Clinical Determinants of Response to Irinotecan-Based Therapy Derived from Cell Line Models


Abstract

Purpose: In an attempt to identify genes that are involved in resistance to SN38, the active metabolite of irinotecan (also known as CPT-11), we carried out DNA microarray profiling of matched HCT116 human colon cancer parental cell lines and SN38-resistant cell lines following treatment with SN38 over time.

Experimental Design: Data analysis identified a list of genes that were acutely altered in the parental cells following SN38 treatment as well as constitutively altered in the SN38-resistant cells.

Results: Independent validation of 20% of these genes by quantitative reverse transcription-PCR revealed a strong correlation with the microarray results: Pearson's correlation was 0.781 ($r^2 = 0.61, P < 0.000001$) for those genes that were acutely altered in the parental setting following SN38 treatment and 0.795 ($r^2 = 0.63, P < 0.000002$) for those genes that were constitutively altered in the SN38-resistant cells. We then assessed the ability of our in vitro-derived gene list to predict clinical response to 5-fluorouracil/irinotecan using pretreatment metastatic biopsies from responding and nonresponding colorectal cancer patients using both unsupervised and supervised approaches. When principal components analysis was used with our in vitro classifier gene list, a good separation between responding and nonresponding patients was obtained, with only one nonresponding and two responding patients separating with the incorrect groups. Supervised class prediction using support vector machines algorithm identified a 16-gene classifier with 75% overall accuracy, 81.8% sensitivity, and 66.6% specificity.

Conclusions: These results suggest that in vitro-derived gene lists can be used to predict clinical response to chemotherapy in colorectal cancer.

In advanced-stage colorectal cancer, combination therapy with 5-fluorouracil (5-FU) and irinotecan (CPT-11) produces inadequate response rates of about 40% to 50% (1, 2). The main reason for the lack of response is drug resistance, which may be either innate or acquired in nature. Many studies have identified markers that may be useful in predicting response to these chemotherapies (3–9), although due to several factors, none of these potential predictive markers have progressed to routine clinical use. Studies have also established that there appears to be a greater ability to predict response to treatment when single predictive markers are combined into cassettes of markers (7). The ultimate aim of predictive marker testing is to tailor treatment according to an individual patient/tumor profile. Most active agents against colorectal cancer have only moderate clinical activity; hence, there is a need to find new predictive markers to identify those patients who are most likely to respond to the treatment.

In the past, many studies have used gene expression profiling to identify novel resistance markers and markers that may represent novel drug targets (10, 11). More recently, gene expression profiling studies have attempted to identify both prognostic (12–16) and predictive (17, 18) signatures of response using patient biopsies, clinical response data, and supervised classification approaches.

The aim of this study was to examine the gene expression changes of HCT116 colorectal cancer cells following acute exposure to SN38 (the active metabolite of CPT-11) over time and also the constitutive gene expression of HCT116 SN38-resistant daughter cells (19). We were most interested in identifying those in vitro genes that were both acutely altered following SN38 treatment in the parental cells and also constitutively altered in the SN38-resistant cells, as we feel that these may represent the key genes mediating SN38 resistance. In addition, we wanted to assess the feasibility of using our in vitro-derived gene sets as determinants of response to chemotherapy in patients.
Translational Relevance

Although there have been significant improvements in the treatment of advanced colorectal cancer, drug resistance remains a major problem limiting the effectiveness of chemotherapy. This study aims to identify genes associated with irinotecan (CPT-11) resistance in vitro and to fully assess the utility of these genes in predicting response in the clinical setting. Firstly, those genes mediating resistance to CPT-11 may represent novel targets for therapeutic intervention; if further phenotypic analysis of these genes proved positive, targeting these genes in combination with traditional chemotherapy should enhance patient response to this treatment. Secondly, the identification of a gene signature predictive of patient response to combination 5-fluorouracil/CPT-11 therapy could potentially be used to predict which patients will not respond to combined 5-fluorouracil/CPT-11 therapy. Following further independent validation, it is hoped such a signature could be routinely used to predict response sparing those patients unlikely to respond to treatment unnecessary toxicity, time, and expense.

Materials and Methods

Materials

SN38 was purchased from Abatara Technology.

Cell culture

The p53 wild-type and p53-null HCT116 human colon cancer cell lines were kindly provided by Prof. Bert Vogelstein (Johns Hopkins University). The SN38-resistant HCT116 subline was generated from the HCT116 p53 wild-type parental cell line in our laboratory by repeated exposure to stepwise increasing concentrations of SN38 over a period of ~9 months and spiked every 4 weeks with 5 nmol/L SN38 to maintain the resistant phenotype (19). The IC_{50} for SN38 in the HCT116 parental cell line was 5 nmol/L, whereas the IC_{50} for SN38 was >100 nmol/L in the HCT116 SN38-resistant subline. All HCT116 cell lines were maintained in McCoy’s 5A medium supplemented with 10% dialyzed FCS, 50 μg/mL penicillin-streptomycin, 2 mmol/L L-glutamine, and 1 mmol/L sodium pyruvate. RKO and HT29 cells (obtained from the National Cancer Institute) were maintained in DMEM and supplemented as above. LoVo cells (kindly provided by AstraZeneca) were maintained in McCoy’s 5A medium supplemented as above minus sodium pyruvate (all media and supplements from Invitrogen Life Technologies). All cell lines were maintained at 37°C in a humidified atmosphere containing 5% CO_{2}.

