Published in final edited form as:
J Proteome Res. 2010 August 6; 9(8): 4123–4130. doi:10.1021/pr100302b.

Gastric Cancer-Specific Protein Profile Identified Using Endoscopic Biopsy Samples via MALDI Mass Spectrometry

Hark Kyun Kim[†,‡], **Michelle L. Reyzer**[§], **Il Ju Choi**[‡], **Chan Gyoo Kim**[‡], **Hee Sung Kim**[‡], **Akira Oshima**[†], **Oleg Chertov**[‖], **Simona Colantonio**[‖], **Robert J. Fisher**[‖], **Jamie L. Allen**[§], **Richard M. Caprioli**[§], and **Jeffrey E. Green**[*,†]
Laboratory of Cancer Biology and Genetics, National Cancer Institute, Bethesda, Maryland 20892, National Cancer Center, Goyang, Gyeonggi, Republic of Korea, 410-769, Mass Spectrometry Research Center, Vanderbilt University, Nashville, Tennessee 37232-8575, and Protein Chemistry Laboratory, SAIC-Frederick, Inc., National Cancer Institute at Frederick, Frederick, Maryland 21702

Abstract

To date, proteomic analyses on gastrointestinal cancer tissue samples have been performed using surgical specimens only, which are obtained after a diagnosis is made. To determine if a proteomic signature obtained from endoscopic biopsy samples could be found to assist with diagnosis, frozen endoscopic biopsy samples collected from 63 gastric cancer patients and 43 healthy volunteers were analyzed using matrix-assisted laser desorption/ionization (MALDI) mass spectrometry. A statistical classification model was developed to distinguish tumor from normal tissues using half the samples and validated with the other half. A protein profile was discovered consisting of 73 signals that could classify 32 cancer and 22 normal samples in the validation set with high predictive values (positive and negative predictive values for cancer, 96.8% and 91.3%; sensitivity, 93.8%; specificity, 95.5%). Signals overexpressed in tumors were identified as α-defensin-1, α-defensin-2, calgranulin A, and calgranulin B. A protein profile was also found to distinguish pathologic stage Ia (pT1N0M0) samples (n = 10) from more advanced stage (Ib or higher) tumors (n = 48). Thus, protein profiles obtained from endoscopic biopsy samples may be useful in assisting with the diagnosis of gastric cancer and, possibly, in identifying early stage disease.

Keywords

direct tissue MALDI; gastric cancer; diagnosis

[*]To whom correspondence should be addressed. Jeffrey E. Green, M.D., Laboratory of Cancer Biology and Genetics, National Cancer Institute, Building 37, Room 4054, 37 Convent Dr., Bethesda, MD 20892 (Phone 301-435-5193, Fax 301-496-8709, jegreen@nih.gov)..
[†]National Cancer Institute.
[‡]National Cancer Center.
[§]Vanderbilt University.
[‖]National Cancer Institute at Frederick.

Supporting Information Available: Supplementary Methods and Table 1, which contains peptide sequences of identified proteins. This material is available free of charge via the Internet at http://pubs.acs.org.

Introduction

Gastric cancer is the second most common cause of cancer death worldwide.[1] Yet, reliable biomarkers for the diagnosis of gastric cancer do not exist.[2–4] Currently available serological tumor markers, such as carcinoembryonic antigen (CEA) or carbohydrate antigen 19–9 (CA19–9), are not sensitive and specific enough for the early detection of gastric cancer.[2] Importantly, precancerous lesions are often difficult to differentiate from gastric carcinomas in biopsy samples by conventional histopathologic analysis. In fact, experienced pathologists often disagree in distinguishing invasive carcinoma from high-grade dysplasia in gastroscopic biopsy specimens.[5]

Matrix-assisted laser desorption/ionization mass spectrometry (MALDI MS) has been demonstrated to be useful for direct molecular profiling of common solid tumors.[6–10] In this approach, thin sections of frozen tissues are obtained from surgical resections or biopsies and mass spectra are obtained from discrete locations on the tissue. The resulting spectra are composed primarily of singly charged ions of proteins present in the tissue at the locations sampled. For example, in a study of human glioma samples, protein profiles were found to accurately classify tumor from nontumor tissue and subclassify tumor grades.[9] A profile consisting of 24 unique signals was also found to be an independent classifier of survival, separating patients into a short-term survival group (mean survival <15 months) and a long-term survival group (mean survival >90 months), even after correcting for other confounding factors (tumor grade, patient age, prior history of radiation, etc.). These promising results prompted us to test the feasibility of using direct tissue MALDI MS to define molecular signatures in gastric cancer.

To date, most of the proteomic analyses on gastric cancer tissue samples have been performed using surgically resected tissue.[11–14] Two dimensional (2D)-gel electrophoresis experiments have identified potential biomarkers by comparing tumor tissues and adjacent noncancerous mucosa.[11,12] Other proteomic technologies, such as surface-enhanced laser desorption and ionization (SELDI),[13] have been applied to gastric cancer tissue samples, but only to surgical tissue samples. Deininger et al. reported a MALDI imaging study on a small number of surgical gastric cancer tissue samples, suggesting its utility for biomarker discovery.[14] These prior proteomic studies on gastric cancer, however, have several limitations. First, in prior studies, only surgically resected tissue samples were evaluated and not endoscopic biopsy samples, which are of great clinical relevance for developing diagnostic profiles. Second, the sample size was inadequate to test the classification performance of any detected proteomic profiles in an independent set of samples. This is critical to assess the potential clinical utility of any obtained molecular panel. Finally, in all cases, adjacent nontumor tissue was used as the control as opposed to true cancer-free tissue. It has been demonstrated that, even in histologically normal-appearing tissue adjacent to tumor tissue, molecular changes may be occurring related to tumor proximity.[15] True cancer-free normal samples, obtained from endoscopic biopsy, are better suited for comparative studies.

