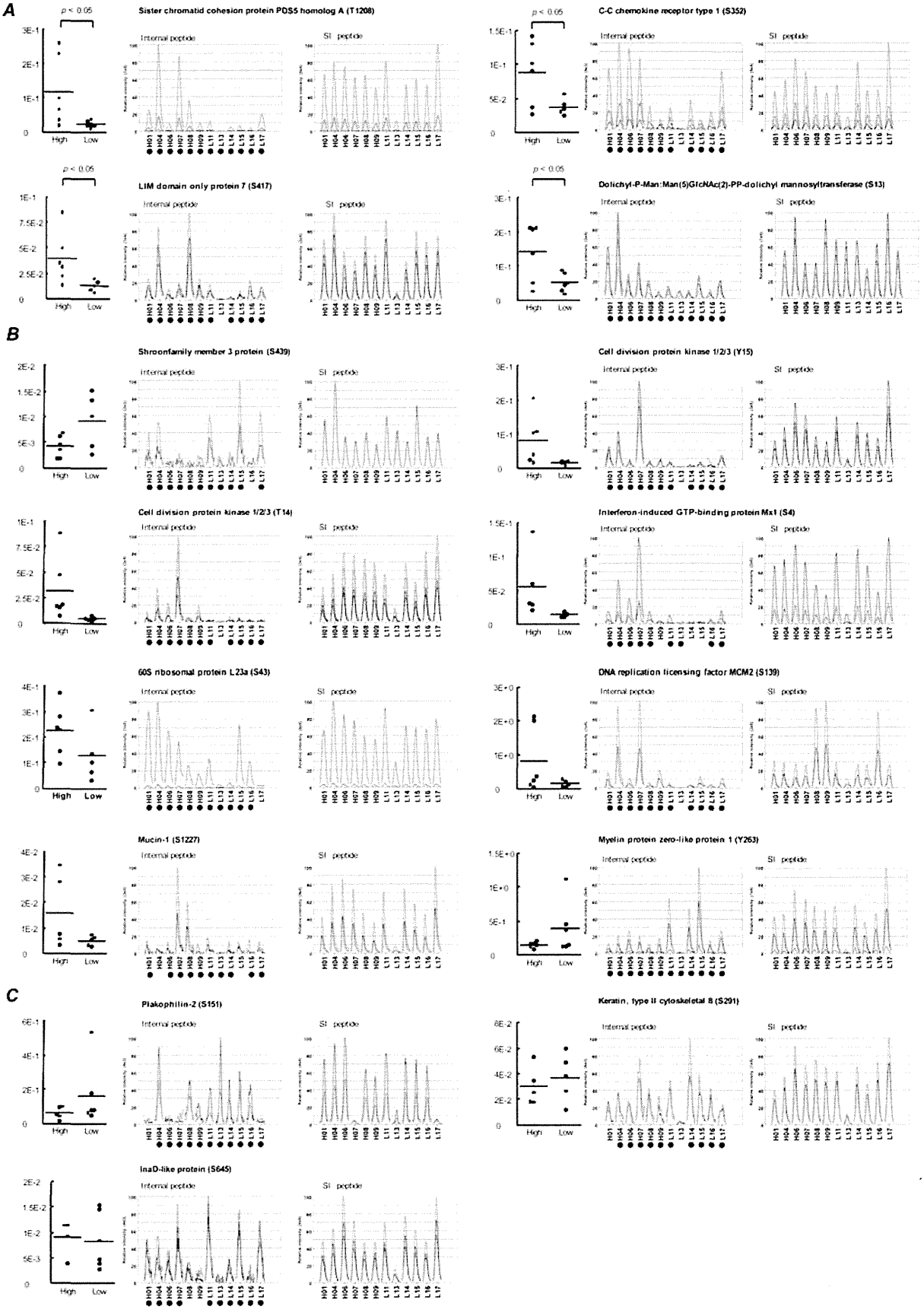


Figure 2. Relative quantitation of phosphopeptides between two groups from breast tissues by SRM. The scatter plots indicate the peak area ratio of the internal peptide to SI peptide and each horizontal bar indicates the mean value. *Y*-axis shows the normalized peak area. The "internal peptide" and "SI peptide" are based on the transition from internal peptides and SI peptides, respectively. (A) indicates significant difference groups ($p < 0.05$), (B) different propensity ($p > 0.05$, <0.2), and (C) no significant difference between two groups ($p > 0.2$). Closed circle indicates samples that were satisfactorily quantified.

dx.doi.org/10.1021/pr3005474 | J. Proteome Res. 2012, 11, 5311−5322

Table 4. SRM-Based Quantification of Phosphopeptides[a]

gene symbol	Uniprot accession	protein name	targeted phosphopeptide	phosphorylated site	high/low ratio	T. TEST	area ratio (unlabeled/stable-isotope labeled peptide) H01	H04	H06	H07	H08	H09	L11	L13	L14	L15	L16	L17
RPL23A	P62750	60S ribosomal protein L23a	IRTpSPTFR	S43	0.20	0.149	3.8×10^{-1}	2.8×10^{-1}	2.4×10^{-1}	2.3×10^{-1}	1.4×10^{-1}	9.5×10^{-2}	1.0×10^{-1}	$6\,5 \times 10^{-2}$	3.2×10^{-2}	3.1×10^{-1}	1.4×10^{-1}	ND
MX1	P20591	Interteron-induced GTP-binding protein Mx1	WpSEVDIAK	S4	3.94	0.123	2.1×10^{-2}	6.0×10^{-2}	2.9×10^{-2}	1.4×10^{-1}	3.2×10^{-2}	ND	1.1×10^{-2}	1.5×10^{-2}	ND	ND	2.0×10^{-2}	$\mathbf{1.0 \times 10^{-2}}$
CDK1 CDK2 CDK3	P06493 P24941 Q00526	Cell division protein kinase 1/2/3	IGEGpTYGWYK	T14	6.99	0.077	1.9×10^{-2}	4.8×10^{-2}	1.7×10^{-2}	8.9×10^{-2}	7.2×10^{-3}	1.6×10^{-2}	5.1×10^{-3}	ND	5.4×10^{-3}	7.5×10^{-3}	2.8×10^{-3}	2.4×10^{-3}
LMO7	Q8WWI1	LIM domain only protein 7	pSYTSDLOK	S417	3.07	0.048	2.2×10^{-2}	5.0×10^{-2}	1.4×10^{-2}	3.5×10^{-2}	8.5×10^{-2}	3.2×10^{-2}	1.3×10^{-2}	ND	6.4×10^{-3}	9.0×10^{-3}	2.0×10^{-2}	1.7×10^{-2}
ALG5	Q92685	Dolichyl-P-Man;Man(5)GlcNAc (2)-PP-dolichyl mannosyltransferase	SGpSAAQAEGLCK	S13	2.74	0.049	2.1×10^{-1}	2.1×10^{-1}	1.4×10^{-1}	2.1×10^{-1}	5.1×10^{-2}	2.6×10^{-2}	4.5×10^{-2}	2.8×10^{-2}	5.1×10^{-2}	6.0×10^{-2}	1.7×10^{-2}	8.7×10^{-2}
PDS5A	Q29RF7−1	Sister chromatid cohesion prolein PDS5 homologue A	IISVpTPVK	T1208	5.14	0.044	6.7×10^{-2}	2.3×10^{-1}	2.1×10^{-2}	2.6×10^{-1}	1.0×10^{-1}	3.6×10^{-2}	2.4×10^{-2}	1.1×10^{-2}	1.7×10^{-2}	3.1×10^{-2}	1.9×10^{-2}	3.7×10^{-2}
CCR1	P32246	C−C chemokine receptor type 1	VSSTSPSTGEHELpSAGF	S352	2.34	0.046	1.3×10^{-1}	1.4×10^{-1}	9.0×10^{-2}	1.0×10^{-1}	3.8×10^{-2}	2.7×10^{-2}	3.4×10^{-2}	ND	3.1×10^{-2}	2.5×10^{-2}	4.2×10^{-2}	5.7×10^{-2}
MCM2	O95297-1	DNA replication licensing faclor MCM2	GLLYDpSDEEDEERPAR	S139	5.30	0.156	3.6×10^{-1}	2.0	2.5×10^{-1}	2.1	3.9×10^{-2}	1.3×10^{-1}	1.3×10^{-1}	ND	9.4×10^{-2}	2.8×10^{-1}	3.3×10^{-2}	2.3×10^{-1}
CDK1 CDK2 CDK3	P60493 P24941 Q00526	Cell division protein kinase 1/2/3	IGEGTpYGWYK	Y15	5.09	0.074	1.0×10^{-1}	1.1×10^{-1}	1.5×10^{-2}	2.1×10^{-1}	2.5×10^{-2}	4.0×10^{-2}	1.5×10^{-2}	ND	6.3×10^{-3}	1.8×10^{-2}	2.1×10^{-2}	2.0×10^{-2}
MPZL1	095297−1	Myelin prolein zero-like protein 1	SESVVpYADIR	Y263	0.39	0.183	1.2×10^{-1}	2.1×10^{-1}	1.7×10^{-1}	1.5×10^{-1}	9.7×10^{-2}	1.7×10^{-1}	4.6×10^{-1}	1.5×10^{-1}	3.8×10^{-1}	1.1	1.3×10^{-1}	1.3×10^{-1}
KRT8	P05787	Keratin, type II cytoskeletal 8	YEELQpSLAGK	S291	0.81	0.533	ND	3.5×10^{-2}	1.8×10^{-2}	5.4×10^{-2}	2.6×10^{-2}	1.9×10^{-2}	3.9×10^{-2}	ND	6.6×10^{-2}	4.9×10^{-2}	2.7×10^{-2}	1.2×10^{-2}
MUC1	P15941−1	Mucin-1	YVPPSSTDRpSPYEK	S1227	3.10	0.170	0.63	0.46	0.72			0.64	1.02	1.1	2.03		1.84	1.28
PKP2	Q99959−1	Plakophilin-2	LELpSPDSSPER	S151	0.37	0.243	0.07	0.27	0.16	0.48	0.34	0.3	1.11	0.32	0.22	5.96	0.84	0.78
INADL	Q8NI35−1	InaD like protein	LFDDEApSVDEPR	S645	1.08	0.834	1.1×10^{-2}	9.3×10^{-3}	4.0×10^{-3}	1.1×10^{-2}	ND	ND	1.5×10^{-2}	4.8×10^{-3}	3.9×10^{-3}	1.5×10^{-2}	2.9×10^{-3}	8.5×10^{-3}
SHROOM3	Q8TF72−1	shroom family member 3 protein	pSPLNSPPVKPK	S439	0.46	0.071	6.9×10^{-3}	4.7×10^{-3}	1.9×10^{-3}	3.7×10^{-3}	1.9×10^{-3}	6.3×10^{-3}	1.0×10^{-3}	2.8×10^{-3}	4.3×10^{-3}	1.3×10^{-3}	ND	1.5×10^{-3}

[a]ND: not detected.

dx.doi.org/10.1021/pr3005474 | J. Proteome Res. 2012, 11, 5311−5322

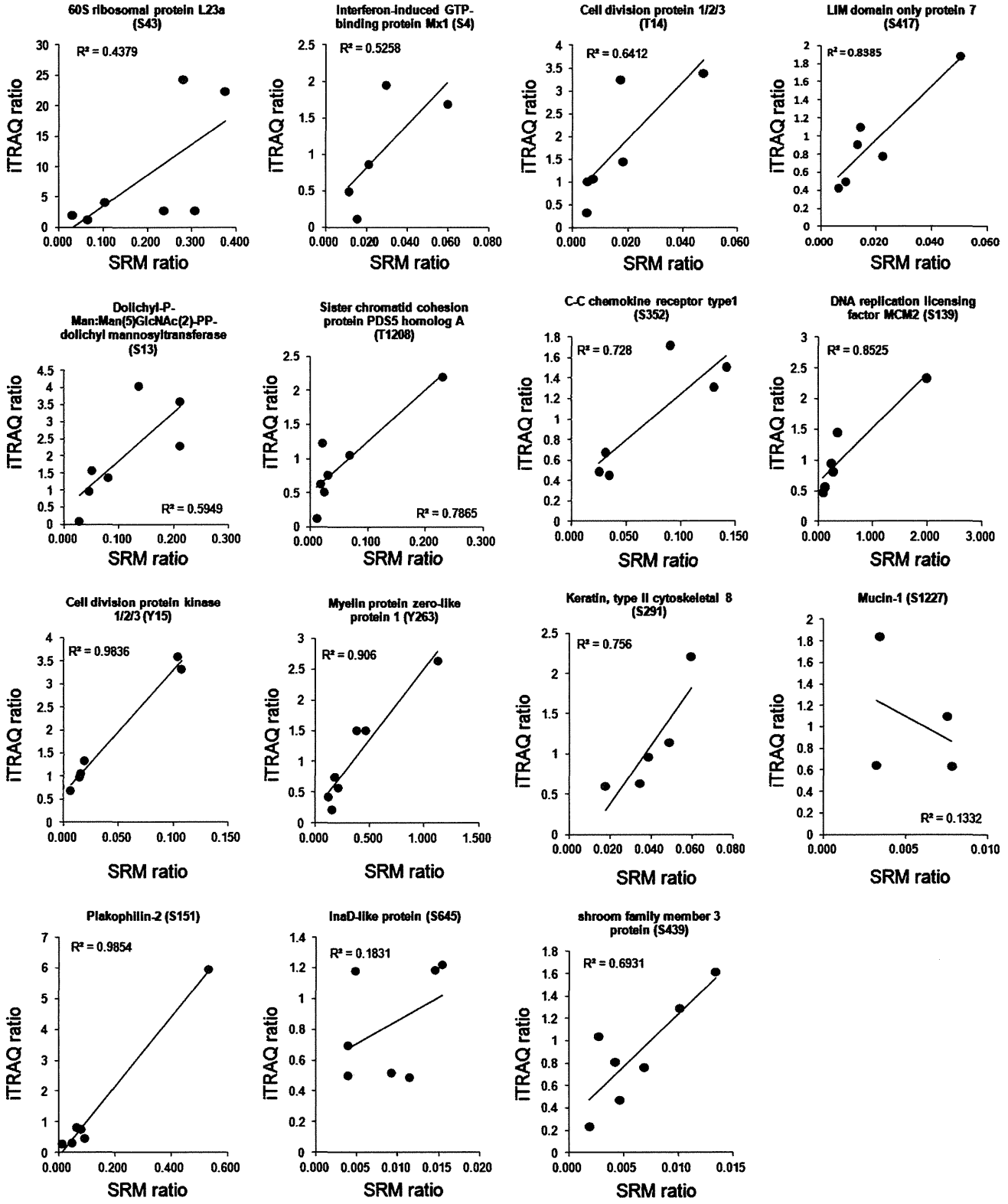


Figure 3. Linear regression comparing peptide ratio results obtained by iTRAQ and SRM assay. The iTRAQ and SRM ratios were plotted on each graph. Each data point represents a given peptide ratio in the same samples, which were quantified by either iTRAQ or SRM assay. Correlation coefficients are shown in plots.

phosphorylation site. For example, FVpSEGDGGR (fold change: 0.26) and RFVpSEGDGGR (0.48) peptides with pS457 of programmed cell death protein 4 and AGGSAALpSPSK (2.49) and AGGSAALpSPSKK (2.08) peptides with pS31 of Histone H1x both showed significant differences between high- and low-risk groups (Supporting Information Table S4), which would indicate that our large-scale phosphoproteomic analysis had sufficient quantitative reproducibility to search for putative phosphoprotein biomarkers.

dx.doi.org/10.1021/pr3005474 | J. Proteome Res. 2012, 11, 5311–5322

Verification of the phosphorylation state is essential in the search for phospho-biomarkers. If specific and well-characterized antibodies for these candidates are available, the validation step could be performed easily using Western blotting and ELISA. However, highly specific antibodies for most phosphoproteins are not available, and the development of good antibodies that recognize a specific phosphorylation frame is a cumbersome, expensive and time-consuming process that requires a priori knowledge of the protein and its phosphorylation sites. On the other hand, SRM does not require antibodies and is able to validate multiple phosphorylation sites within a single run. Recently, SRM analysis was used to validate the evidence for a large-scale proteome;[55−57] however, phosphopeptide SRM has only been performed for specific protein phosphorylation such as Akt,[58] Lyn,[59] EGFR[60] or tyrosine phosphorylated peptides after EGF treatment.[61] In this study, we selected and validated 19 phosphopeptides from 133 biomarker candidate phosphopeptides of breast cancer tissue discovered by iTRAQ-based phosphoproteomics. To our knowledge, this study is the first to validate phosphopeptides discovered by large-scale phosphoproteomic analysis using SRM.

