

Identification of differentially expressed alternative mRNA isoforms associated with chemotherapy resistance in colon cancer cell lines

Malachi Griffith, Michelle J. Tang, Susanna Chan, Jennifer Asano, Adrian Ally, Trevor Pugh, Gregg B. Morin,
Isabella T. Tai and Marco A. Marra

Department of Medical Genetics, University of British Columbia, and Genome Sciences Centre, BC Cancer Research Centre; Vancouver, BC, Canada



ABSTRACT

The drug 5-fluorouracil (5-FU) is a commonly used chemotherapy for colorectal cancer (CRC). Response to this drug among CRCs is variable and is thought to be mediated in part by differences in metabolism. A number of genes such as *UMPS/OPRT*, *DPYD*, *ECGF1*, *UPP1*, *TS* and others are known to be involved in the metabolism of 5-FU. Differential expression and RNA processing of these genes may contribute to 5-FU resistance and poor clinical outcome. The **OBJECTIVE** of this study is to identify differentially expressed alternative mRNA isoforms that are predictive of 5-FU resistance.

METHODS: We developed an approach that uses a combination of Affymetrix Exon arrays and custom splicing microarrays capable of measuring the expression of mRNA isoforms with subtle differences in exon content. We applied this approach to profile RNA extracted from the MIP101, HCT116, and RKO cell lines as well as 5-FU resistant derivatives of each (MIP/5FU, HCT/5FU, and RKO/5FU) ([Tai et al. 2005](#)). Differentially expressed alternative isoforms were predicted from the microarray data and some of these findings have been validated by RT-PCR.

RESULTS: We identified 25 mRNA isoforms that were differentially expressed between MIP101 and MIP/5FU. Interestingly, we found that two isoforms of the gene *UMPS* were reciprocally differentially expressed. In both MIP101 and HCT116 the expression of the canonical isoform was reduced in resistant cells while a previously unstudied isoform lacking exon 2 showed 23.5 ± 1.0 fold over-expression in MIP/5FU ($p=0.00437$) and 6.3 ± 1.4 fold over-expression in HCT/5FU ($p=0.00193$). Global expression and pathway analysis of all three sensitive resistant cell lines pairs is ongoing.

CONCLUSIONS: Our results indicate that differential expression of specific alternative mRNA isoforms of genes involved in the metabolism of 5-FU may be predictive of drug resistance in CRC.

INTRODUCTION

Previous studies have examined the expression of genes with suspected relevance to 5-FU action (Kai et al. 2007; Kidd et al. 2005; Kodera et al. 2007; Matsusaka et al. 2007; and others described below), but have not attempted a genome-wide approach to broadly interrogate predicted members of the 5-FU pathway.

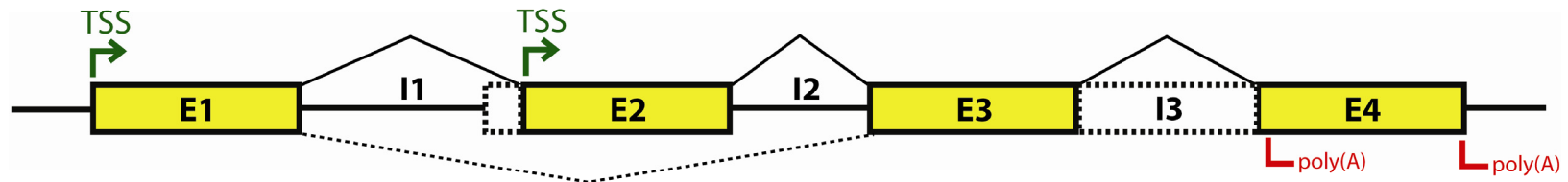
One of the most studied genes related to 5-FU action is *DPYD*, for which a clinical test for 5-FU toxicity now exists (Myriad Genetics, Inc). Although *DPYD* is believed to account for a large percentage (as much as 80%) of the catabolism of 5-FU in the liver, additional genes are required for anti-cancer activity since any drug that does reach tumour cells depends on other enzymes for conversion to its active form (see **Figure 1**) (Maring et al. 2005; Tokunaga et al. 2007). Among these enzymes, *UMPS* (aka *OPRT*) has recently been cited as one of the most critical (Sakamoto et al. 2007), but there is disagreement in the literature as to the relative importance of genes predicted to be involved in 5-FU action.

The use of genomic methods to identify gene variants associated with human diseases such as cancer is an area of rapid development (Griffith and Marra 2007). Techniques such as Expressed Sequence Tag sequencing, full-length cDNA library sequencing (Baross et al. 2004), 454 sequencing (Bainbridge et al. 2006), Serial Analysis of Gene Expression (SAGE), Affymetrix Gene Chip expression analysis and others have proven useful in identifying candidate genes with clinical relevance. However, despite their utility, these methods remain insensitive to some types of events (such as alternative mRNA isoforms) or are limited in their ability to profile large numbers of genes simultaneously. To address these limitations we developed a novel genomic method and applied it to the study of 5-FU resistance in colorectal cancer (CRC). This method uses splicing microarrays to profile the expression of mRNA isoforms (see **Figures 2-3**) (Griffith et al. 2008). **Table 1** and **Figures 4-6** describe the results of microarray analysis and follow-up validations.