We, therefore, performed a prospective clinical study to evaluate the feasibility of using direct tissue MALDI MS analysis on endoscopic biopsy samples for gastric cancer. Due to the high specificity with which matrix could be placed on the tissue sections, tumor-rich areas >200 μm could be specifically targeted for analysis. Thus, we found that the histology-directed MALDI MS approach could generate relevant proteomic information from most of those biopsy samples that contained small tumor-rich areas, owing to its high sensitivity. After statistical analysis, the resulting protein profiles were robust enough to accurately classify tumor from normal tissue, as well as distinguish early stage from more advanced stage cancer.

Materials and Methods

Collecting and Processing Clinical Material and Patient Information

Tissues were obtained, with informed consent and institutional review board approval, from patients and volunteers undergoing endoscopic biopsy at National Cancer Center in Korea from 2005 to 2008. Samples were collected using biopsy forceps at the time of diagnostic gastroscopy, flash frozen in liquid nitrogen, and stored at −80 °C until analysis. Samples were prepared for MALDI analysis as described previously.[16] Briefly, thin (12 μm) sections were obtained from the frozen tissues with a cryostat (Leica CM 3050S, Leica Microsystems Inc., Bannockburn, IL). Two serial sections were obtained from each tissue. One section was affixed to a standard glass slide, stained with hematoxylin and eosin (H&E), and an optical image was acquired via a microscope attached to a digital camera. The other section was thaw-mounted onto a gold-coated stainless steel MALDI plate and washed with graded ethanol solutions (70, 90, 95% ethanol for 30 s each) for subsequent mass spectral analysis.

The samples were processed as described[11] for histology-directed protein profiling (Figure 1). Briefly, the optical image of the H&E stained serial section was evaluated by a pathologist, who digitally marked the picture at discrete locations with 200 μm diameter circles. These circles were intended to cover areas of the tissue enriched with at least 75% of a particular cell type, that is, normal epithelial cells or tumor cells. The image of the H&E section was then overlaid with an image of the serial ethanol fixed section in Photoshop (Adobe Systems Inc., San Jose, CA) to align features of the two serial sections. Distinct *x,y*-coordinates were obtained from each spot and imported into a robotic device for automated matrix deposition.

The robotic device used in this study is a prototype acoustic reagent multispotter (ARM; LabCyte, Sunnyvale, CA) that uses focused acoustic energy to eject matrix droplets onto a target.[17] Sinapinic acid was used as the MALDI matrix and prepared as a 20 mg/mL solution in 50:50 acetonitrile: 0.1% trifluoroacetic acid (TFA). Matrix was deposited in cycles of 13 drops/spot, dispensed at 10 Hz at each designated coordinate. A total of 6 cycles were found to provide optimal analyte extraction and matrix crystals. Thus, a total volume of ~9 nL matrix was deposited per spot, resulting in dried crystal spots of ~200 μm diameter.

Mass spectra were acquired using an Autoflex II (Bruker Daltonics, Billerica, MA) time-of-flight mass spectrometer equipped with a SmartBeam laser (Nd:YAG, 355 nm) and run using a linear-mode acquisition method optimized for 2–40 kDa. Data was acquired in an automated fashion from each discrete matrix spot, with a total of 384 laser shots acquired via random walk over the entire spot for each mass spectrum.

Data Processing and Statistical Analysis

All mass spectra were converted into text files and imported into ProTS Data (Biodesix, Broomfield, CO) for baseline correction, normalization by total ion current, and realignment/recalibration of individual spectra. ProTS software further binned the processed spectra into 236 bins (features) that were used for statistical evaluation of the samples. Binned data was $\log_2$-transformed, median normalized, and subjected to statistical analysis using BRB-ArrayTools (NCI, version 3.6).[18] A principal component analysis (PCA) plot was generated using multidimensional scaling analysis of BRB-ArrayTools, which graphically represents Euclidean distances among samples without forcing the samples into specific clusters. The three primary principal components were used as the axes for the 3-dimensional scaling representation.

Class comparison and class prediction analyses were also performed using BRB-ArrayTools. The class comparison analysis computes a Student *t*-test for each peak and lists peaks differentially expressed among the classes at the statistical significance level selected (typically, a feature selection *P* cutoff of 0.01). Then, it performs random permutations of the class labels (diffuse vs intestinal type, for example). For each random permutation, all of the *t*-tests are recomputed for each peak. The class comparison tool computes the proportion of the random permutations that gave as many peaks significant at the selected level of significance (typically, $P = 0.01$) as were found by comparing the true class labels. Protein profiles of the classes were considered different if this probability (designated as *Permutation P value*) was calculated to be less than 0.05.

For diagnosing cancer, the class prediction was performed using an option of developing the classifiers over a grid of significance levels for peak selection. The 0.632+ bootstrap cross-validated misclassification rate was computed for each significance level in the grid and for all classifier functions (Compound Covariate Predictor, Diagonal Linear Discriminant Analysis, Nearest Neighbor Predictor, Nearest Centroid Predictor, and Support Vector Machine Predictor) in the training set. A classifier with the smallest average cross-validated misclassification rate in the training set was chosen to predict which of 2 classes (cancer vs normal) each sample in the validation set belongs to. For early stage disease identification, the same class prediction was performed using a feature selection $P < 0.01$ and including all samples as a training set.

