Identification of Putative Immunologic Targets for Colon Cancer Prevention Based on Conserved Gene Upregulation from Preinvasive to Malignant Lesions

Elizabeth K. Broussard1, Rachel Kim1, Jesse C. Wiley1, Juan Pablo Marquez1, James E. Annis3, David Pritchard2, and Mary L. Disis1

Abstract

The length of time required for preinvasive adenoma to progress to carcinoma, the immunogenicity of colorectal cancer (CRC), and the identification of high-risk populations make development and testing of a prophylactic vaccine for the prevention of CRC possible. We hypothesized that genes upregulated in adenoma relative to normal tissue, which maintained increased expression in CRC, would encode proteins suitable as putative targets for immunoprevention. We evaluated existing adenoma and CRC microarray datasets and identified 160 genes that were ≥2-fold upregulated in both adenoma and CRC relative to normal colon tissue. We further identified 23 genes that showed protein overexpression in colon adenoma and CRC based on literature review. Silencing the most highly upregulated genes, CDH3, CLDN1, KRT23, and MMP7, in adenoma and CRC cell lines resulted in a significant decrease in viability (P < 0.0001) and proliferation (P < 0.0001) as compared to controls and an increase in cellular apoptosis (P < 0.05 for CDH3, KRT23). Results were duplicated across cell lines representing microsatellite instability, CpG island methylator, and chromosomal instability phenotypes, suggesting immunologic elimination of cells expressing these proteins could impact the progression of all CRC phenotypes. To determine whether these proteins were immunogens, we interrogated sera from early stage CRC patients and controls and found significantly elevated CDH3 (P = 0.006), KRT23 (P = 0.0007), and MMP7 (P < 0.0001) serum immunoglobulin G in cases as compared to controls. These data show a high throughput approach to the identification of biologically relevant putative immunologic targets for CRC and identified three candidates suitable for vaccine development. Cancer Prev Res; 6(7); 666–74. ©2013 AACR.

Introduction

Colorectal carcinoma (CRC) arises from progressive genetic alterations, and the progression from normal epithelium to adenoma to invasive carcinoma occurs over years (1). This time period offers a window for preventative interventions that target high-risk patients, such as those with a prior history of CRC, patients with multiple adenomas, or adenomas with a high risk of malignant transformation (i.e., villous histology or size >1 cm). Chemoprevention targeting adenomas is feasible and effective, as showed in multiple clinical trials of nonsteroidal anti-inflammatory drugs (NSAID) including selective COX-2 inhibitors and aspirin (2–7), but widespread application of the approach has been limited by the lack of patient compliance associated with daily drug dosing, adverse cardiovascular events, and gastrointestinal bleeding (8–11).

Active immunization as a cancer immunoprevention strategy offers several advantages to classic drug-based approaches. First, vaccines are administered over a short period of time without the need for daily dosing. Moreover, if successful, immunologic memory is generated ensuring a destructive adaptive cellular immune response poised to eliminate aberrant cells at the time they arise. Immune elimination of cells that had aberrantly expressed proteins, which were associated with cell growth or survival, could potentially impact CRC development or progression. T-memory cells, once vaccine primed, are active for years and can be boosted periodically with further periodic vaccinations. In addition, vaccines have largely been shown to be nontoxic. There have been numerous clinical studies immunizing cancer bearing patients against proteins expressed in the colon with limited to no adverse events (12–16). Prophylactic vaccines have had remarkable success in...
preventing cancers of viral origin, such as hepatitis B (17) and human papillomavirus (18). A major barrier to the extrapolation of the success seen with prevention of viral malignancies to the prevention of CRC has been the lack of well defined, biologically relevant, and immunogenic proteins expressed early in the malignant transformation of the disease.

Data presented here shows a high throughput approach, combining both expression array analysis and siRNA screening, for the identification of proteins expressed in both adenomas and invasive carcinomas that impact cell growth and senescence. Further evaluation for immunogenicity via serologic screening for immunoglobulin G (IgG) antibodies can identify candidate antigens that are capable of stimulating an adaptive immune response.

