DNA Repair Pathway Profiling and Microsatellite Instability in Colorectal Cancer

Jinsheng Yu,[1,6] Mary A. Mallon,[2] Wanghai Zhang,[1,6] Robert R. Freimuth,[1,4] Sharon Marsh,[1,6] Mark A. Watson,[4,6] Paul J. Goodfellow,[2,3,6] and Howard L. McLeod[1,3,5,6]

Abstract

Background: The ability to maintain DNA integrity is a critical cellular function. DNA repair is conducted by distinct pathways of genes, many of which are thought to be altered in colorectal cancer. However, there has been little characterization of these pathways in colorectal cancer.

Method: By using the TaqMan real-time quantitative PCR, RNA expression profiling of 20 DNA repair pathway genes was done in matched tumor and normal tissues from 52 patients with Dukes' C colorectal cancer.

Results: The relative mRNA expression level across the 20 DNA repair pathway genes varied considerably, and the individual variability was also quite large, with an 85.4 median fold change in the tumor tissue genes and a 127.2 median fold change in the normal tissue genes. Tumor-normal differential expression was found in 13 of 20 DNA repair pathway genes (only *XPA* had a lower RNA level in the tumor samples; the other 12 genes had significantly higher tumor levels, all $P < 0.01$). Coordinated expression of *ERCC6, HMG1, MSH2*, and *POLB* ($R_S \geq 0.60$) was observed in the tumor tissues (all $P < 0.001$). Apoptosis index was not correlated with expression of the 20 DNA repair pathway genes. *MLH1* and *XRCC1* RNA expression was correlated with microsatellite instability status ($P = 0.045$ and 0.020, respectively). An inverse correlation was found between tumor *MLH1* RNA expression and *MLH1* DNA methylation ($P = 0.003$).

Conclusion: Our study provides an initial characterization of the DNA repair pathways for understanding the cellular DNA damage/repair system in human colorectal cancer.

DNA repair is an important process to maintain the integrity of DNA sequence. It seems to contribute to tumorigenesis and is also a mechanism for tumor resistance to chemotherapy (1, 2). Cellular repair of DNA is composed of several distinct pathways, which includes reversion repair, base excision repair, nucleotide excision repair (NER), mismatch repair (MMR), and DNA double-strand break repair (1, 3, 4). Of these DNA repair pathways, NER is a multistep process capable of removing a variety of DNA distorting lesions from prokaryotic and eukaryotic genomes. In eukaryotic cells, the process requires >30 proteins to perform the different steps (1), i.e., recognition of DNA damage, single-strand incisions and excision of the lesion-containing DNA fragment, and DNA repair synthesis/ligation. It has been shown that loss of function of several DNA repair genes is associated with increasing cancer risk, and the resistance to platinum-based drugs has also been associated with altered expression in base excision repair, NER, or MMR (3, 5–7).

In the DNA repair system, the translesional synthetic DNA polymerases POLB and POLH are paradoxically associated with DNA damage recognition protein HMG1 (8, 9). HMG1 binds preferentially to the platinated fork-like synthetic DNA and inhibits the translesional synthesis, which is in contrast to the function of POLB and POLH. Furthermore, defects in DNA mismatch repair system can result in low resistance/high sensitivity of tumor cells to radiation or platinum drug–based chemotherapy (10, 11), and an active MMR system can also lead to tumor cell death via the formation of futile cycles of translesional synthesis past cisplatin-DNA adducts followed by removal of the newly synthesized DNA (7, 11).

The DNA repair pathway is under the regulation of multiple cellular processes, which can be crucial determinants to the fate of tumor cells exposed to DNA-damaging agents: resistance or apoptosis. It has been known that TP53 is a potent mediator of cellular responses against genotoxic insults and exerts its effect on the DNA repair pathway in sequence-specific DNA-binding mode primarily at the transcriptional level (12–14). In addition to its role in regulating cell cycle arrest and apoptosis, TP53 is also known to be involved in regulating DNA repair mechanisms through the induction of P53R2 (14), a ribonucleotide reductase subunit, and may also directly participate in repair by promoting annealing of ssDNAs and rejoining double-stranded breaks. Other sensors or regulators of the DNA repair pathway, such as ATM, ATR, and RAD9,

Authors' Affiliations: Departments of [1]Medicine, [2]Surgery, [3]Genetics, [4]Pathology & Immunology, and [5]Molecular Biology & Pharmacology, Washington University School of Medicine and [6]Siteman Cancer Center, Saint Louis, Missouri

Received 3/6/06; revised 5/23/06; accepted 6/15/06.