Microarray analysis

HCT116 parental cells were treated with 5 nmol/L SN38 for 0, 6, 12, and 24 h as outlined in Fig. 1. Also, untreated SN38-resistant cells were analyzed to identify those genes that are constitutively deregulated between parental and resistant cells. Total RNA was isolated from three independent experiments using the RNA STAT-60 Total RNA isolation reagent (Tel-Test) according to the manufacturer’s instructions. Total RNA (5 μg) was sent to Almac Diagnostics for cDNA synthesis, cRNA synthesis, fragmentation, and hybridization onto Affymetrix HG U133 Plus2.0 microarrays. Detailed experimental protocols and raw expression data are available at ArrayExpress^3 (accession no. E-MEXP-1171).

3www.ebi.ac.uk/arrayexpress/

Data analysis

Generation of gene lists. Microarray data analysis was done using GeneSpring v7.3 (Agilent Technologies UK) as described previously (10). Initially, two independent experiments were created from the microarray raw data. Firstly, to analyze drug-inducible gene expression in HCT116 parental cells, all genes from each of the three replicate arrays were normalized to the median signal intensity of that array. Secondly, each gene on the 6, 12, and 24 h sample arrays was normalized to the median signal intensity of the respective gene on the 0 h (control) array. In addition, the variance for each gene was estimated using the Cross-Gene Error Model (based on the Rocke-Lorenzo model). Genes were filtered using three variables. Firstly, all genes that displayed an Affymetrix present or marginal flag call in all samples were retained. The gene list was subsequently filtered using the cross-gene error model, with genes displaying control values greater than the average base/proportional value being retained (this was required in all samples for any gene). Finally, the list was filtered using a 2-fold cutoff for each gene relative to the 0 h control, with genes meeting this criterion in at least one of the three time points being retained. The genes passing these three filters were considered to be drug-inducible. All data are displayed as log_{2}.

The second experiment aimed to compare basal gene expression in the HCT116 parental cell line relative to the SN38-resistant daughter line. As described above, all genes on each of the three replicate arrays were initially normalized to the median signal intensity of that array. Each gene was then normalized to the median signal intensity of the respective gene on all arrays. The data were filtered as described above, with genes required to show a 2-fold up-regulation or down-regulation in the resistant cell line relative to the parental line. The resultant lists of genes were considered transcriptionally deregulated in drug-resistant cells relative to parental cells.

The two gene lists described above were combined using the Venn diagram functionality in GeneSpring to identify those genes that were both constitutively deregulated in the drug-resistant cell lines and induced or repressed in the parental cell line following acute exposure to chemotherapy. The genes that were common to both lists appeared in the intersection of the diagram and were subsequently retained as the final working gene list.

Unsupervised analysis. The 124 in vitro-derived genes were taken forward to assess if they were able to separate 5-FU/CPT-11 responding and nonresponding advanced colorectal cancer patients using both unsupervised and supervised approaches (Fig. 1C). In unsupervised analysis, principal components analysis (PCA) was used to separate 5-FU/CPT-11 responding and nonresponding patients based on the expression of the in vitro-derived genes identified from the initial microarray analysis.

Supervised analysis. In supervised analysis, the in vitro-derived gene list was used as the starting features to classify 5-FU/CPT-11 responding and nonresponding patients. The classification rule was defined using the support vector machine algorithm (Genespring v7.3; Agilent Technologies) with Fisher’s exact test used as the gene selection method and the polydot product function method as the kernel function. Finally, leave-one-out cross-validation was used to estimate the performance and accuracy of the resultant class prediction rule.

Quantitative reverse transcription-PCR analysis

To validate the microarray results, we measured the expression of a representative number of genes from the resultant crossover gene list (Fig. 1B; Supplementary Table S3) by quantitative reverse transcription-PCR (RT-PCR). We chose 25 (20%) genes from the list of 124 genes for validation in three independent experiments. The genes were selected based on at least 2-fold induction, with both highly and more moderately induced genes chosen, and both up-regulated and down-regulated genes were analyzed. In general, the validation gene list accurately represented the overall expression trends observed in the original gene list.

