An Informatics-assisted Label-free Approach for Personalized Tissue Membrane Proteomics: Case Study on Colorectal Cancer*

Chia-Li Han†, Jinn-Shiun Chen§‖, Err-Cheng Chan¶‖, Chien-Peng Wu‡, Kun-Hsing Yu**, Kuei-Tien Chen¶¶, Chih-Chiang Tsou††, Chia-Feng Tsai‡‡, Chih-Wei Chien§§, Yung-Bin Kuo¶¶, Pei-Yi Lin†, Jau-Song Yu||, Chuen Hsueh|||, Min-Chi Chen§, Chung-Chuan Chan§§, Yu-Sun Chang|||, and Yu-Ju Chen‡‡

We developed a multiplexed label-free quantification strategy, which integrates an efficient gel-assisted digestion protocol, high-performance liquid chromatography tandem MS analysis, and a bioinformatics alignment method to determine personalized proteomic profiles for membrane proteins in human tissues. This strategy provided accurate (6% error) and reproducible (34% relative S.D.) quantification of three independently purified membrane fractions from the same human colorectal cancer (CRC) tissue. Using CRC as a model, we constructed the personalized membrane protein atlas of paired tumor and adjacent normal tissues from 28 patients with different stages of CRC. Without fractionation, this strategy confidently quantified 856 proteins (≥2 unique peptides) across different patients, including the first and robust detection (Mascot score: 22,074) of the well-documented carciinoembryonic antigen 5 by a discovery-type proteomics approach. Further validation of a panel of proteins, annexin A4, neutrophil defensin A1, and claudin 3, confirmed differential expression levels and high occurrences (48–70%) in 60 CRC patients. The most significant discovery is the overexpression of stomatin-like 2 (STOML2) for early diagnostic and prognostic potential. Increased expression of STOML2 was associated with decreased CRC-related survival; the mean survival period was 34.77 ± 2.03 months in patients with high STOML2 expression, whereas 53.67 ± 3.46 months was obtained for patients with low STOML2 expression. Further analysis by ELISA verified that plasma concentrations of STOML2 in early-stage CRC patients were elevated as compared with those of healthy individuals (p < 0.001), suggesting that STOML2 may be a noninvasive serological biomarker for early CRC diagnosis. The overall sensitivity of STOML2 for CRC detection was 71%, which increased to 87% when combined with CEA measurements. This study demonstrated a sensitive, label-free strategy for differential analysis of tissue membrane proteome, which may provide a roadmap for the subsequent identification of molecular target candidates of multiple cancer types. Molecular & Cellular Proteomics 10: 10.1074/mcp.M110.003087, 1–15, 2011.

Colorectal cancer (CRC)1 is one of the most prevalent cancers and the fourth leading cause of cancer mortality worldwide, with an estimated 1,000,000 new cases and ~500,000 related deaths every year (1, 2). Detection of CRC and subsequent intervention at an earlier stage has the potential to reduce both incidence and mortality of the disease (3, 4). Clinically, colonoscopy represents the most sensitive approach for early detection among various screening tests. However, only 39% CRC patients are diagnosed at a localized stage by the invasive colonoscopy (2, 5); most of the patients

1 The abbreviations used are: CRC, colorectal cancer; S.D., standard deviation; ANXA4, annexin A4; DEFA1, neutrophil defensin A1; CLDN3, claudin 3; STOML2, stomatin-like 2; CEA, carciinoembryonic antigen; IHC, immunohistochemical staining; XIC, extracted ion chromatography; ROC, receiver operating characteristic; AUC, the area under the ROC curve.
are diagnosed at advanced stages. At the level of molecular diagnosis, at present, carcinoembryonic antigen (CEA) is the most widely used serum tumor marker for CRC (6–8). Because of its low sensitivity and specificity at early cancer stage, CRC screening based on serum and plasma levels of the CEA biomarker is only recommended for prognostic use (8). Therefore, development of better tumor markers aimed at early detection of CRC would improve the diagnosis and/or monitoring of CRC.

Proteomics technology is a powerful tool for uncovering aberrant protein profiles of clinical samples (9, 10). It has been hypothesized that the concentrations of potential biomarkers are highest in the tumor and its immediate microenvironment (i.e. tissue interstitial fluid) (11). If a protein is up-regulated in cancer cells, cancer tissue should possess higher concentration of the protein than serum does. Therefore, many efforts have been made to map the specific genotyping of expressed genes or proteins in the cancerous tissues (12–14), including the most comprehensive human protein atlas for normal and cancer tissues to date (15, 16). Among published CRC-related proteomics literature, the most widely used proteomics approach for identification of differentially expressed proteins in individual CRC patient was based on two-dimensional (2D) PAGE and the more sophisticated technique of two-dimensional differential gel electrophoresis coupled with MS (17–24). By analyzing paired tumor and adjacent normal tissues from patients, these gel-based methods have led to discovery of a variety of proteins involving in signal transduction, cellular reorganization, and tissue hypoxia as potential biomarkers for CRC. Alfonso et al. (25) analyzed the membrane fractions of six paired CRC mucosal tissues using two-dimensional differential gel electrophoresis analysis, identifying annexin A2, annexin A4 (ANXA4) and VDAC as potential markers for CRC diagnosis and, presumably, therapy. In a recent study, Ma et al. (20) found elevated expression of desmin from fetal colorectal tissues and paired CRC tumor-adjacent normal tissues using 2D-PAGE. Up-regulation of oncofetal desmin is correlated with the severity and differentiation of CRC and with poor survival. Thus, desmin could be considered a potential oncofetal serum tumor marker for CRC (20). Although 2D-PAGE is capable of analyzing differentially expressed proteins, this method has limited dynamic range, lower throughput and lower sensitivity to analyze hydrophobic proteins or low-abundance proteins. To the best of our knowledge, the current clinically used CRC marker, CEA (also named carcinoembryonic antigen-related cell adhesion molecule 5, CEACAM5, a membrane protein), has not been identified using the above proteomics methodologies.

Dysregulation of membrane proteins has been linked to a variety of human cancers (26, 27). Most of the FDA-approved cancer biomarker, such as CEA and CA19–9, and ~70% of all known pharmaceutical drug targets are mainly membrane proteins (28). Thus, precise and in-depth characterization of membrane proteins in tumors and adjacent normal tissues from patients will facilitate our understanding of the roles of these proteins in regulating biological processes, which will ultimately provide more reliable biomarker candidates for diagnosis and prognosis, and may contribute to a personalized treatment approach for individual patients. Despite the promise of membrane proteomics for disease marker discovery, only a few studies (29, 30) have characterized the tissue membrane proteome because of the difficulty and inherent challenges of membrane protein analysis and limited amounts of human tumor samples.