Materials and Methods

Dataset selection for adenoma and colon carcinoma genes

We entered search terms "Human Colorectal Carcinoma" or "Human Colorectal Adenoma" in Gene Expression Omnibus or GEO (19) and Array Express (20). We excluded colorectal carcinoma (CRC; n = 307) and adenoma (n = 47) datasets based on a prospectively defined exclusion algorithm (Supplementary Fig. S1). Of the 9 remaining datasets, 3 examined gene expression in 53 adenoma samples and 8 examined gene expression in 437 CRC samples. The sample distribution included 57% stage I/II, 15% stage III, and 28% stage IV samples (Supplementary Table S1).

Dataset validation

We conducted box plot analyses of each log2-transformed and normalized dataset, and confirmed alignment of median values across the arrays and the similarity of the interquartile range using R (v 2.14.2). Bioconductor limma module (3.11.1), and the OneChannelGUI R interface (v 1.2.2.2). Box plotting of GSE 15960 is shown in Supplementary Fig. S2A. Hierarchical clustering linkage (Supplementary Fig. S2B) and principal component analysis (Supplementary Fig. S2C) showed distinct expression profiles across CRC and adenoma samples and 8 examined gene expression in 437 CRC samples. The sample distribution included 57% stage I/II, 15% stage III, and 28% stage IV samples (Supplementary Table S1).

Cell lines

CRC phenotypes are represented by the following 6 cell lines: microsatellite unstable (MSI) by HCT116 (ATCC CCL-247) and LoVo (ATCC CCL-229) and, CpG island methylator phenotype (CIMP) by RKO (ATCC CRL-2577) and SW48 (ATCC CCL-231), and chromosome instability (CIN) by FET (donated by Dr. W. Grady, Fred Hutchinson Cancer Research Center, Seattle, WA) and SW480 (ATCC CCL-228). Adenomas are represented by the cell line AAC/SB10 (donated by Dr. C. Paraskeva, University of Bristol, Bristol, UK). ATCC conducts authentication testing using short tandem repeat (STR) DNA profiling of all cell lines, and cells received (HCT116, LoVo, RKO, SW48, and SW480) were used in all assays within 6 months of receipt or resuscitation. FET cells were authenticated by a tetranucleotide repeats assay, AAC/SB10 cells were authenticated by genotyping, and all cell lines were tested for mycoplasma (Agilent Technologies) before the use in all assays, and within 6 months of resuscitation. Cells were maintained at 37°C and 5% CO2. Growth media for cell lines: DMEM + 4.5g/L-glucose + 1-glutamine (AAC/ SB10), DMEM + 4.5g/L-glucose + 1-glutamine + NaPyr (FET), McCoy’s 5A medium modified (HCT116), F-12K medium (LoVo), Eagle’s MEM (RKO), Leibovitz’s L-15 medium (SW48, SW480). All media were supplemented with 1% FBS (Benchmark) and penicillin/streptomycin (Cellgro).