Total RNA was isolated from three independent experiments as described above. Reverse transcription was carried out using 2 μg RNA.
using a Moloney murine leukemia virus-based reverse transcriptase kit (Invitrogen) according to the manufacturer’s instructions. Quantitative RT-PCR amplification was carried out in a final volume of 10 μL containing 5 μL of 2 × SYBR Green master mix (Qiagen), 4 μL primers (2 μmol/L), and 1 μL cDNA using an Opticon DNA Engine Thermal Cycler (Bio-Rad Laboratories). All amplifications were primed by pairs of chemically synthesized 18- to 22-mer oligonucleotides designed using freely available primer design software (Primer3)4 (Supplementary Table S1). Initially, temperature gradient analysis was carried out to determine primer-specific annealing temperatures. Reaction conditions were activation at 95°C for 10 min, denaturation at 95°C for 15 s, annealing 54°C to 62°C for 30 s, and extension 72°C for 1 min. All PCR amplifications were carried out for 40 cycles. Melt curve analysis was used to examine the specificity of an amplified product. All primers amplified products of similar size (∼100-200 bp) with similar melting temperatures. Standard curves were generated to quantify the absolute expression levels of each target gene and the 18S RNA reference gene in each sample. The relative expression level of each gene in samples of interest was calculated by dividing the amount of normalized target by the value in an untreated calibrator sample.

To discover whether the changes in gene expression were cell line-dependent phenomena, we examined the expression changes of 6 genes at 24 h following treatment with SN38 (IC50(24 h) dose) in the HCT116 p53-null (5 nmol/L), RKO (7 nmol/L), H630 (25 nmol/L), LoVo (25 nmol/L), and HT29 (25 nmol/L) colorectal cancer cell lines. The genes that we chose to examine in the extended panel of colorectal cancer cell lines were those genes that displayed the highest levels of correlation between microarray and quantitative RT-PCR expression values throughout the time course experiment: IFITM1, CCNG2, TPMT, NUP8-8, DDHD1, and FLJ3871.

**Fresh-frozen clinical samples**

Total RNA extracted from a retrospective series of 20 pretreatment metastatic tumor biopsies from patients with advanced colorectal cancer was used for gene expression profiling studies. All patients provided written fully informed consent as per institutional review board/review ethics committee guidelines in the University of Southern California and approval was granted from this body. These patients underwent biopsy of colorectal liver metastases before commencing CPT-11/5-FU chemotherapy on the IFL schedule: CPT-11 125 mg/m2 i.v. over 90 min and leucovorin 20 mg/m2 as i.v. bolus injection immediately before 5-FU and 5-FU 500 mg/m2 as i.v. bolus injection administered weekly for 4 weeks and repeated every 6 weeks. Of the 20 patients, the majority were male (n = 14). The primary tumor site included both colonic and rectal tumors, with the majority originating in the colon (n = 14).

Computed tomography imaging for response evaluation using WHO criteria was done every 6 weeks. Patients with measurable disease whose

---

4 http://frodo.wi.mit.edu
tumor burden had decreased by ≥50% for at least 6 weeks were classified as responders. Patients whose tumor and all evidence of disease had disappeared were classified as showing complete response. Responders with anything less than complete response were categorized as showing partial response. Progressive disease was defined as a ≥25% increase in tumor burden (compared with the smallest measurement) or the appearance of new metastatic lesions. Of these 20 patients, 1 had a complete response to treatment, 10 had a partial response to treatment, and 9 had progressive disease on treatment. For the purpose of this study, we have further defined “responders” as those patients with either complete or partial response and “nonresponders” as those patients with progressive disease. We have specifically excluded patients with stable disease on chemotherapy treatment.

Statistical analysis

For the 25 genes chosen for quantitative RT-PCR validation, the average fold changes by both microarray and quantitative RT-PCR were log transformed and the correlation between the expression values were examined using Pearson’s product correlation moment ($r$; MS Excel). The significance of the correlation was determined using a two-tailed test of significance. The following definitions of sensitivity and specificity were used: Sensitivity = Number of true positives / Number of true positives + Number of false negatives. Specificity = Number of true negatives / Number of true negatives + Number of false positives.

Results

Gene expression analysis of HCT116 parental and SN38-resistant cell lines. We carried out a gene expression profiling experiment to identify genes that may be mediating resistance to SN38 using the Affymetrix HGU133 Plus2.0 arrays (Fig. 1A). The transcriptional profiles of the HCT116 parental cells following treatment with 5 nmol/L SN38 (approximate IC$_{50}$ dose at 72 h) for 6, 12, and 24 h were examined. In addition, we also examined the constitutive transcriptional profile of an HCT116 SN38-resistant daughter cell line (19). Following normalization and filtering, 357 genes were found to be significantly up-regulated or down-regulated by ≥2-fold in HCT116 parental cells following SN38 treatment (Fig. 1A; Supplementary Table S2). Further analysis found that there were 382 genes whose expression was constitutively altered in the SN38-resistant cells compared with the parental line (Fig. 1A; Supplementary Table S3). When we combined these two gene lists into a Venn diagram, we found that 124 genes were common to both gene lists (Fig. 1B; Supplementary Table S4). This crossover gene list represented the key genes mediating SN38 resistance.