Protein Identification

To obtain a sufficient amount of protein for identification, a surgically removed gastric cancer specimen (not used in the protein profiling analysis) was used for peak identification. Protein was extracted from tumor and adjacent normal tissue of this specimen using hypotonic saline, and fractionated by high-performance liquid chromatography (HPLC). Aliquots of HPLC fractions were analyzed by MALDI-Time of Flight (TOF) MS using sinapinic acid as matrix. Fractions of interest were selected for further identification if they contained an m/z value that matched one of the discriminatory m/z values found from the statistical analysis. Lyophilized aliquots of some of the fractions were additionally fractionated by reverse phase HPLC. For some fractions, MS/MS was performed directly on the ion of interest using a MALDI TOF/TOF instrument. Aliquots of other fractions were analyzed using an automated protein sequencer. To identify more proteins, we also performed LC–MS/MS analyses of the HPLC fractions (treated with trypsin) and then compared the theoretical masses of intact proteins identified in particular fractions with MALDI-TOF MS data of these fractions (Supporting Information for details).

Results

Data Acquisition

This study is the result of direct tissue MALDI analysis for 106 tissue samples that were collected from 63 gastric cancer patients and 43 healthy volunteers. Thirty-one additional cancer patient samples (33%) had to be excluded from this analysis, because of inadequate tumor-rich area in the biopsy samples ($n = 27$), poor matrix spot placement ($n = 2$), or poor spectral quality ($n = 2$). Table 1 summarizes the clinical characteristics of 63 gastric cancer patients whose samples are analyzed in this study. The majority ($n = 35$; 55.5%) of patients had metastatic disease, most of whom received fluorouracil-based chemotherapy after biopsy. There were 10 patients at pathologic stage Ia (pT1N0M0) (15.9%). The median age of healthy volunteers was 47 years (interquantile range, 44–54) with 23 (53.5%) males.

Mass spectra were acquired on individual spots for each tissue section, and these spectra were averaged together after preprocessing to create one average spectrum per patient. The average spectra are composed of 2–19 individual measurements for cancer samples (with a median value of 7), and 2–17 individual measurements for normal samples (with a median value of 8). The individual measurements are averaged to minimize intrasample variability. An example of the average spectra is shown in Figure 2, for one representative cancer and normal sample. As shown, there are numerous signals in the m/z range 2000–20 000. Indeed, postspectral processing identified 236 features across the entire mass range for all of the samples studied. The expanded mass view from m/z 3000–5500 highlights some of the more dramatic differences observed between the groups.

For statistical analysis, the spectra were exported into bins, with each bin designed to contain one peak. In practice, the bins sometimes contained shoulders where peaks were not baseline-resolved or where the mass alignment was poor. Nonetheless, the integrated areas of the bins were used for further statistical evaluation of the samples.

Cancer Patients versus Healthy Volunteers

A PCA plot graphically demonstrates that cancer patient samples and volunteer samples are separately clustered in an unsupervised analysis (Figure 3). When a class comparison analysis was performed using BRB-ArrayTools, the proportion of the random permutations that gave as many significant peaks at a feature selection of $P < 0.01$ as were found by comparing the true class labels (cancer vs normal) was less than 0.001, strongly suggesting that the cancer and normal tissue samples are significantly different in their protein profiles. Class prediction analysis was subsequently performed after dividing the entire set of samples into two groups based upon the chronological order of patient enrollment. The first half of the samples (31 cancer and 21 normal samples) was used as a training set to develop predictors for cancer diagnosis. In the training set, Support Vector Machine (SVM) classifier composed of 73 signals significantly different between the classes at the 0.01 significance level (Table 2) performed best among all of the tested classifiers, with the lowest average cross-validated misclassification rate (3.4%). This SVM predictor was then applied to predict the class of 32 cancer and 22 normal samples in the validation set (the latter half). In this validation set of 54 samples, positive and negative predictive values for cancer were 96.8% and 91.3%, respectively (sensitivity, 93.8%; specificity, 95.5%).

Identification of Early Stage Disease

Next, we evaluated whether protein profiles could distinguish biologic features of the gastric cancer. We were interested in the identification of the American Joint Committee on Cancer (AJCC) pathologic stage Ia (pT1N0M0) lesions, which are potential candidates for nonsurgical management. According to the class comparison analysis, AJCC stage Ia samples (pT1N0M0; $n = 10$) were significantly different from more advanced stage tumors (AJCC stage Ib or higher; $n = 48$) in their protein profiles. A total of 17 peaks were significantly different between these 2 groups using the Student t-test performed at a feature selection of $P < 0.01$ (Table 3). The probability of getting at least 17 significant peaks by chance (at the $p = 0.01$ level) if there are no real differences between the classes was 0.014. Given that this permutation P value of a global test is less than 0.05, this suggests that stage Ia exhibits a different protein profile from more advanced stage samples. The permutations P value was consistently less than 0.05, across P cutoffs for feature selection ranging from 0.001 to 0.05.

We also performed a class prediction analysis to calculate permutation P values for a cross-validated misclassification rate. For this analysis, all 58 samples were used in the training set, since the small number of stage Ia samples did not allow us to withhold some of them in

the model building process as a separate set. When we performed the 0.632+ bootstrap cross-validation at feature selection $P < 0.01$, the average cross-validated misclassification rate ranged from 15 to 23% across several algorithms (Table 4). The probabilities of getting these small misclassification rates in random data sets were estimated to be less than 0.05 with the majority of classifiers tested (Table 4). These further suggest that early (Ia) stage gastric cancer has a distinct protein profile from more advanced stage lesions.

Histopathologic Classification and Outcome Prediction

Class comparison was performed between Lauren intestinal ($n = 25$) and diffuse types ($n = 38$) of gastric cancer, which represent two major histopathologic subtypes of gastric cancer. At a feature selection $P < 0.01$, 8 peaks were significantly different between these 2 classes (Table 5). The probability of getting at least 8 peaks significant by chance (at feature selection $P < 0.01$), if there are no real differences between two classes, was estimated to be 0.038, suggesting that two histological subtypes are different in protein profile.