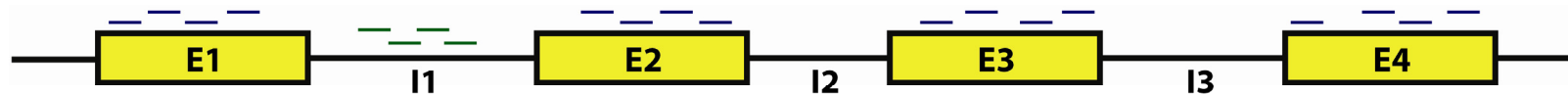


Figure 1. Genes involved in 5-FU metabolism. 5-fluorouracil is metabolized by at least three mechanisms (labeled I, II & III). Catabolism in the liver is thought to primarily occur in the liver via *DPYD* ([Sakamoto et al. 2007](#)). Remaining 5-FU is thought to be activated to FdUMP by several enzymes and lead to the inhibition of *TYMS* (causing inhibition of DNA synthesis via depletion of thymidine in the cell). Inhibition of cell growth also occurs by RNA and DNA damage upon incorporation of fluorine modified substrates.

(A) Alternative expression



(B) Affymetrix array design



(C) ALEXA array design

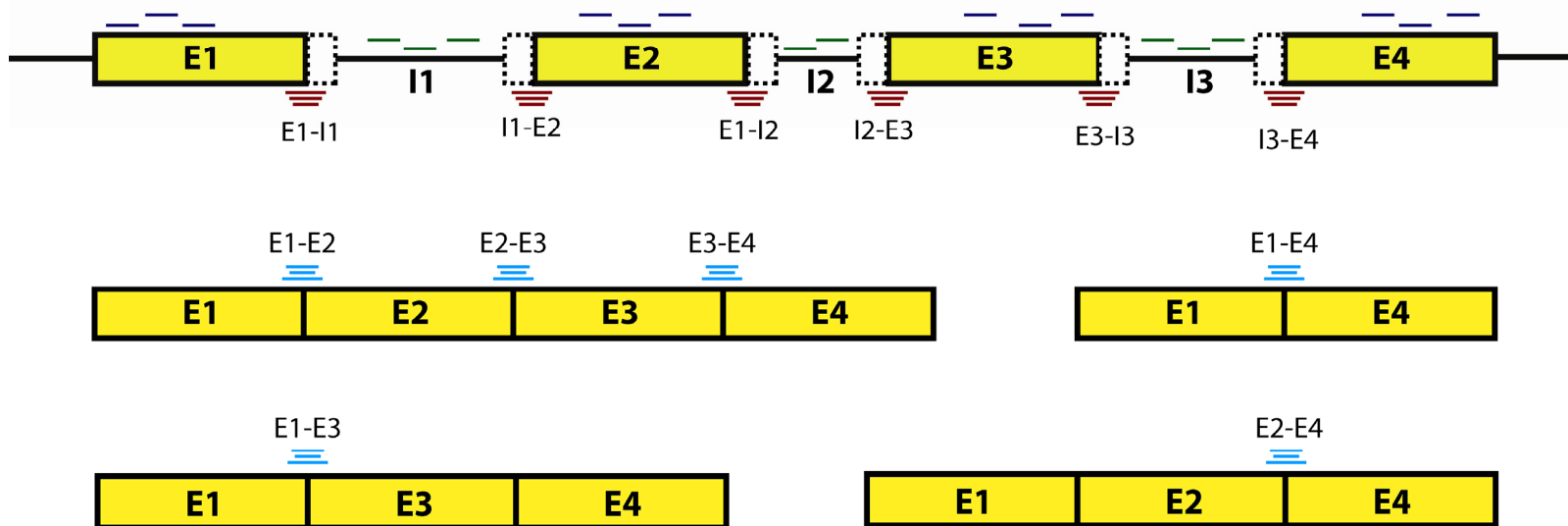
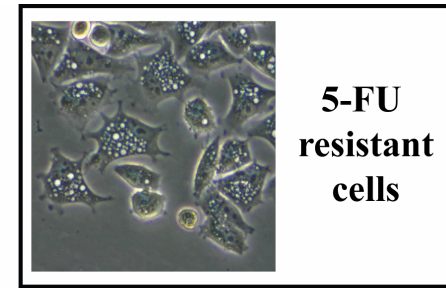


Figure 2. Array design strategy. (A) Alternative expression (alternative transcript initiation, splicing, and polyadenylation). A hypothetical gene locus with four annotated exons (colored rectangles; E1-E4) and three introns (connecting lines; I1-I3) is depicted. Green arrows indicate alternate transcript start sites (TSS). Alternate polyadenylation (polyA) sites are shown in red. Alternative exon boundary usage, exon skipping and intron retention are depicted with black dotted lines. (B) Affymetrix array design. Affymetrix exon arrays use multiple sources of gene annotation and prediction in an attempt to measure expression of every known or predicted expressed region of the genome. The resulting design consists of sets of 4 oligonucleotide probes per exon representing most known and predicted exons. (C) ALEXA array design. The ALEXA approach attempts to profile exon skipping, alternative exon boundary usage, and intron retentions by selecting probes to represent every exon, intron, exon-exon junction and exon-intron boundary. The positions of exon junctions are depicted over the hypothetical processed mRNAs they represent.