Validation of microarray results. To confirm the microarray experiment, 25 genes were chosen for validation by quantitative RT-PCR analysis. For the 25 genes acutely altered in the HCT116 parental cells following SN38 treatment over 6, 12, and 24 h, the Pearson’s correlation ($r$) was 0.781, with $r^2 = 0.61$ ($P = 0.000001$; Fig. 2A). In terms of the genes that were constitutively altered between parental and resistant cells, Pearson’s correlation ($r$) of the 25 genes was 0.795, with $r^2 = 0.67$ ($P = 0.000002$; Fig. 2B). We also examined the correlation of the data in a time-dependent manner and found that the highest correlation occurred at 12 h, with a correlation ($r$) of 0.797, $r^2 = 0.6353$ ($P = 0.000002$; Supplementary Fig. S1B), followed by 24 h ($r = 0.734$, $r^2 = 0.5383$, $P = 0.000003$; Supplementary Fig. S1C) and 6 h ($r = 0.704$, $r^2 = 0.4953$, $P = 0.000086$; Supplementary Fig. S1A). These results compared well with the observed gene expression changes, as the largest changes in expression occurred at 12 h followed by 24 and 6 h as shown by both microarray and quantitative RT-PCR results.

For the 25 genes chosen for validation, 21 (84%) genes showed good overall concordance between microarray and quantitative PCR results. In addition, 64% (16 of 25) showed statistically significant correlations between expression values from the microarray results and the quantitative PCR results. 5 The following definitions of sensitivity and specificity were used: Sensitivity = Number of true positives / Number of true positives + Number of false negatives. Specificity = Number of true negatives / Number of true negatives + Number of false positives.

![Image](http://www.wessa.net/rwasp_correlation.wasp)

Fig. 2. Correlation between average fold changes in gene expression levels for (A) 25 target genes that are acutely altered in the HCT116 parental cells following 5 nmol/L SN38 treatment for 6, 12, and 24 h and (B) 25 genes that are constitutively altered in the HCT116 parental versus SN38-resistant daughter cells. The correlation coefficients were determined by comparing the average DNA microarray fold changes and average quantitative RT-PCR fold changes. All data are displayed as log2.
following SN38 (IC₅₀) treatment in the extended cell line panel, with the exception of the hypothetical protein FLJ3871 in the HT29 cell line. The highest levels of down-regulation for NUP8-8 were observed in the H630 and LoVo cells (Fig. 3B). NUP8-8 is an essential member of the nuclear pore complex and studies have shown that NUP8-8 is overexpressed in tumor cells and that it is associated with a more aggressive phenotype in breast cancer (21, 22). Down-regulation of DDHD1 was observed in each cell line (Fig. 3B). The highest levels of down-regulation for FLJ3871 were observed in the RKO cells (Fig. 3B).

**Gene expression analysis in clinical samples.** We then wanted to assess whether our *in vitro*-derived gene lists were able to predict response to 5-FU/CPT-11 using pretreatment metastatic biopsies from colorectal cancer patients who were subsequently found to have either responded or not responded to 5-FU/CPT-11 therapy. This analysis was done using both unsupervised and supervised approaches (Fig. 1C).

Within the patient-derived gene list, we examined the expression of the 124 genes that we have identified previously from our SN38 *in vitro* microarray experiment (described above). We then applied a fold-change filter to identify which of those genes were changing >1.5-fold between patients who responded to 5-FU/CPT-11 therapy and those who did not. The 1.5-fold gene list contained only two genes, *Drosophila discs* (Table 1).