In this study, we present a multiplexed label-free quantitation strategy to characterize the individual membrane proteomic profiles in human tissues. This strategy integrates an efficient gel-assisted digestion protocol (31), high-performance liquid chromatography tandem MS (LC-MS/MS) analysis, and a bioinformatics alignment method (32) in attempt to comprehensively map and accurately quantify the membrane proteome. Using CRC as a model to discover potential biomarkers for diagnosis, we applied this strategy to the analysis of differentially expressed membrane proteins in tumor and adjacent normal tissues from each CRC patient. Twenty-eight paired tumor and adjacent normal tissues from patients in Dukes’ A (n = 4), Dukes’ B (n = 7), Dukes’ C (n = 11) and Dukes’ D (n = 6) stages were analyzed. We addressed two issues. (1) From the technical prospective, can this approach provide good sensitivity for robust identification of the current CRC protein biomarker, CEA (CEACAM5)? (2) More importantly, can the technological advancement lead to discovery of additional proteins that are involved in colon tumorigenesis and that can serve as new diagnostic and prognostic protein biomarkers? Contrary to most previous studies using pooled tissue samples, the individual paired-tissue comparisons used here provide information concerning biological and genetic variations between different individuals. The high-throughput strategy is also advantageous in that it provides an analysis of individual proteomic patterns among patients to evaluate the heterogeneity of tissues profiles.

In addition to the confident identification of CEA in the 28 patients, this study identified a panel of membrane proteins with high levels of elevated expression levels in CRC patients. To further evaluate the clinical relevance of these candidates, their expression levels in tissue or serum were examined by Western blot, immunohistochemical staining, ELISA, and clinicopathologic analysis from a large cohort of patients with known clinical outcomes. The results demonstrated the power of tissue membrane proteomics for the discovery of valuable biomarker candidates for early diagnosis and prognosis of CRC.

**EXPERIMENTAL PROCEDURES**

*MATERIALS—Protease inhibitor was obtained from Merck (Darmstadt, Germany). Monomeric acrylamide/bisacrylamide solution (40%, 29:1) was purchased from Bio-Rad (Hercules, CA). Trypsin (modified, sequencing grade) was obtained from Promega (Madison, WI). The BCA and Bradford protein assay reagent kits were.*
obtained from Pierce (Rockford, IL). SDS was purchased from GE Healthcare (Central Plaza, Singapore). Ammonium persulfate (APS) and N,N,N',N'-tetramethylethenediamine (TEMED) were purchased from Amersham Biosciences (Piscataway, NJ). Tris-(2-carboxyethyl)phosphine hydrochloride (TCEP), triethylammonium bicarbonate (TEABC), methyl methanethiosulfonate (MMTS), trifluoroacetic acid (TFA), sodium carbonate (Na2CO3), sucrose, carbonate (TEABC), methyl methanethiosulfonate (MMTS), boxyethyl)phosphine hydrochloride (TCEP), triethylammonium bi-

Patients and Specimen Collection—Clinical tissue samples were obtained from Chang Gung Memorial Hospital at Lin-Kou, Taiwan in accordance with approved human subject guidelines authorized by Medical Ethics and Human Clinical Trial Committee at Chang Gung Memorial Hospital. Following surgery, the tumor and adjacent normal tissues were collected in separate tubes, kept on dry ice for 30 min during transportation, and stored at −80 °C before further processing. Adjacent normal tissue was obtained from the distal edge of the resection ≥10 cm from the tumor. In the discovery phase, a total of 28 pairs of cancerous and adjacent normal tissue were collected and analyzed from individual patients with Dukes’ A (n = 4), Dukes’ B (n = 7), Dukes’ C (n = 11) or Dukes’ D (n = 6) stages CRC patients (supplemental Table 1). In the validation phase, 205 colorectal carcinomas patients and 140 blood samples without hemolysis or lipemia, including 70 samples from CRC patient and 70 age-matched individuals without CRC, were collected from the Department of Colorectal Cancer, Chang Gung Memorial Hospital (33). All CRC patients had histologically verified adenocarcinoma of the colon or rectum that was confirmed by pathologists. Patient characteristics were obtained from pathology records. Subjects with a history of other malignant diseases or infectious disease, or who had undergone surgery 6 months prior to the start of this research were excluded for this retrospective study. Fresh plasma samples were obtained before surgery and were stored at −80 °C until use.

Purification of Membrane Protein Fraction from Paired Normal and Tumor Tissues of 28 Patients—Frozen tissues were thawed rapidly at 37 °C, cut into small pieces, and washed by 0.9% NaCl to remove blood. The precleaned tissues were homogenized in STM solution (5 ml/g tissue, 0.25 m sucrose, 10 mM Tris-HCl, 1 mM MgCl2, and HPLC-grade acetonitrile (ACN) were purchased from Sigma-Aldrich (St. Louis, MO). Formic acid (FA) was purchased from Riedel de Haen (Seelze, Germany). Water was obtained from a Milli-Q Ultrapure Water Purification System (Milli-pore, Billerica, MA).

were then applied to the sample to polymerize as a gel directly in the microcentrifuge without electrophoresis. The gel was cut into small pieces, washed several times (0.1 m TEAB in 50% (v/v) ACN) and subjected to tryptic digestion (10:1, protein/trypsin, w/w) in 25 mM TEAB overnight at 37 °C. Peptides were extracted from the gel using sequential extraction with 25 mM TEAB, 0.1% (v/v) TFA in water, 0.1% (v/v) TFA in ACN, and 100% ACN. The extracted peptides were concentrated in a SpeedVac (Thermo Savant SC210A, Holbrook, NY), desalted by using C18 ZipTip (Millipore; Cambridge, Ontario, Canada), and subjected to LC-MS/MS analysis.

LC-MS/MS Analysis—Peptide samples were reconstituted in 0.1% (v/v) FA in H2O and analyzed by Waters Q-TOF Premier (Waters Corp., Milford, MA). Samples were injected into a 20-mm × 180-μm trap column, separated by 200-nm × 75 mm Waters ACQUITY UPLC® system (Waters Corp., Milford, MA), and eluted with a linear gradient of 0–80% of 0.1% (v/v) FA in ACN for 120 min at 300 nL/min. MS was operated in electrospray ionization positive V mode with a resolving power of 10,000. A NanoLockSprayTM source (Waters Corp., Milford, MA) was used for accurate mass measurements, and the lock mass channel was sampled every 30 s. The mass spectrometer was calibrated with a synthetic human [Glu1]-Fibrinopeptide B solution (1 pmol/μl; Sigma Aldrich) delivered through the NanoLockSpray source. Data acquisition was operated in the data-directed analysis mode to automatically switch between a full MS scan (m/z 400–1600, 0.6 s) and three MS/MS scans (m/z 100–1990, 1.2 s for each scan) sequentially on the three most intense ions present in the full MS scan.