siRNA and transfection

On day 1, cells were plated into a 96-well flat bottom plate (Corning). On day 2, cells were transfected with 10 μmol/l. siRNA (Sigma) using lipofectamine RNAimax reagent (Invitrogen). We targeted CDH3, KRT23, MMP7, and CLDN1 by a pool of 3 unique siRNA dimers targeting the same gene, pooled at equa molarity (final concentration 125 nmol/L). A target sequences for the siCDH3 duplex were #1 5’-CCAAUIUUCUCGCUUUGAA[dT, dT]-3’, #2 5’-CAACUUUI[UAUUUUUGA][dA][dT, dT]-3’, #3 5’-GU[UUACUGGUAAUUGAI][dT, dT]-3’; siCLDN1 duplex were #1 5’-CAGUCUAGGCGGAGGAGC[AG][dT, dT]-3’, #2 5’-UGACGAA[GAUUUUGGUCAUGG][dA][dT, dT]-3’, #3 5’-[CAGAUUCAGUCC][AG][dT, dT]-3’; siKRT23 duplex were #1 5’-CUCUGAU[UAUUUCG][dT, dT]-3’, #2 5’-GAUUCA[AA]CUGGACGUA[UAUUUAU][dT, dT]-3’, #3 5’-[AGUIAGGAAGGAGCGG][A][dT, dT]-3’, siMMP7 duplex were #1 5’-CCUIUICUIUUCGGA[GG][dT, dT]-3’, #2 5’-CAAAUUCAGGAGA[U][A][C][A][dT, dT]-3’, #3 5’-GAUGUGAGAGCGAAGGAGGGA[AG][dT, dT]-3’. To assess the impact of targeting more than one gene at a time, we combined all 4 siRNA sequences (siCDH3, siCLDN1, siKRT23, and siMMP7) into Illumina platforms, probes or probe sets were collapsed to the HUGO gene symbol present on the platform annotation. The gene symbols were intersected and we identified 14,915 genes present across all array platforms and experiments, and calculated the average differential gene expression. We identified 631 genes that were upregulated in adenoma and 490 genes that were upregulated in CRC: 160 genes were commonly upregulated in both the adenoma and CRC datasets (Table S2).
a pooled siRNA. All assays were conducted in quadruplicate. Transfections were conducted with negative control PBS and nontargeting control siRNA (MISSION SiRNA Universal Negative Control; Sigma), and positive control Kif11s (Sigma).

**mRNA quantitation**

RNA was collected 48 hours after transfection and reverse transcribed using SuperScript III First-Strand Synthesis System (Invitrogen). Real-time PCR was done on an Applied Biosystems 7900. Primers for CDH3, CLDN1, KRT23, and MMP7 were purchased from Applied Biosystems. All assays were conducted in triplicate, and mRNA quantitated after normalization for β-actin. Significance was calculated using the unpaired 2-tailed Student t test. We confirmed that siRNA targeting all tested genes resulted in significant (all P values < 0.05) mRNA reduction in all cells compared to nontargeting siRNA (Supplementary Fig. S3). Specifically, siCDH3 resulted in mRNA reductions of 58.4 ± 7.7% to 98 ± 0.21% compared to control siRNA (Supplementary Fig. S3A). For siCLDN1, mRNA reductions of 61.1 ± 8.5% to 89.8 ± 2.3% were achieved (Supplementary Fig. S3B). For siKRT23 mRNA reductions of 60.7 ± 20.2% to 97.6 ± 0.25% were achieved (Fig. S3C), and for siMMP7 mRNA reductions of 63.2 ± 4% to 96.1 ± 0.95% were achieved (Supplementary Fig. S3D).

**Cell viability, proliferation, and apoptosis**

FET, LoVo, and SW480 cells (1,000 cells/well), RKO (500 cells/well), SW48 (2,400 cells/well), HCT116 (4,000 cells/well), and AAC/SB10 (8,000 cells/well) were seeded in 96-well plates (Corning). Cell viability was determined at day 7 (Fig. 1) with Resazurin (Sigma) and quantitated using the Perkin-Elmer Wallac EnVision 2104 Multilabel Detector/Plate Reader at 600 nm (21). Proliferation was quantitated at 48 hours by proliferating cell nuclear antigen (PCNA) protein expression (Fig. 2), relative to expression in cells transfected with control nontargeting siRNA. Apoptosis measurements were optimized at 48 hours for AAC/SB10, LoVo, and RKO cells and at 72 hours for FET, HCT116, SW48, and SW480 cells using Caspase-Glo 3/7 (Promega), and luminescence was measured (Fig. 3) using the Perkin-Elmer Wallac EnVision 2104 Multilabel Detector/Plate Reader (21,22). All data are expressed as mean ± SD of cells within the specific phenotype (MSI: HCT116, LoVo; CIN: SW48, RKO; CIMP: FET, SW480; Adenoma: AAC/SB10).