Finally, we asked if protein profiles could distinguish metastatic gastric cancer patients who responded to chemotherapy (partial response according to Response Evaluation Criteria in Solid Tumors (RECIST); $n = 12$) from those who did not (stable or progressive disease; $n = 15$), although the implication of such an analysis is limited given the heterogeneity of chemotherapy regimens. There was no difference in protein profile between good responders and poor responders to chemotherapy (permutation P value = 0.70).

Protein Identification

Discriminatory protein identification was performed using a gastric cancer surgical tissue as described in Materials and Methods. Several signals overexpressed in the tumors include α-defensin-1 ($m/z = 3439$), α-defensin-2 ($m/z = 3368$), calgranulin A ($m/z = 10\,840$), and two forms of calgranulin B ($m/z = 13\,158$ and $12\,694$) (Figure 4, Table 6, and Supplementary Table 1, Supporting Information). Signals underexpressed in tumors were identified as lysozyme C ($m/z = 14\,697$), C-terminal fragment (res 149–175) of anterior gradient protein 2 homologue ($m/z = 2968$), and N-terminal fragment (res 1–70) of histone H2B ($m/z = 7767$).

Discussion

This study is the first work demonstrating that a protein profile obtained from endoscopic biopsy samples via MALDI mass spectrometry can differentiate cancerous gastroscopic samples from normal samples. This technology is capable of generating discriminatory protein profiles from a single section of an endoscopic biopsy piece, suggesting its potential clinical applicability. Notably, with 106 total samples analyzed, this study has considerably more statistical power compared to previous proteomic studies performed using smaller numbers of gastrointestinal tract cancer tissue samples.[7–10] Gastroscopic biopsy tissue samples, especially in diffuse type gastric cancer patients, often contain only small nests of scattered tumor cells, posing a technical challenge for molecular research. Using direct tissue MALDI MS, we could successfully profile almost all the samples that were collected from gastric cancer patients in an unbiased way, except for samples containing practically no tumor cells. Due to the small amount of tissue required for MALDI analysis compared to conventional proteomics technologies, therefore, this approach may provide a clear advantage for initial diagnosis and management of gastrointestinal tract cancer patients. Given its high predictive values, the protein profile identified by this study could also possibly differentiate gastric carcinoma from premalignant lesions in biopsy samples, which needs to be evaluated in future studies specifically designed to address this question.

Signals subsequently identified as alpha-defensin-1 and alpha-defensin-2 were included in the best SVM predictor developed in our training set samples. α-Defensin is expressed in neutrophils and also in the mucosal epithelia of intestine, respiratory tract, and urinary tract. It is found to be expressed in a variety of tumors[13,19]and proposed as a tumor biomarker.[13,19,20] Mohri et al. reported that α-defensin is one of most significantly elevated SELDI peaks in 21 gastric cancer tissue extract samples, compared with paired adjacent normal mucosa samples.[13] These investigators also demonstrated that cancer cells in 5 gastric cancer tissue samples are strongly positive for α-defensin by immunohistochemistry, while matched adjacent normal tissues are negative.[13] Recently, gastric fluid α-defensins have also been reported to be upregulated in gastric cancer patients compared to patients with benign disease.[21] While calgranulins A/B (S100A8/9) are pro-apoptotic molecules produced by immune cells, S100A8/A9 expression in cancer cells has also been associated with tumor development, cancer invasion, or metastasis.[22] S100A8/9 are required for malignant progression including invasion, migration, and proteinase expression in SNU484 human gastric cancer cells.[23]

Interestingly, protein profiles obtained from early stage (pT1N0M0) gastric cancer samples were significantly different from those obtained from more advanced stage tumors. Patients with gastric cancers without lymph node metastasis (pT1N0M0) have an excellent prognosis in gastric cancer, but the prognosis worsens once the tumor invades the proper muscle layer or involves lymph nodes.[24,25] Currently, assessment of nodal involvement of T1 lesions primarily depends upon endoscopic ultrasonography or computed tomography, but none of these imaging modalities can reliably confirm or exclude the presence of lymph node metastases in gastric cancer patients.[26] Prediction of a stage Ia (pT1N0M0) gastric cancer is clinically important, since this lesion is a potential candidate for endoscopic treatment.[27] Although this study was not designed to directly compare pT1N0M0 and pT1N1M0 lesions, we have shown that pT1N0M0 gastric cancers likely have a distinct protein profile compared to more advanced stage tumor samples. Thus, direct tissue MALDI MS may be potentially useful in identifying early stage disease, thereby helping guide disease management. This exciting possibility needs to be validated by more focused studies containing a larger number of stage Ia lesions in the future.

In summary, this study demonstrates that direct tissue MALDI MS analysis on endoscopic biopsy samples may be useful in assisting with the diagnosis of gastric cancer, and, possibly, in identifying early stage disease. The protein profile identified in this study demonstrated a high predictive power for cancer detection. Importantly, the amount of tissue required for these analyses is much smaller than any other available method using molecular profiling techniques, such as array-based gene expression profiling. Therefore, we conclude that direct tissue MALDI MS is a feasible approach to aid in the evaluation of gastric cancer lesions by endoscopy.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported in part by National Institute of Health Intramural Program, Center for Cancer Research, National Cancer Institute; by Korean Ministry of Education, Science, and Technology Grant 0931480; and by Korean National Cancer Center Grant 0910570. This project has been funded in whole or in part with federal funds from the National Cancer Institute, National Institutes of Health, under Contract No. HHSN261200800001E.

