

## Proteomic Consequences of a Single Gene Mutation in a Colorectal Cancer Model

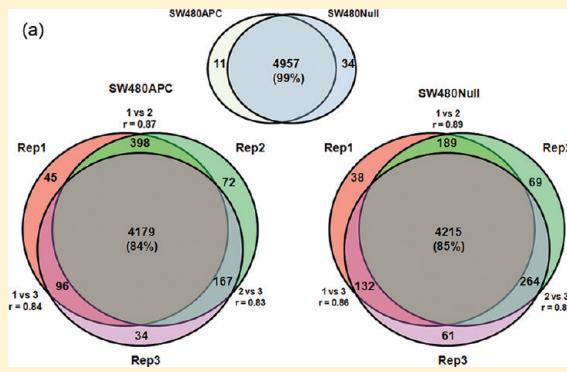
Patrick J. Halvey,<sup>†,‡</sup> Bing Zhang,<sup>§</sup> Robert J. Coffey,<sup>||,⊥</sup> Daniel C. Liebler,<sup>†,‡</sup> and Robbert J. C. Slebos<sup>\*,†,||</sup>

<sup>†</sup>Department of Biochemistry, <sup>‡</sup>Jim Ayers Institute for Precancer Detection and Diagnosis, <sup>§</sup>Department of Biomedical Informatics, <sup>||</sup>Department of Cancer Biology, and <sup>⊥</sup>Department of Medicine, Vanderbilt University School of Medicine, Nashville, Tennessee 37232-6350, United States

Supporting Information

**ABSTRACT:** The proteomic effects of specific cancer-related mutations have not been well characterized. In colorectal cancer (CRC), a relatively small number of mutations in key signaling pathways appear to drive tumorigenesis. Mutations in adenomatous polyposis coli (APC), a negative regulator of Wnt signaling, occur in up to 60% of CRC tumors. Here we examine the proteomic consequences of a single gene mutation by using an isogenic CRC cell culture model in which wildtype APC expression has been ectopically restored. Using LC–MS/MS label free shotgun proteomics, over 5000 proteins were identified in SW480Null (mutant APC) and SW480APC (APC restored). We observed 155 significantly differentially expressed proteins between the two cell lines, with 26 proteins showing opposite expression trends relative to gene expression measurements. Protein changes corresponded to previously characterized features of the APCNull phenotype: loss of cell adhesion proteins, increase in cell cycle regulators, alteration in Wnt signaling related proteins, and redistribution of  $\beta$ -catenin. Increased expression of RNA processing and isoprenoid biosynthetic proteins occurred in SW480Null cells. Therefore, shotgun proteomics reveals proteomic differences associated with a single gene change, including many novel differences that fall outside known target pathways.

**KEYWORDS:** shotgun proteomics, label-free quantitation, multiple reaction monitoring (MRM), LC–MS/MS, colorectal cancer, APC



### INTRODUCTION

Mass spectrometry-based shotgun proteomics is a powerful tool for the interrogation of biological systems. Recent technological advances allow thousands of proteins to be routinely identified from submilligram quantities of tissues or cells.<sup>1</sup> Increased depth of proteome coverage lends itself to the investigation of molecular signatures, which underlie complex disease processes. For instance, in cancer, thousands of mutations have been identified but the precise relationship between genomic variation and cancer phenotype remains largely unclear.<sup>2</sup> Individual mutations may bring about proteomic changes that otherwise would not be predicted based on known gene function. In-depth proteomic analysis is required to define the broader landscape of protein networks and signaling pathways associated with cancer-related mutations.

Determination of the proteomic consequences of individual mutations in tissues presents a major challenge due to the tremendous genetic heterogeneity between individual patient tumors.<sup>3</sup> Isogenic cell lines present an attractive model system, wherein the phenotypic effects of a single gene defect can be assessed in the context of uniform background gene/protein expression. Colorectal cancer (CRC) cell models are ideally suited

for the examination of proteomic effects of single gene changes. Several mutations in key signaling pathways have been identified in CRC. These “driver mutations” have been mapped to specific events during tumor progression from normal epithelia to carcinoma.<sup>4</sup> One essential driver mutation occurs in adenomatous polyposis coli (APC), and is among the most frequently observed genetic aberrations in colorectal adenomas and cancers.<sup>5</sup> APC mutations (or allelic loss) occur in 70–80% of adenomas and carcinomas.<sup>6,7</sup> While many other amino acid-changing mutations are associated with CRC, most can be classified as “passenger” mutations, that is, they do not confer a selective growth advantage on the tumor and are propagated randomly by clonal expansion.<sup>4,8</sup> Therefore, by focusing on proteomic changes associated with a single driver mutation, one may decipher the protein subnetworks that underlie tumor development.

APC occurs in complex with Axin (AXIN1) and glycogen synthase kinase-3 $\beta$  (GSK3B) and plays a role in the proteolytic degradation of  $\beta$ -catenin (CTNNB1). Decreases in CTNNB1 levels result in the inhibition of TCF/LEF-mediated transcription,

Received: September 8, 2011

Published: November 21, 2011

a major component of the Wnt signaling pathway.<sup>9</sup> The colon tumor cell line SW480 contains a mutation at position Q1338 in the APC coding sequence, which generates a premature stop codon.<sup>10</sup> The remaining wild-type allele is deleted. An APC-corrected version of SW480 (SW480APC) was generated by stable transfection with wildtype APC.<sup>11</sup> Relative to APC mutant cells (SW480Null), SW480APC cells exhibit decreased proliferation rates, decreased nuclear localization of  $\beta$ -catenin (CTNNB1), increased translocation of epithelial cadherin (CDH1) to the plasma membrane, and enhanced cell–cell adhesion.<sup>11</sup> Thus, the SW480APC model displays unique phenotypic properties that may map to underlying proteomic expression signatures.

Here we employ two complementary proteomic technology platforms that provide the means to systematically characterize APC-driven proteomic differences. The first, shotgun proteomics, employs liquid chromatography–tandem mass spectrometry (LC–MS/MS) and provides a nondirected, global inventory of proteomes, together with quantitative assessments of protein abundances.<sup>12,13</sup> The second proteomics approach is targeted analysis of individual proteins by multiple reaction monitoring (MRM) mass spectrometry intensity measurement of their constituent peptides.<sup>14,15</sup> The data illustrate mechanisms by which APC inactivation may alter cellular functions by perturbing the composition of cell proteomes. More generally, we demonstrate the ability of label free shotgun proteomics to assess proteomic consequences of a single gene difference.

## ■ EXPERIMENTAL METHODS

### Cell Culture and Subcellular Fractionation

SW480APC and SW480Null were a kind gift from Antony Burgess (Ludwig Institute, Melbourne, Australia). Cells were grown in RPMI 1640 medium, supplemented with 10% fetal bovine serum, 1% penicillin/streptomycin and genetecin (1.5 mg/mL). Three biological replicate cultures were harvested approximately 1 week apart and these replicates were processed separately and independently through the complete analysis. Growth medium was aspirated, cells were washed once in 1  $\times$  PBS and collected in 1  $\times$  PBS, then centrifuged at 300  $\times$  g for 5 min and the supernatant was discarded. Cell pellets were stored at –80 °C until cell lysis could be carried out.

For subcellular fractionation, cell pellets were resuspended in buffer A (10 mM HEPES, pH 7.9, 10 mM KCl, 2 mM MgCl<sub>2</sub>, 0.2 mM EDTA) containing protease inhibitors (aprotinin, leupeptin) and incubated at 4 °C for 20 min. IGEPAL CA-630 (Sigma-Aldrich) was added to a final concentration of 0.5% v/v, samples were incubated at 4 °C for 10 min and centrifuged at 300  $\times$  g for 15 min at 4 °C. The supernatant (cytosolic fraction) was collected in a separate tube. Pellets were resuspended in Buffer A, centrifuged at 2000  $\times$  g for 15 min at 4 °C and the supernatant (cytosolic fraction) was combined with supernatant from the previous step. The nuclear pellet was stored at –80 °C and the cytosolic fraction was lyophilized overnight. Subsequent lys/digestion of cytosolic lyophilates and nuclear pellets was carried out as described.

### Cell Lysis, Protein Digestion, and Isoelectric Focusing of Peptides

Lysis of cell pellets was carried out at ambient temperature. Each biological replicate (one cell pellet from one cell line) was processed in parallel to minimize the effects of systematic errors. Pellets were resuspended in 100  $\mu$ L 100 mM Ammonium

bicarbonate (Ambic) and 100  $\mu$ L trifluoroethanol (TFE) (organic solvent) was added followed by sonication (3  $\times$  20 s). Samples were incubated at 60 °C and resonicated (3  $\times$  20 s). Protein concentration was estimated using a bicinchoninic acid (BCA) assay (Pierce, Rockford, IL). Proteins were reduced and alkylated with 40 mM tris(2-carboxyethyl)phosphine (TCEP)/100 mM dithiothreitol (DTT) and 50 mM iodoacetamide (IAM), respectively. Samples were diluted in 50 mM Ambic, pH 8.0 and trypsinized overnight at 37 °C (1:50, w:w). Subsequently, peptides were lyophilized overnight. Peptides were desalted as described<sup>16</sup> and separated by isoelectric focusing (IEF) using immobiline IPG strips (24 cm, pH 3.5–4.5) (GE Healthcare) as described.<sup>16,17</sup>

### LC–MS/MS

LC–MS/MS shotgun proteomic analyses were performed on LTQ XL mass spectrometer (Thermo Fisher Scientific) equipped with an Eksigent NanoLC AS1 autosampler and Eksigent NanoLC 1D Plus pump, Nanospray source, and Xcalibur 2.0 SR2 instrument control. Peptides were separated on a packed capillary tip (Polymicro Technologies, 100 mm  $\times$  11 cm) with Jupiter C18 resin (5 mm, 300 Å, Phenomenex) using an in-line solid-phase extraction column (100 mm  $\times$  6 cm) packed with the same C18 resin using a frit generated with liquid silicate Kasil 1.<sup>18</sup> Mobile phase A consisted of 0.1% formic acid and mobile phase B consisted of 0.1% formic acid in 90% acetonitrile. A 90-min gradient was carried out with a 30-min washing period (100% A) to allow for solid-phase extraction and removal of any residual salts. Following the washing period, the gradient was increased to 25% B by 35 min, followed by an increase to 90% B by 50 min and held for 9 min before returning 95% A. MS/MS spectra of the peptides are acquired using data-dependent scanning in which one full MS spectrum (mass range 400–2000 *m/z*) is followed by five MS/MS spectra. MS/MS spectra are recorded using dynamic exclusion of previously analyzed precursors for 60 s with a repeat of 1 and a repeat duration of 1. MS/MS spectra were generated by collision-induced dissociation of the peptide ions at normalized collision energy of 35% to generate a series of b- and y-ions as major fragments. Biological samples from 3 independent cell cultures were injected in duplicate for a total of 6 replicate measurements for the SW480null and SW480APC cell lines.