**Table 1.** Correlation analysis of microarray and real-time RT-PCR validation results

<table>
<thead>
<tr>
<th>Common name</th>
<th>Description</th>
<th>Correlation coefficient, r</th>
<th>Coefficient of determination, r²</th>
<th>P</th>
<th>Biological function</th>
</tr>
</thead>
<tbody>
<tr>
<td>SLC30A1</td>
<td>Solute carrier family 30 (zinc transporter), member 1</td>
<td>0.13</td>
<td>0.02</td>
<td>0.068</td>
<td>Cation transport</td>
</tr>
<tr>
<td>NUP8-8 Yeas 2</td>
<td>Nucleoporin 88 kDa YEATS domain containing 2</td>
<td>0.99</td>
<td>0.98</td>
<td>&lt;0.0001</td>
<td>Transporter activity</td>
</tr>
<tr>
<td>MCCC2</td>
<td>Methylcrotonoyl-coenzyme A carboxylase 2 ( regulators)</td>
<td>0.9</td>
<td>0.81</td>
<td>&lt;0.0001</td>
<td>Regulation of transcription, DNA-dependent</td>
</tr>
<tr>
<td>HSPC05-6</td>
<td>Armadillo repeat containing 8</td>
<td>0.68</td>
<td>0.46</td>
<td>0.0149</td>
<td>Binding</td>
</tr>
<tr>
<td>XRCC5</td>
<td>X-ray repair complementing defective repair in Chinese hamster cells 5</td>
<td>0.85</td>
<td>0.72</td>
<td>0.0005</td>
<td>ATP-dependent DNA helicase activity</td>
</tr>
<tr>
<td>DDHD1</td>
<td>DDHD domain containing 1</td>
<td>0.57</td>
<td>0.33</td>
<td>0.0529</td>
<td>Metal ion binding; hydrolase activity</td>
</tr>
<tr>
<td>CCNG2</td>
<td>Cyclin G2</td>
<td>0.93</td>
<td>0.86</td>
<td>&lt;0.0001</td>
<td>Cell cycle checkpoint</td>
</tr>
<tr>
<td>IFITM1</td>
<td>IFN-induced transmembrane protein 1 (9-27)</td>
<td>0.87</td>
<td>0.75</td>
<td>0.0002</td>
<td>Receptor signaling</td>
</tr>
<tr>
<td>ARID2</td>
<td>AT-rich interactive domain 2 (ARID, RFX-like)</td>
<td>0.78</td>
<td>0.61</td>
<td>0.0028</td>
<td>DNA binding; protein binding</td>
</tr>
<tr>
<td>NUP15-3</td>
<td>Nucleoporin 153 kDa</td>
<td>-0.08</td>
<td>0.01</td>
<td>0.8</td>
<td>DNA binding; transporter activity</td>
</tr>
<tr>
<td>MLLT10</td>
<td>Myeloid/lymphoid or mixed-lineage leukemia</td>
<td>0.98</td>
<td>0.96</td>
<td>&lt;0.0001</td>
<td>Transcription factor activity; binding</td>
</tr>
<tr>
<td>ARF1</td>
<td>ADP-ribosylation factor 1</td>
<td>0.26</td>
<td>0.07</td>
<td>0.414</td>
<td>Receptor signaling</td>
</tr>
<tr>
<td>CTGLF1</td>
<td>Centaurin, γ-like family, member 1</td>
<td>0.64</td>
<td>0.41</td>
<td>0.025</td>
<td>GTPase activator activity; binding</td>
</tr>
<tr>
<td>FLJ3871</td>
<td>KH domain containing, RNA binding, signal transduction associated 1</td>
<td>0.73</td>
<td>0.53</td>
<td>0.007</td>
<td>Binding; SH3/SH2 adaptor activity</td>
</tr>
<tr>
<td>KHDRBS1</td>
<td>KH domain containing, RNA binding, signal transduction associated 1</td>
<td>0.62</td>
<td>0.39</td>
<td>0.032</td>
<td>Binding; SH3/SH2 adaptor activity</td>
</tr>
<tr>
<td>MGC4126</td>
<td>Leucine-rich repeats and calponin homology domain containing 3</td>
<td>0.97</td>
<td>0.94</td>
<td>&lt;0.0001</td>
<td>Protein binding</td>
</tr>
<tr>
<td>TPMT</td>
<td>Thiopurine S-methyltransferase</td>
<td>0.93</td>
<td>0.87</td>
<td>&lt;0.0001</td>
<td>Thiopurine S-methyltransferase activity</td>
</tr>
<tr>
<td>BIRC6</td>
<td>Baculoviral IAP repeat-containing 6 (apollon)</td>
<td>0.46</td>
<td>0.21</td>
<td>0.132</td>
<td>Ligase activity; cysteine protease inhibitor activity</td>
</tr>
<tr>
<td>C2orf4</td>
<td>Chromosome 2 open reading frame 4</td>
<td>0.67</td>
<td>0.45</td>
<td>0.0171</td>
<td></td>
</tr>
<tr>
<td>DKFZp547E087 LOC439987</td>
<td>Hypothetical gene LOC283846</td>
<td>0.67</td>
<td>0.45</td>
<td>0.0171</td>
<td></td>
</tr>
<tr>
<td>MBNL1</td>
<td>Muscleblind-like (<em>Drosophila</em>)</td>
<td>0.53</td>
<td>0.28</td>
<td>0.0763</td>
<td>Nucleic acid binding; ion binding</td>
</tr>
<tr>
<td>Sec31L1</td>
<td>SEC31 homologue A (<em>Saccharomyces cerevisiae</em>)</td>
<td>0.42</td>
<td>0.18</td>
<td>0.163</td>
<td>Protein binding</td>
</tr>
<tr>
<td>ZCCHC7</td>
<td>Zinc finger, CCHC domain containing 7</td>
<td>0.54</td>
<td>0.29</td>
<td>0.0699</td>
<td>Nucleic acid binding; ion binding</td>
</tr>
</tbody>
</table>
large homologue 1 (DLG1) and protein phosphatase 3 (formerly 2B), catalytic subunit, α isoform (PPP3CA). Using PCA, the combined expression of these two genes separated clearly responding from nonresponding patients. Only one nonresponding and two responding patients separated with the wrong group (Fig. 4A).

