated nucleophosmin. Proc Natl Acad Sci U S A 105: 3945-3950.

47. Falini B, Nicoletti I, Martelli MF, Mecucci C (2007) Acute myeloid leukemia carrying cytoplasmic/mutated nucleophosmin (NPMc+ AML): Biologic and clinical features. Blood 109: 874-885.

48. Alcalay M, Tiacci E, Bergomas R, Bigerna B, Venturini E, et al. (2005) Acute myeloid leukemia bearing cytoplasmic nucleophosmin (NPMc+AML) shows a distinct gene expression profile characterized by up-regulation of genes involved in stem-cell maintenance. Blood 106: 899-902.

49. Haga A, Ogawara Y, Kubota D, Kitabayashi I, Murakami Y, et al. (2013) Interactomic approach for evaluating nucleophosmin-binding proteins as biomarkers for Ewing's sarcoma. Electrophoresis 34: 1670-1678.

This article was originally published in a special issue, **Proteomics Technologies** handled by Editor(s). Dr. Jie Luo, Institute for Systems Biology, Seattle WA, USA