Grant support: NIH grants U01 GM63340 and P30 CA091842.

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

Note: Data from this study are deposited in the Pharmacogenetics Knowledge Base (http://www.PharmGKB.org).

Requests for reprints: Howard L. McLeod, Washington University School of Medicine, 660 South Euclid Avenue-Campus Box 8069, Saint Louis, MO 63110-1093. Phone: 314-747-5183; Fax: 314-362-3764; E-mail: hmcleod@im.wustl.edu.

doi:10.1158/1078-0432.CCR-06-0547

involved in DNA-damage signaling (15–17), were also included in the present study to determine any potential association between these regulators and tumor cell apoptosis.

In addition, defects in the DNA mismatch repair machinery can lead to instability of short tandem nucleotide repeats, called microsatellite instability (MSI). MSI is typical of hereditary nonpolyposis colorectal cancer that accounts for 2% to 3% of all colorectal carcinomas (18), but MSI has also been identified in 10% to 15% of sporadic colorectal carcinomas in which constitutional mutations of MLH1 and MSH2 are rarely found (19, 20). MSI is frequently due to MLH1 promoter hypermethylation (21, 22). It has been widely shown that loss of MMR protein expression in tumor tissue may correlate with MSI (23, 24). Therefore, the association between MSI status and gene expression of the DNA repair pathway is important to investigate for cellular DNA repair function.

Most studies have been conducted to evaluate specific DNA repair genes. In this study, we have profiled the RNA expression

Table 1. Gene name and primer list for RNA profiling of the DNA repair pathways

Gene symbol	Description	Forward primer 5′ to 3′	Reverse primer 5′ to 3′	TaqMan probe 5′ to 3′
ATM	Ataxia telangiectasia mutated	GCTGAGGCAGGAGAATCTCTTG	TGGAGTGCAATGGCAT GATTT	CCCACAGCAACCTTCACCTCCCA
ATR	Ataxia telangiectasia and Rad3 related	CCAGTGAAAGGGCATTCCA	GGGTCTTGGCCTTTTCATTG	CAACTTCTCCAGTTTC-ATTCAGTGGCGCA
DDB1	Damage-specific DNA binding protein 1, 127 kDa excision repair cross-complementing	GTGGTGGCAAACCTACAGTATGA	TCCGAGTTAGCTCCTCCACAAC	AAGCGAGAGGCCACTGCAGACGACC
ERCC1	Rodent repair deficiency, complementation group 1 excision repair cross-complementing	TACCCCTCGACGAGGATGAG	CAGTGGGAAGGCTCTGTGTAGA	CCTGGAGTGGCCAAGCCCTTATTCC
ERCC2	Rodent repair deficiency, complementation group 2, XPD excision repair cross-complementing	TTGGCGTCCCCTACGTCTAC	CTGGTCCCGCAGGTATTCC	CACAGAGCCGCATTCTCAAGGCG
ERCC3	Rodent repair deficiency, complementation group 3, XPB excision repair cross-complementing	GGAATTTGGCTCCAGATCCA	GATGAGTGGTACTCC-ATGTACACAGTGT	CCAGGCATCTCGGCGCTTTGG
ERCC4	Rodent repair deficiency, complementation group 4, XPF excision repair cross-complementing	TTGTGGATATGCGTGAATTTCG	CACGGGTTCAATGTCAATGC	TGAGCTTCCATCTCTGATCCATCGTCG
ERCC5	Rodent repair deficiency, complementation group 5, XPG excision repair cross-complementing	ACTCACCCCTGGCTTTCCTAA	GAGTCATCCACCACGGGTTT	CCAGCTGTTGCCGAGGCCTACCT
ERCC6	Rodent repair deficiency, complementation group 6, CSB	ACAAGTGCAATTTTTGCAG-GAACT	GCTCCAAAGGCTGGTTGAATC	ATCAGATGTTCAGACACCC-AAATGCCATCTAA
HMG1	High-mobility group (nonhistone chromosomal) protein 1	TCCTGGCCTGTCCATTGG	GCTTGTCATCTGCAGCAGTGTTAT	CCACATCTCTCCCAG-TTTCTTCGCAACA
MLH1	Mut L homologue 1, colon cancer, nonpolyposis type 2 (*Escherichia coli*)	CCATCCGGAAGCAGTACATATCT	ATGGAGCCAGGCACTTCACT	AGGAGTCGACCCTCTCAGGCCAGC
MSH2	Mut S homologue 2, colon cancer, nonpolyposis type 1 (*E. coli*)	TATCAGGTGAAGAAAGGTG-TCTGTGA	GCACACTCTATTACATGC-TTAGGGAAAT	AGCAAGCTCTGCAACATGAATCCCAA
MSH6	Mut S homologue 6 (*E. coli*)	GGTGCTTGTGGATGAATTAGGAA	GCAAGTTCTTTAACA-ACTGCATTTG	TATTGCCGTCCCATCAAA-TGTTGCAGTA
POLB	Polymerase (DNA directed), β	TACATTGCTACAGTCTGT-GGCAGTT	TGGGATGGGTCAGGAGAACA	CCATGTCACCACTGGAC-TCTGCACCTCT
POLH	Polymerase (DNA directed), η	GCATTTAGGGAGGCAGTGTCA	TTCAAGGCCCTCATTGCAA	AACCACGTCTCTTAGA-CTCCAAGGCCCA
RAD9	RAD9 homologue A (*Saccharomyces pombe*)	TCTTACATGATCGCCATGGAAA	CCAGGTGAAAGGGAAATGGA	AGCACCCGCGAGCCCTCATTG
TP53	Tumor protein p53 (Li-Fraumeni syndrome)	AGACTGGGTCTCGCTTTGTTG	AGGCAAAGGCTGCAGTAAGC	AAGATCACGCCACTCCACTCCAGCC
XPA	Xeroderma pigmentosum, complementation group A	TCTGTGATTGCCTTCTTACA-ACAGA	CCTTGGTATCTTGTCCT-CAAATTTG	TGGGAGCTGAGTGCTAGA-GTAGGTGCAGA
XPC	Xeroderma pigmentosum, complementation group C	GAGAGGCTGAAGCGTCGCTA	GAAGAGAGTCCACCTC-CTGCATCT	AAGAGTGAGGCAGCAG-CTCCCCACA
XRCC1	X-ray repair complementing defective repair in Chinese hamster cells 1	GAACACCAGGAGCCTCCTGAT	AAGAAGTGCTTGCCCTGGAA	TGCCAGTCCCTGAGCTCCCAGATTT

