We excluded transition peaks with a signal-to-noise ratio <10, which has been used as empirical LOQ (24), and then compared the profile and proportion of the remaining transition peaks between the SI peptide and endogenous peptide to select appropriate peaks for quantitative analysis. Removing the outliers of transitions due to interference or co-eluting non-specific backgrounds was essential to improve accuracy and reliability. Each transition among the samples had to exhibit a similar peak shape to that with the transition of the SI peptide, which resulted in a minimal CV area ratio (CV<35%) between transitions. We confirmed every transition peak by a manual inspection and removed the peaks that did not conform to the above criteria, which led to accurate and significant quantitation (Supplemental Fig. 2).

We obtained the average of these ratios of more than two transitions as the relative quantitative value of the target peptide. Statistical analysis of the area ratios was performed using the *t* test. In addition, if the expression of one of the two peptides of proteins was significantly different between the sample groups, we considered the protein to be differentially expressed. Using the SRM/MRM method, 172 peptides from 98 proteins were quantified in more than three samples from polyps and cancer with or without metastasis (Supplemental Table 7). Significant differences (ratio >2.0, p-value <0.1; ratio <0.5, p-value <0.1) in at least one of the targeted peptides were detected in 69

proteins (Supplemental Fig. 3, Supplemental Table 7).

The expression of ITGA5, GPRC5A, PDGFRB, and TFRC was shown to be different in colorectal or other cancer tissues (26-29). The results of iTRAQ and SRM/MRM on these proteins are shown in Figure 2A. The expression of these proteins showed very similar patterns on iTRAQ and SRM/MRM (Supplemental Fig. 4). Furthermore, changes in the expression of ITGA5 were confirmed by Western blotting (Fig. 2B). The similar results obtained by SRM/MRM and iTRAQ were further verified by Western blotting, which indicated that the SRM/MRM assay can be used to confirm the candidates identified in the discovery phase.

Verification of biomarker candidates by SRM/MRM

We verified 69 confirmed proteins in an independent set of patient samples (polyps (n=10), cancer without metastasis (n=10), and cancer with metastasis (n=10)) (Table 4, Supplemental Table 1, 9, Supplemental Fig. 5). We performed five technical replicates using sample mixtures prepared from patient tissue samples to evaluate the reproducibility of our SRM/MRM assay, and obtained high reproducibility (CV<11%) (Supplemental Table 8). We did not analyzed process replicates, therefore the actual experimental variability is likely higher than shown by the technical replicate

performance owing to variability in digestion and other sample handling steps. The expression levels of a total of 20 proteins: GPRC5A, PRTN3, CEACAM5, ANTXR1, PXMP4, SLC2A3, ENPEP, PDGFRB, GGT5, MMP14, TFRC, MRC2, SPARC, HSPB1, FCGR1A, THY1, TMEM41A, SLC4A2, FCER1G, and CEACAM1, were significantly higher in cancer without metastasis than in polyps (ratio >2.0, p-value <0.05). In addition, the expression levels of 10 proteins: ITGA11, BST1, LTBP2, ITGA5, TMEM97, TSPAN9, SIGMAR1, C8orf55, UBAC2 and SERPIND1, were significantly higher in cancer without or with metastasis than in polyps (ratio >1.7, p-value <0.05). The expression levels of another five proteins: CEACAM6, LRRC15, GPC6, C5AR1 and TLCD1, were markedly higher in cancer tissues than in polyps. The expression levels of eight proteins: CLCA1, FCGBP, B3GNT6, MUC2, ANXA13, AKAP5, PRG2, and KIAA1324, were lower in cancer with and without metastasis than in polyps (ratio >0.5, p-value <0.05). The expression of EPB41L3 was also shown to be lower in cancer tissues than in polyps. This verification step as well as the discovery step revealed that the expression levels of ITGA5, GPRC5A, PDGFRB, and TFRC were markedly higher in cancer tissues than in polyps (Fig. 3). Overall, the expression patterns of 47 out of 69 confirmed proteins were similar between the confirmation and verification analyses.

Further validation of C8orf55 by Western blotting and immunohistochemistry

We focused on C8orf55 among the biomarker candidates that displayed significant differences in SRM/MRM because it has not been previously reported as a biomarker candidate for cancer and a specific antibody against this protein was available. C8orf55 (also called THEM6) is a 208-amino-acid protein that has one predicted transmembrane domain in the N-terminal region; however, its function is unknown. iTRAQ and subsequent confirmation using the SRM/MRM assay revealed that the expression of C8orf55 was upregulated with cancer progression (Fig. 4A). Furthermore, in the verification step, the expression of this protein was higher in cancer without metastasis than in polyps (ratio=1.92, p-value<0.01). Western blotting was also performed to verify these changes in expression levels (Fig. 4B). Immunohistochemical analysis of colorectal cancer tissue showed that the expression of C8orf55 was high in cancer cells, but was negligible in normal cells (Fig. 4C). These results indicated that the expression of C8orf55 increased in a stepwise fashion with cancer progression.

Examination of C8orf55 expression in various cancer tissues using tissue microarrays

The expression of the tumor markers used in clinical practice, such as CEA and

CA19-9, was shown to be higher in multiple cancer types. Therefore, we investigated whether C8orf55 was expressed in various cancer tissues using tissue microarrays (TMA), which contained 1150 cores from 14 common cancer tissues and 280 cores from corresponding normal tissues (Supplemental Fig. 6). TMA revealed that the expression of C8orf55 was high in many of the cores prepared from colon cancer tissue, but was negligible in those from normal colon tissues (Fig. 5). TMA also showed that that the expression of C8orf55 was significantly higher in colon cancer tissue than in normal tissue. Immunostaining for C8orf55 was stronger in cancer tissues such as those form the stomach and breast than in normal tissues (Fig. 5). These results demonstrated that C8orf55 may be a potential biomarker for colorectal, stomach, and breast cancer.