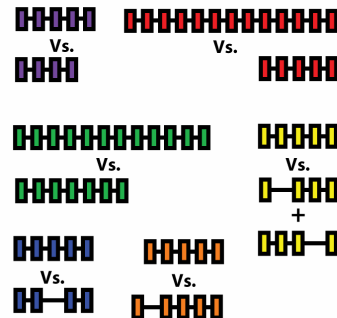


**Isolate total RNA
(triplicates)**

**Affymetrix exon
array analysis**

Profile ~25,000
known genes

Select genes with DE Exons



Purify mRNA

Synthesize cDNA

Custom array analysis

Gene selection

- ~400 known drug resistance genes
- ~2k genes with DE exons

Probe design

- 35k exons; 90k exon junctions;
- 50k exon boundaries, 2k introns

Figure 3. Experimental overview. RNA samples extracted from 5-FU sensitive and resistance colorectal cancer cell lines were analyzed on both Affymetrix Human Exon 1.0 ST arrays and our own ALEXA custom arrays synthesized by NimbleGen Inc. Samples were first analyzed on the Affymetrix platform. A subset of genes were then selected and targeted by a custom ALEXA microarray ([Griffith et al. 2008](#)). The genes on this array included genes known to be involved in 5-FU metabolism as well as those identified as having at least one differentially expressed exon in the Affymetrix analysis.

OBJECTIVE

To identify differentially expressed alternative mRNA isoforms that are predictive of 5-FU resistance in colorectal cancer

MATERIALS AND METHODS

Cell lines

All cell lines (MIP101, HCT116 and RKO; sensitive and resistant) were maintained in DMEM media supplemented with 1% penicillin-streptomycin, 1% kanamycin (Invitrogen) and 10% newborn calf serum at 37°C and 5% CO₂ ([Tai et al. 2005](#)).

RNA Isolation

Total RNA was isolated from cells cultured to ~75% confluence using RNeasy Columns (Qiagen). RNA was DNaseI treated using an RNase free DNaseI kit from Invitrogen. RNA was quantified and tested for degradation using an Agilent 2100 Bioanalyzer. PolyA⁺ RNA was purified from total RNA using an oligoTex kit (Qiagen).

Creation of custom ALEXA arrays

Target genes for the custom ALEXA array design were selected by identifying all genes with 2-fold or greater differential expression of one or more of their exons according to 'Affymetrix GeneChip Human Exon 1.0 ST' arrays. An additional ~400 genes were selected for their potential relevance to drug resistance according to literature reports or their gene ontology terms. Oligonucleotide probes for the custom array were selected and filtered as described in [Griffith et al. \(2008\)](#). The final custom array design consisted of 385,000 probes of 26-46 bp in length corresponding to 2,511 genes. Each exon, exon boundary, exon junction and intron of these genes was represented by 2-4 probes.

Microarray hybridization

Total RNA was used for hybridizations with the Affymetrix exon array platform and processed according to Affymetrix's 'GeneChip Whole Transcript Sense Target Labeling Assay Manual' (Affymetrix Inc.). Custom ALEXA arrays were manufactured by NimbleGen systems Inc. to our specifications. To prepare samples for hybridizations to the custom ALEXA arrays, polyA⁺ RNA was used for double stranded cDNA synthesis with a 'Superscript Choice System' for cDNA Synthesis using random hexamers (Invitrogen). 5 µg of each cDNA sample was shipped to NimbleGen. Labeling, hybridization and scanning was conducted by NimbleGen using their 'ChIP-chip' protocol.

Microarray data analysis

Analysis and visualization of microarray data was conducted using the Ensembl API, Affymetrix's ExACT software, Bioconductor, R, Perl, and the UCSC browser as described in [Griffith et al. \(2008\)](#).

PCR and RT-PCR validation of UMPS isoform expression

Single stranded cDNA was generated from polyA⁺ RNA isolated from each cell line using SuperScript III reverse transcriptase and random hexamer primer (Invitrogen). PCR primers were designed to flank exon 2 (Forward: 5'-TGGGGCCATTGGTGACGGGT-3' and Reverse: 5'-GGGAGCCGGTGGAGCTCATT-3'). PCR was performed with Invitrogen's Platinum Pfx enzyme.

Cloning & sequence validation of UMPS mRNA isoforms

Clones representing the full *UMPS* open reading frame and most of the UTR were generated by TOPO TA cloning (Invitrogen) using primers designed against the *UMPS* reference sequence, NM_000373 (Forward: 5' -CAAACAGGCAGCGCGGACA-3' and Reverse: 5' -GCTGACTTTAGCCTCTTGGTGCCC-3'). PCR was performed with Invitrogen's Platinum Taq, High Fidelity enzyme. Clones were screened for correct insert size and forward orientation relative to the M13F site of the cloning vector by restriction enzyme digestion with EcoRI and NotI/XhoI (double digest) respectively. Clones were end-sequenced by Sanger sequencing with an ABI 3730 device using M13F and M13R primers.