**Western blot analysis**

For detection of serum IgG antibody responses in patients and controls, the serum was diluted 1:200. For quantitation of protein expression in cell lines, nontransfected and transfected cells were seeded in 6-well flat bottom plates (Corning). After 48 hours, the cells were lysed, protein quantitated using BCA Protein Assay Kit (Pierce Biotechnology) and boiled in Laemmli buffer (Bio-Rad). Proteins from cell lines or recombinant proteins were run on 10% SDS-PAGE gels, transferred to nitrocellulose, blocked with 3% milk, and incubated with antibody overnight at 4°C: CDH3 (polyclonal TA302124; Origene), CLDN1 (clone 1C5-D9; Novus), KRT23 (polyclonal 22460002; Novus), MMP7 (polyclonal NB300-1000; Novus), PCNA (polyclonal ab2426; Abcam). Recombinant human proteins were used as positive controls: CDH3 (Origene), CLDN1, KRT23,
and MMP7 (all from Abnova). Untreated HCT116 cells were lysed and loaded as PCNA positive control (23). Membranes were incubated with goat anti-mouse IgG-HRP or goat anti-rabbit IgG-HRP antibody (Invitrogen), and developed (Amersham Biosciences). The intensity of bands was quantitated (Image J), normalized to tubulin (monoclonal 11H10; Cell Signaling Technology), and all assays were done in triplicate. Data are expressed as percent change relative to protein expression in nontargeting csiRNA. Each colon cancer cell line was evaluated individually but results are combined and grouped by phenotype (MSI: LoVo, HCT116; CIMP: RKO, SW48; CIMP: FET, SW480). The Western blot validated that siRNA specifically and markedly reduced protein expression of CDH3, CLDN1, KRT23, and MMP7 (Fig. S4A–G). Specifically, siCDH3 reduced protein expression by 38.8 ± 9.7% to 53.1 ± 3.8%, siCLDN1 by 33 ± 1.1% to 73.4 ± 2.8%, siKRT23 by 19.8 ± 5.3% to 33.8 ± 1.9%, and siMMP7 by 33 ± 7.9% to 76.5 ± 9.3%.

Figure 2. siRNA silencing of CDH3, CLDN1, KRT23, and MMP7 (A–D) in adenoma and CRC cell lines significantly reduces cell proliferation. PCNA protein was quantitated in transfected CRC and adenoma cells. All assays are done in triplicate, cell lines are grouped by phenotype (MSI: HCT116, LoVo; CIMP: RKO, FET, SW480), and results are normalized to tubulin and PCNA expression in csiRNA. Error bars note SD, and calculated P-values are for differences in PCNA expression in csiRNA and each phenotype. csiRNA (control nontargeting siRNA), **P < 0.0001.

Figure 3. siRNA silencing of CDH3, CLDN1, KRT23, and MMP7 (A–D) in adenoma and CRC cell lines induces apoptosis. Transfected CRC and adenoma cells were assayed for cellular apoptosis and results were normalized to nontransfected cells (NT). All assays were conducted in quadruplicate, and cell lines are grouped by phenotype (MSI: HCT116, LoVo; CIMP: RKO, FET, SW480). Calculated P-values are for differences in apoptosis between NT and each phenotype. Error bars note SD. NT, nontransfected cells (transfection with PBS), *P < 0.05, **P < 0.0001.
Evaluation of humoral immunity specific for CDH3, CLDN1, KRT23, and MMP7

Indirect ELISA was conducted as previously described (24) with the following modifications: recombinant proteins CDH3, CLDN1, KRT23, and MMP7 were diluted with carbonate buffer to a concentration of 1.0 mg/mL. A sample was defined as positive when serum IgG value was greater than the mean and 2 SDs of the control sera (n = 25) evaluated for each protein (Fig. 4A, C, and E). Positive responses were verified by Western blot analysis (Fig. 4B, D, and F). The sensitivity and specificity were 100% and 100% for CHD3, 100% and 62.5% for KRT23, and 100% and 100% for MMP7. The CLDN1 ELISA results could not be validated by Western blotting, and therefore the results are not shown.

Human subjects

The CRC patients (n = 25) ranged in age from 45 to 89 (median age 66.5), and 40% were female. Stage 1 (40%) and stage 2 (60%) CRC patient sera were included (HSD# 19394; University of Washington, Seattle, WA and Innovative Research). The normal donors (n = 25) ranged in age from 23 to 84 (median age 61.7), and 48% were female (Puget Sound Blood Bank, Seattle, WA). All donors met criteria for blood donation and informed consent was obtained from each subject. All sera were aliquotted and stored at −80°C.