Recently, several reports identified biomarker candidates by quantitative shotgun proteomics and subsequent validation by SRM,[57,62] but these studies carried out the SRM assay without SI peptides. Although this method has the advantage of reducing the cost and time for SI peptide synthesis, difficulties occur in SRM analysis without internal standards, which provide the correct retention time for target peptides and verify the specificity of the analyte.[56] Also, the use of SI peptide provides the most favorable SRM information, such as the highest intensity fragment ions for each peptide. Thus, inclusion of SI peptides as an internal control is indispensable, especially for quantitation of low-abundance proteins such as phosphopeptide. Whiteaker et al. pointed out that the choice of candidates for quantitative SRM assay development was limited to the most abundant proteins or peptides without internal standards.[56] Our successful quantitation of low abundant phosphopeptides was largely a result of the inclusion of SI peptides.

In this study, only four of 15 potential biomarker candidate phosphopeptides quantified by iTRAQ showed a significant difference between high- and low-risk groups of breast cancer (Figure 2) and quantification of the amount of phosphopeptides by iTRAQ and SRM was not always correlated (Figure 3). Several reasons could be considered for the discrepancy. First, the sample size used for both discovery and verification was very small. In addition to the four phosphopeptides with a significant difference obtained by our SRM analysis, eight candidate phosphopeptides showed quite different expressions, although not significantly, between high- and low-risk groups. If we could increase the number of samples, more biomarker candidate phosphopeptides identified by the discovery approach could be verified. Second, quantitative variation might be generated in the discovery phase of phosphoproteomics. This includes one additional step of phosphopeptide enrichment by IMAC as compared with the usual iTRAQ method, which might create such variation. Moreover, the heterogeneity of cancer tissue samples could further highlight quantitative variation in a step of phosphopeptide enrichment. This is evidenced by the fact that good reproducibility and correlation were obtained between the quantitation of phosphopeptides by iTRAQ and SRM when their analysis was performed using samples prepared from cell lysate (data not shown). Phosphopeptide enrichment might be more sensitive to the composition of the sample and the solution used for lysis or digestion of protein extracts; therefore, validation by SRM analysis is very important for the biomarker candidate phosphopeptides discovered by iTRAQ analysis combined with IMAC. Third, the endogenous phosphopeptide level is near the limit of quantitation so that the number of phosphopeptides quantified even with highly sensitive SRM was not accurate enough. We have observed that iTRAQ-based discovery and SRM-based validation of biomarker candidates of membrane proteins obtained from breast cancer tissues were well correlated.[63] This was probably due to the abundance of membrane proteins as compared with phosphoprotein. Thus, further improvement of the sensitivity of SRM is needed for accurate quantitation of low-abundance protein such as phosphoprotein.

In conclusion, we performed a large-scale phosphoproteome quantification and subsequent SRM-based validation using breast cancer tissue samples. The significance of this study is to provide a strategy for the quantitation and validation of low-abundance phosphopeptides using the most recent proteomic technologies, which might lead to a fundamental shift from traditional validation using antibodies. Quantitation of phosphopeptides by SRM will be applied to examine various kinase activities and signaling pathways in cells in the near future.

■ ASSOCIATED CONTENT

Supporting Information

Figure S1. Schematic workflow of iTRAQ analysis combined with IMAC for identification of potential biomarkers and SRM analysis combined with IMAC for validation. Figure S2. Venn diagram of the phosphopeptides identified in the four experiments of iTRAQ-based proteomic analysis. Figure S3. Expression of Mucin-1 protein in breast cancer tissues. Table S1. Patient information in experiment. Table S2. Identified phosphopeptides. Table S3. Quantified phosphopeptides. Table S4. Phosphopeptides with significant difference between two groups by iTRAQ analysis. Table S5. Transition list of target phosphopeptides. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Tel: +81-72-641-9862. Fax: +81-72-641-9861. E-mail: tomonaga@nibio.go.jp.

Author Contributions

#Ryohei Narumi and Tatsuo Murakami contributed equally to this paper.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by a Grant-in-Aid for Research on Biological Markers for New Drug Development H20-0005 to T.T. from the Ministry of Health, Labour and Welfare of Japan and by Grant-in-Aid 21390354 to T.T. from the Ministry of Education, Science, Sports and Culture of Japan.

■ ABBREVIATIONS

iTRAQ, isobaric peptide tags for relative and absolute quantification; SRM, selected reaction monitoring; IMAC, immobilized metal affinity chromatography; SI peptide, stable isotope-labeled peptide; SCX, strong cation exchange; CID,

collision-induced dissociation; HCD, higher energy collision-induced dissociation; LC–MS/MS, liquid chromatography–tandem mass spectrometry; CE, collision energy; LTQ, linear ion trap; fwhm, full width at half-maximum; FDR, false discovery rate

REFERENCES

(1) Hanahan, D.; Weinberg, R. A. The hallmarks of cancer. *Cell* **2000**, *100* (1), 57–70.

(2) Kaminska, B. MAPK signalling pathways as molecular targets for anti-inflammatory therapy–from molecular mechanisms to therapeutic benefits. *Biochim. Biophys. Acta* **2005**, *1754* (1–2), 253–62.

(3) Peifer, C.; Wagner, G.; Laufer, S. New approaches to the treatment of inflammatory disorders small molecule inhibitors of p38 MAP kinase. *Curr. Top. Med. Chem.* **2006**, *6* (2), 113–49.

(4) White, M. F. Regulating insulin signaling and beta-cell function through IRS proteins. *Can. J. Physiol. Pharmacol.* **2006**, *84* (7), 725–37.

(5) Neville, D. C.; Rozanas, C. R.; Price, E. M.; Gruis, D. B.; Verkman, A. S.; Townsend, R. R. Evidence for phosphorylation of serine 753 in CFTR using a novel metal-ion affinity resin and matrix-assisted laser desorption mass spectrometry. *Protein Sci.* **1997**, *6* (11), 2436–45.

(6) Ong, S. E.; Blagoev, B.; Kratchmarova, I.; Kristensen, D. B.; Steen, H.; Pandey, A.; Mann, M. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol. Cell. Proteomics* **2002**, *1* (5), 376–86.

(7) Ross, P. L.; Huang, Y. N.; Marchese, J. N.; Williamson, B.; Parker, K.; Hattan, S.; Khainovski, N.; Pillai, S.; Dey, S.; Daniels, S.; Purkayastha, S.; Juhasz, P.; Martin, S.; Bartlet-Jones, M.; He, F.; Jacobson, A.; Pappin, D. J. Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. *Mol. Cell. Proteomics* **2004**, *3* (12), 1154–69.

(8) Collins, M. O.; Yu, L.; Coba, M. P.; Husi, H.; Campuzano, I.; Blackstock, W. P.; Choudhary, J. S.; Grant, S. G. Proteomic analysis of in vivo phosphorylated synaptic proteins. *J. Biol. Chem.* **2005**, *280* (7), 5972–82.

(9) Molina, H.; Horn, D. M.; Tang, N.; Mathivanan, S.; Pandey, A. Global proteomic profiling of phosphopeptides using electron transfer dissociation tandem mass spectrometry. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104* (7), 2199–204.

(10) Wissing, J.; Jansch, L.; Nimtz, M.; Dieterich, G.; Hornberger, R.; Keri, G.; Wehland, J.; Daub, H. Proteomics analysis of protein kinases by target class-selective prefractionation and tandem mass spectrometry. *Mol. Cell. Proteomics* **2007**, *6* (3), 537–47.

(11) Villen, J.; Beausoleil, S. A.; Gerber, S. A.; Gygi, S. P. Large-scale phosphorylation analysis of mouse liver. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104* (5), 1488–93.

(12) Ballif, B. A.; Villen, J.; Beausoleil, S. A.; Schwartz, D.; Gygi, S. P. Phosphoproteomic analysis of the developing mouse brain. *Mol. Cell. Proteomics* **2004**, *3* (11), 1093–101.

(13) Beausoleil, S. A.; Jedrychowski, M.; Schwartz, D.; Elias, J. E.; Villen, J.; Li, J.; Cohn, M. A.; Cantley, L. C.; Gygi, S. P. Large-scale characterization of HeLa cell nuclear phosphoproteins. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101* (33), 12130–5.

(14) Ficarro, S. B.; McCleland, M. L.; Stukenberg, P. T.; Burke, D. J.; Ross, M. M.; Shabanowitz, J.; Hunt, D. F.; White, F. M. Phosphoproteome analysis by mass spectrometry and its application to Saccharomyces cerevisiae. *Nat. Biotechnol.* **2002**, *20* (3), 301–5.

(15) Lee, J.; Xu, Y.; Chen, Y.; Sprung, R.; Kim, S. C.; Xie, S.; Zhao, Y. Mitochondrial phosphoproteome revealed by an improved IMAC method and MS/MS/MS. *Mol. Cell. Proteomics* **2007**, *6* (4), 669–76.

(16) Moser, K.; White, F. M. Phosphoproteomic analysis of rat liver by high capacity IMAC and LC–MS/MS. *J. Proteome Res.* **2006**, *5* (1), 98–104.

(17) Trinidad, J. C.; Specht, C. G.; Thalhammer, A.; Schoepfer, R.; Burlingame, A. L. Comprehensive identification of phosphorylation sites in postsynaptic density preparations. *Mol. Cell. Proteomics* **2006**, *5* (5), 914–22.

(18) Li, X.; Gerber, S. A.; Rudner, A. D.; Beausoleil, S. A.; Haas, W.; Villen, J.; Elias, J. E.; Gygi, S. P. Large-scale phosphorylation analysis of alpha-factor-arrested Saccharomyces cerevisiae. *J. Proteome Res.* **2007**, *6* (3), 1190–7.

(19) Matsuoka, S.; Ballif, B. A.; Smogorzewska, A.; McDonald, E. R., 3rd; Hurov, K. E.; Luo, J.; Bakalarski, C. E.; Zhao, Z.; Solimini, N.; Lerenthal, Y.; Shiloh, Y.; Gygi, S. P.; Elledge, S. J. ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. *Science* **2007**, *316* (5828), 1160–6.

(20) Olsen, J. V.; Blagoev, B.; Gnad, F.; Macek, B.; Kumar, C.; Mortensen, P.; Mann, M. Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell* **2006**, *127* (3), 635–48.

(21) Trinidad, J. C.; Thalhammer, A.; Specht, C. G.; Lynn, A. J.; Baker, P. R.; Schoepfer, R.; Burlingame, A. L. Quantitative analysis of synaptic phosphorylation and protein expression. *Mol. Cell. Proteomics* **2008**, *7* (4), 684–96.

(22) Nguyen, V.; Cao, L.; Lin, J. T.; Hung, N.; Ritz, A.; Yu, K.; Jianu, R.; Ulin, S. P.; Raphael, B. J.; Laidlaw, D. H.; Brossay, L.; Salomon, A. R. A new approach for quantitative phosphoproteomic dissection of signaling pathways applied to T cell receptor activation. *Mol. Cell. Proteomics* **2009**, *8* (11), 2418–31.

(23) Zanivan, S.; Gnad, F.; Wickstrom, S. A.; Geiger, T.; Macek, B.; Cox, J.; Fassler, R.; Mann, M. Solid tumor proteome and phosphoproteome analysis by high resolution mass spectrometry. *J. Proteome Res.* **2008**, *7* (12), 5314–26.

(24) Anderson, N. L. The roles of multiple proteomic platforms in a pipeline for new diagnostics. *Mol. Cell. Proteomics* **2005**, *4* (10), 1441–4.

(25) Lange, V.; Malmstrom, J. A.; Didion, J.; King, N. L.; Johansson, B. P.; Schafer, J.; Rameseder, J.; Wong, C. H.; Deutsch, E. W.; Brusniak, M. Y.; Buhlmann, P.; Bjorck, L.; Domon, B.; Aebersold, R. Targeted quantitative analysis of Streptococcus pyogenes virulence factors by multiple reaction monitoring. *Mol. Cell. Proteomics* **2008**, *7* (8), 1489–500.

(26) Kuhn, E.; Wu, J.; Karl, J.; Liao, H.; Zolg, W.; Guild, B. Quantification of C-reactive protein in the serum of patients with rheumatoid arthritis using multiple reaction monitoring mass spectrometry and 13C-labeled peptide standards. *Proteomics* **2004**, *4* (4), 1175–86.

(27) Anderson, L.; Hunter, C. L. Quantitative mass spectrometric multiple reaction monitoring assays for major plasma proteins. *Mol. Cell. Proteomics* **2006**, *5* (4), 573–88.

(28) Keshishian, H.; Addona, T.; Burgess, M.; Kuhn, E.; Carr, S. A. Quantitative, multiplexed assays for low abundance proteins in plasma by targeted mass spectrometry and stable isotope dilution. *Mol. Cell. Proteomics* **2007**, *6* (12), 2212–29.

(29) Keshishian, H.; Addona, T.; Burgess, M.; Mani, D. R.; Shi, X.; Kuhn, E.; Sabatine, M. S.; Gerszten, R. E.; Carr, S. A. Quantification of cardiovascular biomarkers in patient plasma by targeted mass spectrometry and stable isotope dilution. *Mol. Cell. Proteomics* **2009**, *8* (10), 2339–49.

(30) van 't Veer, L. J.; Dai, H.; van de Vijver, M. J.; He, Y. D.; Hart, A. A.; Mao, M.; Peterse, H. L.; van der Kooy, K.; Marton, M. J.; Witteveen, A. T.; Schreiber, G. J.; Kerkhoven, R. M.; Roberts, C.; Linsley, P. S.; Bernards, R.; Friend, S. H. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* **2002**, *415* (6871), 530–6.

(31) Masuda, T.; Tomita, M.; Ishihama, Y. Phase transfer surfactant-aided trypsin digestion for membrane proteome analysis. *J. Proteome Res.* **2008**, *7* (2), 731–40.

(32) Matsumoto, M.; Oyamada, K.; Takahashi, H.; Sato, T.; Hatakeyama, S.; Nakayama, K. I. Large-scale proteomic analysis of tyrosine-phosphorylation induced by T-cell receptor or B-cell receptor activation reveals new signaling pathways. *Proteomics* **2009**, *9* (13), 3549–63.