### LC–MS/MS Data Analysis

MS/MS scans were transferred to mzML file format by “Scansifter”, an in-house-developed software algorithm, which reads tandem mass spectra stored as centroided peak lists from Thermo RAW files and encodes them to universal mzML format.<sup>19</sup> If 90% of the intensity of a tandem mass spectrum appeared at a lower *m/z* than the precursor ion, a single precursor charge was assumed; otherwise, the spectrum was processed under both double and triple precursor charge assumptions. The resulting mzML files were searched against the Human IPI database (v3.64) using the Myrimatch algorithm (version 1.2.11).<sup>20</sup> The database search was configured to look for both fully tryptic and semitryptic peptides with a precursor mass/charge (*m/z*) tolerance of 1.25 and a fragment *m/z* tolerance of 0.5. Carboxamidomethylation of cysteines was included as static modification, and oxidation of methionine as a dynamic modification in the search criteria, while any number of missed cleavages was allowed. A reverse version of the Human IPI database was included in the database search to allow for the calculation of false discovery rates (FDR). The IDpicker algorithm

(version 2.2.2) was used to assemble the set of peptides identified into a minimal list of proteins that could explain the observed spectral data set.<sup>21,22</sup> A minimum of two peptides per protein were required for valid protein identification with a peptide false discovery rate (FDR) of 5%. Proteins that could not be distinguished due to fully shared peptide sequences were grouped together (protein groups).

Statistically significant differences in protein spectral counts between different groups (i.e., SW480Null versus SW480APC) were calculated using the QuasiTel algorithm.<sup>23</sup> QuasiTel is a statistical analysis package which uses quasi-likelihood modeling to compare differences in spectral count data between two groups. The algorithm assigns a log-transformed rate ratio indicating the magnitude of the observed difference in spectral counts, and a probability of observing such differential levels for each protein corrected for multiple comparisons. Complete results from QuasiTel outputs and corresponding summaries from IDpicker reports are provided as Supplementary Table S1, Supporting Information.

### LC–MRM-MS

Cell line samples for LC–MRM-MS were prepared as outlined above for LC–MS/MS proteomics, except peptide extracts were not subjected to further fractionation by IEF. Peptide samples from each cell line were resuspended in 0.1% formic acid at 0.25 µg/µL and analyzed in triplicate (2 µL injection volume) on a TSQ Vantage triple quadrupole mass spectrometer (Thermo-Fisher, San Jose, CA) equipped with an Eksigent nanoLC solvent delivery system, autosampler and a nanospray source. The mobile phase consists of 0.1% formic acid in either HPLC grade water (A) or 90% acetonitrile (B). A 80-min gradient was carried out with a 15-min washing period (100% A). Following the washing period, the gradient was increased to 60% B by 43 min, followed by an increase to 95% B by 49 min and held for 11 min before returning 97% A. Four optimized transitions for each peptide of corresponding proteins were selected using the Skyline software package<sup>24</sup> (see Table S2 in Supporting Information for a list of peptides and corresponding precursor and product *m/z* values). Instrument parameters include Q2 gas 1.5 mTorr, scan width 0.004 *m/z*, scan time 10 ms, and both Q1 and Q3 resolution fwhm 0.7. A stable isotope labeled version of the β-actin peptide GYSFTTTAER was used as an internal standard (60 fmol/injection) for relative quantification of target proteins by the Labeled Reference Peptide (LRP) method.<sup>25</sup> The integrated chromatographic peak areas for the transitions of each targeted peptide were obtained from Skyline,<sup>24</sup> and summed, normalized to summed peak areas for the β-actin internal standard, and multiplied by a factor of 10<sup>6</sup>, as we described in detail elsewhere.<sup>25</sup> An unpaired *t* test was used to test for significant differences between samples (*n* = 3).

### Microarray Transcriptomic Analysis

For total RNA isolation cells were washed once in 1X PBS, lysed in TRIzol, and homogenized. Nucleic acids were isolated by chloroform extraction, and RNA was precipitated with isopropanol and purified by cesium chloride (CsCl) (5.7M) gradient centrifugation. Digoxigenin-UTP labeled cRNA was generated and linearly amplified from 1 µg of total RNA using Applied Biosystems Chemiluminescent RT-IVT Labeling Kit. cRNA was fragmented and hybridized to the Applied Biosystems Human Genome Survey Microarray V2.0 (29,098 genes) and imaged on an AB1700 chemiluminescent microarray analyzer (Applied Biosystems, Foster City, CA). We compared the two groups to identify

genes that were differentially expressed between them. First, expression data were normalized using quartile normalization and then log transformed. Before log transformation, we increased all expression value by a base number to make the lowest expression level equals 1. The normalized and log transformed data was analyzed using the Limma package in Bioconductor.<sup>26</sup> Specifically, we used Limma to fit a linear model to the data using an empirical bayes method to moderate standard errors. Expression data were normalized using quartile normalization and then log transformed. Before log transformation, we increased all expression value by a base number to make the lowest expression level equal to 1.

### Hierarchical Clustering Analysis

We performed a coclustering analysis integrating both mRNA and protein measurements for proteins that showed an adjusted *p*-value of less than 0.1 (QuasiTel analysis) and also had mRNA expression data. Before integration, protein and mRNA expression data sets were standardized protein/gene-wise separately so that each protein/gene had a mean expression value of 0 and standard deviation of 1 across samples in the data set. On the basis of the integrated data, pairwise similarity was calculated using the Pearson's correlation coefficient. Average linkage was used for hierarchical clustering. Genes and samples were clustered separately.

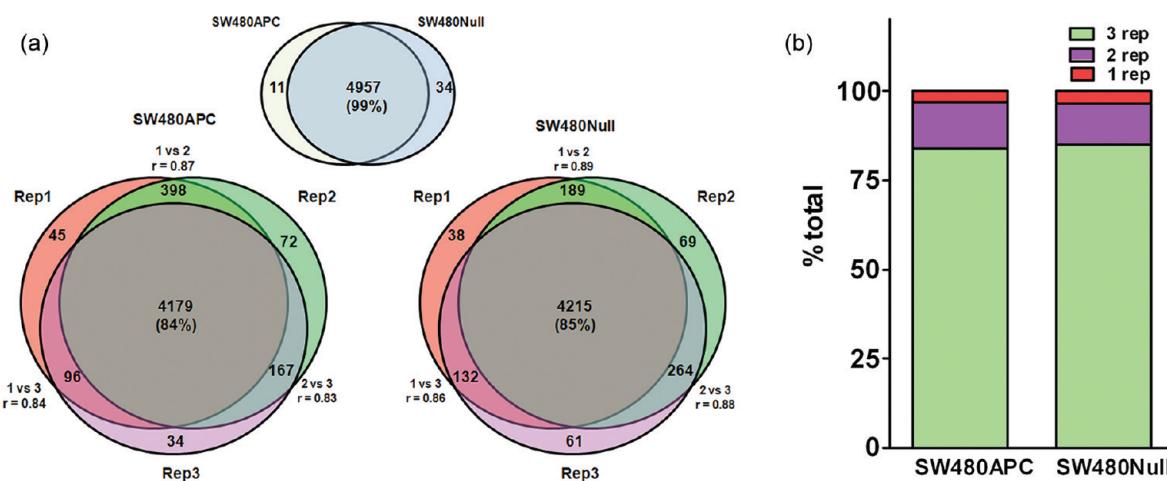
### Webgestalt Enrichment Analysis of Differentially Expressed Proteins

Enrichment analysis of proteomic data was carried out using Webgestalt.<sup>27,28</sup> Webgestalt identifies enriched classes of genes/proteins in large scale data sets by searching from several public resources: NCBI Gene, NCBI Gene Expression Omnibus (GEO) Ensembl, Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG), Pathway Commons, Wikipathways, Molecular Signatures Database (MSigDB). A set of 155 differentially expressed proteins (<0.1 adjusted quasi *p*-value) or 26 discordant proteins/transcripts (proteins with <0.1 adjusted *p*-value showing opposite expression trends with microarray data) were analyzed by Webgestalt using the entire set proteins observed (5002 proteins) as a reference set. Statistical significance was calculated using the Fisher test with Benjamini-Hochberg correction for multiple testing. Each enriched category is assigned an adjusted *p*-value (listed in Results along with each enriched class) and an enrichment factor (ratio of observed to expected for given enrichment class) (see Table 2).

## RESULTS

### Description of LC–MS/MS Shotgun Proteomics

Very few large scale global proteomic studies have examined the consequences of a single genetic mutation on the proteome and it is unclear if such studies are capable of detecting the effects of such small perturbations of the cellular system.<sup>29,30</sup> Previous studies relied on 2-D gel electrophoresis for identification of differentially expressed proteins, thus restricting proteome coverage to less than 1000 proteins.<sup>29,30</sup> Label free shotgun proteomics may provide more a comprehensive approach to detection of the subtle effects of a single gene change. Indeed, proteomic effects may be limited to a discrete number of signaling networks related to the gene of interest, or there may be more widespread changes as a result of secondary events such as transcription factor activation, altered metabolism or growth rates. Furthermore, the proteomic consequences of a single gene mutation may differ from changes in the gene expression profile. Due to differences in



**Figure 1.** Biological variation of LC–MS/MS proteomics. (A) Venn diagrams show protein expression overlap for shotgun proteomic inventories in SW480APC and SW480Null (top). Overlap in three biological replicates is shown for SW480APC (bottom left) and SW480Null (bottom right). Spearman ranked correlations for replicate to replicate comparisons shown in parentheses (B) Stacked plots show percentage of proteins identified in one, two or three biological replicates in SW480APC and SW480Null.

transcript and protein turnover rates as well as post-transcriptional regulation, transcriptomic profiling alone may not provide an accurate assessment of the effects of a single gene change. Therefore, we examined the effects of a single gene alteration on the proteome and transcriptome. We chose to study the effects of restoration of APC in a colorectal cancer cell line because of the central importance of this gene in colorectal carcinoma development.