Statistical analysis

Differences in cellular viability, proliferation, apoptosis, protein and RNA expression, and human serum IgG responses were assessed using the unpaired 2-tailed Student t test. Differences in the incidence of positives in serum IgG responses were quantitated using Fisher exact test. The significance was set at P < 0.05 (GraphPad Prism v 5.0).

Results

Genes that are upregulated encoding overexpressed proteins and are conserved from adenomas to invasive carcinomas can be identified

We identified 631 genes from the adenoma datasets and 490 genes from the CRC datasets. A total of 160 genes (Supplementary Table S2) are present in both datasets. Of note, 2 of the most common CRC antigens, CEA and MUC1, did not meet the selection criterion of at least 2-fold overexpression with a P-value less than 0.05 in any of the
datasets evaluated. We have showed that overexpressed tumor-associated proteins can be immunogenic (24). For each of the 160 genes upregulated in adenoma and carcinoma, we searched PubMed (25) for protein overexpression using the terms "(gene name) human protein colorectal overexpression" or "(gene name) human protein colorectal adenoma overexpression." As of October 1, 2012, 65 published papers (Supplementary Table S3) document protein overexpression. MSIs, CINs, and adenomas showed a 49% to 89% reduction in viability after knockdown of all 4 genes (all P values < 0.0001). Specifically, relative to nontransfected

cells, cells transfected with siCDH3 showed reduced viability: 53.3 ± 8.5% (MSI), 75.5 ± 6.6% (CIMP), 81.0 ± 8.4% (CIN), and 88.9 ± 2.4% (adenoma; Fig. 1A). In cells transfected with siCLDN1 viability decreased by 69.9 ± 6.0% (MSI), 84.1 ± 3.6% (CIMP), 78.6 ± 12.7% (CIN), and 85.6 ± 2.0% (adenoma; Fig. 1B). In cells transfected with siKRT23 viability decreased by 57.8 ± 10.2% (MSI), 58.8 ± 7.4% (CIMP), 69.4 ± 13.3% (CIN), and 88.4 ± 1.0% (adenoma; Fig. 1C). In cells transfected with siMMP7 viability decreased by 49.4 ± 7.0% (MSI), 58.1 ± 10.6% (CIMP), 52.5 ± 13.0% (CIN), and 87.5 ± 2.1% (adenoma; Fig. 1D). Finally, in cells transfected with pooled siRNA from all 4 candidates, viability was similarly decreased by 50.9 ± 8.9% (adenoma), 49.3 ± 5.4% (CIMP), 48.7 ± 23.8% (CIN), and 83.7 ± 1.2% (adenoma; Supplementary Fig. S5).

siRNA silencing of CDH3, CLDN1, KRT23, and MMP7 in adenoma and CRC cell lines significantly reduces cell proliferation

siRNA silencing of all 4 genes significantly (all P-values < 0.0001) inhibited cellular proliferation in adenoma and all CRC phenotypes. Specifically, siCDH3 decreased PCNA expression by 56.9 ± 3.7% (MSI), 67.3 ± 3.0% (CIMP), 55.5 ± 5.7% (CIN), and 65.7 ± 7.2% (adenoma; Fig. 2A) relative to PCNA expression in control siRNA. The siCLDN1 inhibited proliferation by 50 ± 2.9% (MSI), 61.4 ± 3.9% (CIMP), 41.4 ± 3.4% (CIN), and 63.1 ± 3.7% (adenoma; Fig. 2B). The siKRT23 reduced proliferation by 50.7 ± 6.3% (MSI), 51.1 ± 4.1% (CIMP), 49.7 ± 5.9% (CIN), and 50.6 ± 5.9% (adenoma; Fig. 2C). Finally, the siMMP7 inhibited PCNA expression by 49.2 ± 2.8% (MSI), 53.7 ± 4.1% (CIMP), 48.3 ± 36.6% (CIN), and 53.1 ± 5.5% (adenoma) compared to control (Fig. 2D).