Discussion

A number of large-scale proteomic analyses of cancer tissues for biomarker discovery have been reported to date (30-32); however, few studies have validated the candidate proteins identified because of the absence of an appropriate validation method. SRM/MRM was recently shown to be an efficient validation method (3-5) and several studies, including our own, reported the identification of biomarker candidates by quantitative shotgun proteomics using the iTRAQ labeling method and verification

by SRM/MRM (19, 21, 33). In the present study, we performed a proteomic analysis of membrane fractions prepared from colorectal cancer tissue to identify novel biomarker candidates for diagnosis and/or therapeutic targets. We identified membrane proteins, the expression levels of which were altered with the development and progression of colorectal cancer, using comprehensive quantitative analysis with iTRAQ. The most significant achievement of this study was the SRM/MRM-based confirmation and simultaneous large-scale verification using an independent set of tissue samples.. Of the 105 biomarker candidate proteins identified by iTRAQ, changes in the expression of 69 proteins were confirmed by SRM/MRM, with significant differences being verified in 44 proteins between groups. This discovery-confirmation-verification workflow should be able to identify more reliable biomarkers for the clinical diagnosis of colon cancer. To the best of our knowledge, we have performed the largest verification of biomarker candidate membrane proteins to date. This verification process using SRM/MRM enabled us to select more potential candidates and prioritize the subsequent validation, and may represent a rapid and effective method to identify novel biomarkers.

We were able to identify 5566 proteins in the membrane fraction in the present study, 3087 (58.4%) of which were predicted to be membrane proteins. This number was markedly higher than that previously reported (34-38); however, non-membrane

proteins were also identified in addition to membrane proteins, and this was attributed to the preparation of crude membrane fractions using a simple method. One of the reasons for the increased rate of membrane protein identification was the PTS method-based isolation of membrane proteins (12, 13). The PTS method enables the efficient isolation of membrane proteins and allows the use of a high detergent concentration to achieve the efficient solubilization of very hydrophobic membrane proteins in the cleavage procedure of membrane proteins. Thus, this method may provide deeper proteome coverage for the identification of tissue membrane proteins.

We focused on membrane proteins in this study because membrane proteins are not only involved in the regulation of cell signaling and cell-cell interactions, but are also suitable therapeutic targets for cancers (39). One of the greatest advances in the treatment of cancer in recent years has been the discovery of molecular-targeted drugs, which has resulted in the development of many antibody drugs. Membrane proteins are clearly the best targets for antibody drugs. In this study, we identified a number of previously unreported membrane proteins, the expression of which changed with the development and progression of colorectal cancer. These membrane proteins may be novel therapeutic targets for antibody drug discovery.

Membrane proteins are also suitable biomarkers for the screening and

diagnosis of various cancers. Diagnostic biomarkers are ideally detected and quantified in biological fluids such as the plasma and/or urine; however, soluble proteins derived from tissue leakage are often very difficult to detect because there are very few and they are unstable. In contrast, membrane proteins and extracellular proteins are potentially shed and secreted from cells into the circulation; some are actively secreted as microvesicles, such as exosomes, which are very stable and may be potential biomarkers. Several previous studies reported the potential for diagnosing malignant tumors, such as colorectal cancer, melanoma, and glioblastoma, by analyzing exosomal proteins (40-42). Thus, the membrane proteins identified in this study may be promising biomarker candidates for the diagnosis of colorectal cancer.

We observed variations in the quantitative results obtained from iTRAQ and SRM. The samples used for iTRAQ were fractionated with a SCX column, while those for SRM were not. Therefore, variations may have occurred in the quantitative results obtained from iTRAQ and SRM due to differences in the complexities of the samples analyzed. Splicing isoforms or post-translational modifications may also have been involved in these variations because iTRAQ ratios were calculated as the average of all contributing peptide iTRAQ measurements and SRM ratios were obtained by measuring a target peptide.

We investigated differences in the expression levels of proteins between polyps and cancer tissues without metastasis in the present study using proteomic analysis to identify characteristic expression profiles in cancer. Although a number of previous biomarker studies identified hundreds of candidate proteins by comparing cancer tissues with matched normal tissues, many proteins unrelated to malignant properties may also have been included because cancer is generally not directly derived from normal tissues. Thus, the best negative control would be benign tumors, ideally premalignant lesions. In this regard, colorectal polyps are considered to be the best control for colorectal cancer. Moreover, a comparison between different stages of cancer tissues, including benign tumors, is the optimal procedure to identify more useful biomarker candidates.

In our study, C8orf55 was confirmed by SRM/MRM and Western blotting, the findings of which were further verified by multiple cancer tissue microarrays (TMA1150). TMA1150 had 1150 cores from 50 or 100 cases of 14 cancer types and was previously shown to be useful for evaluating changes in protein expression in multiple cancers (25). TMA1150 can also be used to examine the expression of target proteins in various cancer tissues as well as in dozens of cases of colorectal cancer. The extensive validation of the expression of identified candidates in

various types of cancer tissues is important in order to determine their usefulness as biomarkers for diverse cancers. In this regard, multi-cancer TMA is a very effective method that can be used to rapidly and simply evaluate the expression patterns of various cancers. TMA1150 revealed that the expression of C8orf55 was higher not only in colon cancer tissue, but also in other cancer tissues, which suggested that these proteins have the potential to be biomarkers for stomach and breast cancer as well as colon cancer.

In conclusion, we successfully performed a SRM/MRM-based large-scale verification of biomarker candidate membrane proteins for colorectal cancer tissues. The methods described here can be readily applied to any type of cancer tissue and can contribute to the identification of novel biomarkers for the diagnosis and therapeutic targets of diseases.

Acknowledgments

This work was supported by a Grant-in-Aid, Research on Biological Markers for New Drug Development H20-0005, to T.T. from the Ministry of Health, Labour and Welfare of Japan. It was also supported by Grants-in-Aid 21390354 to T.T and 22590545 to H.K. from the Ministry of Education, Science, Sports and Culture of Japan. The mass

spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the PRIDE partner repository (43) with the dataset identifier PXD000851.