(33) Kokubu, M.; Ishihama, Y.; Sato, T.; Nagasu, T.; Oda, Y. Specificity of immobilized metal affinity-based IMAC/C18 tip enrichment of phosphopeptides for protein phosphorylation analysis. *Anal. Chem.* **2005**, *77* (16), 5144–54.
(34) Taus, T.; Kocher, T.; Pichler, P.; Paschke, C.; Schmidt, A.; Henrich, C.; Mechtler, K. Universal and confident phosphorylation site localization using phosphoRS. *J. Proteome Res.* **2011**, *10* (12), 5354–62.
(35) Martins-de-Souza, D.; Guest, P. C.; Vanattou-Saifoudine, N.; Rahmoune, H.; Bahn, S. Phosphoproteomic differences in major depressive disorder postmortem brains indicate effects on synaptic function. *Eur. Arch. Psychiatry Clin. Neurosci.* **2012**.
(36) Jarvinen, T. A.; Tanner, M.; Barlund, M.; Borg, A.; Isola, J. Characterization of topoisomerase II alpha gene amplification and deletion in breast cancer. *Genes, Chromosomes Cancer* **1999**, *26* (2), 142–50.
(37) Nielsen, K. V.; Muller, S.; Moller, S.; Schonau, A.; Balslev, E.; Knoop, A. S.; Ejlertsen, B. Aberrations of ERBB2 and TOP2A genes in breast cancer. *Mol. Oncol.* **2010**, *4* (2), 161–8.
(38) Futreal, P. A.; Liu, Q.; Shattuck-Eidens, D.; Cochran, C.; Harshman, K.; Tavtigian, S.; Bennett, L. M.; Haugen-Strano, A.; Swensen, J.; Miki, Y.; et al. BRCA1 mutations in primary breast and ovarian carcinomas. *Science* **1994**, *266* (5182), 120–2.
(39) O'Donovan, P. J.; Livingston, D. M. BRCA1 and BRCA2: breast/ovarian cancer susceptibility gene products and participants in DNA double-strand break repair. *Carcinogenesis* **2010**, *31* (6), 961–7.
(40) Castilla, L. H.; Couch, F. J.; Erdos, M. R.; Hoskins, K. F.; Calzone, K.; Garber, J. E.; Boyd, J.; Lubin, M. B.; Deshano, M. L.; Brody, L. C.; et al. Mutations in the BRCA1 gene in families with early-onset breast and ovarian cancer. *Nat. Genet.* **1994**, *8* (4), 387–91.
(41) Kim, S. J.; Nakayama, S.; Miyoshi, Y.; Taguchi, T.; Tamaki, Y.; Matsushima, T.; Torikoshi, Y.; Tanaka, S.; Yoshida, T.; Ishihara, H.; Noguchi, S. Determination of the specific activity of CDK1 and CDK2 as a novel prognostic indicator for early breast cancer. *Ann. Oncol.* **2008**, *19* (1), 68–72.
(42) Nasir, A.; Chen, D. T.; Gruidl, M.; Henderson-Jackson, E. B.; Venkataramu, C.; McCarthy, S. M.; McBride, H. L.; Harris, E.; Khakpour, N.; Yeatman, T. J. Novel molecular markers of malignancy in histologically normal and benign breast. *Pathol. Res. Int.* **2011**, *2011*, 489064.
(43) Zheng, M. Z.; Zheng, L. M.; Zeng, Y. X. SCC-112 gene is involved in tumor progression and promotes the cell proliferation in G2/M phase. *J. Cancer Res. Clin. Oncol.* **2008**, *134* (4), 453–62.
(44) Rakha, E. A.; Boyce, R. W.; Abd El-Rehim, D.; Kurien, T.; Green, A. R.; Paish, E. C.; Robertson, J. F.; Ellis, I. O. Expression of mucins (MUC1, MUC2, MUC3, MUC4, MUC5AC and MUC6) and their prognostic significance in human breast cancer. *Mod. Pathol.* **2005**, *18* (10), 1295–304.
(45) Wei, X.; Xu, H.; Kufe, D. MUC1 oncoprotein stabilizes and activates estrogen receptor alpha. *Mol. Cell* **2006**, *21* (2), 295–305.
(46) Mulligan, A. M.; Pinnaduwage, D.; Bane, A. L.; Bull, S. B.; O'Malley, F. P.; Andrulis, I. L. CK8/18 expression, the basal phenotype, and family history in identifying BRCA1-associated breast cancer in the Ontario site of the breast cancer family registry. *Cancer* **2011**, *117* (7), 1350–9.
(47) Busch, T.; Milena; Eiseler, T.; Joodi, G.; Temme, C.; Jansen, J.; Wichert, G. V.; Omary, M. B.; Spatz, J.; Seufferlein, T. Keratin 8 phosphorylation regulates keratin reorganization and migration of epithelial tumor cells. *J. Cell Sci.* **2012**, 2148–59.
(48) Hu, Q.; Guo, C.; Li, Y.; Aronow, B. J.; Zhang, J. LMO7 mediates cell-specific activation of the Rho-myocardin-related transcription factor-serum response factor pathway and plays an important role in breast cancer cell migration. *Mol. Cell. Biol.* **2011**, *31* (16), 3223–40.
(49) Merrell, K. W.; Crofts, J. D.; Smith, R. L.; Sin, J. H.; Kmetzsch, K. E.; Merrell, A.; Miguel, R. O.; Candelaria, N. R.; Lin, C. Y. Differential recruitment of nuclear receptor coregulators in ligand-dependent transcriptional repression by estrogen receptor-alpha. *Oncogene* **2011**, *30* (13), 1608–14.
(50) Hartmaier, R. J.; Tchatchou, S.; Richter, A. S.; Wang, J.; McGuire, S. E.; Skaar, T. C.; Rae, J. M.; Hemminki, K.; Sutter, C.; Ditsch, N.; Bugert, P.; Weber, B. H.; Niederacher, D.; Arnold, N.; Varon-Mateeva, R.; Wappenschmidt, B.; Schmutzler, R. K.; Meindl, A.; Bartram, C. R.; Burwinkel, B.; Oesterreich, S. Nuclear receptor coregulator SNP discovery and impact on breast cancer risk. *BMC Cancer* **2009**, *9*, 438.
(51) Lemeer, S.; Heck, A. J. The phosphoproteomics data explosion. *Curr. Opin. Chem. Biol.* **2009**, *13* (4), 414–20.
(52) Larsen, M. R.; Thingholm, T. E.; Jensen, O. N.; Roepstorff, P.; Jorgensen, T. J. Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns. *Mol. Cell. Proteomics* **2005**, *4* (7), 873–86.
(53) Rush, J.; Moritz, A.; Lee, K. A.; Guo, A.; Goss, V. L.; Spek, E. J.; Zhang, H.; Zha, X. M.; Polakiewicz, R. D.; Comb, M. J. Immunoaffinity profiling of tyrosine phosphorylation in cancer cells. *Nat. Biotechnol.* **2005**, *23* (1), 94–101.
(54) Dephoure, N.; Zhou, C.; Villen, J.; Beausoleil, S. A.; Bakalarski, C. E.; Elledge, S. J.; Gygi, S. P. A quantitative atlas of mitotic phosphorylation. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105* (31), 10762–7.
(55) Ostasiewicz, P.; Zielinska, D. F.; Mann, M.; Wisniewski, J. R. Proteome, phosphoproteome, and N-glycoproteome are quantitatively preserved in formalin-fixed paraffin-embedded tissue and analyzable by high-resolution mass spectrometry. *J. Proteome Res.* **2010**, *9* (7), 3688–700.
(56) Whiteaker, J. R.; Lin, C.; Kennedy, J.; Hou, L.; Trute, M.; Sokal, I.; Yan, P.; Schoenherr, R. M.; Zhao, L.; Voytovich, U. J.; Kelly-Spratt, K. S.; Krasnoselsky, A.; Gafken, P. R.; Hogan, J. M.; Jones, L. A.; Wang, P.; Amon, L.; Chodosh, L. A.; Nelson, P. S.; McIntosh, M. W.; Kemp, C. J.; Paulovich, A. G. A targeted proteomics-based pipeline for verification of biomarkers in plasma. *Nat. Biotechnol.* **2011**, *29* (7), 625–34.
(57) Addona, T. A.; Shi, X.; Keshishian, H.; Mani, D. R.; Burgess, M.; Gillette, M. A.; Clauser, K. R.; Shen, D.; Lewis, G. D.; Farrell, L. A.; Fifer, M. A.; Sabatine, M. S.; Gerszten, R. E.; Carr, S. A. A pipeline that integrates the discovery and verification of plasma protein biomarkers reveals candidate markers for cardiovascular disease. *Nat. Biotechnol.* **2011**, *29* (7), 635–43.
(58) Atrih, A.; Turnock, D.; Sellar, G.; Thompson, A.; Feuerstein, G.; Ferguson, M. A.; Huang, J. T. Stoichiometric quantification of Akt phosphorylation using LC–MS/MS. *J. Proteome Res.* **2010**, *9* (2), 743–51.
(59) Jin, L. L.; Tong, J.; Prakash, A.; Peterman, S. M.; St-Germain, J. R.; Taylor, P.; Trudel, S.; Moran, M. F. Measurement of protein phosphorylation stoichiometry by selected reaction monitoring mass spectrometry. *J. Proteome Res.* **2010**, *9* (5), 2752–61.
(60) Tong, J.; Taylor, P.; Peterman, S. M.; Prakash, A.; Moran, M. F. Epidermal growth factor receptor phosphorylation sites Ser991 and Tyr998 are implicated in the regulation of receptor endocytosis and phosphorylations at Ser1039 and Thr1041. *Mol. Cell. Proteomics* **2009**, *8* (9), 2131–44.
(61) Wolf-Yadlin, A.; Hautaniemi, S.; Lauffenburger, D. A.; White, F. M. Multiple reaction monitoring for robust quantitative proteomic analysis of cellular signaling networks. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104* (14), 5860–5.
(62) Thingholm, T. E.; Bak, S.; Beck-Nielsen, H.; Jensen, O. N.; Gaster, M. Characterization of human myotubes from type 2 diabetic and nondiabetic subjects using complementary quantitative mass spectrometric methods. *Mol. Cell. Proteomics* **2011**, *10* (9), M110 006650.
(63) Muraoka, S.; Kume, H.; Watanabe, S.; Adachi, J.; Kuwano, M.; Sato, M.; Kawasaki, N.; Kodera, Y.; Ishitobi, M.; Inaji, H.; Miyamoto, Y.; Kato, K.; Tomonaga, T. Strategy for SRM-based verification of biomarker candidates discovered by iTRAQ method in limited breast cancer tissue samples. *J. Proteome Res.* **2012**, *11* (8), 4201–10.

Journal of proteome research

Article

pubs.acs.org/jpr

Strategy for SRM-based Verification of Biomarker Candidates Discovered by iTRAQ Method in Limited Breast Cancer Tissue Samples

Satoshi Muraoka,[†] Hideaki Kume,[†] Shio Watanabe,[†] Jun Adachi,[†] Masayoshi Kuwano,[†] Misako Sato,[†] Naoko Kawasaki,[†] Yoshio Kodera,[‡,§] Makoto Ishitobi,[∥] Hideo Inaji,[∥] Yasuhide Miyamoto,[⊥] Kikuya Kato,[⊥] and Takeshi Tomonaga*[,†,§]

[†]Laboratory of Proteome Research, National Institute of Biomedical Innovation, Ibaraki, Japan

[‡]Laboratory of Biomolecular Dynamics, Department of Physics, Kitasato University School of Science, Sagamihara, Japan

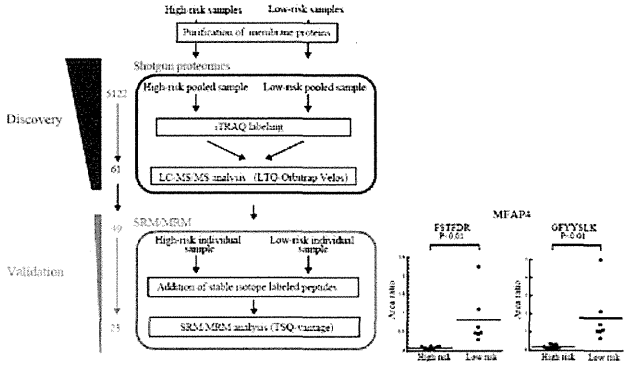
[§]Clinical Proteomics Research Center, Chiba University Hospital, Chiba, Japan

[∥]Department of Breast and Endocrine Surgery, Osaka Medical Center for Cancer and Cardiovascular Diseases, Osaka, Japan

[⊥]Research Institute, Osaka Medical Center for Cancer and Cardiovascular Diseases, Osaka, Japan

Supporting Information

ABSTRACT: Since LC−MS-based quantitative proteomics has become increasingly applied to a wide range of biological applications over the past decade, numerous studies have performed relative and/or absolute abundance determinations across large sets of proteins. In this study, we discovered prognostic biomarker candidates from limited breast cancer tissue samples using discovery-through-verification strategy combining iTRAQ method followed by selected reaction monitoring/multiple reaction monitoring analysis (SRM/MRM). We identified and quantified 5122 proteins with high confidence in 18 patient tissue samples (pooled high-risk (n = 9) or low-risk (n = 9)). A total of 2480 proteins (48.4%) of them were annotated as membrane proteins, 16.1% were plasma membrane and 6.6% were extracellular space proteins by Gene Ontology analysis. Forty-nine proteins with >2-fold differences in two groups were chosen for further analysis and verified in 16 individual tissue samples (high-risk (n = 9) or low-risk (n = 7)) using SRM/MRM. Twenty-three proteins were differentially expressed among two groups of which MFAP4 and GP2 were further confirmed by Western blotting in 17 tissue samples (high-risk (n = 9) or low-risk (n = 8)) and Immunohistochemistry (IHC) in 24 tissue samples (high-risk (n = 12) or low-risk (n = 12)). These results indicate that the combination of iTRAQ and SRM/MRM proteomics will be a powerful tool for identification and verification of candidate protein biomarkers.



KEYWORDS: *iTRAQ, MammaPrint, SRM/MRM, shotgun proteomics, PTS, biomarker, plasma membrane*

INTRODUCTION

In a recently study, LC−MS/MS approaches are being widely used for clinical tissue or cell line analysis.[1,2] To perform quantification of proteins, iTRAQ, isotope-coded affinity tags (ICAT), and stable isotope labeling by amino acids in cell culture (SILAC) are technologies integrated with LC−MS/MS analysis for measuring protein expression levels in complex samples.[3−6] Numerous studies have used such proteomic technologies to discover candidate protein biomarkers for a range of diseases, including cancer. However, as of yet, no protein biomarker identified using proteomics has been introduced into clinical use. Because there are no quantitative assays for the majority of human proteins, assays (typically enzyme-linked immunosorbent assays (ELISA)) must be developed de novo for clinical testing of candidate protein biomarkers, and de novo assay development is prohibitively expensive for testing large numbers of candidate biomarkers. As a result, few putative biomarkers undergo rigorous validation, and the literature is replete with lengthy lists of candidates without follow-up. Recent advances in proteomic technologies have become an integral part of biomarker development workflow, including a discovery phase and subsequent qualification and validation of candidates in bodily fluids.[7,8] SRM mass spectrometry holds the promise to overcome the apparent bottleneck between candidate biomarker discovery and their initial quantitative evaluation. This technology has high reproducibility across complex samples. These attributes, together with its multiplexing capability, have spurred interest in SRM/MRM as an attractive alternative to ELISAs for rapid and multiplexed validation of candidate protein biomarkers.[9,10]

Received: April 3, 2012
Published: June 20, 2012

Plasma membranes are located at the interface between the cell and its environment such as neighboring cells, blood capillaries, or the extracellular matrix. Since plasma membrane proteins are of major significance in many cellular events such as transport of ions and other molecules, signal transduction, and cell-to-cell interaction, it is of importance to analyze these proteins to understand overall cellular function.[11] These proteins are of great interest, particularly because they could be key biomarkers for early diagnosis, progression of diseases, and suitable drug targets.[12] For membrane proteome analysis, however, these approaches cannot be directly applied because of difficulties in protein enrichment/solubilization and also subsequent protease digestion.[13,14] Numerous protocols have been reported to improve them.[15−19] Recently, Masuda et al. reported a new protocol based on the use of phase transfer surfactant as an enhancer for protein extraction, protein solubilization, and trypsin activation. This method was developed for membrane proteome analysis.[20]

Breast cancer is the most frequent malignancy of women in the western world. It is a disease with multiple subtypes, with differing prognostic and therapeutic implications. These two factors drive the need for determining appropriated therapy for patient subsets to avoid both undertreatment and overtreatment. Prognostic molecular tests for patients with breast cancer, including oncotype DX and MammaPrint, which is a prognostic tool to predict the risk (high-risk group with poor prognosis and low-risk group with good prognosis) of breast cancer metastasis using the expression of 70 genes, have already been approved for clinical use.[21,22] However, their cost is orders of magnitude higher than that of histological grading. Accordingly, numerous studies have been performed to discover prognostic protein biomarkers in breast cancer.[23]

In this study, we aimed to establish the strategy for high throughput verification of biomarker candidates identified by proteomic discovery using SRM/MRM. The two-step procedure was carried out for discovery and verification of novel prognostic biomarkers of breast cancer. As a first step, we performed shotgun quantitative proteomics using iTRAQ labeling with pooled high-risk or low-risk breast cancer tissue samples. To verify the candidate proteins, proteins that were differentially expressed among two groups in the iTRAQ discovery study were subsequently verified by SRM/MRM analysis with individual tissue samples. As the results, twenty-three proteins whose expression differs between two groups were identified. Altered expression of GP2 and MFAP4 were further confirmed by Western blot or IHC.

MATERIALS AND METHODS

Human Tissue Samples

Tissue samples were obtained from 27 patients with high-risk or low-risk MammaPrint breast cancer who had surgery at the Osaka Medical Center for Cancer & Cardiovascular Diseases. The medical information of 27 patients is summarized in Supplementary Table 1 (Supporting Information). All samples were frozen by liquid nitrogen and were stored at −80 °C until analysis. Written informed consent was obtained from all subjects. The Ethics Committee of our institute and the Osaka Medical Center for Cancer & Cardiovascular Diseases approved the protocol.