Protein Identification and Label-free Quantification—The peak list resulting from MS/MS spectra was exported to mgf format by Mascot Distiller v2.1.1.0. The datasets were batch-searched and combined-searched by Mascot v2.2 (Matrix science, London, United Kingdom) against International Protein Index (IPI) human database (34) (v3.29, 68161 sequences) from the European Bioinformatics Institute using the following constraints: only tryptic peptides with up to two missed cleavage sites were allowed; 0.3 Da mass tolerances for MS and 0.1 Da mass tolerances for MS/MS fragment ions. Methylthio (Cys) and oxidation (Met) was specified as variable modifications. Only unique peptides with scores ≥35 (p < 0.05) were confidently assigned. In each MS/MS spectrum, a total of at least four b- and y-ions were observed. To evaluate the false discovery rate in protein identification, we performed a decoy database search against a randomized decoy database created by Mascot using identical search parameters and validation criteria. The search results in Mascot were exported in eXtensive Markup Language data format (XML). The MS/MS spectra and assignment for single peptide identification are included in supplemental Fig. 1.

For label-free quantification, data analysis was performed by our recently developed IDEAL-Q software (32). The raw data files, from Waters Q-TOF Premier were converted into files of mzXML format by masswolf v4.0. IDEAL-Q performs quantitation analysis using spectral data in mzXML format and Mascot search result in XML format. The abundance of a peptide was determined by the extracted ion chromatography (XIC) and further normalized by XIC of the internal standard peptide. The protein ratio was determined by a weighted average of the peptide ratios, where the weight of each peptide ratio is determined by the sample abundance of the corresponding peptide.

Protein Annotation—For subcellular localization, molecular function annotations and information linked to cancer, all the proteins identified in this study were analyzed using the Ingenuity Pathway Analysis Knowledge Base, MetaCoreTM, Human Protein Reference Database, and the Gene Ontology consortium. To predict the position of transmembrane helices and signal peptides, the amino acid sequence of every identified protein was individually searched against the transmembrane hidden Markov model (TMHMM) Server v2.0 (Center for
A Label-Free Approach for Tissue Membrane Proteomics

Bartelmais, S. I.; Heier, J. G.; Qi, C.; Jansen, C.; Epp, T.; Hescheler, J.; Lurz, R.; Schütz, G.; Bornkamm, G. W.; Stang, P. W. J. Redox Biol. 2017, 12, 302–310.

Molecular & Cellular Proteomics 10.4

10.1074/mcp.M110.003087–4

Biological Sequence Analysis, Technical University of Denmark and SignalP3.0 server (35) (Center for Biological Sequence Analysis).

Western Blot Assay—Frozen tissues were thawed and resuspended in lysis solution (0.25 M sucrose, 10 mM Tris-HCl, pH 7.6, 1 mM MgCl₂, and 1% SDS) with protease inhibitor mixture (100x, sample/protease inhibitor, v/v, Calbiochem). Following homogenization (Polytron System PT 1200 E, Luzernerstrasse, Switzerland) on ice, tissue lysate (40 μg) was used for following procedures. For analysis of stomatin-like 2 (STOML2) in plasma, 10 μl plasma was applied for each patient. Proteins were fractionated by 12% SDS-PAGE and transferred to a PVDF membrane (Millipore, Billerica, MA). The membrane was blocked with 5% skim milk in TBS-Tween buffer (25 mM Tris, 190 mM NaCl, and 0.5% (v/v) Tween 20, pH 7.5) and then incubated at 25 °C with the appropriate primary antibody (Claudin 3 antibody, Abbccc, Cambridge, MA; Annexin A4 antibody, Abccam; Defensin 1 antibody, Lifespan Bioscience, Seattle, WA; STOML2 antibody, ProteinTech Group Inc., Chicago, IL) at a dilution of 1:1000 for 2 h followed by incubation with peroxidase-conjugated secondary antibody at 25 °C for 1 h. Following washing, the membrane was developed with a chemiluminescence reagent kit (Amersham Biosciences Pharmacia Biotechnology, Inc., Piscataway, NJ) and then exposed to Kodak BioMax light film (Sigma-Aldrich, St. Louis, MO). Immunoblot images were analyzed by ImageMaster (Amersham Biosciences, NJ).

Immunohistochemistry (IHC)—The tumor tissue blocks used for IHC were first fixed in 4% paraformaldehyde and then embedded in paraffin. To reduce selection bias, areas of normal colon epithelium, benign polyps, and adenocarcinoma were reviewed by two independent histopathologists. Sections (5 μm thick) were cut from tissue blocks, mounted on silanized slides (Superfrost, Menzel, Braunschweig, Germany), subsequently deparaffinized with xylene (twice for 10 min each), and rehydrated through ethanol gradient washes as described previously (36). To eliminate the endogenous peroxidase activity, slides were incubated with 3% H₂O₂, at room temperature for 30 min before heating in a microwave oven for antigen retrieval (10 mM citrate buffer, pH 6.0; 20 min, 700 W). To block nonspecific binding, slides were preincubated with 10% nonimmune goat serum at 37 °C for 30 min. Slides were then incubated with anti-human STOML2 primary antibody (mouse, 1:250 dilution; ProteinTech, Inc., Chicago, IL) for 30 min at room temperature. Following washing with PBS (pH 7.4), slides were incubated with HRP-conjugated anti-mouse IgG secondary antibody (rabbit, 1:2000 dilution; Abbccc, Inc., Cambridge, UK) for 30 min at room temperature and then developed using 3,3′-diaminobenzidine (Sigma, St Louis, MO). Sections were counterstained with hematoxylin, washed in running tap water, dehydrated, and mounted in Neo-Mount (Merck, Darmstadt, Germany). Negative control reactions were conducted by omitting the primary antibody.

Scoring of Immunostaining (SI)—Immunostaining was evaluated and scored by two experienced pathologists who were blinded to any knowledge of clinical or pathological parameters and clinical outcome as described previously (37, 38). The percentage of STOML2-positive tumor cells was determined semiquantitatively by assessing the entire tumor section. Expression of the STOML2 protein was categorized as positive or negative and was evaluated according to the simplified H score system (39), which is based on the percentage of cells stained (4, 90–100%; 3, 75–89%; 2, 50–74%; 1, 6–49%; or 0, 0–5%) and the intensity of cell staining (3, strong; 2, moderate; 1, weak; or 0, no cell staining). The two scores are multiplied to obtain the final score. For this study, we defined an SI of 0–9 to indicate negative or low STOML2 protein expression and an SI of 10 or more to indicate positive or high STOML2 expression.