References

1. 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., and Pappin, D. J. (2004) Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents, *Mol. Cell. Proteomics* 3, 1154-1169.
2. Ong, S. E., Blagoev, B., Kratchmarova, I., Kristensen, D. B., Steen, H., Pandey, A., and Mann, M. (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics, *Mol. Cell. Proteomics* 1, 376-386.
3. Whiteaker, J. R., Zhang, H., Zhao, L., Wang, P., Kelly-Spratt, K. S., Ivey, R. G., Piening, B. D., Feng, L. C., Kasarda, E., Gurley, K. E., Eng, J. K., Chodosh, L. A., Kemp, C. J., McIntosh, M. W., and Paulovich, A. G. (2007) Integrated pipeline for mass spectrometry-based discovery and confirmation of biomarkers demonstrated in a mouse model of breast cancer, *J. Proteome Res.* 6, 3962-3975.
4. Keshishian, H., Addona, T., Burgess, M., Mani, D. R., Shi, X., Kuhn, E., Sabatine, M. S., Gerszten, R. E., and Carr, S. A. (2009) Quantification of cardiovascular biomarkers in patient plasma by targeted mass spectrometry and stable isotope dilution, *Mol. Cell. Proteomics* 8, 2339-2349.
5. 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., and Paulovich, A. G. (2011) A targeted proteomics-based pipeline for verification of biomarkers in plasma, *Nat. Biotechnol.* 29, 625-634.
6. Ahram, M., Litou, Z. I., Fang, R., and Al-Tawallbeh, G. (2006) Estimation of membrane proteins in the human proteome, *In Silico Biol.* 6, 379-386.

7. Hopkins, A. L., and Groom, C. R. (2002) The druggable genome, *Nat. Rev. Drug Discovery* 1, 727-730.

8. Chen, E. I., Cociorva, D., Norris, J. L., and Yates, J. R., 3rd. (2007) Optimization of mass spectrometry-compatible surfactants for shotgun proteomics, *J. Proteome Res.* 6, 2529-2538.

9. Mitra, S. K., Gantt, J. A., Ruby, J. F., Clouse, S. D., and Goshe, M. B. (2007) Membrane proteomic analysis of Arabidopsis thaliana using alternative solubilization techniques, *J. Proteome Res.* 6, 1933-1950.

10. Zhang, N., Chen, R., Young, N., Wishart, D., Winter, P., Weiner, J. H., and Li, L. (2007) Comparison of SDS- and methanol-assisted protein solubilization and digestion methods for Escherichia coli membrane proteome analysis by 2-D LC-MS/MS, *Proteomics* 7, 484-493.

11. Zhou, J., Zhou, T., Cao, R., Liu, Z., Shen, J., Chen, P., Wang, X., and Liang, S. (2006) Evaluation of the application of sodium deoxycholate to proteomic analysis of rat hippocampal plasma membrane, *J. Proteome Res.* 5, 2547-2553.

12. Masuda, T., Tomita, M., and Ishihama, Y. (2008) Phase transfer surfactant-aided trypsin digestion for membrane proteome analysis, *J. Proteome Res.* 7, 731-740.

13. Masuda, T., Saito, N., Tomita, M., and Ishihama, Y. (2009) Unbiased quantitation of Escherichia coli membrane proteome using phase transfer surfactants, *Mol. Cell. Proteomics* 8, 2770-2777.

14. Tomonaga, T., Matsushita, K., Yamaguchi, S., Oh-Ishi, M., Kodera, Y., Maeda, T., Shimada, H., Ochiai, T., and Nomura, F. (2004) Identification of altered protein expression and post-translational modifications in primary colorectal cancer by using agarose two-dimensional gel electrophoresis., *Clin Cancer Res.* 10, 2007-2014.

15. Nishimori, T., Tomonaga, T., Matsushita, K., Oh-Ishi, M., Kodera, Y., Maeda, T., Nomura, F., Matsubara, H., Shimada, H., and Ochiai, T. (2006) Proteomic analysis of primary esophageal squamous cell carcinoma reveals downregulation of a cell adhesion protein, periplakin, *Proteomics* 6, 1011-1018.

16. Seimiya, M., Tomonaga, T., Matsushita, K., Sunaga, M., Oh-Ishi, M., Kodera, Y., Maeda, T., Takano, S., Togawa, A., Yoshitomi, H., Otsuka, M., Yamamoto, M., Nakano, M., Miyazaki, M., and Nomura, F. (2008) Identification of novel immunohistochemical tumor markers for primary hepatocellular carcinoma; clathrin heavy chain and formiminotransferase cyclodeaminase, *Hepatology* 48, 519-530.

17. Katada, K., Tomonaga, T., Satoh, M., Matsushita, K., Tonoike, Y., Kodera, Y., Hanazawa, T., Nomura, F., and Okamoto, Y. (2012) Plectin promotes migration and

invasion of cancer cells and is a novel prognostic marker for head and neck squamous cell carcinoma, *J Proteomics.* 75, 1803-1815.

18. Lazebnik, Y. (2010) What are the hallmarks of cancer?, *Nat. Rev. Cancer* 10, 232-233.

19. Muraoka, S., Kume, H., Watanabe, S., Adachi, J., Kuwano, M., Sato, M., Kawasaki, N., Kodera, Y., Ishitobi, M., Inaji, H., Miyamoto, Y., Kato, K., and Tomonaga, T. (2012) Strategy for SRM-based verification of biomarker candidates discovered by iTRAQ method in limited breast cancer tissue samples, *J. Proteome Res.* 11, 4201-4210.

20. Muraoka, S., Kume, H., Adachi, J., Shiromizu, T., Watanabe, S., Masuda, T., Ishihama, Y., and Tomonaga, T. (2013) In-depth membrane proteomic study of breast cancer tissues for the generation of a chromosome-based protein list, *J. Proteome Res.* 12, 208-213.

21. Narumi, R., Murakami, T., Kuga, T., Adachi, J., Shiromizu, T., Muraoka, S., Kume, H., Kodera, Y., Matsumoto, M., Nakayama, K., Miyamoto, Y., Ishitobi, M., Inaji, H., Kato, K., and Tomonaga, T. (2012) A strategy for large-scale phosphoproteomics and SRM-based validation of human breast cancer tissue samples, *J. Proteome Res.* 11, 5311-5322.