Enrichment of Membrane Proteins

For enrichment of membrane proteins, frozen tissue samples were homogenized in PBS containing protease inhibitor mixture (Complete; Roche, Mannheim, Germany) using a Dounce homogenizer (WHEATON, Millville, New Jersey, USA) following centrifugation (1000× *g*) for 10 min at 4 °C. Postnuclear supernatant was centrifuged at 100000× *g* for 1 h at 4 °C. The pellet was suspended in ice-cold 0.1 M Na_2CO_3 solution following centrifugation (100000× *g*) for 1 h at 4 °C. After centrifugation, the pellet was treated using a MPEX PTS reagent kit (GL sciences, Tokyo, Japan) as follows.[20] Briefly, the pellet was solubilized with PTS B buffer at 95 °C for 5 min followed by sonication for 5 min using a Bioruptor sonicator (Cosmo Bio, Tokyo, Japan). The solution was centrifuged at 100000× *g* for 30 min at 4 °C. Supernatant containing membrane proteins was stored at −80 °C. Protein concentration was determined using a DC protein assay kit (Bio-Rad, USA).

iTRAQ Labeling

Membrane proteins from pooled high-risk (n = 9) or low-risk (n = 9) breast cancer tissue samples were digested with trypsin (Proteomics grad; Roche, Swiss) and Lys-C (Wako Pure Chemical Industries, Osaka, Japan) and peptides were labeled with iTRAQ reagents according to the manufacturer's instructions (iTRAQ Reagents Multiplex kit; Applied Biosystems/MDS Sciex, Foster City, CA). Briefly, 90 μg of pooled membrane proteins were reduced with 10 mM dithiothreitol, alkylated with 20 mM iodoacetamide, and digested with 1:100 (w/w) Lys-C and 1:100 (w/w) trypsin. BSA (0.45 μg) was spiked into membrane protein samples as a quality control for iTRAQ labeling. The tryptic digest sample was desalted using C18 stage Tips.[24] Desalted samples were dissolved in 30 μL of dissolution buffer and labeled with two different iTRAQ reagents at room temperature for 1 h and quenched by Milli-Q water. Sample labeling was as follows: high-risk breast cancer tissue samples with 114 tag and low-risk breast cancer tissue samples with 115 tag. Labeled samples were mixed and dried by a Speed-Vac concentrator, dissolved in 100 μL of 2% acetonitrile (ACN), 0.1% formic acid (TFA), and desalted with C18 stage Tips.

Separation with Strong Cation Exchange Chromatography (SCX)

The iTRAQ labeled sample was fractionated using a HPLC system (Shimadzu prominence UFLC) fitted with a SCX column (50 mm × 2.1 mm, 5 mm, 300 Å, ZORBAX 300SCX, Agilent technology). The mobile phases consisted of (A); 25% ACN with 10 mM KH_2PO_4 (pH 3.0) and (B); (A) containing 1 M KCl. The mixed iTRAQ labeled sample was dissolved in 200 μL of buffer A and separated at a flow rate of 200 μL/min using a four-step linear gradient; 0% B for 30 min, 0 to 10% B in 15 min, 10 to 25% B in 10 min, 25 to 40% B in 5 min, and 40 to 100% B in 5 min, and 100% B in 10 min. A total of 36 fractions were collected, dried by the Speed-Vac concentrator, dissolved in 2% ACN and 0.1% TFA, and desalted with C18 stage Tips.

NanoLC−MS/MS

NanoLC−MS/MS analysis was conducted by an LTQ-Orbitrap Velos mass spectrometer (Thermo Fisher Scientific, Bremen, Germany) equipped with a nanoLC interface (AMR, Tokyo, Japan), a nanoHPLC system (Michrom Paradigm MS2), and an HTC-PAL autosampler (CTC, Analytics, Zwingen, Switzerland). L-column2 C18 particles (Chemicals Evaluation and Research Institute (CERI), Japan) were packed into a self-pulled needle (200 mm length × 100 mm inner diameter) using a Nanobaume capillary column packer (Western Fluids Engineering). The mobile

phases consisted of (A) 0.1% TFA and 2% ACN and (B) 0.1% TFA and 90% ACN. The SCX-fractionated peptides dissolved in 2% ACN and 0.1% TFA were loaded onto a trap column (0.3 × 5 mm, L-column ODS; CERI). The nanoLC gradient was delivered at 500 nL/min and consisted of a linear gradient of mobile phase B developed from 5 to 30% B in 135 min. A spray voltage of 2000 V was applied.

Data Acquisition with LTQ-Orbitrap Velos

Full MS scans were performed in the orbitrap mass analyzer of LTQ-Orbitrap Velos (scan range 350−1500 m/z, with 30K fwhm resolution at 400 m/z). In MS scans, the 10 most intense precursor ions were selected for MS/MS scans of LTQ-Orbitrap Velos, in which a dynamic exclusion option was implemented with a repeat count of one and exclusion duration of 60 s. This was followed by collision-induced dissociation (CID) MS/MS scans of the selected ions performed in the liner ion trap mass analyzer and further followed by higher energy collision-induced dissociation (HCD) MS/MS scans of the same precursor ions performed in the orbitrap mass analyzer with 7500 fwhm resolution at 400 m/z. The values of automated gain control (AGC) were set to 1.00×10^6 for full MS, 1.00×10^4 for CID MS/MS, and 5.00×10^4 for HCD MS/MS. Normalized collision energy values were set to 35% for CID and 50% for HCD.

Identification and Quantitation of Membrane Proteins

CID and HCD raw spectra were extracted and searched separately against UniProtKB/Swiss-Prot (release-2010_05) containing 40590 sequences (the forward and reverse-decoy) of *Homo sapiens* using Proteome Discoverer (Thermo Fisher Scientific, Beta Version 1.1) and Mascot v2. 3. 1. Search parameters included trypsin as the enzyme with one missed cleavage allowed; Carbamidomethylation at cysteine and iTRAQ labeling at lysine and the N-terminal residue were set as fixed modifications while oxidation at methionine and iTRAQ labeling at tyrosine were set as variable modifications. Precursor mass tolerance was set to 7 ppm and a fragment mass tolerance was set to 0.6 Da for CID and 0.01 Da for HCD. Peptide data were extracted using Mascot significance threshold 0.05, minimum peptide length 6, and top one peptide rank filters. Protein identification required at least two peptides (Supplementary Table 2 Column F, Supporting Information) and protein quantification required at least one unique peptide (Supplementary Table 2 Column G). FDR was calculated by enabling peptide sequence analysis using a decoy database. High confidence peptide identifications were obtained by setting a target FDR threshold of 0.18% at peptide level.

Bioinformatics Analysis

Mapping of putative transmembrane domains in identified proteins was carried out using the transmembrane hidden Markov model (TMHMM2.0) algorithm, available at http://www.cbs.dtu.dk/services/TMHMM.[25] The subcellular location and function of identified proteins were elucidated by the Ingenuity system, available at www.ingenuity.com and DAVID Bioinformatics Resources 6.7, available at http://david.abcc.ncifcrf.gov/home.jsp.

Stable Isotope-Labeled Peptides

Proteotypic peptides were chosen based on iTRAQ data. For SRM analysis of the 49 target proteins, stable isotope-labeled peptides (SI-peptides, crude peptide: approximately 50% peptide purity and >99% isotope purity, Greiner bio one (Frickenhausen, Germany)) were synthesized. SI-peptides had isotope-labeled Lysine or Arginine at their C-terminus. Each SI-peptide was dissolved in distilled 40% ACN and 0.1% TFA, and stored at −80 °C.

Creation of Spectral Library

Forty-nine kinds of SI-peptides were combined and diluted in the matrix solution composed of 100 fmol/mL BSA digest, 2% ACN and 0.1% TFA. For acquisition of MS/MS spectra of SI- peptides, from which the spectral library for creation of SRM transitions was generated by Pinpoint 1.0 software (Thermo Fisher Scientific, San Jose, CA), the SI-peptide mixture was analyzed by nanoLC−MS/MS on the LTQ-Orbitrap XL mass spectrometer. The spectral library of spectrum-peptide matches was generated by importing output files, called msf files, to Pinpoint 1.0.

Creation of the Preliminary SRM-Transition List

SRM/MRM-transition lists of each SI-peptide were created from the spectral library using Pinpoint 1.0. Fragment ions were selected from the library by picking the eight most intense y-ions other than the fragment ions which were less than 2 amino acids length.

Optimization of SRM/MRM Method and SRM/MRM Assays

SRM methods for each SI-peptide were created by Pinpoint 1.0, which included SRM transition lists and the instrument method with the following parameters; a scan width of 0.002 m/z, a Q1 resolution of 0.7 fwhm, a cycle time of 1 s, and a gas pressure of 1.8 mTorr. The SI-peptide mixture was analyzed by LC-SRM on the TSQ-Vantage triple quadruple mass spectrometer (Thermo Fisher Scientific, Bremen, Germany) equipped with the system mentioned above. The nanoLC gradient was delivered at 500 nL/min and optimized. A spray voltage of 1900−2000 V was applied. Test runs of the SI-peptide mixture were performed to establish the retention time window (±2 min) for each peptide ion and optimize the collision energy (CE) for each transition. Four transitions were chosen for each peptide and all fragment ions are y-ions. When possible, one or two peptides were used per protein and all SRM analyses were run in duplicate. Twenty-five micrograms of membrane proteins from patient tissue was prepared and digested using the PTS method as mentioned above for an SRM assay. Two μg of digested sample was transferred to a new tube and the SI-peptide mixture was added. The spiked volume of SI-peptides in samples was optimized to approximate endogenous ion peaks of each peptide using mixtures of individual membrane fractions. Samples were analyzed by LC-SRM on the TSQ-Vantage using the optimized SRM method.

SRM/MRM Data Analysis

SRM/MRM data were processed using Pinpoint 1.0. As an example, Supplementary Figure 1 (Supporting Information) shows a peak profile of MFAP4 (GFYYSLK). The relative quantification values of each target peptide were determined by calculating the average ratios of peak areas of SI-peptide transitions and endogenous transitions. When the transition profile was different between endogenous peptides and SI-peptides, the transition was excluded from the quantification process. The different transition profile was possibly caused detection of untargeted peptides. In addition, transitions having signal-to-noise ratios (S/N) of <10 were discarded from this study. In such case, only peptides having more than 1 transition were used.

Western Blot Analysis

Membrane proteins were separated by electrophoresis on 5−20% gradient gels (DRC, Tokyo, Japan). Proteins were

transferred to Immobilon-P Transfer membranes (0.45 μm) (Millipore, Bedford, MA) in a tank-transfer apparatus, and membranes were blocked with Immuno Block (DS Pharma Biomedical Co. Ltd.). Anti-GP2 (SIGMA-Aldrich, Japan) diluted 1:200 and anti-MFAP4 (Proteintech Group, Inc.) diluted 1:1000 in blocking buffer, which were used as primary antibodies. Goat antirabbit IgG horseradish peroxidase (Invitrogen Carlsbad, CA) diluted 1:3000 in blocking buffer was used as a secondary antibody. Antigens on membranes were detected with enhanced chemiluminescence detection reagents (GE Healthcare).

Immunohistochemistry

Formalin-fixed paraffin-embedded tissue sections were obtained from the Osaka Medical Center for Cancer & Cardiovascular Diseases. Sections were deparaffinized and rehydrated followed by antigen unmasking in 10 mM citrate buffer (pH 6.0) for 10 min at 15 psi and 120 °C followed by treatment with 0.3% H_2O_2 (Wako). After washing three times with PBS, nonspecific binding of antibodies was blocked with blocking buffer (3% BSA/PBS) for 1 h. Tissues were then incubated for 1 h with anti-GP2 diluted 1:750 in 3% BSA/PBS. After washing with PBS, DAB (DAKO Japan, Kyoto, Japan) was used to visualize tissue antigens according to the manufacturer's instructions. Tissue sections were counterstained with Mayer's Hematoxylin Solution (WAKO Pure Chemical Industries Ltd.) for 30 s and dehydrated with 100% ethanol and xylene (WAKO), and coverslips were mounted with Malinol (Muto Pure Chemicals, Tokyo, Japan). All slides were examined and scored by two of the authors who were blinded to clinical data of patients. Staining intensity was recorded on the following scale: 0, no staining was observed, or cytoplasm staining was observed in less than 10% of tumor cells; 1, faint/barely perceptible cytoplasm staining was detected in more than 10% of tumor cells (cells exhibited incomplete cytoplasm staining); 2, weak or moderate cytoplasm staining was observed in more than 10% of tumor cells or strong cytoplasm staining in less than 30%; and 3, strong cytoplasm staining was observed in more than 30% of tumor cells.

RESULTS

MS-based technologies have been applied to many large-scale proteomics studies and mostly applied to the discovery of protein biomarkers, especially in the field of cancer. These results have produced numerous candidate protein biomarkers. Unfortunately, only a few are currently being validated and applied in clinical prognostics. Consequently, we tried to perform proteomics analysis using a discovery-through-verification strategy. An overview of the method for verification of prognostic biomarkers and its application to the model of breast cancer prognostic biomarkers is shown in Figure 1. In the discovery phase, we analyzed pooled membrane protein fractions isolated from high (n = 9) or low-risk (n = 9) breast cancer tissues by LC–MS/MS using iTRAQ$_4$ to generate data of candidate protein biomarkers. We identified a total of 5122 unique proteins. A list of proteins is presented in Supplementary Table 2 (Supporting Information). The identified 5122 proteins were examined with respect to cellular localization using Gene Ontology annotation analysis in Ingenuity pathway analysis (IPA) or The Database for Annotation, Visualization and Integrated Discovery (DAVID) (Figure 2A). IPA or DAVID revealed that 2480 (48.4%) were annotated to membrane proteins, 826 (16.1%) proteins were plasma membrane, and 340 (6.6%) proteins were extracellular space proteins by GO analysis. Furthermore, of the identified proteins, 1469 (28.7%) proteins were predicted to have transmembrane

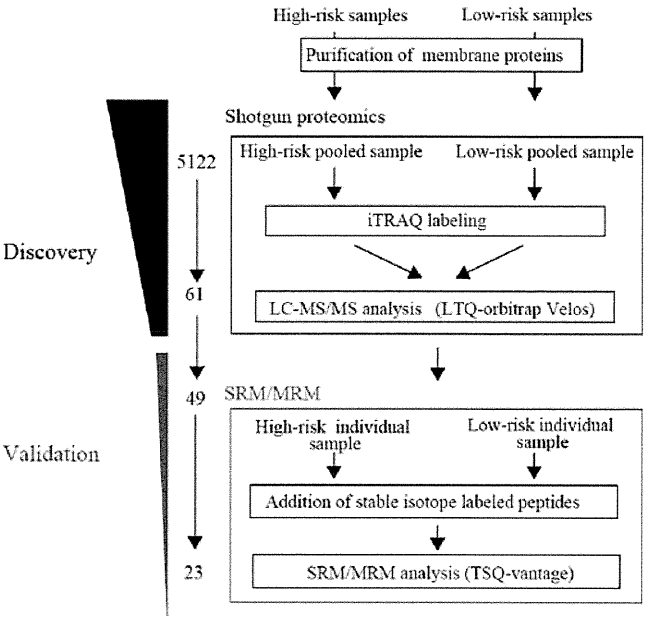


Figure 1. Overview of the workflow used to discovery and verify candidate biomarkers, showing the flux of candidates at each stage of the pipeline. For the discovery step, an iTRAQ$_4$ discovery study in which peptides, and thereby proteins that are differentially expressed among pools of high-risk patients (n = 9) and low-risk patients (n = 9), are identified. For the verification step, a targeted SRM/MRM study was set up based on the information found in the prior study to validate quantitative findings from the iTRAQ$_4$ discovery study.