The induction of these two genes (DLG1 and PPP3CA) following SN38 treatment was examined in the HCT116 parental cells (Fig. 4B). The results showed that both DLG1 and PPP3CA were down-regulated following SN38 treatment over time and that this correlated well with the initial microarray analysis (DLG1 $r^2 = 0.512$ and PPP3CA $r^2 = 0.8992$). We also examined the basal expression of the two genes in the HCT116 parental cells and the SN38-resistant subline. The results showed that both genes were down-regulated in the SN38-resistant daughter line compared with parental and again this correlated well with the microarray data (Fig. 4B).

In the supervised approach, the in vitro-derived gene list was used as the starting features (genes) to build a predictive classifier for patient response to 5-FU/CPT-11 therapy. The classification rule was defined using the support vector machine algorithm, which identified a 16-gene signature (Table 2) that was assessed for overall accuracy using a leave-one-out cross-validation approach. Using this supervised approach, the 16-gene classifier correctly identified 9 of 11 responders (81.8% specificity) and 6 of 9 nonresponders (66.6% sensitivity) resulting in an overall accuracy of 75% as determined by cross-validation. Of note, the 16-gene classifier identified using this supervised method contained the two genes, DLG1 and PPP3CA, identified using the unsupervised (PCA) analysis.

**Discussion**

In advanced colorectal cancer, combination chemotherapy regimens produce disappointing response rates of only 40% to 50%, which is due to either innate or acquired drug resistance. In an effort to identify genomic markers of response to SN38, we have carried out DNA microarray profiling on matched HCT116 parental and SN38-resistant colorectal cancer cell lines.

There have been several studies that have examined the role that various biomarkers play in the cellular response to CPT-11, including ABCG2 (23–25), carbonic anhydrase (26), and TOP1 (9, 27, 28). Yu et al. used a quantitative PCR assay to profile the gene expression changes associated with 24 genes in the CPT-11 pathway. They measured the gene expression changes of these genes in 52 matched tumor and normal tissues. They found that 6 of 24 (ABCB1, ABCG2, CES1, CES2, MLH1, and UGT1A1) genes had lower expression in tumor samples and 8 of 24 genes (ABCC1, CDC45L, DRG1, ERCC1, ERCC2, FXDR, TPDP1, and TP53) had higher expression in tumor samples. Vallbohmer et al. also used quantitative PCR to investigate whether mRNA levels of drug targets (TS and TOP1) enzymes involved in 5-FU metabolism (DPD), angiogenesis (vascular endothelial growth factor, interleukin-8, and epidermal growth factor receptor), and DNA repair/drug detoxification (ERCC1 and GST-P1) were associated with clinical outcome in 33 patients with metastatic colorectal cancer (8). They showed that high intratumor levels of epidermal growth factor receptor and ERCC1 were significantly associated with disease-free survival and furthermore that epidermal growth factor receptor and ERCC1 were primarily responsible for separating responders from nonresponders to CPT-11/5-FU-based treatment (8).

Previously, there have only been a few studies that investigated gene expression following CPT-11 treatment using a global profiling approach (29–31). Minderman et al. studied the changes in gene expression following CPT-11 treatment in pretreatment and post-treatment blood samples from patients with acute myeloid leukemia and chronic myeloid leukemia and also examined the changes in gene expression following SN38 treatment in HL60 cells. The results showed a down-regulation of cell cycle genes that is consistent with the loss of S phase and temporary delay in G1-S-phase transition. An important facet of this study was that it was the first study to show that the gene expression alterations observed following in vitro exposure to SN38 also occur in vivo following CPT-11 treatment (29). Chun et al. also carried out gene expression profiling on pretreatment and post-treatment effusion samples from 8 gastric cancer patients treated with CPT-11 and compared the expression profiles between responders and nonresponders. Five isoforms of the metallothionein family were identified to have significantly higher levels in the 5 responders compared with the 3 nonresponders (31). Many of the studies that are designed to identify clinically relevant genes involved in CPT-11 resistance have been undermined by either small sample sizes or inadequate validation of microarray results.

Our approach was to identify those in vitro genes that were acutely altered following SN38 treatment in the parental cells...
and also constitutively altered in the SN38-resistant cells. Using this novel approach, we identified 124 genes that were both induced following drug treatment in the parental cells and also constitutively altered in the resistant setting. We hypothesize that these 124 in vitro genes may contain targets for therapeutic intervention and that further manipulation of this gene signature may lead to the identification of advanced colorectal cancer patients who may respond to CPT-11-based treatment.

We have previously used this approach to identify novel determinants of response to either 5-FU or oxaliplatin in this model system and identified 116 and 37 genes that correlated with response to either 5-FU or oxaliplatin, respectively (10). Further functional analysis of three candidate genes, spermidine/spermine N1-acetyltransferase, prostate-derived factor, and calretinin, showed that they confer sensitivity or resistance by directly modulating drug response and therefore, importantly, do not represent secondary transcriptional changes that may be downstream and potentially less relevant to the resistance phenotype (10, 32). In confirmation of the microarray experiment, we found a strong correlation between the original microarray and subsequent quantitative RT-PCR results, which is in accordance with previously published data (10, 33, 34).