22. Shiromizu, T., Adachi, J., Watanabe, S., Murakami, T., Kuga, T., Muraoka, S., and Tomonaga, T. (2013) Identification of Missing Proteins in the neXtProt Database and Unregistered Phosphopeptides in the PhosphoSitePlus Database As Part of the Chromosome-Centric Human Proteome Project, *J. Proteome Res.* DOI 10.1021/pr300825v

23. Rappsilber, J., Ishihama, Y., and Mann, M. (2003) Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics, *Anal. Chem.* 75, 663-670.

24. Selevsek, N., Matondo, M., Sanchez Carbayo, M., Aebersold, R., and Domon, B. Systematic quantification of peptides/proteins in urine using selected reaction monitoring, *Proteomics* 11, 1135-1147.

25. Kitano, H., Kageyama, S., Hewitt, S. M., Hayashi, R., Doki, Y., Ozaki, Y., Fujino, S., Takikita, M., Kubo, H., and Fukuoka, J. (2010) Podoplanin expression in cancerous stroma induces lymphangiogenesis and predicts lymphatic spread and patient survival, *Arch Pathol Lab Med.* 134, 1520-1527.

26. Yang, B., Gao, J., Rao, Z., and Shen, Q. (2013) Clinicopathological and Prognostic Significance of alpha5beta1-integrin and MMP-14 Expressions in Colorectal Cancer, *Neoplasma.* DOI 10.4149/neo_2013_034.

27. Zougman, A., Hutchins, G. G., Cairns, D. A., Verghese, E., Perry, S. L., Jayne, D.

G., Selby, P. J., and Banks, R. E. (2013) Retinoic acid-induced protein 3: identification and characterisation of a novel prognostic colon cancer biomarker, *Eur J Cancer* 49, 531-539.

28. Sillars-Hardebol, A. H., Carvalho, B., de Wit, M., Postma, C., Delis-van Diemen, P. M., Mongera, S., Ylstra, B., van de Wiel, M. A., Meijer, G. A., and Fijneman, R. J. (2010) Identification of key genes for carcinogenic pathways associated with colorectal adenoma-to-carcinoma progression, *Tumour Biol.* 31, 89-96.

29. Prutki, M., Poljak-Blazi, M., Jakopovic, M., Tomas, D., Stipancic, I., and Zarkovic, N. (2006) Altered iron metabolism, transferrin receptor 1 and ferritin in patients with colon cancer, *Cancer Lett.* 238, 188-196.

30. Luo, Y., Wang, L., and Wang, J. (2013) Developing proteomics-based biomarkers for colorectal neoplasms for clinical practice: Opportunities and challenges, *Proteomics Clin Appl.* 7, 30-41.

31. Lehtio, J., and De Petris, L. (2010) Lung cancer proteomics, clinical and technological considerations, *J. Proteomics* 73, 1851-1863.

32. Kalinina, J., Peng, J., Ritchie, J. C., and Van Meir, E. G. (2011) Proteomics of gliomas: initial biomarker discovery and evolution of technology, *Neuro Oncol.* 13, 926-942.

33. Thingholm, T. E., Bak, S., Beck-Nielsen, H., Jensen, O. N., and Gaster, M. (2011) Characterization of human myotubes from type 2 diabetic and nondiabetic subjects using complementary quantitative mass spectrometric methods, *Mol. Cell. Proteomics* 10, DOI 10.1074/mcp.M110.006650.

34. Josic, D., and Clifton, J. G. (2007) Mammalian plasma membrane proteomics, *Proteomics* 7, 3010-3029.

35. 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., and Chan, E. C. (2010) Comparison of membrane fraction proteomic profiles of normal and cancerous human colorectal tissues with gel-assisted digestion and iTRAQ labeling mass spectrometry, *FEBS J.* 277, 3028-3038.

36. 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., and Chen, Y. J. (2011) An informatics-assisted label-free approach for personalized tissue membrane proteomics: case study on colorectal cancer, *Mol. Cell. Proteomics* 10, DOI 10.1074/mcp.M110.003087.

37. Besson, D., Pavageau, A. H., Valo, I., Bourreau, A., Belanger, A., Eymerit-Morin, C., Mouliere, A., Chassevent, A., Boisdron-Celle, M., Morel, A., Solassol,

J., Campone, M., Gamelin, E., Barre, B., Coqueret, O., and Guette, C. (2011) A quantitative proteomic approach of the different stages of colorectal cancer establishes OLFM4 as a new nonmetastatic tumor marker, *Mol. Cell. Proteomics* 10, DOI 10.1074/mcp.M111.009712.

38. 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., and Sirdeshmukh, R. (2012) LC-MS/MS analysis of differentially expressed glioblastoma membrane proteome reveals altered calcium signaling and other protein groups of regulatory functions, *Mol. Cell. Proteomics* 11, DOI 10.1074/mcp.M111.013565.

39. Rucevic, M., Hixson, D., and Josic, D. (2011) Mammalian plasma membrane proteins as potential biomarkers and drug targets, *Electrophoresis* 32, 1549-1564.

40. Choi, D. S., Park, J. O., Jang, S. C., Yoon, Y. J., Jung, J. W., Choi, D. Y., Kim, J. W., Kang, J. S., Park, J., Hwang, D., Lee, K. H., Park, S. H., Kim, Y. K., Desiderio, D. M., Kim, K. P., and Gho, Y. S. (2011) Proteomic analysis of microvesicles derived from human colorectal cancer ascites, *Proteomics* 11, 2745-2751.

41. Peinado, H., Aleckovic, M., Lavotshkin, S., Matei, I., Costa-Silva, B., Moreno-Bueno, G., Hergueta-Redondo, M., Williams, C., Garcia-Santos, G., Ghajar, C., Nitadori-Hoshino, A., Hoffman, C., Badal, K., Garcia, B. A., Callahan, M. K., Yuan, J., Martins, V. R., Skog, J., Kaplan, R. N., Brady, M. S., Wolchok, J. D., Chapman, P. B., Kang, Y., Bromberg, J., and Lyden, D. (2012) Melanoma exosomes educate bone marrow progenitor cells toward a pro-metastatic phenotype through MET, *Nat. Med.* 18, 883-891.