Shotgun proteomic analysis was performed on three biological replicates of the isogenic SW480Null (mutant APC) and SW480APC (APC restored) cell lines (15 IEF fractions per replicate, 2 repeat injections per fraction), as described in Experimental Methods. In three biological replicate analyses we observed averages of 26 100 (9.5%CV) and 24 300 (12.6%CV) confidently matched spectra in SW480APC and SW480Null, respectively. These values corresponded to 4855 (7.4%CV) and 4795 (8.0%CV) protein groups in SW480APC and SW480Null, respectively.

Spectral counting was used to determine relative differences in protein expression between SW480APC and SW480Null. This approach to analysis of shotgun proteomic data provides an alternative quantitative approach to isotope labeling approaches, such as SILAC.<sup>12,13,31</sup> Previous studies have shown that spectral counting can achieve reproducible characterization of defined proteomic differences across multiple laboratories and the quasi-likelihood modeling approach to compare data sets can correct for instrument-to-instrument variations and interlaboratory variation.<sup>23,32,33</sup> Moreover, spectral counting-based quantitative proteomic analysis can be universally applied to many types of biospecimens including cell lines, frozen tissues, formalin fixed tissues and blood plasma.<sup>16,34</sup> A single discovery platform provides a straightforward way of comparing protein expression signatures across multiple sample types. Therefore, mutation specific proteomic profiles in cell culture models can easily be weighed against proteomic data sets in more complex clinical samples.

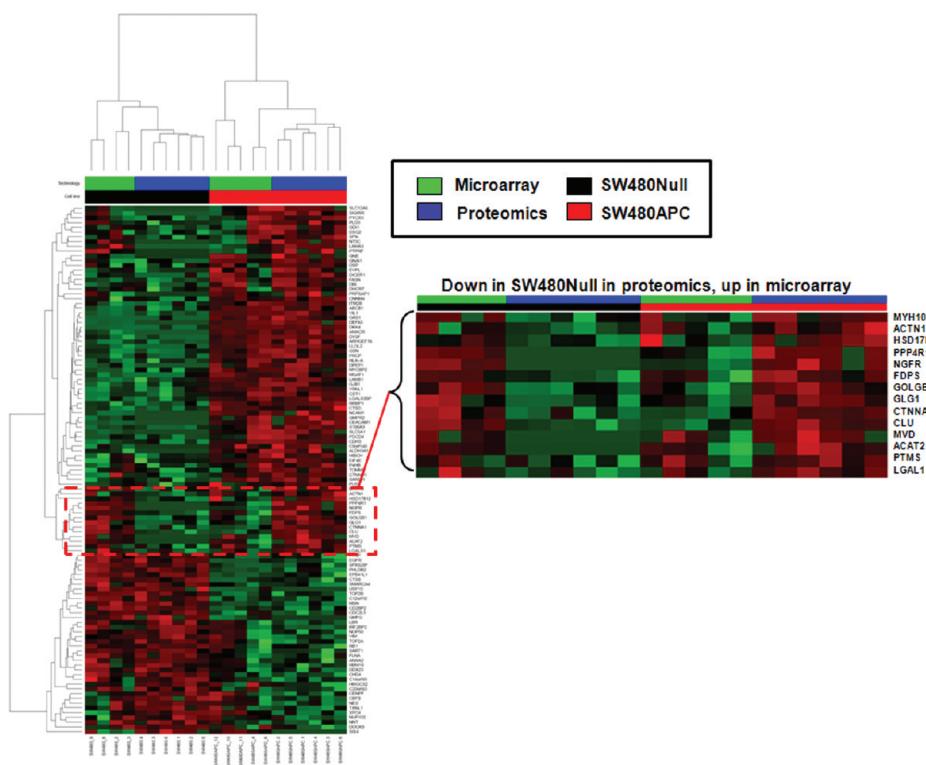
Although conventional search criteria (2 peptides per proteins, 5% peptide-level FDR) were applied to the data set, protein FDR levels remain considerably high (25.7%) after initial database searching. False positive peptides tend to be distributed across a large number of proteins. However, in general only a few spectra map to these false positive proteins (average 1.6 spectral counts per protein per cell line). Therefore, by applying a minimum

spectra-per-protein filtering criterion, protein FDR improved substantially. After removal of proteins with fewer than 6 spectral counts (minimum one spectrum per biological replicate), a total of 5002 proteins (3.8% protein FDR) remained. Of these 99.2% were detected in both SW480APC and SW480Null (Figure 1A). A majority of proteins was observed in all three biological replicates (SW480APC, 84% and SW480Null, 85%), while relatively few proteins were found in just a single biological replicate (SW480APC, 3% and SW480Null, 3.4%) (Figure 1B). Therefore, 97% of all proteins were observed in at least two of the three biological replicates, indicating the considerable reproducibility of the platform. Individual protein spectral counts in biological replicates (average of Rep1 vs Rep2, Rep1 vs Rep3, Rep2 vs Rep3) show good overall correlation (SW480APC,  $r = 0.85$  and SW480Null,  $r = 0.88$ , Spearman ranked correlation).

### Proteomic and Transcriptomic Comparison

To compare proteomic and transcriptomic profiles in the SW480APC model, supervised clustering analysis of shotgun proteomic data and corresponding microarray data was carried out (Figure 2). The analysis was limited to 111 differentially expressed proteins (<0.1 adjusted quasi  $p$ -value) that were also detectable by microarray analysis. Clustering analysis revealed that a plurality of proteins and transcripts (76%, 85/111 proteins) show the same directionality of expression change between SW480APC and SW480Null, that is, most proteins down-regulated in SW480APC are also down-regulated at the transcript level. One distinct cluster of 14 proteins was up-regulated in SW480APC in shotgun proteomics but down-regulated in microarray (Figure 2, red box). This set included cell adhesion proteins CTNNA1 and ACTN1 and phosphatases PPP4R1 and PTPRF. Further analysis of SW480APC/SW80Null log<sub>2</sub> ratios for the set of 111 proteins/genes between microarray and proteomics revealed only modest correlation ( $r = 0.503$ , Spearman ranked correlation). Transcripts found to be significantly differentially expressed by microarray (top 3% ranked by adjusted  $p$ -value, 122 genes/proteins) showed similar correlation with corresponding proteomic log<sub>2</sub> ratios ( $r = 0.63$ , Spearman ranked correlation).

It is likely that differences in turnover rates for proteins and transcripts as well post-transcriptional regulation events



**Figure 2.** Proteomic and transcriptomic profile comparison of SW480APC and SW480Null. Heat map shows supervised clustering analysis of shotgun proteomics data and transcriptomic data for 111 differentially expressed proteins (adjusted quasi  $p$ -value  $<0.1$ ) (red, increased expression; green, decreased expression). Clustering was performed on  $\log_2$  transformed spectral counts from 6 replicate analyses for proteomics and on  $\log_2$  transformed RMA values for microarray.

(e.g., microRNA (miRNA), mRNA processing) contribute to the observed discordance. Interestingly, DICER1 which plays a key role in RNA-induced silencing is up-regulated in SW480APC, presumably resulting in altered microRNA biosynthesis and expression. Altered DICER1 expression is a feature of several cancers including ovarian, pancreatic and lung cancer.<sup>35–37</sup> Previous studies have shown conflicting results for DICER1 expression levels in CRC, with some noting decreased expression in colon adenocarcinomas, while others demonstrated that over-expression is associated with poor prognosis.<sup>38,39</sup> The data here suggest that restoration of wildtype APC correlates with enhanced DICER1 expression which may result in altered miRNA expression profiles. Further inspection of 26 discordant proteins revealed enrichment for several miRNA targets, including miR-338 (adjusted  $p$ -value = 0.01) and miR-193A (adjusted  $p$ -value = 0.07) (Webgestalt enrichment analysis, see Experimental Methods). miR-338 is down-regulated in hepatocellular carcinoma, while overexpression of miR-193A inhibits tumor growth in colon xenografts by targeting KRAS and PLAU.<sup>40,41</sup> These data suggest that changes to specific miRNA levels might result in discordant protein/transcript expression.<sup>42</sup> Therefore, restoration of APC may affect miR-338 and miR-193A expression levels, thereby giving rise to discordant transcript/protein levels in specific targets, e.g. PLD3, SIX4.

#### Protein Networks and Pathways Affected by Reintroduction of APC

Using quasi-likelihood modeling of protein spectral counts we identified a set of 155 proteins showing differential expression between SW480APC and SW480Null (adjusted quasi  $p$ -value  $<0.1$ )