(A)

Location	Numbers	%
total	5122	
Membrane	2480	48.4
Plasma membrane	829	16.2
Extracellular Space	340	6.6

(B)

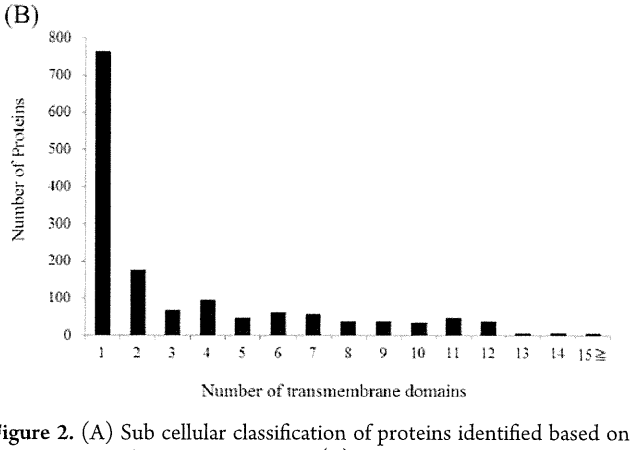


Figure 2. (A) Sub cellular classification of proteins identified based on annotation with gene ontology. (B) Numbers of proteins with transmembrane helices predicted by the TMHMM2.0 algorithm.

segments via TMHMM2.0 algorithm. A histogram of the generated data is presented in Figure 2B. These results support the effectiveness of the method to solubilize and digest integral membrane proteins containing transmembrane domains, allowing large-scale detection and identification of this protein class with no bias against membrane proteins. Fold changes were

dx.doi.org/10.1021/pr300322q | J. Proteome Res. 2012, 11, 4201–4210

Table 1. List of Proteins for which SRM/MRM Assays Were Developed[a,b]

high-risk > low-risk[c]

accession no	protein name	gene name	ratio (low-risk/high-risk)
P26022	pentraxin 3, long	PTX3	0.166
O75762	transient receptor potential cation channel, subfamily A, member 1	TRPA1	0.215
Q5H943	chromosome X open reading frame 61	KKLC1	0.264
Q9NYZ1	family with sequence similarity 18, member B1	FAM18B2	0.269
Q9H313	tweety homologue 1	TTYH1	0.322
Q9NQG1	mannosidase, beta A, lysosomal-like	MANBAL	0.327
P02458	Collegen, type II, alpha 1	COL2A1	0.354
P50443	solute carrier family 26 (Sulfate transporter), member 2	SLC26A2	0.36
P26012	integrin, beta 8	ITGB8	0.367
P15328	folate receptor 1 (adult)	FOLR1	0.381
O43490	prominin 1	PROM1	0.409
Q01650	solute carrier family 7 (cationic amino acid transporter, y+ system), member 5	SLC7A5	0.454
Q9HBA0	transient receptor potential cation channel subfamily V, member 4	TRPV4	0.462
P53985	solute carrier family 16, member 1 (monocarboxylic acid transporter 1)	SLC16A1	0.47
Q9NP58	ATP-binding cassette subfamily B member 6, mitochondrial	ABCB6	0.488

high-risk < low-risk[d]

accession no	protein name	gene name	ratio (low-risk/high-risk)
Q9UBX5	fibulin 5	FBLN5	2.015
O94911	ATP-binding cassette subfamily A (ABC1), member 8	ABCA8	2.018
P01011	serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 3	SERPINA3	2.026
P55058	phospholipid transfer protein	PLTP	2.042
Q92743	HtrA serine peptidase 1	HTRA1	2.043
P51884	lumican	LUM	2.047

high-risk < low-risk[d]

accession no	protein name	gene name	ratio (low-risk/high-risk)
Q9UHI5	solute carrier family 7 (amino acid transporter, L-type), member 8	SLC7A8	2.091
P23142	fibulin 1	FBLN1	2.119
Q05707	collagen, type XIV, alpha 1	COL14A1	2.12
Q9P0K1	ADAM metallopeptidase domain 22	ADAM22	2.148
Q14956	glycoprotein (transmembrane) nmb	GPNMB	2.154
Q99542	matrix metallopeptidase 19	MMP19	2.172
P38185	serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 6	SERPINA6	2.18
Q9BXN1	asporin	ASPN	2.222
O60938	keratocan	KERA	2.223
Q15661	tryptase alpha/beta 1	TPSAB1	2.23
O43927	chemokine (C-X-C motif) ligand 13	CXCL13	2.269
O99983	osteomodulin	OMD	2.405
O43692	peptidase inhibitor 15	PI15	2.42
O76061	stanniocalcin 2	STC2	2.51
P25189	myelin protein zero	MPZ	2.53
P20774	osteoglycin	OGN	2.543
Q53GD3	solute carrier family 44, member 4	SLC44A4	2.623
P27658	collagen, type VIII, alpha 1	COL8A1	2.93
P29120	proprotein convertase subtilsin/kexin type 1	PCSK1	2.951
P01009	serpin peptidase inliibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 1	SERPINA1	3.037
P02743	amyloid P component serum	APCS	3.324
P55083	microfibrillar-associated protein 4	MFAP4	3.392
P25311	alpha-2-glycoprotein 1, zinc-binding	AZGP1	3.773
P55259	glycoprotein 2(zymogen granule membrane)	GP2	3.864
P05060	chromogranin B (secretogranin 1)	CHGB	4.008
P15086	corboxypeptidase B1(tissue)	CPB1	4.306
P05090	apolipoprotein D	APOD	4.887
P12273	prolactin-induced protein	PIP	6.033

[a]The overall protein ratios identified by iTRAQ are indicated for each protein. [b]These values are based on total peptide information obtained for each protein. [c]Protein whose expression in the high-risk group is 2-fold or more than that in the low-risk group. [d]Protein whose expression in the low-risk group is 2-fold or more than that in the high-risk group.

determined based on the ratio of peak areas of iTRAQ reporter ions for the corresponding peptides from high and low-risk samples. Among 2480 membrane proteins, 188 showed >2-fold differences between high and low risk groups (either up-regulated or down-regulated), of which 61 proteins were annotated to plasma membrane and extracellular proteins.

The next step, to perform verification of 61 candidate protein biomarkers in individual patient tissue samples (high-risk; $n = 9$, low-risk; $n = 7$), we carried out relative quantification by SRM/MRM using SI-peptides as internal standards (Figure 1). In a typical SRM/MRM assay, a unique peptide is measured, and its concentration is assumed to be equal to the concentration of its parent protein. In this work, proteotypic peptides were chosen based on shotgun proteomic identification data. Peptides that had modifications, such as partially oxidized methionine, were avoided and when possible, two peptides were used per protein. For SRM/MRM, 49 of 61 differentially expressed proteins were selected by the above criteria. Finally, 86 peptides representing 49 proteins (21 plasma membrane and 28 extracellular proteins) remained in the list (Table 1). To obtain data for making transitions for each peptide in SRM/MRM, we analyzed each SI-peptide using LTQ-Orbitrap XL. Four transitions, y-ions, were chosen for each peptide. A total of 688 transitions were used for targeting 86 peptides of 49 proteins. The complete transition list can be found in Supplementary Table 3 (Supporting Information). Individual membrane fraction samples were digested and analyzed in duplicate. The SRM analysis successfully verified all 49 target proteins. Among them, 23 proteins were differentially expressed between high-risk and low-risk groups; 21 were increased and 2 were decreased in low risk group. Of these 15 of the target analytes showed a significant difference (ratio > 2, $p < 0.05$) between two groups (Figure 3, Supplementary Table 4A, Supporting Information) and 8 showed apparent, but not significant, difference between high and low risk breast cancer tissues (Supplementary Table 4B, Supplementary Figure 2, Supporting Information). On other hands, the 26 proteins showed no significant difference between high and low-risk breast cancer tissues (Supplementary Table 5, Supporting Information). Proteins with significant difference mentioned above included the following 10 proteins that have been reported with altered expression in breast cancer by other methods: SERPINA3, APOD, APCS, SERPINA1, LUM,

dx.doi.org/10.1021/pr300322q | J. Proteome Res. 2012, 11, 4201−4210

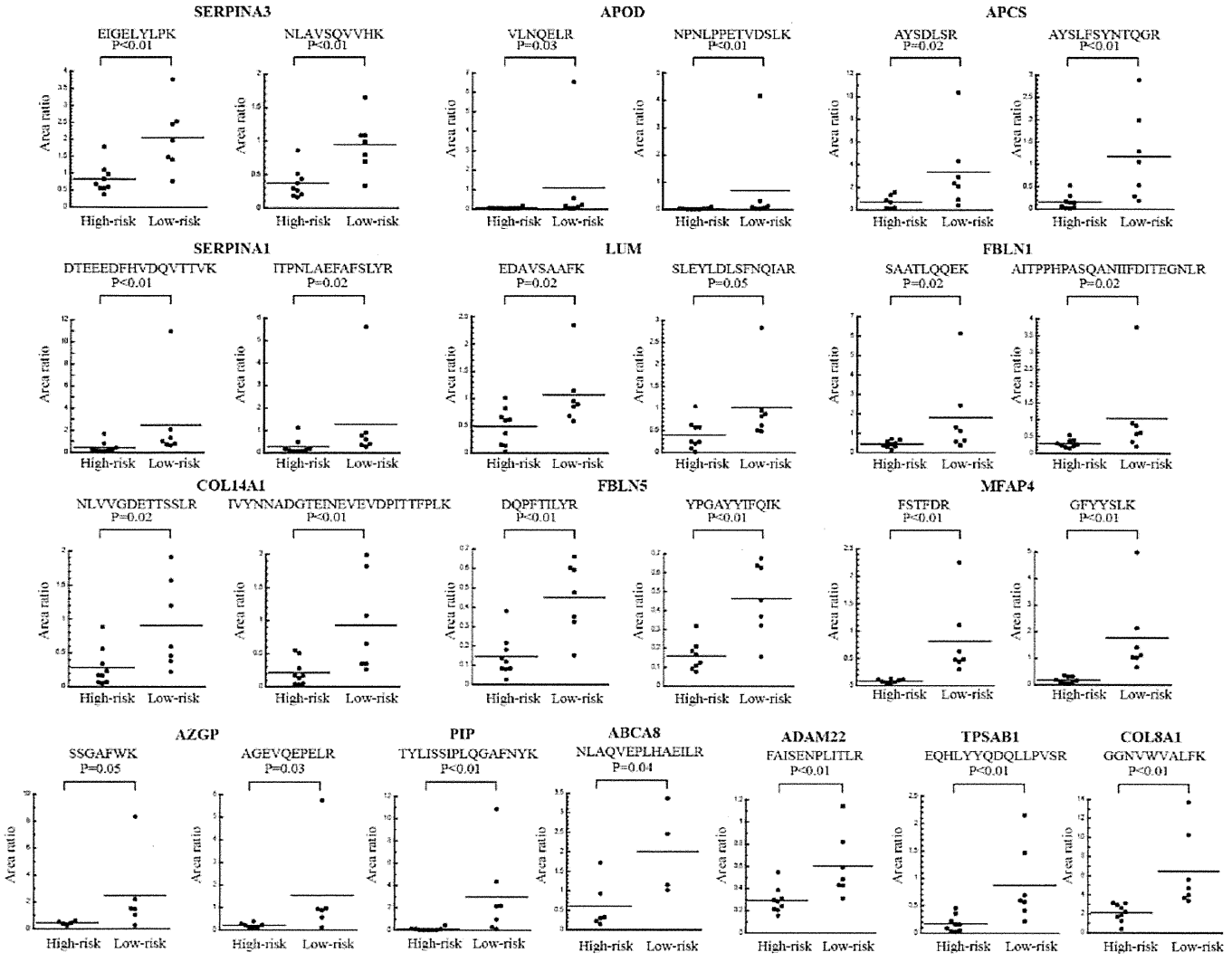


Figure 3. Quantitation of targeted peptide expression in the membrane fraction which showed a significant difference between high-risk and low-risk patients. Area ratio (endogenous peptide/SI-peptide) of each peptide in individual patient tissue samples (high-risk; $n = 9$, low-risk; $n = 7$) were subjected to nonparametric analysis of the Wilcoxon test with a cut off of $p < 0.05$.

FBLN1, FBLN5, AZGP1, PIP, and COL8A1.[26−35] However, the following 5 proteins have not been reported in breast cancer: COL14A1, MFAP4, ABCA8, ADAM22, and TPSAB1. These proteins may be novel prognostic biomarker candidates in breast cancer.

To confirm the differential expression observed by SRM/MRM analysis, the expression level of a selected protein was further examined using Western blotting. We selected candidate proteins on the basis of extent of significant differential expression and antibody availability. MFAP4 was thus selected and Western blotting was performed with membrane fractions of individual patient tissue samples. Western blot results are shown in Figure 4B. Results showed a positive correlation with SRM/MRM analysis results. Thus, SRM/MRM analysis is expected to be an effective method for verification from numerous candidate protein biomarkers.

Among eight proteins with apparent, but not significant, difference between high and low risk breast cancer tissues (Supplementary Table 4B, Supplementary Figure 2, Supporting Information), known markers of breast cancer such as STC2 and ITGB8 has been shown to be differently expressed.[36−38] However, to our knowledge, the other 6 proteins have not been reported in breast cancer: KERA, OMD, SERPINA6, PTX3, PI15, and GP2. Among them, we examined the expression level of the GP2 protein by Western blotting and IHC because a good antibody is available. The expression of GP2 was significantly higher in the low-risk group than that of the high-risk group by either Western blotting or IHC (Figure 5). The above results indicated that GP2 and MFAP4 are potential prognostic biomarkers for breast cancer.

■ DISCUSSION

The aim of this study was to establish a discovery-through-verification pipeline of large scale proteomics analysis using limited tissue samples. In this study, we identified and verified several biomarker candidates predicting recurrence risk of breast cancer by tissue membrane proteomic analysis using an established strategy. In total, 5122 proteins were identified with high confidence by a quantitative proteomics approach using iTRAQ labeling, 2480 protein (48.4%) of them were annotated as membrane proteins, 829 proteins (16.1%) were plasma membrane and 340 proteins (6.6%) were extracellular space proteins by GO analysis. Also 1469 proteins (28.7%) were predicted to have transmembrane domain by TMHMM algorithm.

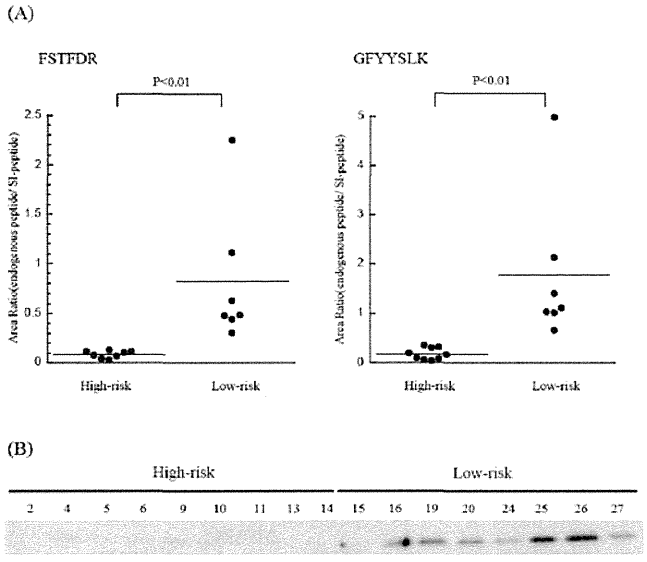


Figure 4. SRM/MRM and Western blot analysis data confirm elevations of MFAP4 in low-risk breast cancer patients. (A) Distributions of area ratio (endogenous peptide/SI-peptide) of MFAP4 in human breast cancer with high-risk and low-risk. The median value is plotted as a line. Results derived from SRM-MS based measurements. (B) Western blot analysis showing higher levels of MFAP4 in low-risk patients than those of high-risk patients. The numbers indicate the patient numbers in Supplementary Table 1, Supporting Information.

A total of 61 proteins were found to be altered by 2-fold or more between high and low-risk breast cancer tissues and 49 of these proteins were subsequently verified with targeted proteomics using SRM/MRM. Twenty-three proteins were shown to be differentially expressed between the two groups. Additionally, two of these proteins, MFAP4 and GP2, were further validated to be differentially expressed between the different groups using Western blotting and IHC.