42. Shao, H., Chung, J., Balaj, L., Charest, A., Bigner, D. D., Carter, B. S., Hochberg, F. H., Breakefield, X. O., Weissleder, R., and Lee, H. (2012) Protein typing of circulating microvesicles allows real-time monitoring of glioblastoma therapy, *Nat. Med.* 18, 1835-1840.

43. Vizcaino, J. A., Cote, R. G., Csordas, A., Dianes, J. A., Fabregat, A., Foster, J. M., Griss, J., Alpi, E., Birim, M., Contell, J., O'Kelly, G., Schoenegger, A., Ovelleiro, D., Perez-Riverol, Y., Reisinger, F., Rios, D., Wang, R., and Hermjakob, H. (2013) The PRoteomics IDEntifications (PRIDE) database and associated tools: status in 2013. *Nucleic Acids Res.* 41, D1063-1069.

Figure Legends

Table 1. Number of predicted membrane proteins

Total identified proteins	5566
number	%
Number of proteins with transmembrane domains	1567 [a]	28.2
GO-annotated	5287	100
Membrane	3087	58.4 [b]
Extracellular	652	12.3 [c]

[a] Number of proteins with transmembrane domains predicted by TMHMM algorithm.

[b, c] The ratio of membrane or extracellular proteins to GO-annotated proteins.

Table 2. Number of proteins with significant difference in expression

		C / P		Cm / C		Cm / P	
ratio	p-value	TM + mem	Extra	TM + mem	Extra	TM + mem	Extra
> 2.0	< 0.1	108	34	21	8	79	21
< 0.5	< 0.1	51	21	11	9	20	16
total		159	55	32	17	99	37

C/P, ratio of cancer without metastasis to polyps. Cm/C, ratio of cancer with metastasis to cancer without metastasis. Cm/P, ratio of cancer with metastasis to polyps. TM + mem, number of proteins with predicted transmembrane domain or annotated as membrane protein. Extra, number of proteins annotated as extracellular protein.

Table 3. List of the proteins analyzed by SRM/MRM and their quantitation data using iTRAQ