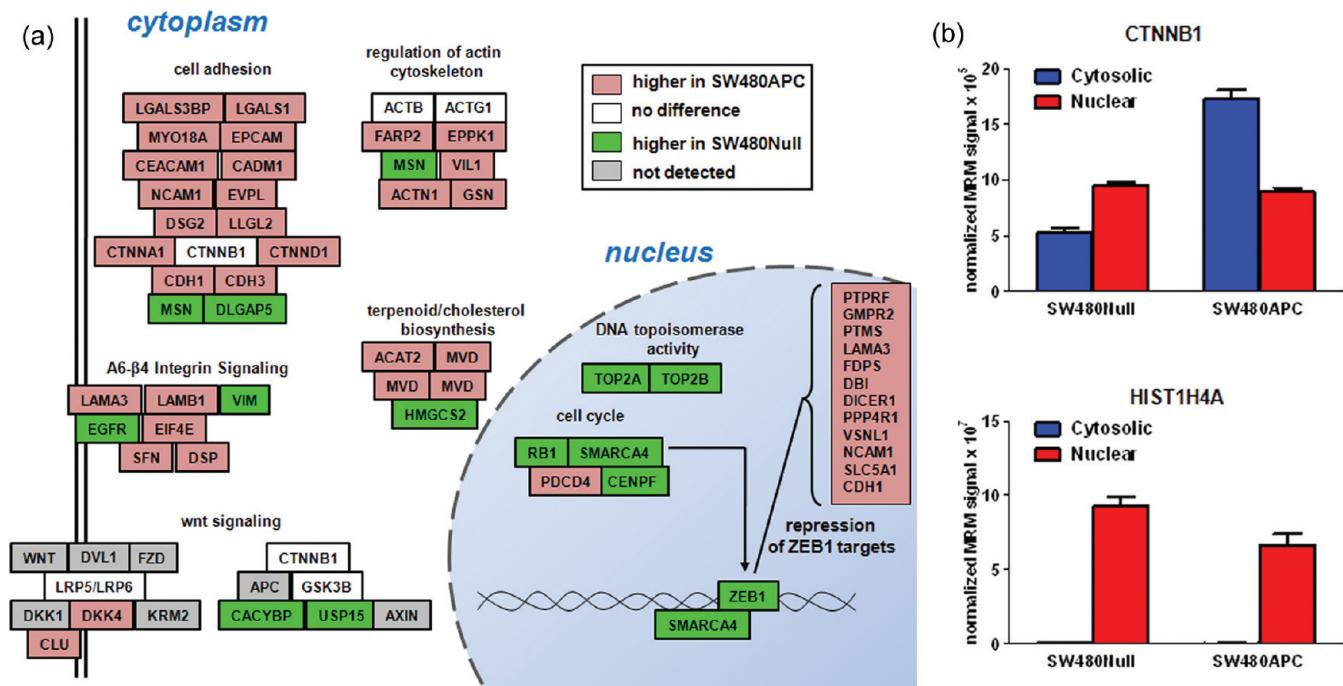
(97 down in SW480Null, 57 up in SW480Null). Within this set of 155 proteins we carried out a systematic search for enrichment of functionally related protein classes using the Webgestalt software tool.<sup>28</sup> Table 1 shows the top five sets of proteins (ranked by adjusted  $p$ -value) from multiple enrichment analyses (GO, KEGG, etc). This approach identified groups of proteins that were down-regulated in SW480Null cells (>75% proteins down in SW480Null, red arrows alone) including cell adhesion (GO:0007155, adjusted  $p$ -value = 0.005) and type II interferon signaling (WP619, adjusted  $p$ -value = 0.031). Increased expression of proteins involved in protein kinase C binding (GO:0005080, adjusted  $p$ -value = 0.133) and DNA topoisomerase activity (GO:0003916, adjusted  $p$ -value = 0.133) was observed in SW480Null (>75% proteins up in SW480Null, green arrows alone). Interestingly, some groups enriched for differentially expressed proteins contained proteins that were either up-regulated or down-regulated in SW480Null (<75% proteins up or <75% down in SW480Null, red and green arrows), including negative regulation of the cell cycle (GO:0045786, adjusted  $p$ -value = 0.015) and terpenoid backbone biosynthesis (KEGG ID 900, adjusted  $p$ -value = 0.003).

Pathways and protein functional classes affected by APC restoration (derived from Webgestalt enrichment analysis) are depicted schematically in Figure 3A. A major proteomic feature of SW480Null cell line is the down-regulation of cell adhesion proteins. Loss of epithelial cadherin (CDH1) and the attendant abrogation of adherens junctions have been previously reported in SW480Null cells.<sup>11</sup> CDH1 was shown by Western blot to have 2-fold higher expression in SW480APC cells.<sup>11</sup> Here we observed 2-fold higher expression of CDH1 in SW480APC by

Table 1. Enrichment Analysis of Differentially Expressed Proteins<sup>a</sup>

Database	Database ID	Description	Up/Down In SW480Null	ER	adj p value
GO (BP)	GO:0007155	cell adhesion	↓	3.1	4.6E-03
GO (BP)	GO:0016337	cell-cell adhesion	↓	5.6	4.6E-03
GO (BP)	GO:0022610	biological adhesion	↓	3.1	4.6E-03
GO (BP)	GO:0045786	negative regulation of cell cycle	↑↓	7.6	1.4E-02
GO (BP)	GO:0007156	homophilic cell adhesion	↓	8.9	2.4E-02
GO (MF)	GO:0032393	MHC class I receptor activity	↓	18.8	4.8E-02
GO (MF)	GO:0003916	DNA topoisomerase activity	↑	12.6	1.3E-01
GO (MF)	GO:0005080	protein kinase C binding	↑	7.9	1.3E-01
GO (MF)	GO:0050839	cell adhesion molecule binding	↑↓	7.3	1.3E-01
GO (MF)	GO:0016628	oxidoreductase activity, acting on the CH-CH group of donors, NAD or NADP as acceptor	↑↓	9.4	1.3E-01
GO (CC)	GO:0005576	extracellular region	↓	3.2	2.0E-04
GO (CC)	GO:0044459	plasma membrane part	↓	2.2	7.7E-03
GO (CC)	GO:0005911	cell-cell junction	↓	3.8	2.1E-02
GO (CC)	GO:0042611	MHC protein complex	↓	13.3	2.1E-02
GO (CC)	GO:0044421	extracellular region part	↓	3.1	2.1E-02
KEGG	4514	Cell adhesion molecules (CAMs)	↓	9.9	3.2E-05
KEGG	900	Terpenoid backbone biosynthesis	↑↓	14.3	2.2E-03
KEGG	5320	Allograft rejection	↓	13.8	1.0E-02
KEGG	5332	Autoimmune thyroid disease	↓	13.8	1.0E-02
KEGG	4520	Type I diabetes mellitus	↓	12.0	1.1E-02
WP	WP244	Alpha6-Beta4 Integrin Signaling Pathway	↑↓	5.6	4.0E-03
WP	WP619	Type II interferon signaling (IFNG)	↓	8.0	3.1E-02
WP	WP197	Cholesterol Biosynthesis	↓	8.8	3.1E-02
WP	WP51	Regulation of Actin Cytoskeleton	↑↓	3.6	3.1E-02
WP	WP306	Focal Adhesion	↑↓	3.2	3.4E-02
PC	DB_ID:43	Cholesterol biosynthesis	↑↓	7.4	2.7E-01
PC	DB_ID:1050	Alpha6Beta4Integrin	↑↓	5.4	2.7E-01
PC	DB_ID:692	Urea synthesis	↓	12.8	2.7E-01
PC	DB_ID:751	Pyrimidine catabolism	↓	12.8	2.7E-01
PC	DB_ID:769	Proline catabolism	↓	12.8	2.7E-01
MSigDB (TF)	DB_ID:1934	hsa_CAGGTA_V\$AREB6_01	↓	3.3	2.1E-02
MSigDB (TF)	DB_ID:2365	hsa_V\$AREB6_03	↓	5.4	4.3E-02
MSigDB (TF)	DB_ID:2002	hsa_WGGAATGY_V\$TEF1_Q6	↑↓	3.3	2.6E-01
MSigDB (TF)	DB_ID:2424	hsa_V\$AP2REP_01	↑↓	2.5	4.2E-01
MSigDB (TF)	DB_ID:2159	hsa_V\$HMGIY_Q6	↓	2.6	4.2E-01
MSigDB (miRNA)	DB_ID:872	hsa_TTGGGAG,MIR-150	↓	2.5	7.2E-01
MSigDB (miRNA)	DB_ID:742	hsa_GGCCAGT,MIR-193A,MIR-193B	↑↓	2.1	7.2E-01
MSigDB (miRNA)	DB_ID:666	hsa_ACTGTGA,MIR-27A,MIR-27B	↑	1.5	7.2E-01
MSigDB (miRNA)	DB_ID:692	hsa_TGCAC TG,MIR-148A,MIR-152,MIR-148B	↑	2.0	7.2E-01
MSigDB (miRNA)	DB_ID:838	hsa_CTATGCA,MIR-153	↑	1.6	7.2E-01
MSigDB (PPI)	DB_ID:934	hsa_MODULE_23	↓	14.3	9.0E-04
MSigDB (PPI)	DB_ID:720	hsa_MODULE_83	↑	5.4	1.5E-01
MSigDB (PPI)	DB_ID:1031	hsa_MODULE_20	↑↓	4.6	1.5E-01
MSigDB (PPI)	DB_ID:1036	hsa_MODULE_7	↑↓	5.4	1.5E-01
MSigDB (PPI)	DB_ID:902	hsa_MODULE_3	↓	3.8	1.7E-01
MSigDB (CP)	DB_ID:11292	hsachr17q25	↓	4.7	3.1E-02
MSigDB (CP)	DB_ID:11320	hsachr18q	↑↓	5.6	1.3E-01
MSigDB (CP)	DB_ID:11078	hsachr7q21	↑↓	6.9	1.4E-01
MSigDB (CP)	DB_ID:11007	hsachr16q24	↓	5.4	2.2E-01
MSigDB (CP)	DB_ID:11311	hsachr16q	↓	2.4	3.8E-01

<sup>a</sup> The Webgestalt algorithm<sup>27,28</sup> was used to identify enriched sets of proteins in a set of 155 differentially expressed proteins. The top five ranked categories (based on adjusted *p*-value) from each database are listed. Enriched categories with >75% proteins up-regulated in SW480Null were classified as up in SW480Null (green arrow alone) and enriched classes with >75% down regulated in SW480Null were classified as down in SW480Null (red arrow alone). All others classified as a mix of up-regulated and down-regulated proteins. The entire data set of 5002 proteins was used as a reference set in the analysis. Databases are abbreviated as follows; GO, gene ontology (BP, biological process, MF, molecular function, CC, cellular component); KEGG, Kyoto Encyclopedia of Genes and Genomes; WP, Wiki Pathways; PC, Pathway Commons; MsigDB, Molecular Signatures Database (TF, transcription factor, miRNA, microRNA, PPI, protein protein interaction, CP, chromosome position). For a given category the enrichment ratio (ER) is the ratio of observed proteins in the 155 protein set to the number expected based on chance.



**Figure 3.** Proteomic changes associated with APC restoration. (A) Schematic figure represents a subset of enriched classes from 155 differentially expressed proteins ( $<0.1$  adjusted quasi  $p$ -value) between SW480Null and SW480APC (see Table 2 for details on all enriched classes). APCnull status results in decreased expression (red symbols) of proteins associated with cell adhesion and the actin cytoskeleton and up-regulation (green symbols) of proteins involved in cell cycle control cholesterol biosynthesis. (B) SW480APC cells display redistribution of CTNNB1 from cytoplasm to nucleus. Cells were fractionated and nuclear and cytosolic fractions were analyzed by LC–MRM for CTNNB1 (top panel) and HIST1H4A (nuclear protein marker) (bottom panel).