References

(1). Alberts S, Cervantes A, van de Velde C. Gastric cancer: epidemiology, pathology, and treatment. Ann. Oncol. 2003; 14:31–6.

(2). Carpelan-Holmstrom M, Louhimo J, Stenman UH, Alfthan H, Haglund C. CEA, CA19–9 and CA72–4 improve the diagnostic accuracy in gastrointestinal cancers. Anticancer Res. 2002; 22:2311–6. [PubMed: 12174919]

(3). Ohuchi N, Takahashi K, Matoba N, Sato T, Taira Y, Sakai N, Masuda M, Mori S. Comparison of serum assays for TAG-72, CA19–9 and CEA in gastrointestinal carcinoma patients. Jpn. J. Clin. Oncol. 1989; 19(3):242–8. [PubMed: 2810823]

(4). Ychou M, Duffour J, Kramar A, Gourgou S, Grenier J. Clinical significance and prognostic value of CA72–4 compared with CEA and CA19–9 in patients with gastric cancer. Dis. Markers. 2000; 16(3–4):105–10. [PubMed: 11381189]

(5). Sarela AI, Scott N, Verbeke CS, Wyatt JI, Dexter SP, Sue-Ling HM, Guillou PJ. Diagnostic variation and outcome for high-grade gastric epithelial dysplasia. Arch. Surg. 2005; 140(7):644–9. [PubMed: 16027328]

(6). Caprioli RM, Farmer TB, Gile J. Molecular imaging of biological samples: Localization of peptides and proteins using MALDI-TOF MS. Anal. Chem. 1997; 69:4751–60. [PubMed: 9406525]

(7). Stoeckli M, Chaurand P, Hallahan DE, Caprioli RM. Imaging mass spectrometry: A new technology for the analysis of protein expression in mammalian tissues. Nature Med. 2001; 7:493–6. [PubMed: 11283679]

(8). Yanagisawa K, Shyr Y, Xu BJ, Massion PP, Larsen PH, White BC, Roberts JR, Edgerton M, Gonzalez A, Nadaf S, Moore JH, Caprioli RM, Carbone DP. Proteomic patterns of tumour subsets in non-small-cell lung cancer. Lancet. 2003; 362(9382):433–9. [PubMed: 12927430]

(9). Schwartz SA, Weil RJ, Thompson RC, Shyr Y, Moore JH, Toms SA, Johnson MD, Caprioli RM. Proteomic-based prognosis of brain tumor patients using direct-tissue matrix-assisted laser desorption ionization mass spectrometry. Cancer Res. 2005; 65(17):7674–81. [PubMed: 16140934]

(10). Cornett DS, Mobley JA, Dias EC, Andersson M, Arteaga CL, Sanders ME, Caprioli RM. A novel histology-directed strategy for MALDI-MS tissue profiling that improves throughput and cellular specificity in human breast cancer. Mol. Cell. Proteomics. 2006; 5:1975–83. [PubMed: 16849436]

(11). Chen CD, Wang CS, Huang YH, Chien KY, Liang Y, Chen WJ, Lin KH. Overexpression of CLIC1 in human gastric carcinoma and its clinicopathological significance. Proteomics. 2007; 7(1):155–67. [PubMed: 17154271]

(12). Li W, Li JF, Qu Y, Qin JM, Gu QL, Yan M, Zhu ZG, Liu BY. Comparative proteomics analysis of human gastric cancer. World J. Gastroenterol. 2008; 14(37):5657–64. [PubMed: 18837081]

(13). Mohri Y, Mohri T, Wei W, Qi YJ, Martin A, Miki C, Kusunoki M, Ward DG, Johnson PJ. Identification of macrophage migration inhibitory factor and human neutrophil peptides 1–3 as potential biomarkers for gastric cancer. Br. J. Cancer. 2009; 101(2):295–302. [PubMed: 19550422]

(14). Deininger SO, Ebert MP, Fütterer A, Gerhard M, Röcken C. MALDI imaging combined with hierarchical clustering as a new tool for the interpretation of complex human cancers. J. Proteome Res. 2008; 7(12):5230–6. [PubMed: 19367705]

(15). Herring KD, Oppenheimer SR, Caprioli RM. Direct tissue analysis by matrix-assisted laser desorption ionization mass spectrometry: application to kidney biology. Semin. Nephrol. 2007; 27(6):597–608. [PubMed: 18061842]

(16). Schwartz SA, Reyzer ML, Caprioli RM. Direct tissue analysis using matrix-assisted laser desorption/ionization mass spectrometry: practical aspects of sample preparation. J. Mass Spectrom. 2003; 38:699–708. [PubMed: 12898649]

(17). Aerni HR, Cornett DS, Caprioli RM. Automated acoustic matrix deposition for MALDI sample preparation. Anal. Chem. 2006; 78:827–34. [PubMed: 16448057]

(18). Simon R, Lam A, Li MC, Ngan M, Menenzes S, Zhao Y. Analysis of gene expression data using BRB-Array Tools. Cancer Inf. 2007; 2:11–7.