A. The list of proteins increased in expression between polyps and cancer without metastasis (n=66)

Accession	protein name	gene name	TM	GO (mem)	GO (extra)	C/P	p-value	Cm/C	p-value	Cm/P	p-value
Q12884	Seprase	FAP	1	mem		5.98	<0.01	0.67	0.190	4.03	0.029
P32926	Desmoglein-3	DSG3	0	mem		4.54	<0.01	0.41	0.083	1.87	0.323
Q6P5W5	Zinc transporter ZIP4	SLC39A4	7	mem		4.35	0.075	0.42	0.189	1.84	0.217
Q8NFJ5	Retinoic acid-induced protein 3	GPRC5A	7	mem		3.99	<0.01	0.77	0.359	3.06	0.012
P40199	Carcinoembryonic antigen-related cell adhesion molecule 6	CEACAM6	0	mem		3.69	0.029	0.85	0.690	3.12	0.031
O95832	Claudin-1	CLDN1	4	mem		3.47	0.054	0.51	0.180	1.77	0.127
Q8TF66	Leucine-rich repeat-containing protein 15	LRRC15	1	mem		3.40	0.032	0.58	0.193	1.96	0.060
P24158	Myeloblastin	PRTN3	0	mem	extra	3.35	0.098	0.38	0.134	1.28	0.526
P50150	Guanine nucleotide-binding protein G(I)/G(S)/G(O) subunit gamma-4	GNG4	0	mem		3.31	0.074	0.77	0.570	2.56	0.051
P80511	Protein S100-A12	S100A12	0	mem	extra	3.28	0.068	1.06	0.857	3.46	0.070
P06731	Carcinoembryonic antigen-related cell adhesion molecule 5	CEACAM5	0	mem		3.27	<0.01	0.79	0.275	2.57	<0.01
Q9UKX5	Integrin alpha-11	ITGA11	1	mem		3.23	<0.01	0.62	0.081	2.00	0.016
Q10588	ADP-ribosyl cyclase 2	BST1	1	mem		3.16	0.023	0.55	0.102	1.75	0.033
P08253	72 kDa type IV collagenase	MMP2	0	mem	extra	3.02	<0.01	0.39	<0.01	1.19	0.599
Q9H6X2	Anthrax toxin receptor 1	ANTXR1	1	mem		2.98	<0.01	0.70	0.027	2.09	<0.01
Q9Y6I8	Peroxisomal membrane protein 4	PXMP4	0	mem		2.97	<0.01	0.74	0.084	2.19	<0.01
Q12805	EGF-containing fibulin-like extracellular matrix protein 1	EFEMP1	0	mem	extra	2.97	<0.01	0.57	0.037	1.71	0.040
Q14767	Latent-transforming growth factor beta-binding protein 2	LTBP2	0	mem	extra	2.91	<0.01	0.69	0.182	2.01	0.141
P16444	Dipeptidase 1	DPEP1	0	mem		2.85	0.033	1.06	0.854	3.02	<0.01
P84157	Matrix-remodeling-associated protein 7	MXRA7	1	mem		2.82	0.069	0.45	0.112	1.26	0.569
P11169	Solute carrier family 2, facilitated glucose transporter member 3	SLC2A3	10	mem		2.74	0.051	0.58	0.181	1.60	0.064
P08648	Integrin alpha-5	ITGA5	1	mem		2.59	<0.01	0.66	0.056	1.70	0.044
P55001	Microfibrillar-associated protein 2	MFAP2	0		extra	2.56	<0.01	0.46	<0.01	1.16	0.322
Q9ULK5	Vang-like protein 2	VANGL2	4	mem		2.55	0.098	0.39	0.086	1.00	0.989
Q5BJF2	Transmembrane protein 97	TMEM97	4	mem		2.54	<0.01	0.73	0.250	1.85	0.040
Q07075	Glutamyl aminopeptidase	ENPEP	1	mem		2.53	<0.01	0.70	0.201	1.77	0.104
Q9UGT4	Sushi domain-containing protein 2	SUSD2	1	mem		2.46	0.013	0.58	0.066	1.43	0.062
Q8N6Q3	CD177 antigen	CD177	0	mem		2.45	0.031	0.50	0.055	1.23	0.378
P07093	Glia-derived nexin	SERPINE2	0	mem	extra	2.43	0.059	0.85	0.699	2.06	0.132
Q96KR6	Transmembrane protein C20orf108	C20orf108	3	mem		2.39	0.020	0.73	0.287	1.75	0.041
P09619	Beta-type platelet-derived growth factor receptor	PDGFRB	1	mem		2.38	<0.01	0.85	0.423	2.01	0.014
Q7L4E1	Protein FAM73B	FAM73B	0	mem		2.34	<0.01	0.50	<0.01	1.17	0.289
O75954	Tetraspanin-9	TSPAN9	4	mem		2.31	<0.01	0.70	0.088	1.61	<0.01
Q9Y625	Glypican-6	GPC6	0	mem	extra	2.31	<0.01	0.63	0.055	1.45	0.179
Q8IUS5	Epoxide hydrolase 4	EPHX4	1	mem		2.29	0.043	1.13	0.614	2.59	<0.01
P36269	Gamma-glutamyltransferase 5	GGT5	1	mem		2.28	<0.01	0.71	0.172	1.63	0.047
Q8IWU6	Extracellular sulfatase Sulf-1	SULF1	0		extra	2.28	<0.01	0.82	0.445	1.88	0.074
Q6ZMP0	Thrombospondin type-1 domain-containing protein 4	THSD4	0		extra	2.26	0.042	0.59	0.278	1.35	0.656
P21730	C5a anaphylatoxin chemotactic receptor	C5AR1	7	mem		2.22	0.090	0.42	0.065	0.93	0.776
P35555	Fibrillin-1	FBN1	0	mem	extra	2.22	0.039	0.38	0.022	0.84	0.363
P98095	Fibulin-2	FBLN2	0		extra	2.20	<0.01	0.68	0.206	1.49	0.300
P31997	Carcinoembryonic antigen-related cell adhesion molecule 8	CEACAM8	0	mem	extra	2.20	0.090	0.51	0.118	1.11	0.592
Q14766	Latent-transforming growth factor beta-binding protein 1	LTBP1	0	mem	extra	2.19	<0.01	0.62	0.015	1.35	0.087
Q99720	Sigma non-opioid intracellular receptor 1	SIGMAR1	1	mem		2.19	<0.01	0.86	0.439	1.88	<0.01
P50281	Matrix metalloproteinase-14	MMP14	1	mem	extra	2.19	<0.01	0.70	0.096	1.53	0.078
P02786	Transferrin receptor protein 1	TFRC	1	mem	extra	2.18	<0.01	1.09	0.579	2.38	<0.01
P31431	Syndecan-4	SDC4	1	mem	extra	2.16	0.082	0.55	0.143	1.20	0.172
Q9UBG0	C-type mannose receptor 2	MRC2	1	mem		2.15	<0.01	0.68	0.078	1.47	0.230
Q9P121	Neurotrimin	NTM	0	mem		2.15	0.058	0.56	0.082	1.20	0.368
P09486	SPARC	SPARC	0		extra	2.14	<0.01	0.85	0.321	1.81	0.025
P05106	Integrin beta-3	ITGB3	1	mem		2.13	0.023	0.70	0.245	1.49	0.165
P04792	Heat shock protein beta-1	HSPB1	0	mem		2.11	<0.01	1.34	0.421	2.83	0.035
Q9NVM1	Protein FAM176B	FAM176B	1	mem		2.08	0.046	1.03	0.951	2.13	0.276
P08514	Integrin alpha-IIb	ITGA2B	1	mem		2.08	0.083	0.88	0.759	1.83	0.130
Q8WUY1	UPF0670 protein C8orf55	C8orf55	1		extra	2.07	<0.01	1.40	0.158	2.90	<0.01
P12314	High affinity immunoglobulin gamma Fc receptor I	FCGR1A	1	mem		2.07	<0.01	0.68	0.105	1.41	0.149
P04216	Thy-1 mem glycoprotein	THY1	0	mem		2.06	<0.01	0.77	0.092	1.59	0.023
P08174	Complement decay-accelerating factor	CD55	0	mem	extra	2.05	<0.01	1.04	0.879	2.13	0.020
Q96HV5	Transmem protein 41A	TMEM41A	6	mem		2.04	<0.01	0.76	0.060	1.54	<0.01
Q9ULS5	Transmem and coiled-coil domains protein 3	TMCC3	2	mem		2.04	0.040	0.61	0.195	1.25	0.434
Q01628	Interferon-induced transmem protein 3	IFITM3	2	mem		2.04	0.021	1.07	0.770	2.18	<0.01
P04920	Anion exchange protein 2	SLC4A2	11	mem		2.04	0.044	0.82	0.441	1.66	<0.01
Q9Y289	Sodium-dependent multivitamin transporter	SLC5A6	14	mem		2.03	<0.01	0.65	0.030	1.31	0.086
P30273	High affinity immunoglobulin epsilon receptor subunit gamma	FCER1G	1	mem		2.02	<0.01	0.71	0.140	1.43	0.141
P08473	Neprilysin	MME	1	mem		2.01	0.097	0.87	0.626	1.74	0.030
P13688	Carcinoembryonic antigen-related cell adhesion molecule 1	CEACAM1	1	mem	extra	2.00	0.014	0.74	0.169	1.48	0.173