LC–MS/MS and 3-fold higher expression by LC–MRM (see Table 2). SW480Null cells had decreased levels of many other proteins involved in adherens junction formation and cell adhesion, including CDH3, CTNNA1, CTNND1, NCAM1, EVPL, DSG2, CECAM1, and EPCAM (Figure 3A). Another group of proteins down-regulated in SW480Null cells includes regulators of the actin cytoskeleton,<sup>11</sup> which undergoes dramatic reorganization associated with the APCnull phenotype. Levels of the actins ACTB and ACTG1 were unchanged in the SW480Null cells, but levels of the actin regulators FARP2, EPPK1, VIL1, ACTN1 and GSN all were decreased. Thus, proteomic findings confirm a previously identified phenotypic feature of SW480Null cells.

The effect of APC loss on the Wnt signaling pathway itself is relatively modest and appears to reflect adaptation to APC loss. The negative Wnt regulators DKK4 and CLU were decreased. CACYBP, which participates in CTNNB1 degradation and USP15, a deubiquitylation enzyme which prevents APC degradation, both were increased in SW480ull cells (Figure 3A). We observed a trend of elevated CTNNB1 in SW480APC, although this trend did not meet the 0.1 adjusted quasi  $p$ -value threshold (adjusted quasi  $p$ -value = 0.18,  $\log_2$  SW480APC/SW480Null rate ratio = 0.52). However, measurement of cytosolic and nuclear CTNNB1 by MRM revealed a 3-fold increase in cytosolic CTNNB1 in SW480APC ( $p < 0.001$ , unpaired  $t$  test), suggesting that restoration of APC results in reduced Wnt activation (Figure 3B, top panel). MRM analysis confirmed that total CTNNB1 (cytosolic and nuclear) is higher in SW480APC ( $\log_2$  SW480APC/SW480Null = 0.81,  $p = 0.001$ , unpaired  $t$  test).

Changes to proteomic expression profiles in APC restored cells may be the result of altered transcription factor activities. Enrichment analysis revealed that targets of the transcription

repressor ZEB1 (hsa\_CAGGTA\_V\$AREB6\_01, adjusted  $p$ -value = 0.022) were measured at lower levels in SW480Null, implying elevated ZEB1 activity in this cell line. Activation of ZEB1 is thought to play a key role in epithelial to mesenchymal transition (EMT), a process characterized by loss of cell polarity and adhesion and increased motility. A prominent feature of EMT is loss of CDH1 expression. ZEB1 represses CDH1 by binding to E-box elements in the promoter region. Therefore, the observed decrease in CDH1 expression in SW480Null may be a direct result of ZEB1 activation.<sup>43</sup> Furthermore, the ZEB1 corepressor SMARCA4 is up-regulated in SW480Null (Figure 3A). Others have shown that disruption of the ZEB1/SMARCA4 binding causes an increase in CDH1 expression and a decrease in the mesenchymal marker VIM, trends seen here in SW480APC.<sup>44</sup> Therefore, SMARCA4-dependent activation of ZEB1 may play a role in conferring key phenotypic features in SW480Null.

Functional class enrichment analysis identified several groups with both up-regulated and down-regulated proteins in SW480Null (Table 1, red and green arrows). This result suggests that APC restoration may affect expression of proteins in a common pathway, but changes do not necessarily occur in a uniform direction. For instance, 2/6 proteins (PDCD4, NGFR) were down in SW480-Null and 4/6 proteins (SMARCA4, RB1, EGFR, CENPF) were up in SW480Null for the GO category negative regulation of the cell cycle (GO:0045786, adjusted  $p$ -value = 0.014). This observation may reflect the multiple roles played by individual proteins in the cell cycle. Similarly, for  $\alpha$ 6 $\beta$ 4 integrin signaling (WP244, adjusted  $p$ -value = 0.004) 5/7 proteins (LAMA3, LAMB1, DSP, EIF4E, SFN) were down in SW480Null and 2/7 proteins (VIM, EGFR) were up in SW480Null. Elevated VIM and EGFR levels in SW480Null may result in changes to integrin-mediated cell

**Table 2.** Summary of LC–MS/MS and LC–MRM-MS Data for 22 Proteins Validated by Targeted Proteomics

HGNC gene symbol	SW480APC count	LC–MS/MS			LC–MRM		
		SW480Null count	$\log_2$ (SW480Null/ SW480APC)	QL p-value	AdjP	$\log_2$ (SW480Null/ SW480APC)	t test, p-value
ADD3	39	40	0.1	$5.15 \times 10^{-1}$	$8.36 \times 10^{-1}$	-0.3	$9.80 \times 10^{-3}$
ASS1	63	30	-1	$9.10 \times 10^{-6}$	$4.04 \times 10^{-3}$	-0.5	$1.37 \times 10^{-1}$
CADM1	25	5	-2.2	$1.75 \times 10^{-3}$	$7.21 \times 10^{-2}$	-1.3	$1.00 \times 10^{-4}$
CDH1	39	18	-1	$3.44 \times 10^{-3}$	$1.06 \times 10^{-1}$	-1.7	$1.00 \times 10^{-4}$
CLU	40	12	-1.6	$1.70 \times 10^{-5}$	$5.67 \times 10^{-3}$	-2.1	$1.00 \times 10^{-4}$
CST1	40	7	-2.4	$3.83 \times 10^{-5}$	$8.00 \times 10^{-3}$	-1.8	$6.00 \times 10^{-4}$
CTNNND1	85	47	-0.7	$8.30 \times 10^{-5}$	$1.39 \times 10^{-2}$	-1.2	$1.00 \times 10^{-4}$
DKK4	100	2	-5.5	$1.64 \times 10^{-6}$	$1.37 \times 10^{-3}$	-6	$1.00 \times 10^{-4}$
DSG2	41	19	-1	$6.49 \times 10^{-4}$	$3.77 \times 10^{-2}$	-1.2	$1.80 \times 10^{-3}$
DYSF	124	8	-3.8	$2.54 \times 10^{-9}$	$6.37 \times 10^{-6}$	-0.9	$1.14 \times 10^{-2}$
EGFR	26	70	1.5	$3.10 \times 10^{-4}$	$2.38 \times 10^{-2}$	1.6	$3.05 \times 10^{-2}$
FDPS	41	19	-1	$3.83 \times 10^{-4}$	$2.78 \times 10^{-2}$	-1	$2.00 \times 10^{-4}$
LCP1	21	23	0.2	$3.80 \times 10^{-1}$	$7.67 \times 10^{-1}$	-0.2	$9.33 \times 10^{-2}$
LGALS3BP	59	22	-1.3	$1.59 \times 10^{-4}$	$1.63 \times 10^{-2}$	-1.5	$9.60 \times 10^{-3}$
LLGL2	18	1	-4.1	$3.79 \times 10^{-4}$	$2.78 \times 10^{-2}$	-1.4	$4.00 \times 10^{-2}$
MGAT1	14	2	-2.7	$2.98 \times 10^{-3}$	$9.84 \times 10^{-2}$	-1	$2.21 \times 10^{-2}$
NES	112	258	1.3	$9.44 \times 10^{-5}$	$1.41 \times 10^{-2}$	1.2	$4.57 \times 10^{-2}$
PDCD4	48	12	-1.9	$1.69 \times 10^{-4}$	$1.66 \times 10^{-2}$	-2.3	$5.00 \times 10^{-4}$
PPL	173	106	-0.6	$4.09 \times 10^{-3}$	$1.16 \times 10^{-1}$	-1.2	$2.00 \times 10^{-4}$
PYCR1	101	37	-1.3	$5.15 \times 10^{-4}$	$3.47 \times 10^{-2}$	-1.9	$1.60 \times 10^{-3}$
VIL1	82	26	-1.6	$5.09 \times 10^{-6}$	$2.55 \times 10^{-3}$	-1.5	$1.53 \times 10^{-2}$
VSNL1	75	30	-1.2	$1.57 \times 10^{-3}$	$6.67 \times 10^{-2}$	-1.1	$1.77 \times 10^{-2}$

adhesion and motility, thereby contributing to a more tumorigenic phenotype APCNull cells.

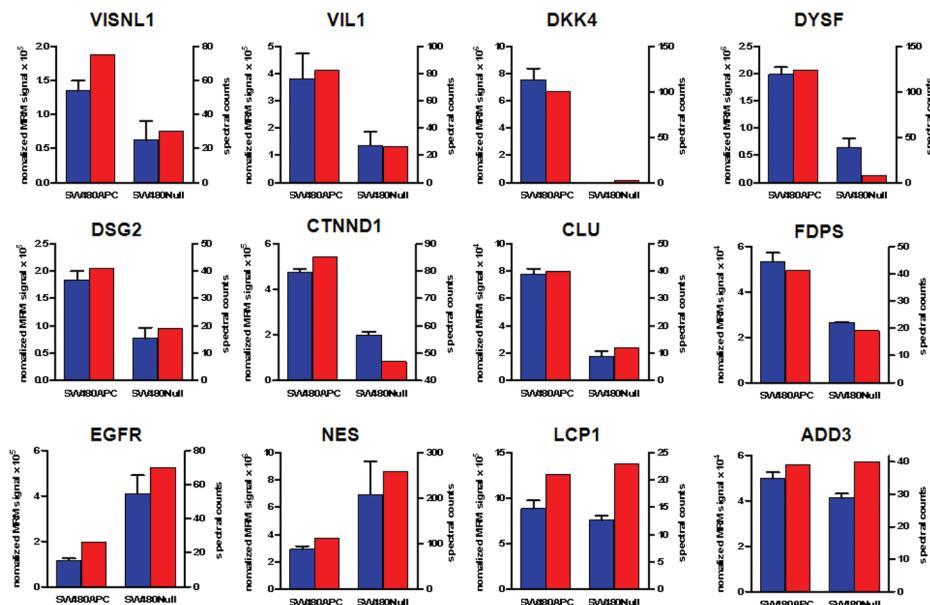
Proteins up-regulated in SW480Null cells included cell cycle-related proteins RB1, SMARCA4, CENPF and DNA topoisomerases TOP2A, TOP2B (Figure 3A). Among other proteomic differences observed, SW480Null cells display up-regulated EGFR, but down-regulated NGFR, which suggests a shift in growth factor response. SW480Null cells display down-regulation of SLC5A1 (sodium glucose cotransporter) and SLC12A6 (potassium chloride cotransporter). Down-regulation of MVD and FDPS, two key regulators of terpenoid backbone synthesis synthesis, may affect the prenylation of signaling enzymes. These changes may be the result of increased growth rates in SW480Null cells. To broaden the search for enriched functional protein groups, we carried out a limited manual inspection of significantly differentially expressed proteins. This approach reveals functional groups which fall outside the Webgestalt enrichment criteria (i.e., those groups not in the top five ranked categories based on adjusted  $p$ -values), as well as proteins which share a common function but lack a common GO designation. Manual interrogation ( $\geq 3$  proteins, all up in SW480Null) revealed elevated levels of proteins involved in mRNA processing (DDX23, RBM10, SART1, CD2BP2, SFRS2IP) and nuclear transport (NUP50, NUP155, XPO4). Assessment of the impact of these protein expression changes would require further experiments. However, the magnitude of these protein changes is similar to those associated with well-documented effects on cell adhesion, cytoskeletal organization and cell cycle regulation. Thus, these newly identified proteomic changes provide new avenues to explore the APCNull phenotype.