We showed that the combination of iTRAQ shotgun and SRM/MRM proteomics is a powerful technique for identification and verification of numerous biomarker candidates. At the discovery stage of this study, human patient tissue samples were pooled within each group to perform effective screening of differentially expressed proteins across breast cancers of each group using limited amounts of specimen. Additionally, for verification from numerous candidate protein biomarkers, we performed SRM/MRM analysis using individual patient tissue samples. This combination approach was able to discover and verify a novel candidate biomarker with high-throughput. Thingholm and co-workers published a large-scale quantitative proteomic analysis using iTRAQ combined with SRM for discovery and validation of biomarkers for type 2 diabetes.[8] However, their targeted proteomics were performed without SI-peptides. The difficulty with SRM/MRM analysis may occur without internal standards, which provide reference signals for the verification of analyte specificity.[7] The use of SI-peptide provides the most favorable SRM/MRM information for each peptide, such as highest intensity fragment ions and peptide elution time. The target peptide area ratio is then determined by measuring the observed area value for the target peptide relative to that of the SI-peptide. Therefore, we used crude SI-peptides that can be synthesized at a lower cost than purified synthesized peptides. As a consequence, our methods to rapidly

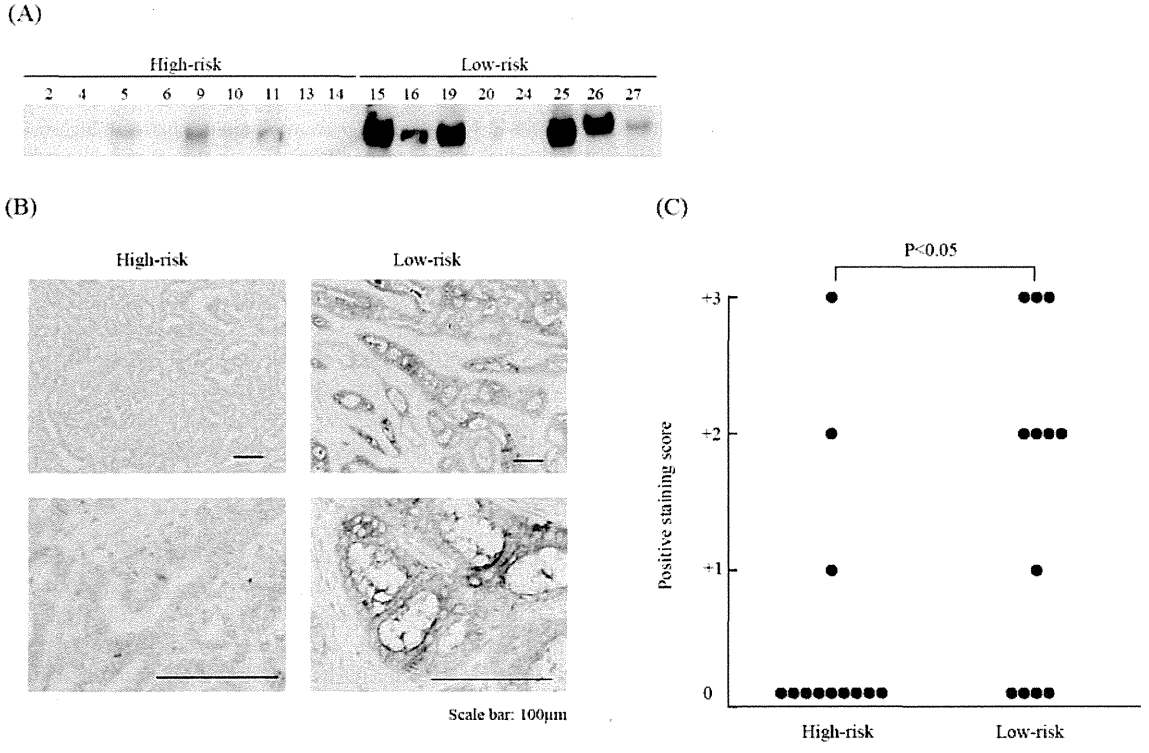


Figure 5. Western blot and IHC analysis data confirm elevations of GP2 in low-risk breast cancer patients. (A) Western blot analysis showing higher levels of GP2 in low-risk patients than those in high-risk patients. The numbers indicate the patient numbers in Supplementary Table 1 (Supporting Information). (B) Immunohistochemistry staining of GP2 on high-risk and low-risk breast cancer specimens. Top is ×200 and bottom is ×400. (C) Staining intensity of 24 samples (12 high-risk groups and 12 low-risk groups) was examined and scored according to Materials and Methods. Statistical analysis was performed using the Wilcoxon–Mann–Whitney test with a cut off of $p < 0.05$ in (B).

generate a SRM/MRM assay using crude synthetic peptides expands the application of SRM based targeted MS for high-throughput protein detection and quantification.[39]

Several previously published reports have described membrane proteome analysis.[13,40,41] In a recent study, Polisetty and co-workers carried out a large-scale discovery study in which they utilized a technology combining iTRAQ and LC−MS/MS and identified 1834 distinct proteins from membrane fractions of glioblastoma multiforme patient specimens, 56% of them (1027) are annotated as membrane protein.[12] In this study, we identified a total of 5122 proteins in the membrane fraction, 48% of them (2480) are known membrane proteins associated with major cellular processes and this number of membrane proteins was much greater than those previously reported. This was because we performed a method that utilized a combination of PTS method-based isolation of membrane proteins and iTRAQ method. In the membrane preparation stage, efficient isolation of membrane proteins was achieved using ultra centrifugation and subsequent PTS. In the cleavage procedure of membrane proteins, PTS method allowed the use of a high detergent concentration to achieve efficient solubilization of very hydrophobic membrane proteins while avoiding interference with subsequent iTRAQ-LC−MS/MS analysis.[42] In this SRM/MRM assay, we found that a number of proteins were detected in the membrane fraction, but not in the total fraction (data not shown). Thus, this method may provide deeper proteome coverage for identification of tissue membrane proteins.

In our study, eight proteins were not detected in several patient samples with this SRM/MRM analysis (Supplementary Table 4B, Supplementary Figure 2, Supporting Information). Consequently, they could not be subjected to the Wilcoxon test. However, GP2 showed a significant difference between high-risk and low-risk tissues in breast cancer using Western blotting and IHC (Figure 5). In SRM/MRM, there are several possible reasons why not all targeted proteins were detected. They include 1) the absence of targeted protein in samples, that is no expression in the patient tissues used, (2) the targeted protein was present but its expression level was below the limit of quantitation, and (3) the MS peak of the target protein was hard to detect due to high background noise. In this study, expression of undetected proteins may be below the limit of quantitation. Thus, although the expression of GP2 protein could not be verified by SRM/MRM, Western blotting and IHC results showed a significant difference between high-risk and low-risk tissues in breast cancer.

A number of commercialized multigene prognostic and predictive tests have entered the complex and expanding landscape of breast cancer prognostics. The two assays that have achieved the most practical success are oncotype DX and MammaPrint. However, it has a high cost. In contrast, commercialized IHC multigene predictors are less expensive.[43] Cheang et al. reported that the immunohistochemical determination of ER, PgR, HER2, and Ki67 indexes are able to distinguish the luminal B subgroup with poor prognosis from the luminal A subgroup with good prognosis.[23] This assay was recommended at St-Gallen in 2011.[44] To our knowledge, most prognostic and predictive biomarkers show stronger staining in the high-risk group than that of the low-risk group. However, we found GP2 and MFAP4 were highly stained in the low-risk group, which could define this group. Consequently, combining this biomarker with existing prognostic tools (ER/PgR, Ki67, HER2) should be able to predict prognosis of breast cancer more accurately.

MFAP4 and GP2 show a significant differential expression between high and low-risk groups in breast cancer. MFAP4 was initially identified as a gene commonly deleted in the contiguous gene syndrome Smith-Magenis.[45] MFAP4 was involved in calcium-dependent cell adhesion or interactions with Integrin, Pulmonary surfactant protein A, or Lung surfactant protein D.[46−48] To our knowledge, there has been no report to show the expression and function of MFAP4 on cancer. GP2 is the major membrane protein present in pancreatic zymogen granules, and is cleaved and released into the pancreatic duct along with exocrine secretions.[49−51] Recently, Hase et al. reported that GP2, specifically expressed on the apical plasma membrane of M cells among enterocytes, serves as a transcytotic receptor for mucosal antigen.[52] However, there has been no report about GP2 protein expression on cancer tissues until now. Further investigation will be necessary to clarify the function of MFAP4 and GP2 on cancer.

In conclusion, although numerous studies have identified hundreds of biomarker candidate proteins for various diseases, few of them have performed large scale verification due to the unavailability of antibodies with high quality. This study showed that the iTRAQ and SRM based discovery-through-verification strategy contributes to generate clinically useful biomarkers for various diseases, including cancer. It may be possible to replace current clinical examinations by this latest proteomic technology in the near future.

■ ASSOCIATED CONTENT

Supporting Information

Supplementary Figure 1. A peak profile for SRM transitions of MFAP4: GFYYSLK. Supplementary Figure 2. Quantitation of eight proteins expression which showed apparent, but not significant, difference between high and low risk breast cancer tissues. Supplementary Table 1. Clinical features of breast cancer patients. Supplementary Table 2. A list of identified and quantified proteins by iTRAQ proteomics. Supplementary Table 3. Sequences and SRM/MRM transitions of target peptide for 49 proteins. Supplementary Table 4. Summary of SRM/MRM data; (A) proteins with significant different expression between high and low risk breast cancer tissues; (B) proteins with apparent, but not significant, difference between high and low risk breast cancer tissues. Supplementary Table 5. Summary of SRM/MRM data; proteins with no significant difference between high and low risk breast cancer tissues. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Laboratory of Proteome Research, National Institute of Biomedical Innovation 7-6-8 Saito-Asagi, Ibaraki City, Osaka 567-0085, Japan. Tel.: +81-72-641-9862. Fax: +81-72-641-9861. E-mail: tomonaga@nibio.go.jp.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by Grants-in-Aid, Research on Biological Markers for New Drug Development H20-0005 to K.Y. from the Ministry of Health, Labour, and Welfare of Japan. This work was supported by Grants-in-Aid 21390354 to T.T. and 22800095 to S.M. from the Ministry of Education, Science,

Sports, and Culture of Japan. We thank Dr. Yasuko Nishizawa for advice with immunohistochemistry.

ABBREVIATIONS

iTRAQ, isobaric peptide tags for relative and absolute quantification; SRM, selected reaction monitoring; MRM, multiple reaction monitoring; PTS, phase-transfer surfactants; SI-peptide, stable isotope-labeled peptide; CID, collision-induced dissociation; HCD, higher energy collision-induced dissociation; IHC, Immunohistochemistry; LC–MS/MS, liquid chromatography tandem mass spectrometry; CE, Collision energy; LTQ, linear ion trap; fwhm, full wide at half maximum; FDR, false discovery rate.

REFERENCES

(1) Zeng, G. Q.; Zhang, P. F.; Deng, X. M.; Yu, F. L.; Li, C.; Xu, Y.; Yi, H.; Li, M. Y.; Hu, R.; Zuo, J. H.; Li, X. H.; Wan, X. X.; Qu, J. Q.; He, Q. Y.; Li, J. H.; Ye, X.; Chen, Y.; Li, J. Y.; Xiao, Z. Q. Identification of candidate biomarkers for early detection of human lung squamous cell cancer by quantitative proteomics. *Mol. Cell. Proteomics* **2012**, *11* (6), M111013946.

(2) Ghosh, D.; Yu, H.; Tan, X. F.; Lim, T. K.; Zubaidah, R. M.; Tan, H. T.; Chung, M. C.; Lin, Q. Identification of key players for colorectal cancer metastasis by iTRAQ quantitative proteomics profiling of isogenic SW480 and SW620 cell lines. *J. Proteome Res.* **2011**, *10* (10), 4373–87.

(3) Li, Z.; Adams, R. M.; Chourey, K.; Hurst, G. B.; Hettich, R. L.; Pan, C. Systematic comparison of label-free, metabolic labeling, and isobaric chemical labeling for quantitative proteomics on LTQ Orbitrap velos. *J. Proteome Res.* **2012**, *11* (3), 1582–90.

(4) Gygi, S. P.; Rist, B.; Gerber, S. A.; Turecek, F.; Gelb, M. H.; Aebersold, R. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat. Biotechnol.* **1999**, *17* (10), 994–9.

(5) Ross, P. L.; Huang, Y. N.; Marchese, J. N.; Williamson, B.; Parker, K.; Hattan, S.; Khainovski, N.; Pillai, S.; Dey, S.; Daniels, S.; Purkayastha, S.; Juhasz, P.; Martin, S.; Bartlet-Jones, M.; He, F.; Jacobson, A.; Pappin, D. J. Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. *Mol. Cell. Proteomics* **2004**, *3* (12), 1154–69.

(6) Collier, T. S.; Sarkar, P.; Franck, W. L.; Rao, B. M.; Dean, R. A.; Muddiman, D. C. Direct comparison of stable isotope labeling by amino acids in cell culture and spectral counting for quantitative proteomics. *Anal. Chem.* **2010**, *82* (20), 8696–702.

(7) Whiteaker, J. R.; Lin, C.; Kennedy, J.; Hou, L.; Trute, M.; Sokal, I.; Yan, P.; Schoenherr, R. M.; Zhao, L.; Voytovich, U. J.; Kelly-Spratt, K. S.; Krasnoselsky, A.; Gafken, P. R.; Hogan, J. M.; Jones, L. A.; Wang, P.; Amon, L.; Chodosh, L. A.; Nelson, P. S.; McIntosh, M. W.; Kemp, C. J.; Paulovich, A. G. A targeted proteomics-based pipeline for verification of biomarkers in plasma. *Nat. Biotechnol.* **2011**, *29* (7), 625–34.

(8) Thingholm, T. E.; Bak, S.; Beck-Nielsen, H.; Jensen, O. N.; Gaster, M. Characterization of human myotubes from type 2 diabetic and nondiabetic subjects using complementary quantitative mass spectrometric methods. *Mol. Cell. Proteomics* **2011**, *10* (9), M110 006650.

(9) Addona, T. A.; Abbatiello, S. E.; Schilling, B.; Skates, S. J.; Mani, D. R.; Bunk, D. M.; Spiegelman, C. H.; Zimmerman, L. J.; Ham, A. J.; Keshishian, H.; Hall, S. C.; Allen, S.; Blackman, R. K.; Borchers, C. H.; Buck, C.; Cardasis, H. L.; Cusack, M. P.; Dodder, N. G.; Gibson, B. W.; Held, J. M.; Hiltke, T.; Jackson, A.; Johansen, E. B.; Kinsinger, C. R.; Li, J.; Mesri, M.; Neubert, T. A.; Niles, R. K.; Pulsipher, T. C.; Ransohoff, D.; Rodriguez, H.; Rudnick, P. A.; Smith, D.; Tabb, D. L.; Tegeler, T. J.; Variyath, A. M.; Vega-Montoto, L. J.; Wahlander, A.; Waldemarson, S.; Wang, M.; Whiteaker, J. R.; Zhao, L.; Anderson, N. L.; Fisher, S. J.; Liebler, D. C.; Paulovich, A. G.; Regnier, F. E.; Tempst, P.; Carr, S. A. Multi-site assessment of the precision and reproducibility of multiple reaction monitoring-based measurements of proteins in plasma. *Nat. Biotechnol.* **2009**, *27* (7), 633–41.

(10) Drabovich, A. P.; Jarvi, K.; Diamandis, E. P. Verification of male infertility biomarkers in seminal plasma by multiplex selected reaction monitoring assay. *Mol. Cell. Proteomics* **2011**, *10* (12), M110 004127.

(11) Yates, J. R., 3rd; Gilchrist, A.; Howell, K. E.; Bergeron, J. J. Proteomics of organelles and large cellular structures. *Nat. Rev. Mol. Cell. Biol.* **2005**, *6* (9), 702–14.

(12) Polisetty, R. V.; Gautam, P.; Sharma, R.; Harsha, H. C.; Nair, S. C.; Gupta, M. K.; Uppin, M. S.; Challa, S.; Puligopu, A. K.; Ankathi, P.; Purohit, A. K.; Chandak, G. R.; Pandey, A.; Sirdeshmukh, R. LC-MS/MS analysis of differentially expressed glioblastoma membrane proteome reveals altered calcium signalling and other protein groups of regulatory functions. *Mol. Cell. Proteomics* **2012**, *11* (6), M111013565.

(13) Josic, D.; Clifton, J. G. Mammalian plasma membrane proteomics. *Proteomics* **2007**, *7* (16), 3010–29.

(14) Russell, W. K.; Park, Z. Y.; Russell, D. H. Proteolysis in mixed organic-aqueous solvent systems: applications for peptide mass mapping using mass spectrometry. *Anal. Chem.* **2001**, *73* (11), 2682–5.

(15) Blonder, J.; Goshe, M. B.; Moore, R. J.; Pasa-Tolic, L.; Masselon, C. D.; Lipton, M. S.; Smith, R. D. Enrichment of integral membrane proteins for proteomic analysis using liquid chromatography-tandem mass spectrometry. *J. Proteome Res.* **2002**, *1* (4), 351–60.

(16) Chick, J. M.; Haynes, P. A.; Molloy, M. P.; Bjellqvist, B.; Baker, M. S.; Len, A. C. Characterization of the rat liver membrane proteome using peptide immobilized pH gradient isoelectric focusing. *J. Proteome Res.* **2008**, *7* (3), 1036–45.