PCR & sequence validation of UMPS genomic DNA

To investigate the mechanism of *UMPS* alternative isoform expression events involving *UMPS* exon 2, the genomic region of the *UMPS* locus was sequenced by generating 22 amplicons overlapping the region from 1kb upstream of *UMPS* exon 1 to the end of exon 3. Each primer contained either an M13F or M13R linker which were used for direct sequencing of PCR products. Genomic DNA was isolated from cells grown to ~75% confluence using a Gentra PureGene kit (Gentra Systems). PCR was performed with Platinum Taq, High Fidelity (Invitrogen) enzyme and each amplicon was column or gel purified and Sanger sequenced with an ABI 3730 device using M13F and M13R primers. Sequence analysis for identification of mutations was conducted with Mutation Surveyor (SoftGenetics).

RESULTS

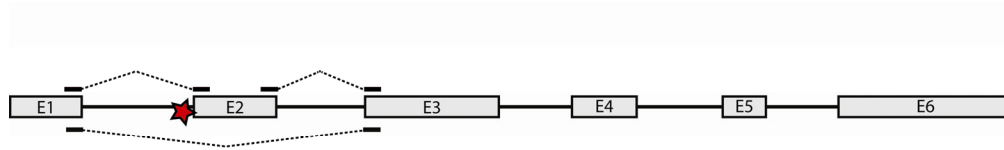
Gene Symbol	Gene name	Event type	Fold change (Isoform A)	Fold change (Isoform B)	# and type of sequences supporting alternative expression event
<i>AKAP7</i>	A kinase (PRKA) anchor protein 7	Alternative TSS/polyA	-3.07	1.01	Multiple mRNAs
<i>APLP1</i>	Amyloid beta (A4) precursor-like protein 1	Complex	N/A	N/A	mRNA and EST evidence indicates several AS events similar to those observed
<i>ATP6AP1</i>	ATPase, H ⁺ transporting, lysosomal accessory protein 1	Alternative exon boundary	2.85	-1.67	Multiple ESTs
<i>C12orf63</i>	Chromosome 12 open reading frame 63	Alternative TSS/polyA	-21.17	1.23	None
<i>C5</i>	Complement component 5	Alternative TSS/polyA	13.22	-1.75	None
<i>CDC25B</i>	Cell division cycle 25 homolog B	Alternative TSS/polyA	6.31	1.00	Multiple mRNAs and ESTs support the use of alternative 5' exons
<i>COL21A1</i>	Collagen, type XXI, alpha 1	Intron retention	2.72	-3.10	Single EST
<i>DIS3</i>	DIS3 mitotic control homolog (S. cerevisiae)	Exon skipping	-3.44	1.02	Single mRNA and ~50 ESTs
<i>EIF4A2</i>	Eukaryotic translation initiation factor 4A, isoform 2	Alternative exon boundary	3.07	-1.20	Multiple ESTs.
<i>ENO2</i>	Enolase 2 (gamma, neuronal)	Intron retention	2.22	-3.44	Single EST (cloned from Cerebellum)
<i>EPB41L3</i>	Erythrocyte membrane protein band 4.1-like 3	Alternative TSS/polyA	-9.27	1.08	None
<i>FGD5</i>	FYVE, RhoGEF and PH domain containing 5	Alternative TSS/polyA	3.33	1.08	Single mRNA
<i>HHIP</i>	Hedgehog interacting protein	Alternative TSS/polyA	8.08	1.10	Single mRNA
<i>IL10RB</i>	Interleukin 10 receptor, beta	Exon skipping	4.69	-2.55	None
<i>KLK6</i>	Kallikrein-related peptidase 6	Complex	N/A	N/A	Multiple mRNAs which could contribute to observed expression
<i>LAMA3</i>	Laminin, alpha 3	Alternative TSS/polyA	-3.83	1.28	Two mRNAs
<i>MLPH</i>	Melanophilin	Exon skipping	2.02	-4.59	Multiple mRNAs and ESTs indicate skipping of this exon
<i>MYT1</i>	Myelin transcription factor 1	Complex	N/A	N/A	Multiple mRNAs which could contribute to observed expression
<i>PLCB4</i>	Phospholipase C, beta 4	Exon skipping	-5.03	-1.20	Two mRNAs and two ESTs
<i>PPP2R1B</i>	Protein phosphatase 2, regulatory subunit A, beta isoform	Exon skipping	-3.40	1.14	Single mRNA and ~10 ESTs
<i>RC74</i>	Integrator complex subunit 9	Exon skipping	-12.66	1.06	Single EST (cloned from hepatocellular carcinoma cell line).
<i>RCC1</i>	Regulator of chromatin condensation 1	Alternative TSS/polyA	-3.72	-1.10	None (some mRNA support for similar alternate TSS usage)
<i>SSBP2</i>	Single-stranded DNA binding protein 2	Alternative exon boundary	3.11	-3.21	None
<i>TPST1</i>	Tyrosylprotein sulfotransferase 1	Exon skipping	5.25	1.10	Multiple ESTs
UMPS *	Uridine monophosphate synthetase	Exon skipping	-5.77	5.21	Multiple mRNAs and ESTs in human, mouse and rat

* Additional examination of this event is described in the following results

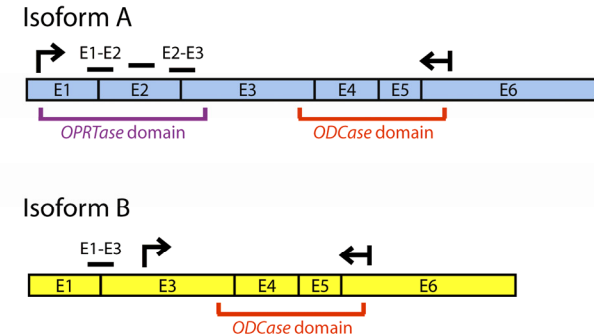
Table 1. Candidate differential isoform expression events associated with 5-FU resistance.