(19). Droin N, Hendra J-B, Ducoroy P, Solary E. Human defensins as cancer biomarkers and antitumour molecules. J. Proteomics. 2009; 72(6):918–27. [PubMed: 19186224]

(20). Alahou A, Schellhammer PF, Mendrinos S, Patel K, Kondylis FI, Gong L, Nasim S, Wright GL Jr. Development of a novel proteomic approach for the detection of transitional cell carcinoma of the bladder in urine. Am. J. Pathol. 2001; 158(4):1491–1502. [PubMed: 11290567]

(21). Kon OL, Yip TT, Ho MF, Chan WH, Wong WK, Tan SY, Ng WH, Kam SY, Eng AK, Ho P, Viner R, Ong HS, Kumarasinghe MP. The distinctive gastric fluid proteome in gastric cancer reveals a multi-biomarker diagnostic profile. BMC Med. Genomics. 2008; 1:54–67. [PubMed: 18950519]

(22). Ghavami S, Chitayat S, Hashemi M, Eshraghi M, Chazin WJ, Halayko AJ, Kerkhoff C. S100A8/A9: a Janus-faced molecule in cancer therapy and tumorgenesis. Eur. J. Pharmacol. 2009; 625(1–3):73–83. [PubMed: 19835859]

(23). Yong HY, Moon A. Roles of calcium-binding proteins, S100A8 and S100A9, in invasive phenotype of human gastric cancer cells. Arch. Pharm. Res. 2007; 30(1):75–81. [PubMed: 17328245]

(24). Tachibana M, Takemoto Y, Monden N, Nakashima Y, Kinugasa S, Dhar DK, Kotoh T, Kubota H, Kohno H, Nagasue N. Clinicopathological features of early gastric cancer: Results of 100 cases from a rural general hospital. Eur. J. Surg. 1999; 165:319–25. [PubMed: 10365832]

(25). Wang CS, Hsueh S, Chao TC, Jeng LB, Jan YY, Chen SC, Hwang TL, Chen PC, Chen MF. Prognostic study of gastric cancer without serosal invasion: reevaluation of the definition of early gastric cancer. J. Am. Coll. Surg. 1997; 185(5):476–80. [PubMed: 9358093]

(26). Kwee RM, Kwee TC. Imaging in assessing lymph node status in gastric cancer. Gastric Cancer. 2009; 12(1):6–22. [PubMed: 19390927]

(27). Gotoda T, Yamamoto H, Soetikno RM. Endoscopic submucosal dissection of early gastric cancer. J. Gastroenterol. 2006; 41:929–42. [PubMed: 17096062]

(28). Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc., Ser. B: Stat. Methodol. 1995; 57:289–300.

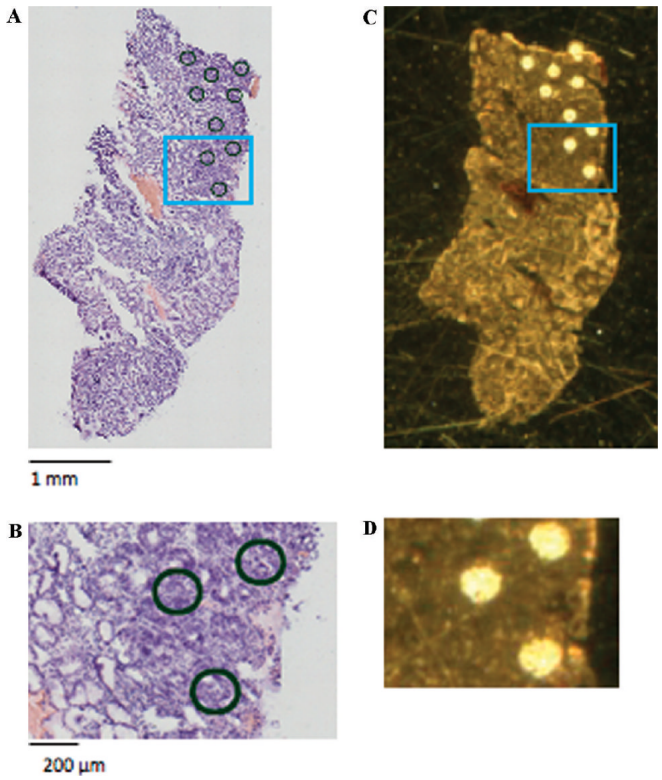

Figure 1.
General procedure for preparing sections for MALDI analysis. (A) Optical image of an H&E section from a cancer endoscopic biopsy tissue cut serially to the section shown in (C). The optical section was evaluated by a pathologist, and marked at discrete locations (enriched in tumor cells) with 200 μm diameter circles. (C) Optical image of the serial section with matrix applied at the locations marked in (A). A total of ~9 nL sinapinic acid was applied at each spot, resulting in dried matrix spots of ~200 μm diameter. (B) and (D) are magnified areas of the H&E stained section and MALDI section, respectively.

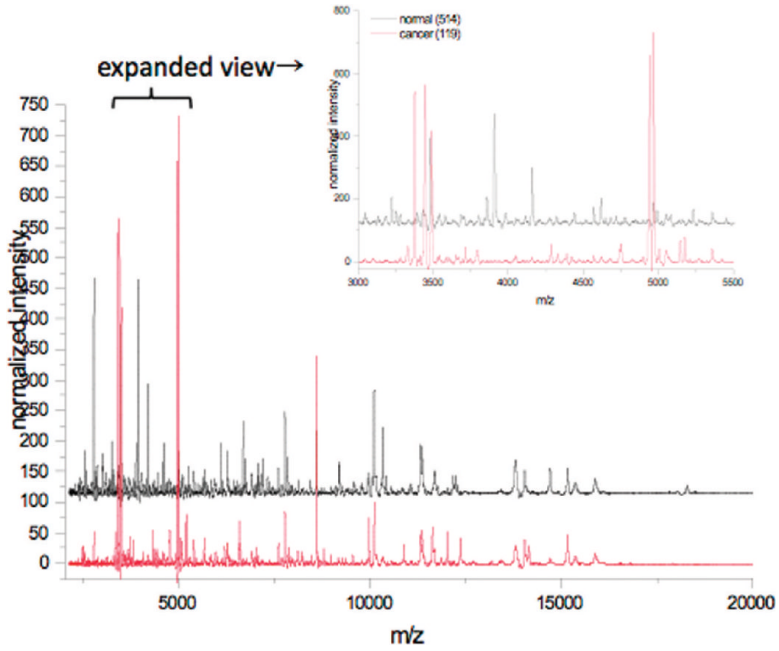

Figure 2.