### Verification of Proteomic Differences by LC–MRM-MS

A subset of the proteins found to be differentially expressed by LC–MS/MS shotgun proteomics were assessed by LC–MRM-MS. This approach provides a quantitative measurement of proteotypic peptides and an accurate assessment of the relative protein abundance in SW480APC and SW480Null.<sup>25</sup> LC–MRM-MS measurements showed substantial reproducibility across three replicate runs. For 22 proteotypic peptides analyzed, a median CV value of 11.5% was observed. Correlation between log ratios (SW480APC/SW480Null) from each approach was also evaluated ( $r = 0.63$ , Spearman ranked correlation). For all proteins, LC–MRM-MS measurements agree with LC–MS/MS shotgun proteomics, and identical trends in protein expression were observed between the two platforms, that is, proteins with high levels in SW480APC by shotgun proteomics are also seen to be high by LC–MRM-MS (e.g., DKK4, CDH1) (Figure 4 and Table 2). Similarly, proteins which were not found to be differentially expressed by shotgun proteomics, showed no difference by LC–MRM-MS (e.g., LCP1, ADD3). Thus, the data demonstrate that for a subset of differentially expressed proteins, label free shotgun proteomics data and LC-MRM data are broadly concordant.

### DISCUSSION

Extensive information is currently available on genomic variation in cancer. Large data repositories such as the Cancer Genome Atlas (TCGA) and Catalogue of Somatic Mutations in Cancer (COSMIC) provide vast knowledge on a range of mutations across a spectrum of tumor types.<sup>2,6</sup> However, the global proteomic effects



**Figure 4.** Validation of proteomic differences by LC–MRM-MS. Shotgun proteomics data are plotted as spectral counts for triplicate analyses (red bars), whereas MRM data are plotted as summed signal intensity for measured transitions normalized to summed intensity for transitions measured for a reference peptide (blue bars). ( $n = 3$ ). A list of peptides and corresponding precursor and product  $m/z$  values is provided in Supplementary Table S2.

of cancer-related mutations have not been well characterized. Here we examined the proteomic consequences of a single gene difference in a CRC cell line model in which APC expression has been restored. Label free LC–MS/MS shotgun proteomics provided a robust and reproducible means to inventory over 5000 proteins in SW480APC and SW480Null. Differentially expressed proteins included cell adhesion and actin binding proteins and cell cycle regulators all known to be impacted by restoration of wildtype APC expression.<sup>11</sup> Additionally, several protein classes not previously thought to be related to Wnt signaling were differentially expressed, for example, cholesterol biosynthesis, ZEB1 transcription targets.

Although most canonical Wnt signaling proteins did not show altered expression (e.g., GSK3B, CTNNB1, DVL1), DKK4, a known negative regulator of Wnt signaling, was significantly up-regulated in SW480APC.<sup>45</sup> DKK4, in conjunction with Kremen2 (KRM2), inhibits the interaction between Wnt and its coreceptor low-density lipoprotein receptor-related protein 5/6 (LRP5/6).<sup>46</sup> Interestingly, DKK4 is down-regulated in primary colorectal tumors and several colorectal cancer cell lines, and re-expression of DKK4 resulted in attenuation of cell cycle progression.<sup>47</sup> Our data agree with these earlier findings that DKK4 plays an important role in APC-dependent tumor formation. In SW480Null cells we observed increased expression of CACYBP which plays a role in degradation of CTNNB1.<sup>48</sup> CACYBP is up-regulated in pancreatic cancer and suppression by siRNA results in decreased cellular proliferation rates.<sup>49</sup> In mutant APC expressing cells, enhanced CACYBP expression may be an adaptive response to increased levels of unbound CTNNB1.

A distinct group of proteins involved in RNA processing is up-regulated in SW480Null cells (e.g., DDX23, SART1 and HNRNPL). Several genes are known to undergo abnormal splicing in cancer, including KLF6, DNMT3B and BRCA1.<sup>50–52</sup> Changes in the levels or the activities of regulatory splice proteins may contribute to tumor formation.<sup>53</sup> For instance serine/arginine rich

(SR) proteins play a key role in the specification of splice sites in eukaryotic mRNA and are regulated in part by their phosphorylation status. SFRS2 is an SR protein involved in the formation of an early ATP-dependent splicing complex.<sup>54</sup> Here we see an increase in its interacting protein SFRSIP in SW480Null. SR proteins may also interact with heterogeneous nuclear ribonucleoproteins (hnRNPs).<sup>55</sup> In the current data set HNRNPL is up-regulated in mutant APC cells. Changes to the stoichiometric ratios of SR proteins and hnRNPs may influence the expression profiles of splice variants in APCNull cells. Further experiments will be required to establish if increases to RNA processing proteins are merely a byproduct of the faster growth rate in SW480Null cells or the result of another aspect of APC deficiency.

Altered expression was observed in proteins associated with isoprenoid biosynthetic pathways (ACAT2, MVD, FDPS, HMGCS2). Prenylation and farnesylation of key signal transduction proteins may modify pathways activities in cancer. Modulations to PTP4A3 and CDC42 have been shown to contribute to colon cancer and breast cancer, respectively.<sup>55,56</sup> Given the important role of prenylation in tumor formation, anticancer agents have been developed which suppress key enzymes in this pathway, including RCE1 and IMCT.<sup>57,58</sup> Isoprenoid biosynthetic proteins identified here may represent alternative therapeutic targets in mutant APC-dependent colorectal cancer. Interestingly, many of these proteins showed discordant expression patterns between transcriptomic and proteomic data sets. The isoprenoid pathway provides metabolic precursors to steroid biosynthesis, a process subject to feedback regulation. For example, elevated sterol levels result in decreased HMGCR, the rate-limiting enzyme of sterol biosynthesis, via ER-associated degradation and proteolysis.<sup>59</sup> It is unclear if such a feedback mechanism may explain discordant expression of isoprenoid synthesis proteins in this study.

In general, our results illustrate how proteomic analysis may provide biological insights that would otherwise not be predicted by transcriptomic analysis alone or by previous descriptions of

the APCNull phenotype. Several novel pathways are implicated in the current proteomic data set including DNA topoisomerase activity, terpenoid backbone biosynthesis and  $\alpha 6\beta 4$  integrin signaling. Functional enrichment results from transcriptomic data alone (data not shown) showed considerable divergence from those of proteomics data. Known APCNull phenotypic properties were either not observed (e.g., negative regulation of the cell cycle) or showed only modest enrichment in transcriptomic only analyses (e.g., actin-binding, adjusted *p*-value = 0.371; cell–cell adherens junction, adjusted *p*-value = 0.378). Loss of CDH1 expression, a previously validated hallmark of APCNull cells, is weakly corroborated at the transcript level ( $\log_2$  ratio = 0.2, adjusted *p*-value = 0.45) but clearly demonstrated at the proteomic level ( $\log_2$  ratio = 1, adjusted *p*-value = 0.1). Furthermore, target enrichment of the EMT-related ZEB1 transcriptional repressor in SW480APC is only observed in the proteomics data set. Therefore, proteomics appears to present a more reliable approach to biological discovery in the SW480APC model and underscores the added value of proteomic techniques over transcriptomic approaches.

Label free shotgun proteomics provides significant depth of proteome coverage and aids in the detection of subtle proteomic differences in the SW480APC model. Nonetheless, there are some important limitations associated with the approach. The proteomic consequences of APC restoration were only examined in a single SW480APC clone, raising the possibility that clonal cell line features contribute to observed proteomic differences. However, Faux and colleagues showed that multiple SW480APC clones display a consistent set of phenotypic properties, for example, decreased growth rates, loss of anchorage independent growth, increased CDH1 levels.<sup>11</sup> Our proteomic data recapitulate many of these features. In particular, we document decreased expression of over 20 cell adhesion proteins (e.g., DSP, CTNND1, CADM1) and changes to cell cycle regulatory proteins (e.g., RB1, SMARCA4) in SW480Null. LC–MS/MS and LC–MRM proteomic data for CDH1 (2- and 3-fold higher in SW480APC, respectively) agree with the previously reported 2-fold decrease.<sup>11</sup> These specific findings suggest that proteomic differences primarily reflect the biological consequences of APC restoration.

As with most proteomic technologies, there is a bias toward detection of high abundant proteins and many lower abundant proteins are either not detected at all, or do not produce enough signal to allow a confident comparison of relative expression levels between samples. In the SW480APC model several proteins from the Wnt signaling pathway (e.g., APC, AXIN2, TCF7, LEF1, DVL1) were not detected. Consequently, only a partial evaluation of the proteomic effects of APC restoration on the Wnt signaling pathway can be made. Likewise many transcription factors, which may play key regulatory roles in the SW480APC model, are not detected. However, the high number proteins confidently identified in this data set (>5000) still allows many significant differences to be observed and meaningful biological information may be inferred from these differences. For instance, enhanced activity of the transcriptional repressor ZEB1 in SW480Null can be deduced from elevated levels of ZEB1 protein targets in SW480APC. Furthermore, verification of 22 proteins by targeted LC–MRM confirmed the direction and magnitude of the differences observed in the LC–MS/MS proteomics data, thereby demonstrating the general reliability of the approach.