Production of Recombinant STOML2—STOML2 cDNA was obtained by reverse transcription-PCR using total RNA isolated from Colo205 cells (ATCC no. CCL-22) and the following STOML2-specific primers, synthesized by MWG-Biotech (Seminole Drive, Huntsville, AL) and based on the STOML2 gene sequence of GenBank accession number AF190167; STOML2 forward primer, 5′-TTA- CATATGTTTTTTGCTAGGGGCTCTCCT-3′; STOML2 reverse primer, 5′-TTTCCTCAGGTGTACCCTGGACATCTCCTGCTG-3′. The PCR amplification product containing nucleotides 97 - 1089 of STOML2 was subcloned into a pET30b (Novagen, EMD Chemicals Inc., Darmstadt, Germany) expression vector. Sequence analysis was performed using a T7 promoter primer to confirm the STOML2 sequence. The STOML2 expression plasmid was then transformed into Escherichia coli BL21 (DE3) (pLysS for recombinant expression. Isopropyl β-d-thiogalacto-pyanoside (IPTG) was used to induce protein expression. The cells were harvested by centrifugation and disrupted by sonication in lysis buffer containing 0.2 M Tris–HCl, pH 7.5; 20% (v/v) glycerol; and 1 mM phenylmethylsulfonyl fluoride at 4 °C. r-STOML2 was purified using a HiTrap™ Chelating HP column (GE Healthcare, Munich, Germany). Expression and purification of the recombinant protein were confirmed with immunoblotting using a mouse monoclonal antibody against the STOML2 protein (ProteinTech Group Inc., Chicago, IL).

ELISA—Briefly, monoclonal anti-human STOML2 (Protein Tech Group Inc., Chicago, IL) in coating buffer (pH 9.5; 15 mmol/l Na₂CO₃, 35 mmol/l NaHCO₃ in distilled water) was added to 96-well microtiter plates (Costar, Illinois, IL) at 100 ng/well and incubated overnight at 4 °C. The plates were washed three times with PBS buffer (PBS with 0.05% Tween-20) and then blocked in blocking buffer (PBST with 2.5% BSA) overnight at 4 °C. For following, a twofold dilution of individual plasma samples (50 μl) in blocking buffer (50 μl) was added to each well and incubated at 37 °C for 1 h. Plates were washed three times with PBST buffer and incubated at 37 °C for 1 h with polyclonal rabbit anti-human STOML2 (1:2000; Abbccc, Inc.). Plates were washed and incubated at 37 °C for 30 min with HRP-conjugated goat anti-rabbit IgG(1:10,000; Santa Cruz Biotechnology, Santa Cruz, CA). The signal was developed using 3,3',5,5'-tetramethylbenzidine (TMB; 100 μl/well; KBL, Gaithersburg, MD) as the substrate. To stop the reaction, 1 N HCl (100 μl/well; Sigma-Aldrich) was added in each well, and the absorbance was determined at 450 nm with 650 nm as reference filter by spectrophotometer ( Molecular Devices, Sunnyvale, CA). All samples were tested in duplicate, and the mean was used for data analysis. A subset of plasma samples and purified recombinant STOML2 proteins (50, 40, 30, 20, 10, and 5 ng/ml) was assayed in every ELISA batch for quality control and quantification, respectively (supplemental Fig. 2). For CEA assay, the plasma CEA levels were determined with a commercially available CEA ELISA kit and protocol (Roche Diagnostics, Mannheim, Germany).

Statistical Analysis—For tissue IHC analysis, associations between protein expression and clinicopathological characteristics were analyzed using the Spearman rank correlation test. For plasma STOML2 levels, relationships were analyzed using the Kruskal-Wallis test and the Mann-Whitney U test to analyze differences between individual variables from two groups. The Pearson’s Chi-square test was used to analyze the relationship between categorical variables. To determine factors related to overall survival, Cox proportional hazard models (both univariate and multivariate) were used, and the probability was calculated using the Log-rank test by the Kaplan-Meier method. All p values were derived from two-tailed statistical tests, and p < 0.05 was regarded as statistically significant. The diagnostic potential of a specific marker was evaluated by performing the receiver operating characteristic (ROC) curve analysis and the discriminative efficacy of an individual biomarker was calculated by the area under the ROC curve (AUC).

RESULTS

Analytical Workflow and Performance Assessments

Workflow and Experimental Design—In this study, a simple label-free strategy was designed to quantitatively analyze the
individual membrane proteome of paired tumor and adjacent normal tissues from 28 CRC patients with Dukes’ A, B, C, or D stage. The analytical workflow for individual patients consisted of three steps: (1) membrane protein purification, (2) gel-assisted digestion, and (3) LC-MS/MS analysis (Fig. 1A). Membrane proteins were first purified from paired tumor and adjacent normal tissues, mixed with internal standard protein (BSA), and then subjected to our recently reported gel-assisted digestion (31). When sample along with the spiked BSA was incorporated into a polyacrylamide gel matrix, any interfering components can be removed by in-gel washing steps to ensure efficient enzymatic digestion and LC-MS/MS analysis. To improve reproducibility and reliability of label-free quantitation of individual samples, each batch of extracted peptides was analyzed in triplicate by LC-MS/MS without the fractionation step.

The quantitation analysis was performed by our recently developed software, IDEAL-Q (Fig. 1B) (32), an automated tool for label-free quantitation analysis using an efficient peptide alignment approach and spectral data validation. The software first integrates all protein and peptide identification results from all LC-MS/MS runs, and performed peptide alignment and identification assignment according to the commonly and confidently (p < 0.05) identified peptides. The unidentified peptides, which were the result of either a low identification score or an absence of MS/MS sequencing data, can be retrieved based on predicted elution times and m/z values. To ensure accurate peptide detection and assignment, IDEAL-Q further validated the detected peak clusters by three criteria including signal-to-noise ratio of >3, a correct charge state, and a good isotope pattern. Peptide abundance was determined by peak area of XIC and normalized by XIC area of internal standard peptide in the same LC-MS/MS run. The calculated peptide ratios were further normalized by central tendency normalization (40), and the protein ratios were determined by the weighted average of normalized peptide ratios.

Selection of Internal Standard and Assessment of Reproducibility and Accuracy—To evaluate the reproducibility and accuracy of our method for tissues membrane proteomics, three replicate sets of membrane proteins (20 μg) were purified independently from the same CRC tissue and analyzed. A total of 2152 peptides corresponding to 513 proteins were confidently identified (p < 0.05, false discovery rate = 1.1%). We evaluated the quantitation accuracy and precision obtained using different peptides from spiked internal standard protein, BSA (HLVDEPQNLK, LGEYGFQNALIVR and DA-FLGSFLYEYSR), which were chosen based on their different elution times (52.58 ± 0.33 min, 78.99 ± 0.19 min, and 89.13 ± 0.12 min, respectively). In principle, we hypothesized that internal standard peptides that elute at different times can be used to correct chromatographic shifts during long LC gradient. To our surprise, similar accuracy (error = 3–4%) and precision (S.D. = 0.27–0.37) were obtained using single or
combined (two or three) internal standard peptides (supplemental Table 2). Thus, the use of one internal standard peptide, HLVDEPQNLIK, was sufficient to normalize systematic errors.

As shown in Fig. 2, all three pair-wise ratios showed narrow normal distributions with mean values of $-0.05 \pm 0.27$, $0.01 \pm 0.30$, and $0.11 \pm 0.37$ (95% confidence interval), which are consistent with an expected ratio of 1.0 and relative S.D. of 20.6–29.2% on a linear scale. Even with the additional membrane protein purification step, the obtained S.D. were comparable with those from previous studies of label-free quantitation of protein expression levels (41, 42). Based on the two-standard deviation model (confidence interval $\pm 95.5\%$), we considered a difference in abundance of 1.5-fold to indicate a statistically significant degree of higher or lower expression, respectively.