(17) Wu, C. C.; MacCoss, M. J.; Howell, K. E.; Yates, J. R., 3rd A method for the comprehensive proteomic analysis of membrane proteins. *Nat. Biotechnol.* **2003**, *21* (5), 532–8.

(18) Schindler, J.; Lewandrowski, U.; Sickmann, A.; Friauf, E.; Nothwang, H. G. Proteomic analysis of brain plasma membranes isolated by affinity two-phase partitioning. *Mol. Cell. Proteomics* **2006**, *5* (2), 390–400.

(19) Rahbar, A. M.; Fenselau, C. Integration of Jacobson's pellicle method into proteomic strategies for plasma membrane proteins. *J. Proteome Res.* **2004**, *3* (6), 1267–77.

(20) Masuda, T.; Tomita, M.; Ishihama, Y. Phase transfer surfactant-aided trypsin digestion for membrane proteome analysis. *J. Proteome Res.* **2008**, *7* (2), 731–40.

(21) van 't Veer, L. J.; Dai, H.; van de Vijver, M. J.; He, Y. D.; Hart, A. A.; Mao, M.; Peterse, H. L.; van der Kooy, K.; Marton, M. J.; Witteveen, A. T.; Schreiber, G. J.; Kerkhoven, R. M.; Roberts, C.; Linsley, P. S.; Bernards, R.; Friend, S. H. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* **2002**, *415* (6871), 530–6.

(22) Paik, S.; Shak, S.; Tang, G.; Kim, C.; Baker, J.; Cronin, M.; Baehner, F. L.; Walker, M. G.; Watson, D.; Park, T.; Hiller, W.; Fisher, E. R.; Wickerham, D. L.; Bryant, J.; Wolmark, N. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N. Engl. J. Med.* **2004**, *351* (27), 2817–26.

(23) Cheang, M. C.; Chia, S. K.; Voduc, D.; Gao, D.; Leung, S.; Snider, J.; Watson, M.; Davies, S.; Bernard, P. S.; Parker, J. S.; Perou, C. M.; Ellis, M. J.; Nielsen, T. O. Ki67 index, HER2 status, and prognosis of patients with luminal B breast cancer. *J. Natl. Cancer Inst.* **2009**, *101* (10), 736–50.

(24) Rappsilber, J.; Mann, M.; Ishihama, Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat. Protoc.* **2007**, *2* (8), 1896–906.

(25) Krogh, A.; Larsson, B.; von Heijne, G.; Sonnhammer, E. L. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* **2001**, *305* (3), 567–80.

(26) Ishitobi, M.; Goranova, T. E.; Komoike, Y.; Motomura, K.; Koyama, H.; Glas, A. M.; van Lienen, E.; Inaji, H.; Van't Veer, L. J.;

Kato, K. Clinical utility of the 70-gene MammaPrint profile in a Japanese population. *Jpn. J. Clin. Oncol.* **2010**, *40* (6), 508–12.
(27) Soiland, H.; Janssen, E. A.; Korner, H.; Varhaug, J. E.; Skaland, I.; Gudlaugsson, E.; Baak, J. P.; Soreide, J. A. Apolipoprotein D predicts adverse outcome in women >or=70 years with operable breast cancer. *Breast Cancer Res. Treat.* **2009**, *113* (3), 519–28.
(28) Zeng, Z.; Hincapie, M.; Pitteri, S. J.; Hanash, S.; Schalkwijk, J.; Hogan, J. M.; Wang, H.; Hancock, W. S. A proteomics platform combining depletion, multi-lectin affinity chromatography (M-LAC), and isoelectric focusing to study the breast cancer proteome. *Anal. Chem.* **2011**, *83* (12), 4845–54.
(29) Hamrita, B.; Chahed, K.; Trimeche, M.; Guillier, C. L.; Hammann, P.; Chaieb, A.; Korbi, S.; Chouchane, L. Proteomics-based identification of alpha1-antitrypsin and haptoglobin precursors as novel serum markers in infiltrating ductal breast carcinomas. *Clin. Chim. Acta* **2009**, *404* (2), 111–8.
(30) Shipp, C.; Watson, K.; Jones, G. L. Associations of HSP90 client proteins in human breast cancer. *Anticancer Res.* **2011**, *31* (6), 2095–101.
(31) Pupa, S. M.; Argraves, W. S.; Forti, S.; Casalini, P.; Berno, V.; Agresti, R.; Aiello, P.; Invernizzi, A.; Baldassari, P.; Twal, W. O.; Mortarini, R.; Anichini, A.; Menard, S. Immunological and pathobiological roles of fibulin-1 in breast cancer. *Oncogene* **2004**, *23* (12), 2153–60.
(32) Lee, Y. H.; Albig, A. R.; Regner, M.; Schiemann, B. J.; Schiemann, W. P. Fibulin-5 initiates epithelial-mesenchymal transition (EMT) and enhances EMT induced by TGF-beta in mammary epithelial cells via a MMP-dependent mechanism. *Carcinogenesis* **2008**, *29* (12), 2243–51.
(33) Freije, J. P.; Fueyo, A.; Uria, J.; Lopez-Otin, C. Human Zn-alpha 2-glycoprotein cDNA cloning and expression analysis in benign and malignant breast tissues. *FEBS Lett.* **1991**, *290* (1–2), 247–9.
(34) Baniwal, S. K.; Little, G. H.; Chimge, N. O.; Frenkel, B. Runx2 controls a feed-forward loop between androgen and prolactin-induced protein (PIP) in stimulating T47D cell proliferation. *J. Cell Physiol.* **2012**, *227* (5), 2276–82.
(35) Ma, X. J.; Dahiya, S.; Richardson, E.; Erlander, M.; Sgroi, D. C. Gene expression profiling of the tumor microenvironment during breast cancer progression. *Breast Cancer Res.* **2009**, *11* (1), R7.
(36) Joensuu, K.; Heikkila, P.; Andersson, L. C. Tumor dormancy: elevated expression of stanniocalcins in late relapsing breast cancer. *Cancer Lett.* **2008**, *265* (1), 76–83.
(37) Raulic, S.; Ramos-Valdes, Y.; DiMattia, G. E. Stanniocalcin 2 expression is regulated by hormone signalling and negatively affects breast cancer cell viability in vitro. *J. Endocrinol.* **2008**, *197* (3), 517–29.
(38) Culhane, A. C.; Quackenbush, J. Confounding effects in "A six-gene signature predicting breast cancer lung metastasis". *Cancer Res.* **2009**, *69* (18), 7480–5.
(39) Picotti, P.; Rinner, O.; Stallmach, R.; Dautel, F.; Farrah, T.; Domon, B.; Wenschuh, H.; Aebersold, R. High-throughput generation of selected reaction-monitoring assays for proteins and proteomes. *Nat. Methods* **2010**, *7* (1), 43–6.
(40) Chen, J. S.; Chen, K. T.; Fan, C. W.; Han, C. L.; Chen, Y. J.; Yu, J. S.; Chang, Y. S.; Chien, C. W.; Wu, C. P.; Hung, R. P.; Chan, E. C. Comparison of membrane fraction proteomic profiles of normal and cancerous human colorectal tissues with gel-assisted digestion and iTRAQ labeling mass spectrometry. *FEBS J.* **2010**, *277* (14), 3028–38.
(41) Han, C. L.; Chen, J. S.; Chan, E. C.; Wu, C. P.; Yu, K. H.; Chen, K. T.; Tsou, C. C.; Tsai, C. F.; Chien, C. W.; Kuo, Y. B.; Lin, P. Y.; Yu, J. S.; Hsueh, C.; Chen, M. C.; Chan, C. C.; Chang, Y. S.; Chen, Y. J. An informatics-assisted label-free approach for personalized tissue membrane proteomics: case study on colorectal cancer. *Mol. Cell. Proteomics* **2011**, *10* (4), M110 003087.
(42) Iwasaki, M.; Masuda, T.; Tomita, M.; Ishihama, Y. Chemical cleavage-assisted tryptic digestion for membrane proteome analysis. *J. Proteome Res.* **2009**, *8* (6), 3169–75.
(43) Ross, J. S.; Hatzis, C.; Symmans, W. F.; Pusztai, L.; Hortobagyi, G. N. Commercialized multigene predictors of clinical outcome for breast cancer. *Oncologist* **2008**, *13* (5), 477–93.
(44) Goldhirsch, A; Wood, W. C.; Coates, A. S.; Gelber, R. D.; Thurlimann, B; Senn, H. J. Strategies for subtypes--dealing with the diversity of breast cancer: highlights of the St. Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2011. *Ann. Oncol.* **2011**, *22* (8), 1736–47.
(45) Zhao, Z.; Lee, C. C.; Jiralerspong, S.; Juyal, R. C.; Lu, F.; Baldini, A.; Greenberg, F.; Caskey, C. T.; Patel, P. I. The gene for a human microfibril-associated glycoprotein is commonly deleted in Smith-Magenis syndrome patients. *Hum. Mol. Genet.* **1995**, *4* (4), 589–97.
(46) Schlosser, A.; Thomsen, T.; Shipley, J. M.; Hein, P. W.; Brasch, F.; Tornoe, I.; Nielsen, O.; Skjodt, K.; Palaniyar, N.; Steinhilber, W.; McCormack, F. X.; Holmskov, U. Microfibril-associated protein 4 binds to surfactant protein A (SP-A) and colocalizes with SP-A in the extracellular matrix of the lung. *Scand. J. Immunol.* **2006**, *64* (2), 104–16.
(47) Lausen, M.; Lynch, N.; Schlosser, A.; Tornoe, I.; Saekmose, S. G.; Teisner, B.; Willis, A. C.; Crouch, E.; Schwaeble, W.; Holmskov, U. Microfibril-associated protein 4 is present in lung washings and binds to the collagen region of lung surfactant protein D. *J. Biol. Chem.* **1999**, *274* (45), 32234–40.
(48) Toyoshima, T.; Ishida, T.; Nishi, N.; Kobayashi, R.; Nakamura, T.; Itano, T. Differential gene expression of 36-kDa microfibril-associated glycoprotein (MAGP-36/MFAP4) in rat organs. *Cell Tissue Res.* **2008**, *332* (2), 271–8.
(49) MacDonald, R. J.; Ronzio, R. A. Comparative analysis of zymogen granule membrane polypeptides. *Biochem. Biophys. Res. Commun.* **1972**, *49* (2), 377–82.
(50) Ronzio, R. A.; Kronquist, K. E.; Lewis, D. S.; MacDonald, R. J.; Mohrlok, S. H.; O'Donnell, J. J., Jr. Glycoprotein synthesis in the adult rat pancreas. IV. Subcellular distribution of membrane glycoproteins. *Biochim. Biophys. Acta* **1978**, *508* (1), 65–84.
(51) Scheele, G. A.; Fukuoka, S.; Freedman, S. D. Role of the GP2/THP family of GPI-anchored proteins in membrane trafficking during regulated exocrine secretion. *Pancreas* **1994**, *9* (2), 139–49.
(52) Hase, K.; Kawano, K.; Nochi, T.; Pontes, G. S.; Fukuda, S.; Ebisawa, M.; Kadokura, K.; Tobe, T.; Fujimura, Y.; Kawano, S.; Yabashi, A.; Waguri, S.; Nakato, G.; Kimura, S.; Murakami, T.; Iimura, M.; Hamura, K.; Fukuoka, S.; Lowe, A. W.; Itoh, K.; Kiyono, H.; Ohno, H. Uptake through glycoprotein 2 of FimH(+) bacteria by M cells initiates mucosal immune response. *Nature* **2009**, *462* (7270), 226–30.

IJC
International Journal of Cancer

$CD34^+/CD38^-$ acute myelogenous leukemia cells aberrantly express CD82 which regulates adhesion and survival of leukemia stem cells

Chie Nishioka[1,2], Takayuki Ikezoe[3], Mutsuo Furihata[4], Jing Yang[3], Satoshi Serada[5], Tetsuji Naka[5], Atsuya Nobumoto[6], Sayo Kataoka[7], Masayuki Tsuda[6], Keiko Udaka[1] and Akihito Yokoyama[3]

[1] Department of Immunology, Kochi Medical School, Kochi University, Nankoku, Kochi, Japan
[2] Japanese Society for the Promotion of Science (JSPS), Chiyoda-ku, Tokyo, Japan
[3] Hematology and Respiratory Medicine, Kochi Medical School, Kochi University, Nankoku, Kochi, Japan
[4] Tumor Pathology, Kochi Medical School, Kochi University, Nankoku, Kochi, Japan
[5] Laboratory for immune Signal, National Institute of Biomedical Innovation, Ibaraki, Osaka
[6] The Facility for Animal Research, Kochi Medical School, Kochi University, Nankoku, Kochi, Japan
[7] Medical Research Center, Kochi Medical School, Kochi University, Nankoku, Kochi, Japan

To identify molecular targets in leukemia stem cells (LSCs), this study compared the protein expression profile of freshly isolated $CD34^+/CD38^-$ cells with that of $CD34^+/CD38^+$ counterparts from individuals with acute myelogenous leukemia (n = 2, AML) using isobaric tags for relative and absolute quantitation (iTRAQ). A total of 98 proteins were overexpressed, while six proteins were underexpressed in $CD34^+/CD38^-$ AML cells compared with their $CD34^+/CD38^+$ counterparts. Proteins overexpressed in $CD34^+/CD38^-$ AML cells included a number of proteins involved in DNA repair, cell cycle arrest, gland differentiation, antiapoptosis, adhesion, and drug resistance. Aberrant expression of CD82, a family of adhesion molecules, in $CD34^+/CD38^-$ AML cells was noted in additional clinical samples (n = 12) by flow cytometry. Importantly, down-regulation of CD82 in $CD34^+/CD38^-$ AML cells by a short hairpin RNA (shRNA) inhibited adhesion to fibronectin *via* up-regulation of matrix metalloproteinases 9 (MMP9) and colony forming ability of these cells as assessed by transwell assay, real-time RT-PCR, and colony forming assay, respectively. Moreover, we found that down-regulation of CD82 in $CD34^+/CD38^-$ AML cells by an shRNA significantly impaired engraftment of these cells in severely immunocompromised mice. Taken together, aberrant expression of CD82 might play a role in adhesion of LSCs to bone marrow microenvironment and survival of LSCs. CD82 could be an attractive molecular target to eradicate LSCs.

Key words: AML, leukemia stem cells, bone marrow microenvironment, CD82, MMP9
Additional Supporting Information may be found in the online version of this article.
*Takayuki Ikezoe contributed to the concept and design, interpreted and analyzed the data and wrote an article. Chie Nishioka performed all experiments and wrote an article. Mutsuo Furihata, Jing Yang, Satoshi Serada, and Tetsuji Naka, Sayo Kataoka and Atsuya Nobumoto performed the experiments. Akihito Yokoyama and Keiko Udaka provided important intellectual content
Grant sponsors: The Kochi University President's Discretionary Grant, Setsuro Fujii Memorial, The Osaka Foundation for Promotion of Fundamental Medical Research, Certificate of Kochi Shin-kin/Anshin-tomo-no-kai Prize and Japan Society for the Promotion of Science
DOI: 10.1002/ijc.27904
History: Received 15 Mar 2012; Accepted 25 Sep 2012; Online 11 Oct 2012
Correspondence to: Takayuki Ikezoe, Department of Hematology and Respiratory Medicine, Kochi University, Oko-cho, Nankoku, Kochi 783-8505, Japan, Tel.: 81-88-880-2345, Fax: 81-88-880-2348, E-mail: ikezoet@kochi-u.ac.jp

Acute myelogenous leukemia (AML) is characterized by a cellular hierarchy, and is initiated and maintained by a subset of self-renewing leukemia stem cells (LSCs).[1] To produce cure in individuals with AML, development of a novel treatment strategy targeting LSCs is urgently required. LSCs share some antigenic features with normal hematopoietic stem cells (HSCs). For example, both LSCs and HSCs express CD34 but not CD38. However, LSCs can be phenotypically distinguished from HSCs by several disparate markers, including $CD117^-$ and $CD123^+$.[1–3] LSCs exist in a quiescent state and are capable of self-renewal and differentiation, and are able to perpetuate leukemic cell growth in long-term culture assays and in the murine nonobese diabetic/severe combined immunodeficiency (NOD/SCID) model system.[1–4] $CD34^+/CD38^-$ AML cells were shown to fulfill the criteria for LSCs *in vivo*.[5,6] Although, recent studies employed more severely immunocompromised mice found that even $CD34^-$ or $CD38^+$ AML cells in some cases were able to reconstitute AML.[7,8]

The regulation of stem cell self-renewal and differentiation requires a specific microenvironment of surrounding cells known as the stem cell niche. The concept of the stem cell niche was first proposed for the human hematopoietic system in the 1970s.[9] The HSC niche in mouse bone marrow (BM) is

What's new?