Representative average mass spectra from one normal patient (black) and one cancer patient (red). The area from *m*/*z* 3000–5500 has been expanded to highlight differences. The normal spectrum is an average of 8 independent measurements; the cancer spectrum is an average of 9 independent measurements.

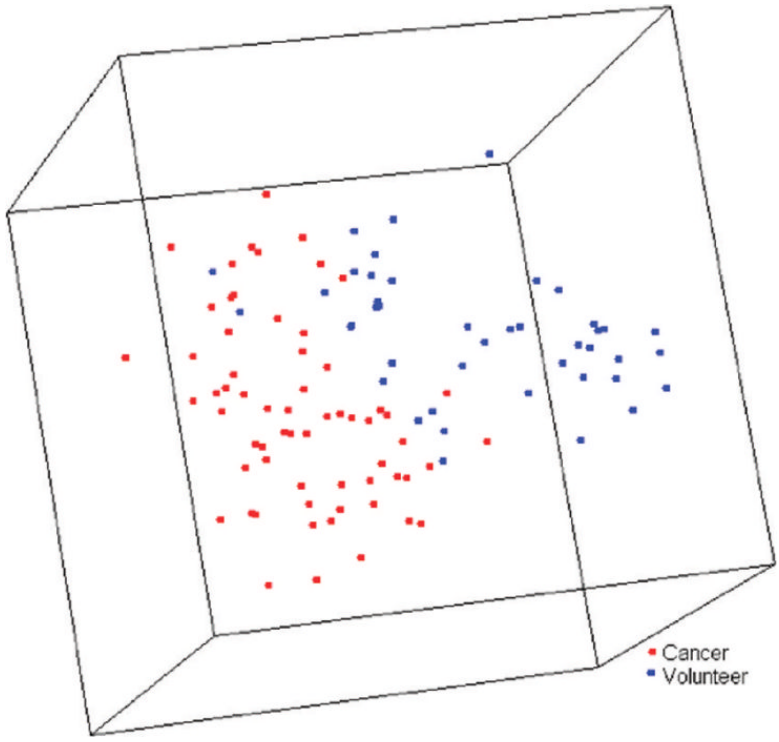

Figure 3.

Principal component analysis plot for 63 gastric cancer patient samples (red) and 43 samples from healthy volunteers (blue), which graphically represents Euclidean distances among samples. Each sphere represents a single sample, and samples whose protein expression profiles are very similar are shown close together.

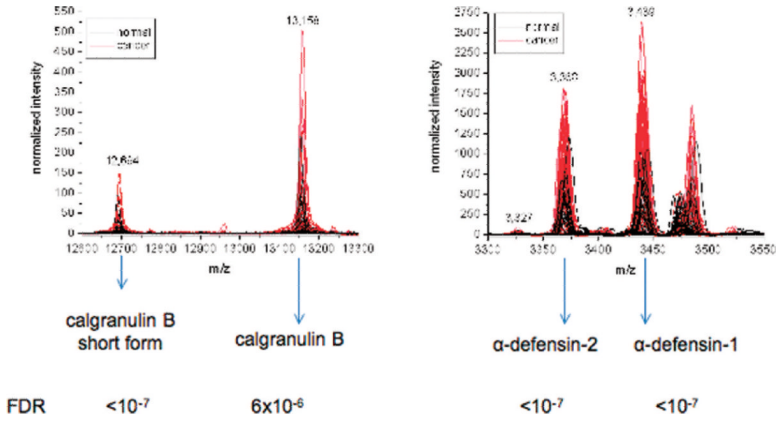

Figure 4.

Intensity profile for identified discriminatory signals among 63 cancer patient samples (red) and 43 samples from healthy volunteers (black). The false discovery rate (FDR) was estimated as the proportion of the signal with univariate *P* values less than or equal to the *P* value of each identified signal, using the method of Benjami and Hochberg.[28]

Table 1

Clinical Characteristics of Patients[a]

characteristics	no patients (%)
Age (years)	
Median	62
Range	29–82
Sex-no (%)	
Male	36 (57.1%)
Female	27 (42.9%)
ECOG PS-no (%)	
0–1	54 (85.7%)
2	6 (9.5%)
3	3 (4.8%)
Histological type (Lauren)-no (%)	
Intestinal	25 (39.7%)
Diffuse	38 (60.3%)
Location of primary lesion-no (%)	
Upper 1/3	15 (23.8%)
Middle 1/3	12 (19%)
Lower 1/3	28 (44.4%)
Entire stomach	8 (12.7%)
Clinical stage (AJCC)	
Ia	10 (15.9%)
Ib	7 (11.1%)
II	6 (9.5%)
IIIa	3 (4.8%)
IIIb	2 (3.2%)
IV	35 (55.5%)
Pathologic stage (AJCC)	
Ia	10 (15.9%)
Ib	7 (11.1%)
II	4 (6.3%)
IIIa	1 (1.6%)
IIIb	1 (1.6%)
IV	35 (55.5%)
Unknown	5 (7.9%)
Primary treatment	
Surgical resection	23 (36.5%)
Chemotherapy	32 (50.8%)
Refusal to treat	8 (12.7%)
Number of metastatic sites	
0	28 (44.4%)

characteristics	no patients (%)
1	18 (28.6%)
2	11 (17.5%)
3	4 (6.3%)
4	2 (3.2%)

[a]ECOG PS, Eastern Cooperative Oncology Group Performance Status. AJCC, American Joint Committee on Cancer staging system (6th edition).