**Differential Membrane Proteomics Profiles of Paired Tumor and Adjacent Normal Tissues in CRC Patients**

**Overall Identification and Quantitation Results**—We applied the label-free quantitation strategy to membrane proteomic analysis of paired tumor and adjacent normal tissues from 28 patients. Supplemental Table 1 lists the clinical characteristics information of the 28 patients. Based on conventional single LC-MS/MS analysis and database searching, individual patients had only 261–591 protein identifications (Fig. 3A). Because patient heterogeneity leads to the diverse protein expression profiles, the peptide alignment and cross-assignment strategy in IDEAL-Q tool effectively detected almost every identified peptide among LC-MS/MS dataset from different patients. As shown in the gray bars of Fig. 3A, thus, the number of identified proteins increased from 453 proteins to 1482 proteins (false discovery rate $=2.24\%$, see detailed identification information in supplemental Table 3) when the number of patient increased.

Under more stringent quantitation criteria by which at least two peptides are required for reliable quantification, 856 proteins were confidently quantified from the 56 tissue samples by IDEAL-Q (Fig. 3B). Among the 856 quantified proteins, 642 were annotated as membrane proteins or membrane-associated proteins including 327 plasma membrane proteins by Gene Ontology, Ingenuity Pathway Analysis Knowledge Base, and TMHMM prediction.

**Differentially Expressed Proteins in the Four Different Stages of CRC**—To overcome intrinsic intra- and interspecimen variability associated with differing patient characteristics and tissue heterogeneity, quantitative comparison between the tumor and adjacent normal tissues from the same patient is important to minimize the genetic variations. With the criterion of a 1.5-fold change, the quantitation result revealed 216 (Dukes’ A), 176 (Dukes’ B), 96 (Dukes’ C), and 94 (Dukes’ D) proteins with higher expression (Ratio $>1.5$) and 103 (Dukes’ A), 111 (Dukes’ B), 45 (Dukes’ C), and 54 (Dukes’...
D) proteins with lower expression (Ratio < 0.67). Among the proteins previously reported as related to CRC, we confidently identified and quantified several carcinoembryonic antigen-related cell adhesion molecules including CEACAM5 (CEA) and CEACAM6; the former is the current prognostic marker used for CRC (8). The representative XICs of CEACAM5 between normal and tumor samples are shown in supplemental Fig. 3. The overexpression of CEA was commonly observed by quantitation of 10 peptides with a total Mascot score of 22,074. CEA showed 1.5-fold higher expression in 80% of CRC patients (80% of Dukes’ A, 85.7% of Dukes’ B, 90.9% of Dukes’ C, and 100% of Dukes’ D).

We further extracted proteins with high occurrence (> 20 patients) of overexpression from four stages (Fig. 4A; Table I). Table I lists their detailed information for these proteins including Mascot score, subcellular localization, and number of transmembrane helices as well as expression in cell lines, tissues, and serum. The color and size of dot shown indicates the different extent of up-regulation and down-regulation. The larger size of the dot signifies the greater discrepancy in expression level between tumor and normal tissues. All the patients are organized according to their tumor stage.

The comparison of individual ratio and average ratio indicated that the traditional way of pooling samples from different patients may identify misleading biomarker candidate. The protein represented in Fig. 4B, have significantly high average ratios obtained from the 28 paired tissues. However, only a few patients had dramatically elevated levels, which influenced the overall average ratios of the pooled sample; the expression levels in the remaining samples were unchanged or even slightly down-regulated. Therefore, pooled results do not adequately represent each individual, emphasizing the superiority of personalized membrane proteomics. Personalized membrane proteomics profiles of individual patients can differentiate among biomarker candidates by avoiding the use of potentially misleading average ratios.

Validation of Selected Proteins by Western Blotting—To further validate the differential membrane proteomic profiles
# Table I

## Top 23 proteins with 1.5-fold higher expression in ≥20 CRC patients and their detailed information including identification confidence, subcellular localization, and number of transmembrane helices (TMH) as well as expression in cell lines, tissues and serum

| No | Symbol | Description | No. of Patients with up-regulation | Location<sup>a</sup> | TMH<sup>b</sup> | Signal peptide<sup>c</sup> | Plasma/ Serum<sup>d</sup> | Large Intestine<sup>e</sup> | COLO205<sup>f</sup> | HCT-116<sup>f</sup> | HCT-15<sup>f</sup> | HT22<sup>f</sup> | KM-12<sup>f</sup> | SW-620<sup>f</sup> | Expressed in CRC Tissue<sup>e</sup> |
|----|--------|-------------|----------------------------------|----------------------|---------------|--------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| 1  | CEACAM6 | Carcinoembryonic antigen-related cell adhesion molecule 5 precursor | 25 | PM | 0 | Yes | x | x | x | x | x | x |
| 2  | STOML2 | Stomatrin-like protein 2 | 24 | PM | 1 | Yes | x | x | x | x | x | x | x | x | x |
| 3  | CEACAM6 | Carcinoembryonic antigen-related cell adhesion molecule 6 precursor | 24 | PM | 0 | Yes | x | x | x | x | x | x | x | x | x |
| 4  | DEFA1 | Neutrophil defensin 1 precursor | 24 | Extracellular region | 0 | Yes | x | x | x | x | x | x | x | x | x |
| 5  | RPL36 | 60S ribosomal protein L36 | 23 | OM | 3 | Yes | x | x | x | x | x | x | x | x | x |
| 6  | TPSO | Peripheral benzodiazepine receptor | 23 | OM | 0 | No | x | x | x | x | x | x | x | x | x |
| 7  | TSHD2 | OCA domain containing 2 isoform 1 | 22 | OM | 0 | No | x | x | x | x | x | x | x | x | x |
| 8  | IKP2 | IkappaB | 22 | OM | 0 | No | x | x | x | x | x | x | x | x | x |
| 9  | CYBA | Cytochrome b-245 light chain | 22 | PM | 3 | Yes | x | x | x | x | x | x | x | x | x |
| 10 | GOLT1B | Vesicle transport protein G0T1B | 22 | OM | 3 | Signal anchor | x | x | x | x | x | x | x | x | x |
| 11 | ELA2 | Leukocyte elastase precursor | 22 | Extracellular region | 1 | Yes | x | x | x | x | x | x | x | x | x |
| 12 | CYB5B | Cytochrome b5 outer mitochondrial membrane precursor | 21 | OM | 1 | No | x | x | x | x | x | x | x | x | x |
| 13 | SEC61B | Protein transport protein Sec61 subunit beta | 21 | OM | 1 | No | x | x | x | x | x | x | x | x | x |
| 14 | TOMM40 | Isoform 1 of Probable mitochondrial import receptor subunit TOM40 homolog | 21 | OM | 0 | No | x | x | x | x | x | x | x | x | x |
| 15 | TMEM63A | Transmembrane protein 63A | 21 | PM | 11 | Signal anchor | x | x | x | x | x | x | x | x | x |
| 16 | TSPAN8 | Tetraspanin-8 | 21 | PM | 4 | Yes | x | x | x | x | x | x | x | x | x |
| 17 | SLC16A3 | Monocarboxylate transporter 4 | 21 | PM | 12 | Yes | x | x | x | x | x | x | x | x | x |
| 18 | LAM1 | Lysosomal-associated membrane protein 1 | 21 | OM | 2 | Yes | x | x | x | x | x | x | x | x | x |
| 19 | ANX4 | ANX4 protein | 21 | PM | 0 | No | x | x | x | x | x | x | x | x | x |
| 20 | TOMM5 | Uncharacterized protein C9orf105 | 21 | Unknown | 0 | No | x | x | x | x | x | x | x | x | x |
| 21 | SLC2A1 | Solute carrier family 2, facilitated glucose transporter member 1 | 20 | PM | 12 | Yes | x | x | x | x | x | x | x | x | x |
| 22 | TMCO1 | Isoform 1 of Transmembrane and coiled-coil domain-containing protein 1 | 20 | OM | 2 | Yes | x | x | x | x | x | x | x | x | x |
| 23 | OPRS1 | Isoform 3 of Sigma 1-type opioid receptor | 20 | PM | 1 | Yes | x | x | x | x | x | x | x | x | x |