The most important facet of the present study was to use an in vitro-based approach to identify markers of resistance to SN38 and to further assess whether any of these in vitro-derived genes could predict for response to CPT-11 therapy in patients. If this in vitro approach proves successful, it will reduce the number of patient biopsies that are currently needed to build predictive signatures of response to therapy. The limited number of available biopsies can then be best used as independent validation tools, with the result that the number of patient biopsies required for these studies will be significantly reduced. In our study, using the in vitro gene lists, PCA showed a clear separation between responding and nonresponding patients using the combined expression of two genes, DLG1 and PPP3CA.

Using the in vitro-derived gene lists, we have constructed an in vitro-derived predictive signature of response to therapy using the support vector machine algorithm. This 16 in vitro gene signature contains 14 genes that are down-regulated and 2 genes that are up-regulated in patients that respond to 5-FU/CPT-11 therapy. On internal cross-validation, the signature displayed an overall accuracy of 75%. The genes contained within the list were involved in cellular processes such as transcription, transport, metabolism, cell proliferation, and Ca²⁺ ion binding. The 16-gene signature also contained PPP3CA and DLG1, which were shown previously to separate responding and nonresponding patients using PCA. PPP3CA

Fig. 4. A, PCA plot in two dimensions showing patient separation using the 1.5-fold changing (between the responding and nonresponding patients) in vitro gene list (2 genes; DLG1 and PPP3CA). Gray circles, responding patients; black circles, nonresponding patients. B, quantitative RT-PCR validation analysis of the two genes identified from the PCA analysis, DLG1 and PPP3CA. Fold induction compared with untreated control cells in the parental setting following SN38 treatment and also the fold change in untreated SN38-resistant cells compared with untreated HCT116 parental cells. The correlation coefficients were determined by comparing the average DNA microarray fold changes in the parental cells with average quantitative RT-PCR fold changes in the parental cells.
is a Ca²⁺/calmodulin-dependent protein phosphatase, which is important for Ca²⁺-mediated signal transduction (35). Dlg1 is a mammalian homologue of the Drosophila discs large tumor suppressor protein (36). Dlg1 along with Scrib and lethal giant larvae forms the Scribble complex, which regulates epithelial polarity (37). In this type of "predictive analysis," the aim of the study is to identify genes that have a high discriminatory power to separate the two patient subgroups; therefore, single genes contained within the signature may not be either functionally related to the development of drug resistance or represent a viable "druggable" target.

Recently, Del Rio et al. generated a 14-gene signature that was 95% accurate in identifying FOLFIRI responding and non-responding patients (38). When comparing our 16 gene signature with this 14-gene signature, there was no overlap in composition. Several reasons could account for this. Firstly, their study used the primary tumor to predict for response in the metastatic disease and we were analyzing the metastatic disease site. Secondly, our aim was to use in vitro models to identify genes that could predict response to chemotherapy in patients, whereas Del Rio et al. generated a gene list directly from the patient samples. However, it is widely accepted that many different gene lists can predict response in patients at a similar level (18).

In conclusion, we have identified 124 genes through in vitro transcriptional profiling studies that correlate with SN38 response/resistance and have shown strong correlations between microarray data and quantitative RT-PCR validation data. We have also shown that the regulation of these genes is not cell line dependent, as we have positively validated several of these genes in a panel of 5 additional colorectal cancer cell lines. We have assessed whether genes identified using an in vitro-based approach are able to predict treatment response in patient samples. The results show that several of the in vitro genes were differentially regulated between these two patient populations; in addition, unsupervised (PCA) and supervised (support vector machine) classification analysis revealed that several of these in vitro genes were able to separate responding and nonresponding patients based on their expression profiles. In our novel in vitro-based approach, we have shown that we are able to gain meaningful data using in vitro model systems. Although this study has shown a good separation of patient samples, it contained only 20 patients; therefore, the current analysis will require further independent validation in a much larger patient cohort. There have only been a few reported studies in the literature that have tried to identify predictive signatures of response to chemotherapy in colorectal cancer. The recent study by Del Rio et al. identified a 14-gene signature that was predictive for 5-FU/CPT-11 therapy from only 21 primary colorectal cancer patients (38). Ghadimi et al. used oligonucleotide and cDNA-based microarray profiling of 30 rectal adenocarcinoma biopsies to identify a 54-gene signature predictive of response to preoperative chemoradiotherapy based on a training set of 23 samples and a test set of 7 samples (39). In addition, Graudens et al. identified 679 genes that predicted sensitivity to 5-FU/CPT-11 therapy from only 13 patients, which included the following samples: tumor (n = 8), metastatic (n = 12), and adjacent normal (n = 6). They further validated a subset of these genes in two independent patients and showed a good concordance with the original experiment (40). These two studies, although hampered by their low sample numbers, have shown that it is possible to generate signatures of response to chemotherapy in the advanced