The top 25 candidate differential isoform expression events associated with 5-FU resistance are reported (listed alphabetically). These events were selected by applying a series of filters to identify events that were likely to involve differential expression (reciprocal expression in most cases) of specific isoforms as opposed to the entire gene. Fold change 'A' and 'B' refer to values for probes capable of distinguishing expression of putative alternative isoforms ([Griffith et al. 2008](#)). Positive values indicate over-expression in 5-FU sensitive cells (MIP101). Negative values indicate over-expression in 5-FU resistant cells (MIP/5FU). TSS refers to 'transcription start site'. Fold-change values for 'complex' events are listed as 'N/A' because the number of potential isoforms is too large to assign particular probes to particular isoforms.

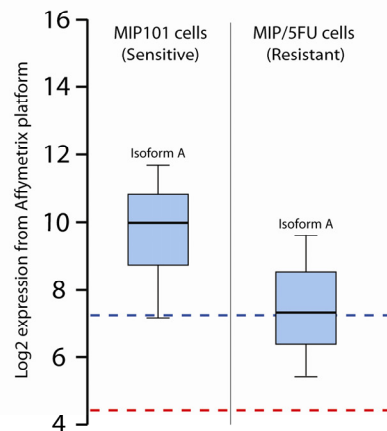
(A) *UMPS* gene model showing alternative splicing of exon 2



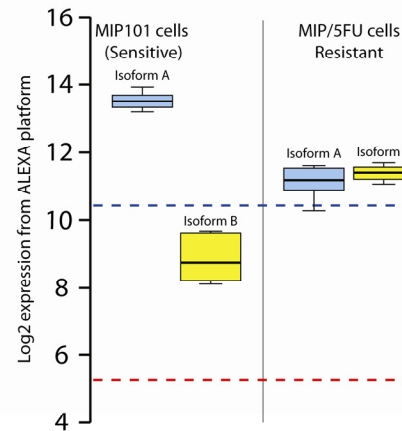
(B) *UMPS* isoforms, protein domains, probesets



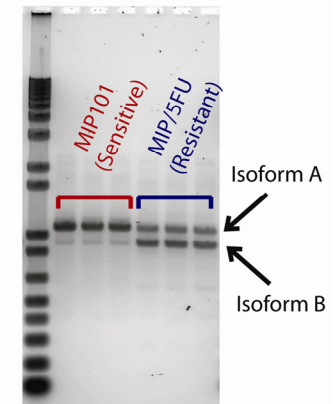
(C) Affymetrix microarray data



(D) ALEXA microarray data



(E) RT-PCR data



(F) Genomic sequence data

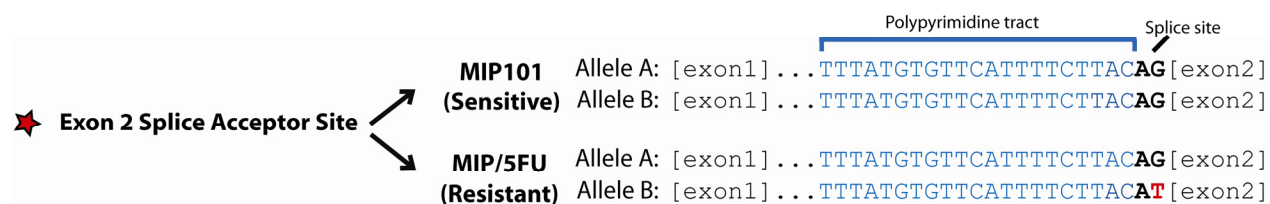
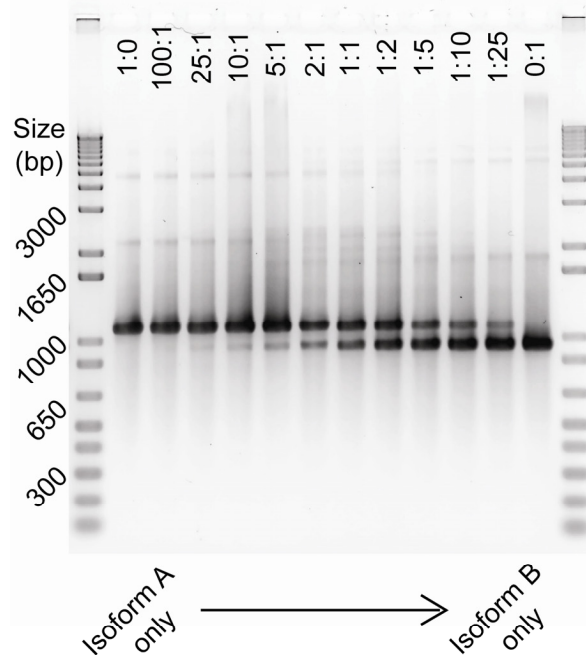