---

<sup>a</sup>The location is annotated by Gene Ontology and Ingenuity Pathway Analysis Knowledge Base. PM indicates plasma membrane and OM indicates organelle membrane.

<sup>b</sup>TMH (transmembrane helix) is predicted by TMHMM Server v.2.0.

<sup>c</sup>Signal peptide is predicted by SignalP 3.0 Server.

<sup>d</sup>The data are obtained from Ingenuity Pathway Analysis.

<sup>e</sup>The expression of proteins in colorectal cancer tissues are obtained in Human Protein Reference Database.
in tissues obtained from our label-free strategy, we examined the expression levels of several candidate proteins in tissues from 60 patients by Western blotting. Neutrophil defensin 1 (DEFA1), annexin A4 (ANXA4), and claudin 3 (CLDN3) were selected based on several criteria including (1) structural characterization as plasma membrane proteins, which may enable their use as serological biomarkers; (2) their potential roles in contributing to CRC diagnosis, and (3) the availability of commercial antibodies.

As shown in Fig. 5A–C (supplemental Fig. 5), the expression levels were 9.1 ± 10.1 for DEFA1 in normal tissue, and 27.4 ± 42.2 in tumor tissue. The expression levels of ANXA4 were 90.2 ± 83.5 in normal tissue and 154.4 ± 82.7 in tumor. The CLDN3 expression levels were 27.1 ± 31.8 in normal and 94.5 ± 49.0 in tumor tissue. As shown in Fig. 5D–F, statistical analysis revealed that the mean of expression level of DEFA1, ANXA4, and CLDN3 was higher (threelfold, \( p = 0.0004 \), 1.7-fold, \( p < 0.0001 \), and 3.5-fold, \( p < 0.0001 \), respectively; by two-sided \( t \) test) in tumor versus normal tissues. The fold-change in expression levels of the three proteins obtained by Western blot for these patients were similar to the higher expression in cancerous tissue (DEFA1: 5.1 ± 4.7, ANXA4: 2.4 ± 1.2, CLDN3: 2.6 ± 2.5) as measured by the MS-based label-free quantification strategy.

Validation of STOML2 as a Candidate Marker in Early CRC Detection and Prognosis—To search for proteins with potential utility in early detection and/or prognosis of CRC, in addition to the above-mentioned criteria, we further filter proteins with criteria of (1) high frequency of overexpression in both early stages and advanced stages of CRC and (2) structural characteristic as secreted proteins, which may enable their use as serological biomarkers. Stomatin-like 2 (STOML2) was overexpressed in 24 CRC patients (Fig. 4 and Table I). Elevated levels of STOML2 had been reported in various tumor types (37, 43–45), suggesting that STOML2 might play an important role in tumorigenesis. As for its association with CRC, however, it was only found in the large intestine, and expression levels of STOML2 in CRC patients have not been investigated.

To confirm the high expression of STOML2, tumor tissues from 205 CRC patients were collected for immunohistochemical (IHC) analysis. Fig. 6A–C shows representative cases that harbored both tumor and nontumor cells. Little or weak staining of STOML2 on nontumor epithelial cells was found in most of the samples examined. In contrast, the tumor tissues showed moderately to strongly positive reactions to STOML2. Elevated levels of STOML2 had been reported in various tumor types (37, 43–45), suggesting that STOML2 might play an important role in tumorigenesis. As for its association with CRC, however, it was only found in the large intestine, and expression levels of STOML2 in CRC patients have not been investigated.

A high level of STOML2 expression, i.e. a SI value of \( \geq 10 \) (see Experimental procedures), occurred consistently in 96% of the 205 patients (Table II). Among the 205 CRC tissue sections examined, 205 sections harbored nontumor epithelial cells, of which 43% (88/205) were positive for STOML2 expression. The percentage of positive staining in malignant tissues was significantly increased as compared with that in normal tissues (Fisher exact test, \( p < 0.001 \)). To further determine whether the overexpression of STOML2 occurred in premalignant lesions of CRC, we compared the SI value between normal and benign polyp tissues. A total of 184 sections contained benign polyp adenoma; positive STOML2 expression was significantly higher (62%, 114/184; Fisher exact test, \( p < 0.001 \)) in benign polyp tissue versus normal tissue. Based on this observation, we hypothesize that overexpression of STOML2 is an early event in CRC development, which needed larger-scale clinical studies to verify this speculation. It is noted that there was no

Fig. 5. Western blot validation of A, DEFA1, B, ANXA4, and C, CLDN3 expression levels in tissue samples. The fold-changes of D) DEFA1, E) ANXA4, and F) CLDN3 in tumor versus normal tissues were evaluated by two-sided \( t \) test (\( p < 0.05 \)). The raw data points were presented as scatter plot and the mean values were indicated.

---

A Label-free Approach for Tissue Membrane Proteomics

Molecular & Cellular Proteomics 10.4

10.1074/mcp.M110.003087–9
statistical relationship between STOML2 expression and the clinicopathological parameters examined in the present study (Table III), including histological grade \((p = 0.451)\), tumor stage \((p = 0.232)\), lymph node metastasis \((p = 0.343)\), distal metastasis \((p = 1.000)\).