Table 2

Signals Differentially Expressed between Cancer Patient and Volunteer Samples of the Training Set, at Feature Selection $P < 0.01$ (Listed in Order of Increasing P Value)

overexpressed in cancer			underexpressed in cancer		
m/z	FC[a]	*P*	*m/z*	FC[a]	*P*
8454	8.62	<0.001	3902	0.12	<0.001
4224	4.46	<0.001	3216	0.09	<0.001
10 881	4.79	<0.001	8115	0.14	<0.001
11 609	3.21	<0.001	2968	0.13	<0.001
10 840	5.01	0.0001	6720	0.23	<0.001
3327	2.53	0.0001	2789	0.36	<0.001
11 077	2.47	0.0001	14 697	0.14	<0.001
6174	3.14	0.0001	18 025	0.18	<0.001
8093	2.92	0.0002	2830	0.26	<0.001
6547	2.18	0.0004	7348	0.19	<0.001
5825	1.65	0.0004	6669	0.24	<0.001
12 694	3.61	0.0006	9165	0.27	0.0001
5418	2.85	0.0009	9188	0.25	0.0001
4324	2.80	0.0009	4843	0.25	0.0002
3594	3.32	0.0015	3247	0.17	0.0002
13 158	4.97	0.0017	18 269	0.16	0.0003
3439	4.45	0.0020	12 159	0.28	0.0004
5205	1.66	0.0021	6076	0.40	0.0004
3707	3.58	0.0023	5905	0.29	0.0005
12 654	1.89	0.0024	2751	0.21	0.0008
4936	3.71	0.0026	7437	0.29	0.0008
3368	4.05	0.0025	6491	0.27	0.0009
4044	1.81	0.0026	3976	0.36	0.0010
6578	3.05	0.0030	7835	0.18	0.0010
10 275	1.80	0.0039	3854	0.23	0.0010
4134	3.12	0.0054	7143	0.44	0.0014

overexpressed in cancer			underexpressed in cancer		
m/z	FC[a]	*P*	*m/z*	FC[a]	*P*
14 425	2.45	0.0057	9552	0.22	0.0016
8657	1.63	0.0069	3001	0.48	0.0017
3656	2.25	0.0074	3040	0.44	0.0017
4383	1.57	0.0078	18 434	0.47	0.0018
			7278	0.46	0.0020
			6693	0.36	0.0021
			4710	0.34	0.0023
			2425	0.57	0.0029
			14 916	0.19	0.0035
			4152	0.24	0.0039
			12 229	0.22	0.0048
			17 481	0.46	0.0050
			20 265	0.48	0.0058
			6139	0.63	0.0063
			2708	0.58	0.0066
			14 635	0.41	0.0078
			5226	0.43	0.0079

[a]FC, fold change (cancer patient to volunteer).

Table 3

Signals Differentially Expressed between Stage Ia Tumors ($n = 10$) and More Advanced (Ib-IV) Stage Tumors ($n = 48$) at $P < 0.01$

overexpressed in stage Ia			underexpressed in stage Ia		
m/z	FC[a]	*P*	*m/z*	FC[a]	*P*
6693	16.27	<0.001	13 158	0.15	<0.001
4988	2.67	0.0024	5418	0.31	<0.001
9520	2.82	0.0031	10 840	0.29	0.0012
8093	2.90	0.0042	4324	0.35	0.0019
6720	3.00	0.0054	12 694	0.30	0.0033
8115	4.07	0.0075	5535	0.48	0.0052
5205	1.82	0.0087	11 043	0.44	0.0059
9929	1.87	0.0097	10 881	0.41	0.0092
9188	2.76	0.01			

[a]FC, fold change (pathologic stage la to advanced stage).

Table 4

Proportion of the Random Permutations that Gave as Small 0.632+ Bootstrap Cross-Validated Misclassification Rate as Could Be Obtained with the True Class Labels (Stage Ia *vs* More Advanced (Ib-IV) Stage Tumors)

classifier	average cross-validated misclassification rate	permutation *P* value
Compound Covariate Predictor	0.23	0.021
Diagonal Linear Discriminant Analysis	0.23	0.034
1-Nearest Neighbor Predictor	0.15	0.003
3-Nearest Neighbor Predictor	0.15	0.006
Nearest Centroid Predictor	0.21	0.009
Support Vector Machine Predictor	0.20	0.114

Table 5

Signals Differentially Expressed between Intestinal and Diffuse Histological Types at $P < 0.01$

overexpressed in intestinal type			overexpressed in diffuse type		
m/z	FC[a]	*P*	*m/z*	FC[a]	*P*
4936	3.98	0.0003	10 796	0.41	0.0083
3142	1.94	0.0004	11 043	0.57	0.0086
8093	2.36	0.0011			
6693	3.34	0.0060			
4963	3.22	0.0078			
8774	1.50	0.0096			

[a]FC, fold change (intestinal to diffuse).

J Proteome Res. Author manuscript; available in PMC 2012 September 13.

Table 6

Protein Identification

m/z	accession	name
Overexpressed in cancer		
3368	P59665	Alpha-defensin-2
3439	P59665	Alpha- defensin-1
10 840	P05109	S100-A8, Calgranulin A
13 158	P06702	S100-A9, Calgranulin B
12 694	P06702	S100-A9, Calgranulin B, short form
Underexpressed in cancer		
2968	O95994	Anterior gradient protein 2 homologue, C-terminal (res 150–175)
2830	O95994	Anterior gradient protein 2 homologue, C-terminal (res 151–175)
14 697	P61626	Lysozyme C