Figure 4. Differential expression of alternative *UMPS* isoforms in 5-FU sensitive and resistant cells. (A) The uridine monophosphate synthetase (*UMPS*) locus (chromosome 3q13). Alternative splicing of exon 2 is depicted by dotted lines. (B) The positions of ALEXA splicing microarray probesets (each consisting of 2-4 oligonucleotide probes) specific to *UMPS* isoforms A and B are depicted. Probes are labeled according to the exons or junctions they profile (e.g., E1-E3 detects the connection of exon 1 to exon 3). Black arrows indicate the predicted Open Reading Frame (ORF) of each isoform and the position of protein domains is indicated beneath each isoform. (C) Box plots depict expression values for oligonucleotide probes from triplicate samples profiled on the Affymetrix exon microarray platform. The results indicated a loss of expression for exon 2 in 5-FU resistant cells (MIP/5FU) relative to sensitive cells (MIP101). The median expression value of all exons (blue dotted line) and all negative controls (red dotted line) on the Affymetrix exon microarray platform are also shown. (D) Expression values for oligonucleotide probes from the ALEXA splicing microarray platform specific to each isoform are shown as box plots. The median log2 expression value of all exons (blue dotted line) and all negative controls (red dotted line) on the ALEXA microarray are also shown. Isoform A was ~5-fold over-expressed in 5-FU sensitive cells relative to resistant cells. Isoform B was ~6-fold over-expressed in 5-FU resistant cells relative to sensitive cells. (E) RT-PCR results verified the predicted shift in isoform expression values between 5-FU sensitive and resistant cells. (F) Sequencing of the genomic region of *UMPS* revealed a heterozygous splice site mutation at the splice acceptor site of exon 2 which was present in resistant (MIP/5FU) cells but not in sensitive (MIP101) cells. This splice site mutation, acquired in the resistant cells is predicted to prevent pre-mRNA splicing of exon 2 and favor production of isoform B from the mutated allele.

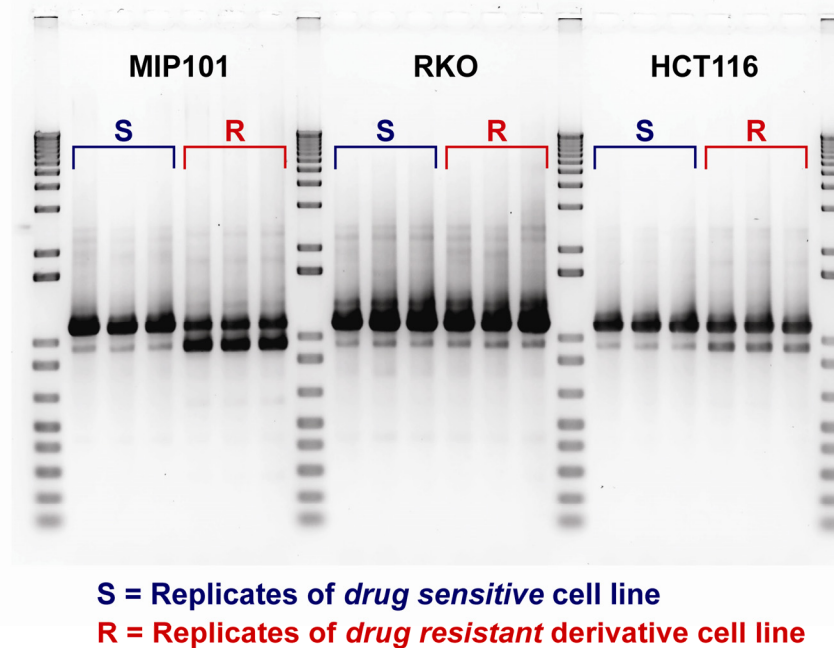
(A)

RT-PCR using molar ratios of
plasmid DNA clones
representing *UMPS* Isoforms A & B



(B)

RT-PCR using cDNA generated from polyA⁺ RNA
isolated from 5-FU sensitive and resistant cell lines