Acute myelogenous leukemia (AML) is maintained by a subset of self-renewing leukemia stem cells (LSCs). Thus, to effectively treat AML, treatments targeting LSCs are needed. AML cells expressing CD34 but not CD38 (CD34+/CD38-) contain abundant LSCs and were found in this study to express a greater amount of CD82 than CD34+/CD38+ AML cells. CD82 was further found to regulate the survival of CD34+/CD38- AML cells and their adhesion to the bone marrow microenvironment, suggesting that this glycoprotein could be an attractive target for LSC eradication.

composed of an endosteal lining of stromal cells, extracellular matrix proteins, and osteoblasts.[10–12] Specific adherens junction molecules such as N-cadherin mediate adhesion between HSCs and niche cells in the adult hematopoietic system.[11]

Recent work has shown that interaction between CXCR4 on leukemic cells and its ligand stromal cell-derived factor-1 (SDF-1) in the niche is necessary for proper homing and *in vivo* growth of leukemic cells.[13] Moreover, interaction between LSCs and the niche mediated by adhesion molecule CD44 is required for maintenance of LSCs behavior.[14] CD44 mediates adhesive cell-cell and cell-extracellular matrix interactions by binding its main ligand, hyaluronan, a glycosaminoglycan that is highly concentrated in the endosteal region.[14,15] All together, adhesion molecules play an important role in maintaining the characteristics of LSCs.

CD82/KAI-1, a member of the tetraspanin superfamily, was originally identified as an accessory molecule in T-cell activation.[16] The most well-characterized function of CD82 in nonimmune cells is integrin-mediated cell adhesion to extracellular matrix.[17] Forced expression of CD82 up-regulated tissue inhibitors of metalloproteinase 1 (TIMP1) and inactivated matrix metalloproteinases 9 (MMP9) in the H1299 human lung carcinoma cells, resulting in suppression of tumor invasion and metastasis.[18] Cell adhesion to collagen I, which is one of the major proteins in the bone marrow (BM) niche, is mostly mediated by three integrin receptors $\alpha1\beta1$, $\alpha2\beta1$, and $\alpha11\beta1$ expressed on cell surface of mesenchymal stem cells.[19] Integrin may associate with CD82 in $CD34^+/CD38^-$ AML cells to promote adhesion to the endosteal niche. However, the roles of CD82 in hematopoietic cells remain to be elucidated.

In this study, we analyzed the protein expression profile of freshly isolated $CD34^+/CD38^-$ AML cells from individuals with AML and compared it with the expression profile of their $CD34^+/CD38^+$ counterparts using isobaric tags for relative and absolute quantitation (iTRAQ) and found the aberrant expression of CD82 in $CD34^+/CD38^-$ AML cells. This study also explored the function of CD82 in $CD34^+/CD38^-$ AML cells *in vitro* as well as *in vivo* by utilizing NOD.Cg-*Rag1*^{tm1Mom} *Il2rg*^{tm1Wjl}/SzJ mice.

Material and Methods

Sample collection and isolation of $CD34^+/CD38^-$ AML cells and their $CD34^+/CD38^+$ counterparts

Leukemia cells were freshly isolated from AML patients (n = 18) with World Health Organization (WHO) classification system subtype minimally differentiated AML (case 6), AML without maturation (cases 1 and 10), AML with maturation (cases 2, 7, and 12), acute myelomonocytic leukemia (cases 4, 14, and 15), AML with myelodysplasia changes (cases 3, 5, 8, 9,16, 17, and 18), and therapy-related AML (cases 11 and 13) after obtaining informed consent with Kochi University Institutional Review Board approval (Supporting Information Table S1). The informed consent was obtained in accordance with the Declaration of Helsinki. $CD34^+/CD38^-$ AML cells and $CD34^+/CD38^+$ counterparts were purified by magnetic cell sorting utilizing a CD34 MultiSort kit and a CD38 MicroBead kit (Miltenyi Biotec GmbH, Germany), as previously described (Supporting Information Fig. S1*a*).[20]

Cells

Chronic eosinophilic leukemia (CEL) EOL-1 cells were obtained from RIKEN BRC Cell Bank (Tsukuba, Japan). Imatinib-resistant EOL-1R cell line was established by culturing with increasing concentrations of imatinib (from 1 to 100 nM) for 6 months.[21] Most of EOL-1R cells expressed CD34 ($92 \pm 9\%$) on their cell surface. On the other hand, CD34 was rarely detectable on cell surface of parental EOL-1 cells ($0.1 \pm 0.1\%$) (figure not shown).

Isolation and culture of primary mesenchymal stromal cells (MSCs)

MSCs were isolated from a BM of healthy donors. BM cells were subjected to centrifugation over a Ficoll-Hypaque gradient to separate mononuclear cells. These cells were resuspended in α-minimal essential medium (Gibco BRL, Rockville, MD) containing 20% fetal bovine serum (FBS) and plated at an initial density of 10^6 cells.[22]

Protein extraction

Proteins were extracted using the complete mammalian proteome kit (539779, Calbiochem, Darmstadt, Germany), according to the manufacturer's instructions.

iTRAQ labeling of peptides

Each protein sample (100 μg) was digested with trypsin and labeled with iTRAQ reagents (Applied Biosystems, Framingham, MA) according to the manufacturer's instructions. Briefly, the proteins extracted from $CD34^+/CD38^-$ AML cells were labeled with iTRAQ reagents 114 (case 1) or 116 (case 2), and proteins extracted from $CD34^+/CD38^+$ counterparts

Table 1. PCR primers

Gene	Direction	Primer
MMP9	Forward	5′-CTCGAACTTTGACAGCGACA-3′
	Reverse	5′-GCCATTCACGTCGTCCTTAT-3′
MMP2	Forward	5′-ACCCAGATGTGGCCAACTAC-3′
	Reverse	5′-TCATGATGTCTGCCTCTCCA-3′
18S	Forward	5′-AAACGGCTACCACATCCAAG-3′
	Reverse	5′-CCTCCAATGGATCCTCGTTA-3′

were labeled with iTRAQ reagents 115 (case 1) or 117 (case 2). Labeled peptide samples were mixed and fractionated as described previously.[23]

Mass spectrometric analysis

NanoLC-MS/MS analyses were performed on an LTQ-Orbitrap XL (Thermo Fisher Scientific, Waltham, MA) equipped with a nano-ESI source and coupled to a Paradigm MG2 pump (Michrom Bioresources, Auburn, CA) and autosampler (HTC PAL, CTC Analytics, Zwingen, Switzerland).[23]

iTRAQ data analysis

Protein identification and quantitation for iTRAQ analysis was carried out using SEQUEST (Bioworks version 3.3.1, Thermo Fisher Scientific) searching against the International Protein Index (IPI) human protein database (version 3.26).[23] Relative protein abundances were determined by comparing the ratio of iTRAQ reporter ion intensities in the MS/MS scan.[23]

Quantitation of CD82-expressing cells using flow cytometry (FACS)

Leukemic peripheral blood (PB) ($n = 3$) and BM ($n = 9$) cells were collected from 12 AML patients (case numbers 1–12) after obtaining informed consent. Leukemic cells were stained with a fluorescein isothiocyanate (FITC)-conjugated monoclonal antibody (mAb) against CD34 (Beckman coulter, CA), a phycoerythrin (PE)-conjugated mAb against CD82 (Abcam, Cambridge, UK), and a PE Cy5-conjugated mAb against CD38 (BioLegend, San Jose, CA). Cells were stained for 30 min on ice. Isotypo-matched immunoglobulins were used as controls. Cells were then analyzed using flow cytometry (FACS Calibur, Becton Dickinson, San Jose, CA) following data analysis by FlowJo software (TreeStar, San Carlos, CA).

RNA isolation and real-time reverse transcription-polymerase chain reaction (RT-PCR)

RNA isolation and cDNA preparation were performed as described previously.[24] Real-time RT-PCR was carried out by using Power SYBR Green PCR Master Mix (Applied Biosystems, Warrington, UK) as described previously.[24] Primers for PCR are shown in Table 1.

Small interfering RNA

Control small interfering (si) RNA and an siRNA against CD82 were purchased from Santa Cruz Biotechnology and Sigma (Deisenhofen, Germany), respectively.

Transfections

EOL-1 and EOL-1R cells were transiently transfected with either control or CD82 siRNA (300 nM) using an Amaxa Nucleofector II electroporator (Wako Pure Chemical Industries, Ltd., Osaka, Japan) with a Nucleofector Kit V (program U-001) as previously described.[25] The preliminary experiments using the green fluorescence protein-expressing vector found that efficacy of transfection with this program was approximately 70% with nearly 70% cell viability, as measured by FACS and Annexin V/PI staining, respectively (figure not shown).

CD82 shRNA lentiviral vector, production and infection

The short hairpin (sh) RNA sequence used to target human CD82 corresponded to the following sequence on the human CD82 transcript variant 2, NCB I accession number NM_001024844. Lentiviral shRNA particles were produced using the viral power packaging system (Invitrogen, CA) with the 293FT packaging cell line (Invitrogen), and lentiviral CD82 shRNA particles (>10^8 titer unit (TU)/ml) were prepared using ultracentrifugation. 5×10^4 $CD34^+/CD38^-$ AML cells were seeded in 24-well plates in 500 μl of Iscove's modified Dulbecco's medium (IMDM) (Invitrogen) containing 10% heat inactivated fetal bovine serum (FBS). After overnight incubation, 5×10^5 TU lentiviral CD82 shRNA particles and polybrene (10 μg/ml) were added per well with serum free medium containing IMDM. After overnight, 1 ml of full media supplemented with FBS, 2-mercaptoethanol, stem cell factor, granulocyte-macrophage colony-stimulating factor (GM-CSF), granulocyte-colony-stimulating factor (G-CSF), IL-3, and erythropoietin (EPO) was added and incubated for 7 days. The control and CD82 shRNA lentiviral vectors co-expressed green fluorescence protein (GFP). Quantification of GFP-positive cells using FACS analysis indicated that the lentiviral transduction efficiency was nearly 70% (Supporting Information Fig. S1*b*). GFP-positive cells were sorted using JSAN (Bay bioscience Co., Ltd., Kobe, Japan).

CD82 lentiviral vector

CD82 cDNA was purchased from Mammalian gene collection (BC000726) and was used as the template for PCR. PCR products were cloned into pLenti6.3/V5-TOPO vector (Invitrogen). CD82-transfected lentiviral particles were produced using the viral power packaging system (Invitrogen) with the 293FT packaging cell line (Invitrogen). The pLenti6.3/V5-TOPO vector was designed to co-express V5 epitope; 5×10^4 $CD34^+/CD38^+$ AML cells were seeded in 24-well plates in 500 μl of IMDM (Invitrogen) containing 10% FBS. After overnight incubation, 5×10^5 TU lentiviral CD82-transfected

lentiviral particles and polybrene (10 μg/ml) were added per well. After overnight, supernatant was removed and 1 ml of full media was added and incubated for 7 days. FACS analysis utilizing an anti-V5 antibody (Invitrogen, R960-25) indicated that the efficiency of transduction into $CD34^+/CD38^+$ AML cells was nearly 80% (Supporting Information Fig. S1*b*).

Migration assays

Freshly isolated either $CD34^+/CD38^-$ (5×10^5 cells) or $CD34^+/CD38^+$ AML cells (1×10^5 cells) were transduced by either CD82 shRNA or CD82 cDNA, respectively, and then seeded in the upper inserts with 3 μm pores coated by fibronectin (Cat. No. 354543, Becton Dickinson Biosciences, Bedford, MA) and mesenchymal stromal cells (MSCs) established from healthy donors, while the lower wells were filled with Iscove's Modified Dulbecco's Medium (IMDM) containing 10% heat inactivated FBS. Similarly, EOL-1 and EOL-1R cells were transiently transfected with either control or CD82 siRNA. After 48 hr, these cells (5×10^5 cells in 100 μl RPMI-1640) were seeded in the upper biocoat cell culture inserts coated by fibronectin, and the lower well was filled to the top with RPMI-1640 containing 10% heat inactivated FBS as a chemoattractant. After incubation for 48 hr, the supernatant was discarded and the cells that had adhered to the fibronectin were gently washed in phosphate-buffered saline (PBS). Cells were then fixed for 1 hr in 4% paraformaldehyde, washed twice in PBS, stained with 4'6-diamidino-2-phenylindole (DAPI), and counted under a microscope (OLYMPUS FV1000-D). The cells that had passed through the membrane filter were collected and the number of viable cells was counted under light microscope after staining with trypan blue.

Gelatin zymography

The culture supernatant as well as whole cell proteins of EOL-1 and EOL-1R cells were harvested. Gelatinolytic activities were carried out by utilizing a gelatin-zymography kit (Primarycell, Hokkaido, Japan). Each lane was loaded with 30 μg of whole protein lysates or 20 μl of supernatant.

Colony forming assay

The colony-forming assay was performed with methylcellulose medium H4034 (StemCell Technologies, Vancouver, BC, Canada), as previously described.[20]

Bone marrow transplantation and engraftment assay

NOD.Cg-*Rag1*^{tm1Mom} *Il2rg*^{tm1Wjl}/SzJ mice (Stock number: 007799) were purchased from the Jackson Laboratory for experimental animals (Bar Harbor)[26] and bred in a pathogen-free environment in accordance with the guidelines of the Kochi University School of Medicine. The 6-week-old mice were utilized for experiments. $CD34^+/CD38^-$ AML cells (1×10^4 cells) transfected with either scrambled control or CD82 shRNA were injected to each mouse intravenously *via* the tail vein. At 9 weeks after transplantation, mice were euthanized and BM were removed. BM cells were flushed from the femurs using 25-gauge needles (Becton Dickinson Biosciences) and then fixed in formalin. The human cell engraftment was analyzed by using flow cytometry after staining of spleen cells with human CD45 PE Cy5-conjugated mAb (Dako, Glostrup, Denmark) and human CD33 PE-conjugated mAb (Becton Dickinson Biosciences).

Single cell RT-PCR

A single $CD34^+/CD38^-$ AML cell was isolated by BD FACS AriaII (Becton Dickinson Biosciences) and subjected to RT-PCR by AmpliSpeed slide cycler (Beckman Coulter, Munich, Germany) and ABI StepOnePlus (Applied Biosystems) to measure the levels of CD82 and MMP9.

Immunohistochemistry of CD82 in BM sections

Immunohistochemical staining of CD82 was performed with a Ventana DISCOVERYTM autostainer system (Ventana Japan, Osaka, Japan) as previously described.[25] The anti-CD82 antibody (Santa Cruz Biotechnology, Santa Cruz, CA) was used.

Cell cycle analysis by flow cytometry (FACS)

Cell cycle distribution of $CD34^+/CD38^-$ AML cell was measured as previously described after transduction of either scrambled control or CD82 shRNA. Briefly, the cells were stained with Ki-67 (Santa Cruz Biotechnology) and propidium iodide and subjected to FACS.[20]

Homing analysis

The human AML cells isolated from patients were treated with anti-human CD82 (Santa Cruz Biotechnology) or control IgG (eBiosScience, San Diego, CA) antibody on ice for 1 hr. This anti-CD82 antibody worked as a neutralizing antibody (Nishioka *et al.* unpublished data). Cells were washed gently with PBS and were injected to each mouse intravenously *via* the tail vein. At 16 hr after transplantation, mice were euthanized and BMs were removed and analyzed by flow cytometry with anti-human CD45 and CD34 antibodies. We acquired in the range of 1×10^6 to 3×10^6 events per sample.

Statistical analysis

When comparing two groups, Student's *t*-test was used. For demonstration of association, the Pearson's correlation coefficient test was applied. All statistical analyses were carried out using SPSS software (Version 11.03; spss, Tokyo, Japan) and the results were considered to be significant when the p value was <0.05, and highly significant when the p value was <0.01.

Results

Protein expression profiles of $CD34^+/CD38^-$ AML cells and $CD34^+/CD38^+$ counterparts

Four samples (two from $CD34^+/CD38^-$ AML cells lysates and two from $CD34^+/CD38^+$ counterparts lysates, approximately 10 million cells per sample) were trypsinized and