**Prognostic Implications of Tissue STOML2 Expression**—Among the 205 CRC patients, 154 patients showed valid survival data. We conducted univariate and multivariate analyses of several prognostic factors (gender, age at onset, tumor stage, nodal stage, metastasis stage, overall stage, differentiation grade, CEA level, and STOML2 expression) for these 154 patients on overall survival using the Cox proportional regression models. Several factors displayed statistically significance in univariate analysis (Table IV); however, following multivariate analysis, only the STOML2 expression \((p = 0.029)\) and the metastasis stage \((p < 0.001)\) were found to be independent poor prognostic indicators for overall survival (Table IV). Further analysis using Kaplan-Meier plot showed that the 5-year overall survival for patient subgroups stratified by STOML2 expression levels \((SI \leq 10\) and \(\geq 10)\) were 75 and 42%, respectively; this difference was statistically significant using a log-rank test \((p = 0.047; \text{supplemental Fig. 4})\). Mean survival period was 34.77 ± 2.03 months in patients with high STOML2 expression, whereas 53.67 ± 3.46 months was obtained for patients with low STOML2 expression.

**The Diagnostic Value of Plasma STOML2 Protein**—We hypothesized that STOML2, a transmembrane protein with a signal peptide, may be secreted into the blood stream as a potential serologic CRC biomarker, although its presence in serum had not been confirmed (Table I). Therefore, we developed an in-house ELISA (see supplemental Fig. 2) to analyze the plasma levels of STOML2 protein from 70 CRC patients and 70 healthy controls (see “Experimental Procedures”). The mean plasma level of STOML2 was twofold higher in the CRC patients (5.61 ng/ml) as compared with that in the healthy controls (2.77 ng/ml) (Student’s \(t\) test, \(p < 0.01\), Fig. 7A). To further evaluate the diagnostic potential, an ROC curve was constructed for the plasma STOML2 and CEA level (Fig. 7B). STOML2 and CEA had statistically significant AUC values of 0.74 and 0.69, respectively, demonstrating the potential of STOML2 as a novel candidate CRC biomarker. When 3.5 ng/ml, determined by the optimal Youden’s index, was set as a cutoff value, overexpression of STOML2 was found in 37.1% (26 of 70) of normal controls and 71% (50 of 70) of CRC patients. Combining data from both markers led to a further increase in the AUC to 0.77 (Fig. 7B).

Interestingly, the analysis of blood samples from CRC patients revealed that STOML2 and CEA are highly expressed to varying degrees among CRC stages. Many patients (20/29; 69%) with early-stages CRC exhibited high STOML2 plasma levels above the cutoff limit of 3.5 ng/ml, whereas the CEA levels of these patients were mostly below the standard cutoff limit of 5 ng/ml (Fig. 7C). The mean expression level of CEA increased substantially in patients with early-stage CRC as compared with in normal controls (Student \(t\) test, \(p < 0.001\); Fig. 7D). Elevated STOML2 plasma levels, which were similarly distributed among patients with early-stage and ad-
When cutoff values of 5 ng/ml for CEA and 3.5 ng/ml for STOML2 were applied, the overall diagnostic sensitivity of STOML2 and CEA at all stages were 71% and 40% (Table V), respectively. In contrast to CEA, which exhibited higher sensitivity values in advanced tumor stages (24/41, 59%), the superior performance of STOML2 was particularly evident for the early-stage carcinoma (TNM stage I and II). The sensitivity of STOML2 was as high as 69%, whereas CEA has low sensitivity of 14%, which is comparable to previous literature (8). By combining STOML2 and CEA, the sensitivity for detecting early-stage CRC increased to 76%. Based on the observation, overexpression and secretion of STOML2 into blood may be an early event in colorectal carcinogenesis, which may be utilized for noninvasive early detection of CRC.

**DISCUSSION**

One of the current bottlenecks in clinical membrane proteomics technologies is the small amount of biomaterial and low throughput when dealing with tissue biopsy specimens. Because the complex composition of proteome is incompat-
ible with the limited acquisition speed of mass spectrometer, only a small fraction of highly abundant peptides will be targeted for MS/MS analysis. Thus, large volumes of biomaterials are necessary when additional fractionation strategies are used to profile less-abundant proteins (46, 47). Although these strategies increase the numbers of confidently identified proteins, nevertheless, they create a tradeoff for overall analytical throughput and may also cause lower quantitation accuracy because of the analytical challenge in confident peptide assignment and peak matching when the same protein/peptide is present in different fractions (32).

In this study, our label-free strategy was designed to balance analytical sensitivity and throughput with depth of proteome coverage for personalized membrane proteomic profiles. Because of the intra- and interstage diversity among patients, the individual membrane proteomics profiles had a small overlap (17.6–39.9%) among all patients. When the proposed SEMI label-free strategy was used for effective peptide alignment and cross-assignment, the patient-to-patient diversity increased the number of quantified proteins without fractionation prior to LC-MS/MS. For each patient, <10 μg of isolated membrane proteins was used to obtain ≈1500 proteins (Q-TOF MS) in the membrane proteomics profiles of tumor and adjacent normal tissue, which is comparable to the result generated using multidimensional peptide separation approaches (48). The high sensitivity of our approach was also demonstrated by the confident identification of the current prognostic marker CEACAM 5 (CEA), with a Mascot score of 22,074. To our knowledge, it is the first report regarding the discovery of CEACAM5 using a proteomics strategy. In addition, the new label-free quantitation strategy provided good accuracy (6% error) and reasonable reproducibility (34% relative S.D.) comparable to isotope labeling strategies (31).

AXNA4, DEFA1, and CLDN3 were identified as potential CRC biomarkers using the label-free strategy. Overexpression of ANXA4 occurs in renal clear cell carcinoma (49), clear cell carcinoma of the ovary (50), papillary thyroid carcinoma (51), and pancreatic adenocarcinoma (52). ANXA4 is also regulated by Tiam1 (T-cell lymphoma invasion and metastasis-inducing protein 1), which is a CRC metastasis-related gene (53); however, its functional role in CRC development is yet to be defined. Defensins (including DEFA1 and DEFA3) levels in stool samples were elevated in CRC, adenoma, and upper gastrointestinal cancers (54). Overexpression of defensins was also found in tissues of colorectal adenomas and carcinomas (55). Aberrant claudin expression has been identified in various cancers, such as those originating from the pancreas, bladder, thyroid, fallopian tubes, ovary, stomach, colon, breast, uterus, and the prostate (56). Thus, we hypothesize that elevated levels of these proteins may be a typical observation in various types of cancer.