B. The list of proteins increased in expression between cancer without and with metastasis (n=10)

Accession	protein name	gene name	TM	GO (mem)	GO (extra)	C/P	p-value	Cm/C	p-value	Cm/P	p-value
Q96HR9	Receptor expression-enhancing protein 6	REEP6	2	mem		1.13	0.651	3.18	0.070	3.61	0.035
P05451	Lithostathine-1-alpha	REG1A	0		extra	0.20	0.164	3.08	<0.01	0.60	0.379
Q8N323	Protein FAM55A	FAM55A	1		extra	0.22	0.102	2.98	0.057	0.65	0.416
O95395	Beta-1,3-galactosyl-O-glycosyl-glycoprotein beta-1,6-N-acetylglucosaminyltransferase 3	GCNT3	1	mem		0.82	0.595	2.85	0.086	2.33	0.089
O95994	Anterior gradient protein 2 homolog	AGR2	0		extra	0.44	0.012	2.56	0.094	1.12	0.727
Q9NRD8	Dual oxidase 2	DUOX2	6	mem		0.43	0.081	2.51	0.045	1.07	0.843
Q8TD06	Anterior gradient protein 3 homolog	AGR3	0		extra	0.51	0.028	2.49	0.017	1.26	0.301
Q09327	Beta-1,4-mannosyl-glycoprotein 4-beta-N-acetylglucosaminyltransferase	MGAT3	1	mem		0.89	0.694	2.24	0.046	1.99	0.024
Q9Y5L3	Ectonucleoside triphosphate diphosphohydrolase 2	ENTPD2	2	mem	extra	1.36	0.369	2.09	0.028	2.83	<0.01
Q8NCC5	Sugar phosphate exchanger 3	SLC37A3	12	mem		0.94	0.838	2.06	0.016	1.92	0.030

Table 3. List of the proteins analyzed by SRM/MRM and their quantitation data using iTRAQ (continued)

C. The list of proteins decreased in expression between polyps and cancer without metastasis (n=13)

Accession	protein name	gene name	TM	GO (mem)	GO (extra)	C/P	p-value	Cm/C	p-value	Cm/P	p-value
A8K7I4	Calcium-activated chloride channel regulator 1	CLCA1	0	mem	extra	0.14	<0.01	0.85	0.673	0.12	<0.01
Q01524	Defensin-6	DEFA6	0		extra	0.18	0.053	0.69	0.428	0.12	0.043
Q9Y6R7	IgGFc-binding protein	FCGBP	0	mem	extra	0.23	<0.01	1.12	0.830	0.25	<0.01
Q6ZMB0	UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 6	B3GNT6	1	mem		0.26	<0.01	1.10	0.713	0.29	<0.01
Q02817	Mucin-2	MUC2	0		extra	0.26	<0.01	1.31	0.537	0.34	<0.01
Q07654	Trefoil factor 3	TFF3	0		extra	0.32	<0.01	1.14	0.651	0.37	<0.01
Q9HC84	Mucin-5B	MUC5B	0		extra	0.36	0.038	1.29	0.352	0.47	0.042
P27216	Annexin A13	ANXA13	0	mem		0.37	<0.01	1.24	0.574	0.46	0.018
P24588	A-kinase anchor protein 5	AKAP5	0	mem		0.43	0.017	1.12	0.606	0.48	0.011
Q7Z3J2	UPF0505 protein C16orf62	C16orf62	0	mem		0.46	<0.01	1.39	0.072	0.65	<0.01
P13727	Bone marrow proteoglycan	PRG2	0		extra	0.48	0.045	0.89	0.695	0.43	0.021
Q6UXG2	UPF0577 protein KIAA1324	KIAA1324	1	mem		0.49	0.051	1.25	0.390	0.61	0.085
Q9Y2J2	Band 4.1-like protein 3	EPB41L3	0	mem		0.49	0.062	1.35	0.086	0.66	0.167

D. The list of proteins decreased in expression between cancer without and cancer with metastasis (n=6)

Accession	protein name	gene name	TM	GO (mem)	GO (extra)	C/P	p-value	Cm/C	p-value	Cm/P	p-value
P08123	Collagen alpha-2(I) chain	COL1A2	0	mem		1.87	0.086	0.34	0.010	0.63	0.213
O75015	Low affinity immunoglobulin gamma Fc region receptor III-B	FCGR3B	1	mem		2.14	0.236	0.34	0.057	0.73	0.673
P02452	Collagen alpha-1(I) chain	COL1A1	0	mem	extra	1.86	0.109	0.36	0.025	0.66	0.252
P02461	Collagen alpha-1(III) chain	COL3A1	0	mem		1.70	0.182	0.39	0.039	0.66	0.152
Q15063	Periostin	POSTN	0	mem		1.57	0.214	0.43	<0.01	0.67	0.406
O43934	UNC93-like protein MFSD11	MFSD11	10	mem		1.27	0.434	0.45	0.067	0.58	0.110

E. The list of proteins increased in expression between polyps and cancer with metastasis (n=10)

Accession	protein name	gene name	TM	GO (mem)	GO (extra)	C/P	p-value	Cm/C	p-value	Cm/P	p-value
P21589	5'-nucleotidase	NT5E	2	mem		1.69	0.104	1.41	0.447	2.39	0.082
Q92968	Peroxisomal membrane protein PEX13	PEX13	0	mem		1.35	0.012	1.73	0.031	2.34	0.012
O43291	Kunitz-type protease inhibitor 2	SPINT2	1	mem	extra	1.63	<0.01	1.39	0.419	2.27	0.087
Q8N4S7	Progestin and adipoQ receptor family member 4	PAQR4	3	mem		1.82	<0.01	1.23	0.342	2.25	0.019
Q8NBM4	Ubiquitin-associated domain-containing protein 2	UBAC2	4	mem		1.95	<0.01	1.14	0.441	2.22	<0.01
Q96CP7	TLC domain-containing protein 1	TLCD1	5	mem		1.33	0.288	1.67	0.226	2.21	0.100
P05546	Heparin cofactor 2	SERPIND1	0		extra	2.15	0.124	0.99	0.978	2.13	0.024
P11166	Solute carrier family 2, facilitated glucose transporter member 1	SLC2A1	12	mem		1.92	<0.01	1.10	0.716	2.11	0.031
Q9BQD7	Protein FAM173A	FAM173A	1	mem		1.55	<0.01	1.36	0.236	2.11	0.050
Q96B21	Transmembrane protein 45B	TMEM45B	5	mem		1.29	0.132	1.61	0.100	2.07	0.018

P-values were calculated by t-test. TM, number of transmembrane domain. C/P, average ratio of cancer without metastasis to polyps. Cm/C, average ratio of cancer with metastasis to cancer without metastasis. Cm/P, average ratio of cancer with metastasis to polyps.