---

Table 2. Sixteen in vitro-derived predictive signature of response to 5-FU/CPT-11 therapy

<table>
<thead>
<tr>
<th>Common name</th>
<th>Description</th>
<th>Fold change in responders</th>
<th>Predictive value</th>
<th>Gene Ontology</th>
</tr>
</thead>
<tbody>
<tr>
<td>SP1</td>
<td>Homo sapiens Sp1 transcription factor mRNA</td>
<td>-1.55</td>
<td>7.68</td>
<td>Transcription</td>
</tr>
<tr>
<td>KIAA1267</td>
<td>KIAA1267</td>
<td>2.04</td>
<td>7.43</td>
<td>Nucleus</td>
</tr>
<tr>
<td>TPMT</td>
<td>H. sapiens thiopurine S-methyltransferase mRNA</td>
<td>-1.68</td>
<td>5.9</td>
<td>Nucleobase, nucleoside, nucleotide and nucleic acid metabolic process</td>
</tr>
<tr>
<td>ARF1</td>
<td>H. sapiens ADP-ribosylation factor 1 transcript variant</td>
<td>-1.55</td>
<td>5.9</td>
<td>Endoplasmic reticulum to Golgi vesicle-mediated transport</td>
</tr>
<tr>
<td>SP1</td>
<td>H. sapiens Sp1 transcription factor mRNA</td>
<td>-1.33</td>
<td>5.74</td>
<td>Transcription</td>
</tr>
<tr>
<td>NCOR1</td>
<td>H. sapiens nuclear receptor corepressor 1 mRNA</td>
<td>-1.44</td>
<td>5.74</td>
<td>Transcription from RNA polymerase II promoter</td>
</tr>
<tr>
<td>CLTC</td>
<td>H. sapiens clathrin heavy chain mRNA</td>
<td>-1.25</td>
<td>5.46</td>
<td>Intracellular protein transport</td>
</tr>
<tr>
<td>ARF1</td>
<td>H. sapiens ADP-ribosylation factor 1 transcript variant</td>
<td>-1.39</td>
<td>4.81</td>
<td>Endoplasmic reticulum to Golgi vesicle-mediated transport</td>
</tr>
<tr>
<td>CTBP2</td>
<td>COOH-terminal binding protein 2</td>
<td>-1.61</td>
<td>4.81</td>
<td>Ï€-Serine biosynthetic process</td>
</tr>
<tr>
<td>MCC2</td>
<td>H. sapiens methylcoenzyme A carboxylase 2 (m)</td>
<td>-1.35</td>
<td>4.81</td>
<td>Leucine catabolic process</td>
</tr>
<tr>
<td>DLG1</td>
<td>H. sapiens mRNA for presynaptic protein SAP97 variant protein</td>
<td>-1.43</td>
<td>4.81</td>
<td>Endothelial cell proliferation</td>
</tr>
<tr>
<td>SNRPA1</td>
<td>Small nuclear ribonucleoprotein polypeptide A</td>
<td>-1.29</td>
<td>4.69</td>
<td>Nuclear mRNA splicing, via spliceosome</td>
</tr>
<tr>
<td>STK4</td>
<td>Serine/threonine kinase 4</td>
<td>1.84</td>
<td>4.69</td>
<td>Cell morphogenesis</td>
</tr>
<tr>
<td>REPS1</td>
<td>H. sapiens RALBP1-associated Eps domain containing 1</td>
<td>1.83</td>
<td>4.43</td>
<td>Calcium ion binding</td>
</tr>
<tr>
<td>PPP3CA</td>
<td>H. sapiens protein phosphatase 3 (formerly 2B) catalytic</td>
<td>-1.42</td>
<td>4.43</td>
<td>G1-S transition of mitotic cell cycle</td>
</tr>
<tr>
<td>MALAT1</td>
<td>H. sapiens metastasis-associated lung adenocarcinoma transcript 1</td>
<td>-1.37</td>
<td>4.43</td>
<td></td>
</tr>
</tbody>
</table>
References


Clinical Determinants of Response to Irinotecan-Based Therapy Derived from Cell Line Models

Wendy L. Allen, Vicky M. Coyle, Puthen V. Jithesh, et al.


Updated version
Access the most recent version of this article at:
http://clincancerres.aacrjournals.org/content/14/20/6647

Supplementary Material
Access the most recent supplemental material at:
http://clincancerres.aacrjournals.org/content/suppl/2008/10/17/14.20.6647.DC1

Cited articles
This article cites 39 articles, 24 of which you can access for free at:
http://clincancerres.aacrjournals.org/content/14/20/6647.full.html#ref-list-1

Citing articles
This article has been cited by 3 HighWire-hosted articles. Access the articles at:
/content/14/20/6647.full.html#related-urls

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.