Stomatin is a member of the highly conserved family of stomatin proteins (57). Overexpression of stomatin-like 2 (STOML2) has been previously reported in several cancers including esophageal squamous cell carcinoma (43), lung cancer (44), laryngeal cancer (37), and endometrial adenocarcinoma (45), but not in CRC. Transfection of a human esoph-

### Table IV

| Univariate Cox regression analysis | Overall survival | Multiivariate Cox regression analysis | Overall survival |
|-----------------------------------|-----------------|--------------------------------------|-----------------|
| **Gender**                        | Hazard Ratio    | 95% confidence interval              | **P**           |
| Male vs Female                    | 0.610           | 0.397–0.938                          | 0.024           |
| **Age**                           |                 |                                      |                 |
| ≥60 vs <60                        | 1.318           | 0.826–2.012                          | 0.246           |
| **Tumor stage**                   |                 |                                      |                 |
| T1-T2 vs T3-T4                    | 4.264           | 1.561–11.649                         | 0.005           |
| **Nodal stage**                   |                 |                                      |                 |
| N1–3 vs N0                        | 2.521           | 1.584–4.011                          | <0.001          |
| **Tumor metastasis**              |                 |                                      |                 |
| M1 vs M0                          | 4.460           | 2.862–6.951                          | <0.001          |
| **Overall TNM stage**             |                 |                                      |                 |
| Stage 3-4 vs stage 1–2            | 3.591           | 2.062–6.008                          | <0.001          |
| **Differentiation grade**         |                 |                                      |                 |
| Moderate, poor vs well            | 1.934           | 0.843–4.436                          | 0.120           |
| **CEA (5 ng/ml)**                 |                 |                                      |                 |
| Positive vs negative              | 2.038           | 1.318–3.153                          | 0.001           |
| **STOML2 expression**             |                 |                                      |                 |
| High vs Low                       | 3.211           | 1.014–10.170                         | 0.047           |
| **Tumor metastasis**              | Hazard Ratio    | 95% confidence interval              | **P**           |
| M1 vs M0                          | 4.603           | 2.949–7.186                          | <0.001          |
| **STOML2 expression**             |                 |                                      |                 |
| High vs Low                       | 3.626           | 1.143–11.502                         | 0.029           |
A Label-free Approach for Tissue Membrane Proteomics

Fig. 7. A, ELISA of the plasma levels of STOML2 in 70 CRC patients and 70 age-matched healthy controls. B, ROC curves of CEA and STOML2 based on the ELISA data shown in (A). C, Comparison of STOML2 and CEA plasma levels in early- and advanced-stage CRC patients. D, ELISA analysis of STOML2 expression in plasma from CRC patients of different stages (early stage: TNM stage I and II, n = 29; advanced stage: TNM stage III and IV, n = 41) and from healthy controls (n = 70). *, p < 0.01, **, p < 0.001; Student’s t test.

Table V
Detection sensitivity of CEA and STOML2 as CRC biomarkers

| CRC Stage   | Number of patients | (A) CEA* | (B) STOML2b | Combination |
|-------------|--------------------|----------|-------------|-------------|
| Early Stage | 29                 | > 5 ng/ml| > 3.5 ng/ml | (A + B)     |
| Advance stage| 41                | 4 (14%)  | 20 (69%)    | 22/29 (76%) |
| All stages  | 70                 | 24 (50%) | 30 (73%)    | 39/41 (95%) |

*a Cutoff value was set as 5 ng/ml.
*b Cutoff value was set as 3.5 ng/ml.

ageal squamous cell carcinoma cell line with antisense STOML2 leads to decreased cell growth, cell adhesion, and tumorigenesis, demonstrating its potential role in controlling tumor growth (58). As one of the most up-regulated proteins in superinvasive cancer cells (59), STOML2 is also associated with decreased overall survival in breast cancer patients when it is expressed at high levels (44). It is regarded as an important player in T-cell activation by ensuring sustained T-cell receptor signaling, making it a potential target for immunomodulation (60). The role of STOML2 in CRC development is, however, still not known (61).

In the study, we presented the first report on the link between STOML2 and its potential applicability for CRC detection and survival prediction. High-level STOML2 expression, i.e., an SI value of 10, occurred consistently in almost all CRC patients (94%; 192/205) in every stage (Table II), suggesting the correlation between STOML2 and neoplastic transformation of epithelial cells in colorectum. Elevated level of STOML2 was observed in both CRC tissues and plasma from CRC
patients. The potential use of STOML2 might be best described in comparison with CEA, the most well-established single tumor marker for CRC (7). The overall sensitivity of STOML2 was superior to that of CEA and was comparable to combination of CEA with u-PA or CA 19–9 for screening of CRC (62). CRC detection sensitivity was further improved by the combined use of STOML2 and CEA (87%). Most significantly, more than 70% of patients with early-stage CRC were found to have increased plasma levels of STOML2, suggesting its potential role as an early detection marker in screening asymptomatic patients with CRC.

In conclusion, we demonstrated the feasibility of a simple label-free strategy as a multiplexed quantitative membrane proteomic platform to analyze individual paired tissues in discovery-driven translational cancer research. The proteomic analysis of CRC tissues conducted here provides data for novel candidate biomarkers with potential diagnostic and prognosis value in CRC. Further validation of several plasma membrane proteins in serum samples indicated the potential for tissue membrane proteomics to identify serum markers for noninvasive diagnosis of cancer. It is anticipated that this efficient strategy can be applied to other types of cancer.

REFERENCES

1. Weitz, J., Koch, M., Debus, J., Höhler, T., Galle, P. R., and Büchler, M. W. (2005) Colorectal cancer. Lancet 365, 153–165
2. Jemal, A., Siegel, R., Ward, E., Murray, T., Xu, J., and Thun, M. J. (2007) Cancer statistics, 2007. CA Cancer J. Clin. 57, 43–66
3. Kronborg, O., Fenger, C., Olsen, J., Jørgensen, O. D., and Sondergaard, O. (1996) Randomised study of screening for colorectal cancer with faecal occult-blood test. Lancet 348, 1467–1471
4. Walsh, J. M., and Terdiman, J. P. (2003) Colorectal cancer screening: scientific review. JAMA 289, 1298–1298
5. Smith, R. A., Cokkinides, V., and Eyre, H. J. American Cancer Society guidelines for the early detection of cancer, 2006. CA Cancer J. Clin. 56:11–25, 2006; quiz 49–50
6. Thompson, J. A., Grunert, F., and Zimmermann, W. (1991) Carcinoembryonic antigen gene family: molecular biology and clinical perspectives. J. Clin. Lab. Anal. 5, 344–366
7. Kim, H. J., Yu, M. H., Kim, H., Byun, J., and Lee, C. (2008) Noninvasive molecular biomarkers for the detection of colorectal cancer. BMB Rep. 41, 685–692
8. Wanebo, H. J., Rao, B., Pinsky, C. M., Hoffman, R. G., Stearns, M., Schwartz, M. K., and Oettgen, H. F. (1978) Preoperative carcinoembryonic antigen level as a prognostic indicator in colorectal cancer. N. Engl. J. Med. 298, 448–451
9. Minton, O., and Stone, P. C. (2010) Review: the use of proteomics as a research methodology for studying cancer-related fatigue: a review. Palliat. Med. 24, 310–316
10. Leth-Larsen, R., Lund, R. R., and Ditzel, H. J. (2010) Plasma membrane proteomics and its application in clinical cancer biomarker discovery.
In order to cite this article properly, please include all of the following information: 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(4):M110.003087. DOI: 10.1074/mcp.M110.003087.