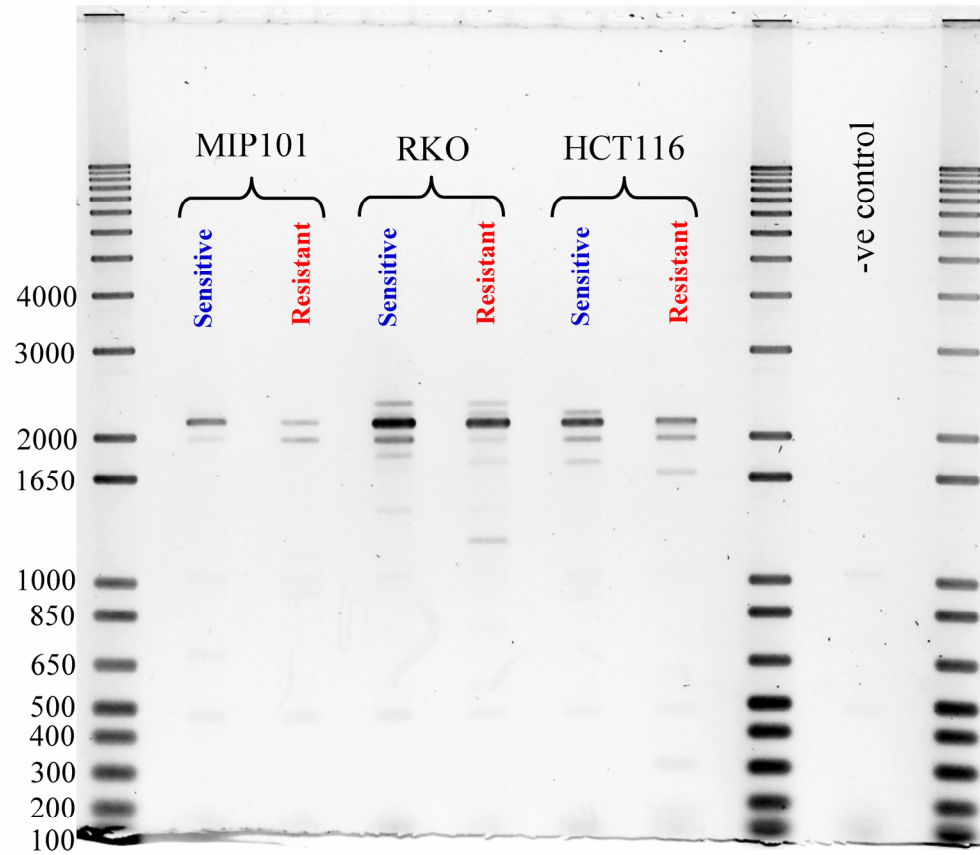


(C)

Cell Line Comparison	Isoform A fold change	Isoform B fold change
MIP101 (sensitive) versus MIP/5FU (resistant)	$-4.92 \pm 1.40 \uparrow$	$23.45 \pm 1.05 \uparrow$
RKO (sensitive) versus RKO/5FU (resistant)	1.14 ± 1.23	1.34 ± 1.29
HCT116 (sensitive) versus HCT/5FU (resistant)	-1.24 ± 1.20	$6.27 \pm 1.40 \uparrow$

Figure 5. RT-PCR of *UMPS* isoforms in additional cell lines. (A) Test of RT-PCR primers using mixtures of plasmid DNA clones representing *UMPS* isoforms A & B. These plasmids were mixed in the indicated molar ratios and used as templates for PCR reactions. The expected product sizes for this primer pair are 1107 bp for Isoform A and 953 bp for Isoform B (missing exon 2). (B) The same PCR primers were used to amplify *UMPS* isoforms from ss-cDNA generated from polyA⁺ RNA extracted from 5-FU sensitive and resistant CRC cell lines. Both MIP101 and HCT116 showed an increase in the presence of *UMPS* isoform B in the 5-FU resistant derivative compared to the sensitive line. (C) RT-PCR products depicted in the previous figure were quantified with an Agilent 2100 bioanalyzer. † indicates a significant fold change difference between sensitive and resistant cells ($p < 0.05$ by two-tailed Student's t-test).

(A)



(B)

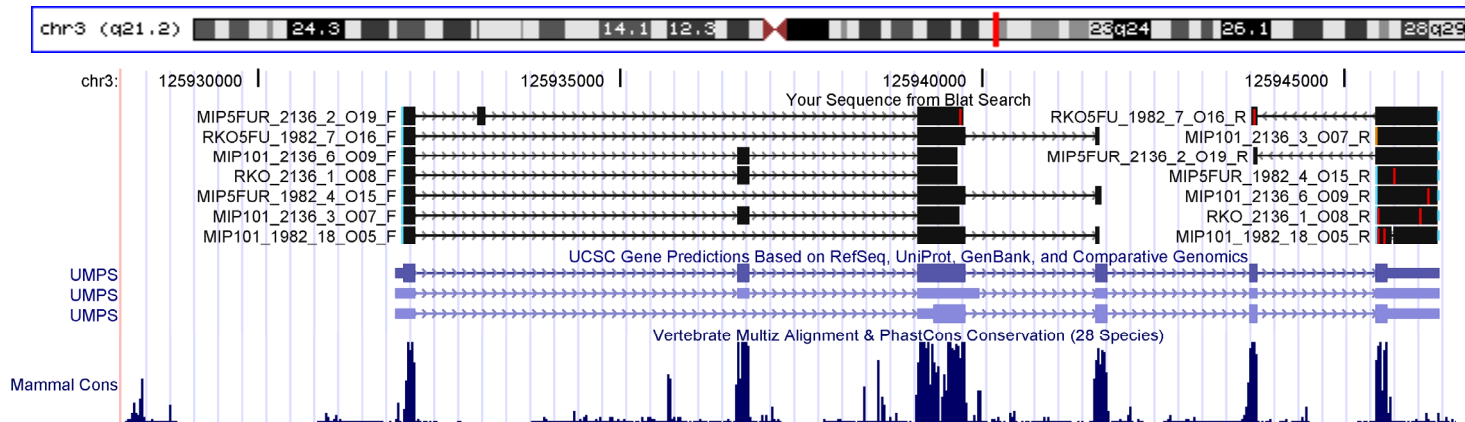


Figure 6. Cloning and sequencing of selected *UMPS* isoforms from 5-FU sensitive and resistant cell lines. (A) RT-PCR using primers located near the 5' and 3' ends of the *UMPS* mRNA sequence resulted in several amplicons for each cell line. The amplicons represent the known *UMPS* isoforms A and B and possibly additional novel isoforms. 18 amplicons from 6 source cell lines were resolved by gel electrophoresis, individually extracted and used in cloning reactions. (B) ~300 *UMPS* clones have been generated to date from the amplicons depicted above. End-sequencing of a preliminary selection of 24 of these clones confirmed the presence of at least three distinct alternative isoforms of *UMPS*.