siRNA silencing of CDH3, CLDN1, KRT23, and MMP7 in adenoma and CRC cell lines induces apoptosis

siRNA silencing of CDH3 increased cellular apoptosis in MSI (1.48 ± 0.2, P = 0.008), CIMP (1.66 ± 0.2, P = 0.0007), CIN (1.16 ± 0.05, P = 0.009), and adenoma cells (2.54 ± 0.4, P < 0.0001; Fig. 3A) relative to control of nontransfected cells. Knockdown of CLDN1 increased apoptosis in MSI (1.19 ± 0.1, P = 0.13), CIMP (1.57 ± 0.1, P < 0.0001), CIN (1.35 ± 0.1, P = 0.0001), and adenoma (1.68 ± 0.1, P < 0.0001; Fig. 3B). Silencing KRT23 increased apoptosis in MSI (1.35 ± 0.1, P = 0.03), CIMP (1.54 ± 0.1, P < 0.0001), CIN (1.31 ± 0.1, P = 0.0001), and adenoma (2.04 ± 0.1, P < 0.0001; Fig. 3C). Similarly, knockdown of MMP7 induced apoptosis in MSI (1.42 ± 0.3, P = 0.26), CIMP (1.98 ± 0.1, P < 0.0001), CIN (1.45 ± 0.1, P < 0.0001), and adenoma cells (2.45 ± 0.1, P < 0.0001) compared to control (Fig. 3D).

CDH3, KRT23, and MMP7 are immunogenic in patients with early stage CRC

We next determined whether these proteins were immunogenic by determining whether antigen-specific IgG could be detected. The serum responses in CRC patients to CDH3 were higher than serum responses in control patients (mean 2.75 ± 0.4 µg/mL vs. 1.58 ± 0.2 µg/mL, P = 0.006; Fig. 4A).

Table 1. Genes showing protein overexpression in colon CRC or AD

<table>
<thead>
<tr>
<th>Gene</th>
<th>Protein overexpression</th>
<th>Fold increase (CRC)</th>
<th>Fold increase (AD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDH3</td>
<td>AD, CRC</td>
<td>21</td>
<td>31</td>
</tr>
<tr>
<td>KRT23</td>
<td>CRC</td>
<td>16</td>
<td>3</td>
</tr>
<tr>
<td>MMP7</td>
<td>AD, CRC</td>
<td>13</td>
<td>23</td>
</tr>
<tr>
<td>CLDN1</td>
<td>AD, CRC</td>
<td>12</td>
<td>5</td>
</tr>
<tr>
<td>ETV4</td>
<td>AD, CRC</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>CLDN2</td>
<td>AD, CRC</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>LGR5</td>
<td>AD, CRC</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>SLC01B3</td>
<td>CRC</td>
<td>5</td>
<td>13</td>
</tr>
<tr>
<td>TGFBI</td>
<td>CRC</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>SLC7A5</td>
<td>CRC</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>STC2</td>
<td>CRC</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>FABP6</td>
<td>AD, CRC</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>LCN2</td>
<td>CRC</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>TROP2</td>
<td>CRC</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>BMP4</td>
<td>AD, CRC</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>PTP4A3</td>
<td>AD, CRC</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>SLC6A14</td>
<td>CRC</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>S100A11</td>
<td>AD, CRC</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>TNS4</td>
<td>CRC</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>WNT2</td>
<td>CRC</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>CA9</td>
<td>AD, CRC</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>FKBPI0</td>
<td>CRC</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>IFITM3</td>
<td>AD, CRC</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

Abbreviation: AD, adenoma.
The incidence in CRC patients was higher than in controls (52% vs. 0%, \( P < 0.0001 \)). The serum responses in CRC patients to KRT23 were higher than serum responses in controls (mean 1.42 \pm 0.2 \mu g/mL vs. 0.54 \pm 0.1 \mu g/mL, \( P = 0.0007 \); Fig. 4B). The incidence in CRC patients was higher than in controls (24% vs. 8%, \( P < 0.0001 \)). The serum responses in CRC patients to MMP7 were higher than serum responses in controls (mean 1.74 \pm 0.3 \mu g/mL vs. 0.43 \pm 0.1 \mu g/mL, \( P < 0.0001 \); Fig. 4C). The incidence in CRC patients was higher than in controls (40% vs. 4%, \( P < 0.0001 \)). Antigen-specific antibody responses detected in either control donors or CRC patients could be validated by Western blot (Fig. 4B, D, F).