The approach used here to analyze the effects restoration of wildtype APC can be applied to other model systems where known

phenotypic differences arise from a specific genetic alteration. In colorectal cancer mutations in TP53, KRAS and SMAD4 contribute significantly to tumor development.<sup>5</sup> Elucidation of the global proteomic consequences of these mutations should yield new insights into basic tumor biology. In conclusion, we have demonstrated the utility of label free shotgun proteomics to assess the proteomic effects of a single gene difference in a colorectal cancer cell model. Our findings reflect previously reported phenotypic features of the APCNull phenotype while also identifying new aspects of the underlying biology associated with APC deficiency.

## ASSOCIATED CONTENT

### Supporting Information

Supplemental tables. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## AUTHOR INFORMATION

### Corresponding Author

\*Robbert J. C. Slebos, Ph.D., Vanderbilt University School of Medicine, Medical Research Building III, U1213, 465 21st Avenue South, Nashville, TN 37232. Tel: 615-936-3063. Fax: 615-936-1001. E-mail: r.slebos@vanderbilt.edu.

## ACKNOWLEDGMENT

For purposes of clarity and brevity, proteins have been abbreviated to HGNC gene symbols. This work was supported in part by NIH grants US4CA126479 (to D.C.L.) and R01GM088822 (to B.Z.). We thank Misti Martinez, Kristin Cheek, and Sarah Stuart for technical assistance. We thank Zhiaoxi Shi for programming assistance.

## REFERENCES

- (1) Ahrens, C. H.; Brunner, E.; Qeli, E.; Basler, K.; Aebersold, R. Generating and navigating proteome maps using mass spectrometry. *Nat. Rev. Mol. Cell Biol.* **2010**, *11* (11), 789–801.
- (2) The Cancer Genome Atlas, <http://cancergenome.nih.gov/> (accessed Aug 31, 2011).
- (3) Hanash, S.; Taguchi, A. The grand challenge to decipher the cancer proteome. *Nat. Rev. Cancer* **2010**, *10* (9), 652–60.
- (4) Bozic, I.; Antal, T.; Ohtsuki, H.; Carter, H.; Kim, D.; Chen, S.; Karchin, R.; Kinzler, K. W.; Vogelstein, B.; Nowak, M. A. Accumulation of driver and passenger mutations during tumor progression. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107* (43), 18545–50.
- (5) Fearon, E. R.; Vogelstein, B. A genetic model for colorectal tumorigenesis. *Cell* **1990**, *61* (5), 759–67.
- (6) Catalogue of Somatic Mutations in Cancer, <http://www.sanger.ac.uk/genetics/CGP/cosmic/> (accessed Aug 31, 2011).
- (7) Fearon, E. R. Molecular genetics of colorectal cancer. *Annu. Rev. Pathol. Mol. Biol.* **2011**, *6*, 479–507.
- (8) Wood, L. D.; Parsons, D. W.; Jones, S.; Lin, J.; Sjöblom, T.; Leary, R. J.; Shen, D.; Boca, S. M.; Barber, T.; Ptak, J.; Silliman, N.; Szabo, S.; Dezso, Z.; Ustyankovsky, V.; Nikolskaya, T.; Nikolsky, Y.; Karchin, R.; Wilson, P. A.; Kaminker, J. S.; Zhang, Z.; Croshaw, R.; Willis, J.; Dawson, D.; Shipitsin, M.; Willson, J. K.; Sukumar, S.; Polyak, K.; Park, B. H.; Pethiyagoda, C. L.; Pant, P. V.; Ballinger, D. G.; Sparks, A. B.; Hartigan, J.; Smith, D. R.; Suh, E.; Papadopoulos, N.; Buckhaults, P.; Markowitz, S. D.; Parmigiani, G.; Kinzler, K. W.; Velculescu, V. E.; Vogelstein, B. The genomic landscapes of human breast and colorectal cancers. *Science* **2007**, *318* (5853), 1108–13.
- (9) MacDonald, B. T.; Tamai, K.; He, X. Wnt/beta-catenin signaling: components, mechanisms, and diseases. *Dev. Cell* **2009**, *17* (1), 9–26.

- (10) Nishisho, I.; Nakamura, Y.; Miyoshi, Y.; Miki, Y.; Ando, H.; Horii, A.; Koyama, K.; Utsunomiya, J.; Baba, S.; Hedge, P. Mutations of chromosome 5q21 genes in FAP and colorectal cancer patients. *Science* **1991**, *253* (5020), 665–9.
- (11) Faux, M. C.; Ross, J. L.; Meeker, C.; Johns, T.; Ji, H.; Simpson, R. J.; Layton, M. J.; Burgess, A. W. Restoration of full-length adenomatous polyposis coli (APC) protein in a colon cancer cell line enhances cell adhesion. *J. Cell Sci.* **2004**, *117* (Pt 3), 427–39.
- (12) Liu, H.; Sadygov, R. G.; Yates, J. R., 3rd A model for random sampling and estimation of relative protein abundance in shotgun proteomics. *Anal. Chem.* **2004**, *76* (14), 4193–201.
- (13) Old, W. M.; Meyer-Arendt, K.; Aveline-Wolf, L.; Pierce, K. G.; Mendoza, A.; Sevinsky, J. R.; Resing, K. A.; Ahn, N. G. Comparison of label-free methods for quantifying human proteins by shotgun proteomics. *Mol. Cell. Proteomics* **2005**, *4* (10), 1487–502.
- (14) Picotti, P.; Bodenmiller, B.; Mueller, L. N.; Domon, B.; Aebersold, R. Full dynamic range proteome analysis of *S. cerevisiae* by targeted proteomics. *Cell* **2009**, *138* (4), 795–806.
- (15) Gerber, S. A.; Rush, J.; Stenman, O.; Kirschner, M. W.; Gygi, S. P. Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100* (12), 6940–5.
- (16) Sprung, R. W., Jr.; Brock, J. W.; Tanksley, J. P.; Li, M.; Washington, M. K.; Slebos, R. J.; Liebler, D. C. Equivalence of protein inventories obtained from formalin-fixed paraffin-embedded and frozen tissue in multidimensional liquid chromatography-tandem mass spectrometry shotgun proteomic analysis. *Mol. Cell. Proteomics* **2009**, *8* (8), 1988–98.
- (17) Slebos, R. J.; Brock, J. W.; Winters, N. F.; Stuart, S. R.; Martinez, M. A.; Li, M.; Chambers, M. C.; Zimmerman, L. J.; Ham, A. J.; Tabb, D. L.; Liebler, D. C. Evaluation of strong cation exchange versus isoelectric focusing of peptides for multidimensional liquid chromatography-tandem mass spectrometry. *J. Proteome Res.* **2008**, *7* (12), 5286–94.
- (18) Licklider, L. J.; Thoreen, C. C.; Peng, J.; Gygi, S. P. Automation of nanoscale microcapillary liquid chromatography-tandem mass spectrometry with a vented column. *Anal. Chem.* **2002**, *74* (13), 3076–83.
- (19) Kessner, D.; Chambers, M.; Burke, R.; Agus, D.; Mallick, P. ProteoWizard: open source software for rapid proteomics tools development. *Bioinformatics* **2008**, *24* (21), 2534–6.
- (20) Tabb, D. L.; Fernando, C. G.; Chambers, M. C. MyriMatch: highly accurate tandem mass spectral peptide identification by multivariate hypergeometric analysis. *J. Proteome Res.* **2007**, *6* (2), 654–61.
- (21) Zhang, B.; Chambers, M. C.; Tabb, D. L. Proteomic parsimony through bipartite graph analysis improves accuracy and transparency. *J. Proteome Res.* **2007**, *6* (9), 3549–57.
- (22) Ma, Z. Q.; Dasari, S.; Chambers, M. C.; Litton, M. D.; Sobecki, S. M.; Zimmerman, L. J.; Halvey, P. J.; Schilling, B.; Drake, P. M.; Gibson, B. W.; Tabb, D. L. IDPicker 2.0: Improved protein assembly with high discrimination peptide identification filtering. *J. Proteome Res.* **2009**, *8* (8), 3872–81.
- (23) Li, M.; Gray, W.; Zhang, H.; Chung, C. H.; Billheimer, D.; Yarbrough, W. G.; Liebler, D. C.; Shyr, Y.; Slebos, R. J. Comparative shotgun proteomics using spectral count data and quasi-likelihood modeling. *J. Proteome Res.* **2010**, *9* (8), 4295–305.
- (24) MacLean, B.; Tomazela, D. M.; Shulman, N.; Chambers, M.; Finney, G. L.; Frewen, B.; Kern, R.; Tabb, D. L.; Liebler, D. C.; MacCoss, M. J. Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics* **2010**, *26* (7), 966–8.
- (25) Zhang, H.; Liu, Q.; Zimmerman, L. J.; Ham, A. J.; Slebos, R. J.; Rahman, J.; Kikuchi, T.; Massion, P. P.; Carbone, D. P.; Billheimer, D.; Liebler, D. C. Methods for Peptide and protein quantitation by liquid chromatography-multiple reaction monitoring mass spectrometry. *Mol. Cell. Proteomics* **2011**, *10* (6), No. M110006593.
- (26) Reimers, M.; Carey, V. J. Bioconductor: an open source framework for bioinformatics and computational biology. *Methods Enzymol.* **2006**, *411*, 119–34.
- (27) Kirov, S. A.; Zhang, B.; Snoddy, J. R. Association analysis for large-scale gene set data. *Methods Mol. Biol.* **2007**, *408*, 19–33.
- (28) Zhang, B.; Kirov, S.; Snoddy, J. WebGestalt: an integrated system for exploring gene sets in various biological contexts. *Nucleic Acids Res.* **2005**, *33* (Web Server issue), W741–8.
- (29) Kim, D. W.; Chae, J. I.; Kim, J. Y.; Pak, J. H.; Koo, D. B.; Bahk, Y. Y.; Seo, S. B. Proteomic analysis of apoptosis related proteins regulated by proto-oncogene protein DEK. *J. Cell Biochem.* **2009**, *106* (6), 1048–59.
- (30) Mao, L.; Zabel, C.; Herrmann, M.; Nolden, T.; Mertes, F.; Magnol, L.; Chabert, C.; Hartl, D.; Herault, Y.; Delabar, J. M.; Manke, T.; Himmelbauer, H.; Klose, J. Proteomic shifts in embryonic stem cells with gene dose modifications suggest the presence of balancer proteins in protein regulatory networks. *PLoS One* **2007**, *2* (11), e1218.
- (31) Collier, T. S.; Sarkar, P.; Franck, W. L.; Rao, B. M.; Dean, R. A.; Muddiman, D. C. Direct comparison of stable isotope labeling by amino acids in cell culture and spectral counting for quantitative proteomics. *Anal. Chem.* **2010**, *82* (20), 8696–702.
- (32) Paulovich, A. G.; Billheimer, D.; Ham, A. J.; Vega-Montoto, L.; Rudnick, P. A.; Tabb, D. L.; Wang, P.; Blackman, R. K.; Bunk, D. M.; Cardasis, H. L.; Clauser, K. R.; Kinsinger, C. R.; Schilling, B.; Tegeler, T. J.; Variyath, A. M.; Wang, M.; Whiteaker, J. R.; Zimmerman, L. J.; Fenyo, D.; Carr, S. A.; Fisher, S. J.; Gibson, B. W.; Mesri, M.; Neubert, T. A.; Regnier, F. E.; Rodriguez, H.; Spiegelman, C.; Stein, S. E.; Tempst, P.; Liebler, D. C. Interlaboratory study characterizing a yeast performance standard for benchmarking LC-MS platform performance. *Mol. Cell. Proteomics* **2010**, *9* (2), 242–54.
- (33) Tabb, D. L.; Vega-Montoto, L.; Rudnick, P. A.; Variyath, A. M.; Ham, A. J.; Bunk, D. M.; Kilpatrick, L. E.; Billheimer, D. D.; Blackman, R. K.; Cardasis, H. L.; Carr, S. A.; Clauser, K. R.; Jaffe, J. D.; Kowalski, K. A.; Neubert, T. A.; Regnier, F. E.; Schilling, B.; Tegeler, T. J.; Wang, M.; Wang, P.; Whiteaker, J. R.; Zimmerman, L. J.; Fisher, S. J.; Gibson, B. W.; Kinsinger, C. R.; Mesri, M.; Rodriguez, H.; Stein, S. E.; Tempst, P.; Paulovich, A. G.; Liebler, D. C.; Spiegelman, C. Repeatability and reproducibility in proteomic identifications by liquid chromatography-tandem mass spectrometry. *J. Proteome Res.* **2010**, *9* (2), 761–76.
- (34) Tu, C.; Rudnick, P. A.; Martinez, M. Y.; Cheek, K. L.; Stein, S. E.; Slebos, R. J.; Liebler, D. C. Depletion of abundant plasma proteins and limitations of plasma proteomics. *J. Proteome Res.* **2010**, *9* (10), 4982–91.
- (35) Chiosea, S.; Jelezova, E.; Chandran, U.; Acquafonda, M.; McHale, T.; Sobol, R. W.; Dhir, R. Up-regulation of dicer, a component of the MicroRNA machinery, in prostate adenocarcinoma. *Am. J. Pathol.* **2006**, *169* (5), 1812–20.
- (36) Karube, Y.; Tanaka, H.; Osada, H.; Tomida, S.; Tatematsu, Y.; Yanagisawa, K.; Yatabe, Y.; Takamizawa, J.; Miyoshi, S.; Mitsudomi, T.; Takahashi, T. Reduced expression of Dicer associated with poor prognosis in lung cancer patients. *Cancer Sci.* **2005**, *96* (2), 111–5.
- (37) Merritt, W. M.; Lin, Y. G.; Han, L. Y.; Kamat, A. A.; Spannuth, W. A.; Schmandt, R.; Urbauer, D.; Pennacchio, L. A.; Cheng, J. F.; Nick, A. M.; Deavers, M. T.; Mourad-Zeidan, A.; Wang, H.; Mueller, P.; Lenburg, M. E.; Gray, J. W.; Mok, S.; Birrer, M. J.; Lopez-Berestein, G.; Coleman, R. L.; Bar-Eli, M.; Sood, A. K. Dicer, Drosha, and outcomes in patients with ovarian cancer. *N. Engl. J. Med.* **2008**, *359* (25), 2641–50.
- (38) Faber, C.; Horst, D.; Hlubek, F.; Kirchner, T. Overexpression of Dicer predicts poor survival in colorectal cancer. *Eur. J. Cancer* **2011**, *47* (9), 1414–9.
- (39) Chiosea, S.; Acquafonda, M.; Luo, J.; Kuan, S.; Seethala, R. DICER1 and PRKRA in Colon Adenocarcinoma. *Biomark Insights* **2008**, *3*, 253–58.
- (40) Huang, X. H.; Wang, Q.; Chen, J. S.; Fu, X. H.; Chen, X. L.; Chen, L. Z.; Li, W.; Bi, J.; Zhang, L. J.; Fu, Q.; Zeng, W. T.; Cao, L. Q.; Tan, H. X.; Su, Q. Bead-based microarray analysis of microRNA expression in hepatocellular carcinoma: miR-338 is downregulated. *Hepatol. Res.* **2009**, *39* (8), 786–94.
- (41) Iliopoulos, D.; Rotem, A.; Struhl, K. Inhibition of miR-193a expression by Max and RXR $\alpha$  activates K-Ras and PLA2U to mediate distinct aspects of cellular transformation. *Cancer Res.* **2011**, *71* (15), 5144–53.
- (42) Selbach, M.; Schwanhausser, B.; Thierfelder, N.; Fang, Z.; Khanin, R.; Rajewsky, N. Widespread changes in protein synthesis induced by microRNAs. *Nature* **2008**, *455* (7209), 58–63.