Table 4. SRM/MRM analysis of biomarker candidate proteins

gene name	C / P	p-value	Cm / C	p-value	Cm / P	p-value
FAP	1.59	0.515	2.21	0.052	3.52	0.198
GPRC5A	4.31	0.040	1.30	0.514	5.59	<0.01
CEACAM6	13.41	0.096	0.87	0.822	11.61	<0.01
LRRC15	2.51	0.084	1.83	0.237	4.59	0.037
PRTN3	2.68	0.014	1.67	0.098	4.47	<0.01
CEACAM5	7.29	0.044	0.85	0.737	6.22	<0.01
ITGA11	1.90	0.019	0.82	0.408	1.55	0.066
BST1	1.93	0.012	1.84	0.064	3.55	<0.01
MMP2	1.18	0.601	1.11	0.761	1.30	0.396
ANTXR1	3.23	<0.01	1.08	0.818	3.48	<0.01
PXMP4	2.29	<0.01	0.80	0.385	1.82	0.025
EFEMP1	1.30	0.478	1.04	0.900	1.35	0.358
LTBP2	1.83	0.036	1.10	0.676	2.02	<0.01
SLC2A3	3.56	0.030	0.92	0.817	3.28	<0.01
ITGA5	1.83	<0.01	1.79	0.162	3.28	0.031
MFAP2	1.41	0.394	1.17	0.496	1.65	0.128
TMEM97	2.00	<0.01	0.83	0.411	1.67	0.064
ENPEP	3.83	<0.01	1.13	0.445	4.32	<0.01
CD177	1.17	0.581	1.42	0.224	1.66	0.144
C20orf108	1.23	0.368	0.94	0.823	1.16	0.560
PDGFRB	2.22	<0.01	1.00	0.995	2.22	<0.01
FAM73B	1.22	0.207	0.51	<0.01	0.62	0.013
TSPAN9	1.75	<0.01	0.99	0.968	1.74	<0.01
GPC6	1.89	0.072	1.20	0.614	2.26	0.044
GGT5	2.06	0.034	1.24	0.432	2.56	<0.01
C5AR1	1.48	0.120	1.49	0.167	2.21	0.016
FBN1	1.37	0.443	1.34	0.257	1.84	0.072
FBLN2	1.75	0.102	1.02	0.946	1.79	0.087
SIGMAR1	1.74	<0.01	0.98	0.914	1.71	0.013
MMP14	2.43	<0.01	1.00	0.988	2.42	<0.01
TFRC	2.32	0.018	1.01	0.973	2.35	0.027
MRC2	2.09	<0.01	1.13	0.631	2.36	<0.01
SPARC	2.49	<0.01	0.82	0.317	2.03	0.027
HSPB1	2.73	0.016	1.50	0.231	4.10	<0.01
C8orf55	1.92	<0.01	0.74	0.123	1.42	0.024
FCGR1A	2.47	<0.01	1.40	0.277	3.45	<0.01
THY1	2.14	<0.01	1.00	0.983	2.15	<0.01
TMEM41A	2.04	<0.01	0.90	0.593	1.84	<0.01
SLC4A2	2.41	<0.01	0.92	0.746	2.21	0.014
FCER1G	2.23	<0.01	0.97	0.888	2.17	<0.01
MME	5.21	0.058	0.97	0.959	5.05	0.058
CEACAM1	5.95	0.025	0.83	0.646	4.92	<0.01
REEP6	1.21	0.509	0.97	0.934	1.18	0.608
GCNT3	1.75	0.078	1.46	0.306	2.55	0.063
AGR3	0.20	<0.01	1.89	0.073	0.38	0.021
ENTPD2	1.11	0.800	0.88	0.778	0.98	0.942
CLCA1	0.17	0.022	1.32	0.739	0.22	0.019
FCGBP	0.22	<0.01	1.15	0.782	0.25	<0.01
B3GNT6	0.32	<0.01	1.48	0.359	0.48	0.036
MUC2	0.14	<0.01	2.02	0.279	0.29	0.013
TFF3	0.33		2.80		0.93	
ANXA13	0.23	<0.01	1.41	0.259	0.32	<0.01
AKAP5	0.19	0.016	0.83	0.487	0.16	0.013
C16orf62	0.76	0.442	0.59		0.45	
PRG2	0.34	0.018	1.10	0.744	0.38	0.021
KIAA1324	0.32	<0.01	1.13	0.657	0.36	<0.01
EPB41L3	0.55	0.060	0.67	0.142	0.37	<0.01
COL1A2	1.55	0.438	1.22	0.650	1.90	0.031
COL1A1	1.39	0.590	1.19	0.737	1.65	0.093
COL3A1	1.24	0.642	1.34	0.517	1.67	0.212
POSTN	0.90	0.687	1.94	0.018	1.75	0.033
NT5E	1.05	0.802	1.21	0.473	1.27	0.329
PEX13	1.73	<0.01	0.81	0.224	1.40	0.025
UBAC2	1.87	<0.01	0.95	0.834	1.78	0.044
TLCD1	2.13	<0.01	0.76	0.335	1.63	0.113
SERPIND1	1.57	0.018	1.21	0.371	1.90	<0.01
SLC2A1	2.57	0.175	1.18	0.758	3.03	0.051
FAM173A	1.18	0.433	0.81	0.259	0.96	0.860
TMEM45B	1.35	0.255	0.97	0.918	1.30	0.400

P-values were calculated by t-test.
C/P, average ratio of cancer without metastasis to polyps.
Cm/C, average ratio of cancer with metastasis to cancer without metastasis.
Cm/P, average ratio of cancer with metastasis to polyps.