Discussion

Few immunogenic proteins have been identified for immunologic interventions in CRC. Furthermore, the biologic relevance of some of the most commonly exploited antigens in CRC, such as MUC-1 and CEA, is not well elucidated (13,14,16). Data presented here shows that existing microarray datasets provide a rapid method to identify genes upregulated in adenomas that maintain increased expression in carcinomas and that these genes encode proteins that are overexpressed in both adenoma and CRC. Silencing adenoma–CRC gene expression with siRNA in multiple cell lines suggests these proteins may impact cell viability, proliferation, and survival across all CRC phenotypes. Finally, we show that overexpressed proteins conserved from adenoma to CRC have the potential to stimulate an adaptive immune response in either control donors or CRC patients. This latter observation suggests that tolerance has been circumvented to these self-proteins and that boosting immunity against these proteins with active immunization may be possible.

The development of CRC is a complex process involving multiple molecular pathways. Selective pathways active in advanced disease are already initiated in early tumorogenesis (i.e., adenomas) and are, thus, rational targets for chemoprevention as well as therapy (26). Prior published studies have primarily focused on expression of genes in CRC relative to normal colon tissue. Of the 27 studies of gene expression profiling in CRC published between 1998 and 2008 (27), only 4 studies included colon adenoma (28–31) and only 2 studies examined gene expression across the continuum of normal tissue, adenoma, and CRC (28,30). Taken together these 2 studies included 13 adenoma, 29 CRC, and 42 normal colon samples, and identified only 56 genes that were upregulated in adenoma and CRC although MMP7 and CDH3, 2 of the proteins presented here, were identified in each of the studies. More recent publications have combined expression data from normal, adenoma, and CRC, but these investigations examined similarly small samples sizes (\( n = 16 \)) or focused on the differential expression of individual genes (32–34). Here we present the largest sample size analyzed to date that includes normal, adenoma, and CRC tissues (153 normal, 53 adenoma, 437 CRC), filtered through a prospectively defined exclusion algorithm and stringent quality control metrics, resulting in the identification of 160 genes as potential chemoprevention targets (Supplementary Table S2). Neither CEA nor MUC1, 2 common CRC targets used for immune modulation, met the expression criteria outlined. These 160 candidate genes may have great potential as prevention or therapeutic targets. Of the 23 proteins (Table 1) we identified from published studies as overexpressed in adenoma and/or CRC, overexpression of the following 8 proteins was an independent predictor of significantly decreased overall survival in CRC: CA9, CDH3, ETV4, LCN2, MMP7, PTP4A3, TNS4, and TROP2 (35–42). Association of protein expression with poor prognosis suggests that such expression confers a more aggressive phenotype. The potential biologic relevance of these genes was further supported by silencing gene expression in adenoma and CRC cell lines showing a significantly reduced viability, decreased proliferation, and increased apoptosis. The observation that these functional effects were showed across all 3 CRC phenotypes suggests that immune elimination of cells expressing these proteins could potentially have some impact on tumor growth or progression. MSI is present in 15% of CRC, CIMP tumors represent approximately 20% of CRC, and CIN is the most common phenotype, representing 50% to 85% of all CRC (43).