- (43) Vandewalle, C.; Van Roy, F.; Berx, G. The role of the ZEB family of transcription factors in development and disease. *Cell. Mol. Life Sci.* **2009**, *66* (5), 773–87.
- (44) Sanchez-Tillo, E.; Lazaro, A.; Torrent, R.; Cuatrecasas, M.; Vaquero, E. C.; Castells, A.; Engel, P.; Postigo, A. ZEB1 represses E-cadherin and induces an EMT by recruiting the SWI/SNF chromatin-remodeling protein BRG1. *Oncogene* **2010**, *29* (24), 3490–500.
- (45) Krupnik, V. E.; Sharp, J. D.; Jiang, C.; Robison, K.; Chickering, T. W.; Amaravadi, L.; Brown, D. E.; Guyot, D.; Mays, G.; Leiby, K.; Chang, B.; Duong, T.; Goodearl, A. D.; Gearing, D. P.; Sokol, S. Y.; McCarthy, S. A. Functional and structural diversity of the human Dickkopf gene family. *Gene* **1999**, *238* (2), 301–13.
- (46) Niehrs, C. Function and biological roles of the Dickkopf family of Wnt modulators. *Oncogene* **2006**, *25* (57), 7469–81.
- (47) Baehs, S.; Herbst, A.; Thieme, S. E.; Perschl, C.; Behrens, A.; Scheel, S.; Jung, A.; Brabletz, T.; Goke, B.; Blum, H.; Kolligs, F. T. Dickkopf-4 is frequently down-regulated and inhibits growth of colorectal cancer cells. *Cancer Lett.* **2009**, *276* (2), 152–9.
- (48) Matsuzawa, S. I.; Reed, J. C.; Siah-1, S. I. P. and Ebi collaborate in a novel pathway for beta-catenin degradation linked to p53 responses. *Mol. Cell* **2001**, *7* (5), 915–26.
- (49) Chen, X.; Mo, P.; Li, X.; Zheng, P.; Zhao, L.; Xue, Z.; Ren, G.; Han, G.; Wang, X.; Fan, D. CacyBP/SIP protein promotes proliferation and G1/S transition of human pancreatic cancer cells. *Mol. Carcinog.* **2011**, *50* (10), 804–10.
- (50) Wang, L.; Wang, J.; Sun, S.; Rodriguez, M.; Yue, P.; Jang, S. J.; Mao, L. A novel DNMT3B subfamily, DeltaDNMT3B, is the predominant form of DNMT3B in non-small cell lung cancer. *Int. J. Oncol.* **2006**, *29* (1), 201–7.
- (51) DiFeo, A.; Feld, L.; Rodriguez, E.; Wang, C.; Beer, D. G.; Martignetti, J. A.; Narla, G. A functional role for KLF6-SV1 in lung adenocarcinoma prognosis and chemotherapy response. *Cancer Res.* **2008**, *68* (4), 965–70.
- (52) Chen, X.; Truong, T. T.; Weaver, J.; Bove, B. A.; Cattie, K.; Armstrong, B. A.; Daly, M. B.; Godwin, A. K. Intronic alterations in BRCA1 and BRCA2: effect on mRNA splicing fidelity and expression. *Hum. Mutat.* **2006**, *27* (5), 427–35.
- (53) Fackenthal, J. D.; Godley, L. A. Aberrant RNA splicing and its functional consequences in cancer cells. *Dis. Model Mech.* **2008**, *1* (1), 37–42.
- (54) Long, J. C.; Caceres, J. F. The SR protein family of splicing factors: master regulators of gene expression. *Biochem. J.* **2009**, *417* (1), 15–27.
- (55) Sahai, E.; Marshall, C. J. RHO-GTPases and cancer. *Nat. Rev. Cancer* **2002**, *2* (2), 133–42.
- (56) Bardelli, A.; Saha, S.; Sager, J. A.; Romans, K. E.; Xin, B.; Markowitz, S. D.; Lengauer, C.; Velculescu, V. E.; Kinzler, K. W.; Vogelstein, B. PRL-3 expression in metastatic cancers. *Clin. Cancer Res.* **2003**, *9* (15), 5607–15.
- (57) Winter-Vann, A. M.; Baron, R. A.; Wong, W.; dela Cruz, J.; York, J. D.; Gooden, D. M.; Bergo, M. O.; Young, S. G.; Toone, E. J.; Casey, P. J. A small-molecule inhibitor of isoprenylcysteine carboxyl methyltransferase with antitumor activity in cancer cells. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102* (12), 4336–41.
- (58) Jang, G. F.; Gelb, M. H. Substrate specificity of mammalian prenyl protein-specific endoprotease activity. *Biochemistry* **1998**, *37* (13), 4473–81.
- (59) Sever, N.; Song, B. L.; Yabe, D.; Goldstein, J. L.; Brown, M. S.; DeBose-Boyd, R. A. Insig-dependent ubiquitination and degradation of mammalian 3-hydroxy-3-methylglutaryl-CoA reductase stimulated by sterols and geranylgeraniol. *J. Biol. Chem.* **2003**, *278* (52), 52479–90.