CONCLUSIONS

Exon tiling and splicing microarray analysis of a panel of three colorectal cancer cell lines and their 5-FU resistant derivatives revealed candidate differentially expressed genes and mRNA isoforms associated with acquired 5-FU resistance. Additional PCR, RT-PCR, cloning, and sequencing experiments suggest that the pyrimidine metabolism gene, *UMPS*, which has known relevance to 5-FU metabolism may be consistently perturbed in acquired resistance either by changes in the expression of alternative isoforms or mutations that affect the protein sequence. Previous studies that have attempted to correlate the expression of this gene with 5-FU resistance in cancer have relied on methods that are insensitive to expression of individual transcript variants (typically by design of probes corresponding to the 3' end of the gene only). Given the apparent diversity of transcripts expressed from the *UMPS* locus it is possible that these studies may have missed important insights into the role of this enzyme in mediating response to 5-FU. This kind of diversity is also likely affect other genes known to be involved in 5-FU action. Our goal is to use broad genomic approaches capable of profiling the transcript diversity of all relevant loci to identify those variants with clinical value in predicting resistance to 5-FU in colorectal cancer.

ACKNOWLEDGEMENTS

A large empty rectangular box with a blue border, intended for acknowledgements.

We are grateful for funding support from the Canadian Institutes of Health Research (CIHR), the Canadian Digestive Health Foundation (CDHF), the Michael Smith Foundation for Health Research (MSFHR), the National Cancer Institute of Canada (NCIC), the Natural Sciences and Engineering Research Council (NSERC) and the British Columbia Cancer Foundation. MG is supported by scholarships from NSERC, Genome British Columbia and MSFHR. MJT is supported by the CDHF/CIHR Doctoral Research Award. TP is supported by MSFHR. ITT and MAM are MSFHR Scholars.

REFERENCES

- Bainbridge, M. N., R. L. Warren, et al. (2006). "Analysis of the prostate cancer cell line LNCaP transcriptome using a sequencing-by-synthesis approach." Submitted to BMC Genomics.
- Baross, A., Y. S. Butterfield, et al. (2004). "Systematic recovery and analysis of full-ORF human cDNA clones." Genome Res **14**(10B): 2083-92.
- Griffith, M. and M. A. Marra (2007). Alternative expression analysis: Experimental and bioinformatic approaches for the analysis of transcript diversity. Genes, Genomes & Genomics. D. Thangadurai, W. Tang and T. Pullaiah. New Delhi, Regency Publications. **2**: 201-242.
- Griffith, M., M. J. Tang, et al. (2008). "ALEXA: a microarray design platform for alternative expression analysis." Nat Methods **5**(2): 118.
- Kai, K., Y. Kitajima, et al. (2007). "Quantitative double-fluorescence immunohistochemistry (qDFIHC), a novel technology to assess protein expression: a pilot study analyzing 5-FU sensitive markers thymidylate synthase, dihydropyrimidine dehydrogenase and orotate phosphoribosyl transferases in gastric cancer tissue specimens." Cancer Lett **258**(1): 45-54.
- Kidd, E. A., J. Yu, et al. (2005). "Variance in the expression of 5-Fluorouracil pathway genes in colorectal cancer." Clin Cancer Res **11**(7): 2612-9.
- Kodera, Y., S. Ito, et al. (2007). "Gene expression of 5-fluorouracil metabolic enzymes in primary gastric cancer: Correlation with drug sensitivity against 5-fluorouracil." Cancer Lett **252**(2): 307-13.
- Maring, J. G., H. J. Groen, et al. (2005). "Genetic factors influencing pyrimidine-antagonist chemotherapy." Pharmacogenomics J **5**(4): 226-43.
- Matsusaka, S., H. Yamasaki, et al. (2007). "Upregulation of enzymes metabolizing 5-fluorouracil in colorectal cancer." Chemotherapy **53**(1): 36-41.
- Sakamoto, E., H. Nagase, et al. (2007). "Orotate phosphoribosyltransferase expression level in tumors is a potential determinant of the efficacy of 5-fluorouracil." Biochem Biophys Res Commun **363**(1): 216-22.
- Tai, I. T., M. Dai, et al. (2005). "Genome-wide expression analysis of therapy-resistant tumors reveals SPARC as a novel target for cancer therapy." J Clin Invest **115**(6): 1492-502.
- Tokunaga, Y., H. Sasaki, et al. (2007). "Clinical role of orotate phosphoribosyl transferase and dihydropyrimidine dehydrogenase in colorectal cancer treated with postoperative fluoropyrimidine." Surgery **141**(3): 346-53.