Active immunization against overexpressed cancer-related proteins can result in elimination of premalignant cells. Immunizing against ductal carcinomas in situ (DCIS) with a vaccine targeting HER2, an overexpressed protein present on the majority of DCIS, resulted in resolution of lesions in some women or eradication of HER2-expressing cells (44). Clinical data in the prevention or high-risk setting does not yet exist for CRC; however, preclinical studies targeting a single antigen MUC-1 suggest the immune prevention of CRC is feasible. MUC1 is a cell surface associated glycoprotein overexpressed in the premalignant environment of inflammatory bowel disease (IBD) and in CRC. In MUC-1 transgenic mice, immunization against MUC1 generated both anti-MUC1 IgG and MUC1-specific cytotoxic T cells. Immunized animals showed a significant delay in the development of IBD as measured by the development of rectal prolapse (\( P = 0.043 \) as compared to controls), and 80% of immunized animals had complete protection against the development of CRC (45). Studies by our group have showed that overexpression of a self protein is an independent predictor of immunogenicity in multivariate analysis (46). Immunogenicity of CDH3, KRT23, and MMP7 in either cancer patients or control donors shows, for some reason, that tolerance has been circumvented to these antigens in some individuals. The observation that CDH3, KRT23, and MMP7 antibodies are found in higher levels in CRC patients than controls suggests that the presence of malignancy may increase exposure to these antigens resulting in further stimulation of specific immunity. If high levels of immunity were induced earlier in the transformative process from adenoma to carcinoma, perhaps the progression to invasive disease could be prevented or slowed. These 3 genes...
function in the pathogenesis of intestinal tumor development, suggesting that they could be appropriate targets for immunoprevention. CDH3 is involved in maintaining cellular localization and tissue integrity. Epigenetic demethylation of the CDH3 promoter permits its ectopic expression very early in the colorectal adenoma–carcinoma sequence and persistence during invasive cancer. Induced expression of CDH3 in mucosal damage leads to an increased rate of crypt fission, a common feature of clonal expansion in gastrointestinal dysplasia (47). Upregulation of CDH3 in cancer is associated with increased proliferation (48). KRT23 is responsible for the structural integrity of epithelial cells, and important in modulating and controlling cellular signaling processes and apoptosis (49). KRT23 expression differentiates between microsatellite-stable (MSS) and microsatellite-instable (MSI) colon cancers (50), with 88% of MSI tumors negative for KRT23 and 70% of MSS tumors with KRT23 overexpression. MMP7 is involved in the breakdown of extracellular matrix by degradation of basement membrane proteins laminin and collagen IV. MMP7 expression is correlated with tumor malignancy and liver metastasis of CRC (51).

Immunization programs against hepatitis B and human papillomavirus to prevent hepatocellular carcinoma and cervical, vulvar, and vaginal cancers, respectively, have been implemented worldwide and show significant clinical efficacy (17,18). Vaccines are important to prophylaxis, as they can generate immunologic memory, which would result in the elimination of cells that begin to develop a cancer "phenotype." To date targeting such self proteins, including MUC1, CEA, and HER2, has been nontoxic and safe (13,14,16,52,53). The identification of biologically relevant antigens expressed early in the oncogenic process lays the foundation for the further testing of immunoprevention for CRC.

Disclosure of Potential Conflicts of Interest
M.L. Disis has ownership interest (including patents) in University of Washington. No potential conflicts of interest were disclosed by the other authors.

Authors’ Contributions
Conception and design: E.K. Broussard, J.C. Wiley, D.K. Pritchard, M.L. Disis
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): E.K. Broussard, R.Y. Kim, M.L. Disis
Writing, review, and/or revision of the manuscript: E.K. Broussard, R.Y. Kim, J.C. Wiley, D.K. Pritchard, M.L. Disis
Study supervision: E.K. Broussard, M.L. Disis

Acknowledgments
The authors thank Dr. W. Grady for donating the FET cell line, and Dr. C. Paraskeva for donating the AAC5B10 cell line.

Grant Support
This publication was supported by HHSN261-200531000C N01-CN-53300 (M. Disis) and by the National Center For Advancing Translational Sciences of the National Institutes of Health under Award Number K12 TR000421 (E. Broussard). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. Additional funding was provided from the Sonora Cancer Research Center SON-2011-1-02/CEN and the Marcos & Saide Cancer Foundation RASAMA-112MEX/PC1976 (J.P. Marquez).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received December 13, 2012; revised April 30, 2013; accepted May 1, 2013; published OnlineFirst May 16, 2013.

References
Broussard et al.


20. http://www.ebi.ac.uk/arrayexpress/