Table 2. *MLH1* DNA methylation primer list

Primers*	Round 1 (external)	Round 2 (nested)
Forward	5′-tttTtTaaTtTtgtgggttgTtggg-3′	5′-Bio-TtgTTcgTtaTTtagaaggatatg-3′
Reverse	5′-AAaAAccacaaAaAcaAAAccaa-3′	5′-tctActcctattAActAAatatttc-3′
Pyrosequencing	5′-TcgTtaTTtagaaggatatg-3′	

*The capital letters in the sequence are bisulfite converted.

of 20 genes in the DNA repair pathways by a quantitative real-time reverse transcription-PCR assay, and we have also evaluated the potential associations between DNA repair gene expression and three distinct phenotypes: DNA methylation, genomic microsatellite status, and tumor cell apoptosis. This study aims to further understand the DNA repair pathway and to provide insights into chemosensitivity to DNA-damaging agents in human colorectal cancer.

Materials and Methods

Patients and samples. In this study, genomic DNA and total RNA samples were prepared from tumor specimens and paired nonmalignant tissues of 52 consecutive patients with Dukes' C colorectal cancer undergoing colorectal surgery at Barnes Jewish Hospital. The age of the patients ranged from 32 to 96 years (median 69 years); 29 males and 23 females were included. Dissected specimens were snap frozen in liquid nitrogen immediately after surgery and stored at −80°C. None of the patients had received preoperative radiation or chemotherapy. Histologic examination was done in all of the cases to evaluate tumor histotype (41 enteric and 11 mucinous) and grade of differentiation (1, 38, and 13 in grade 1, 2, and 3, respectively) according to WHO criteria. Twenty-seven tumors were localized in the right colon, 19 in the left colon, and the remaining 6 were localized in the rectum. Written informed consent was obtained from all patients to bank tumor tissue and to perform genetic analysis. This study was approved by the Washington University Human Subjects Committee.

Extraction of genomic DNA and cellular total RNA. The tumor specimens selected for DNA and RNA isolation had high tumor cellularities (median 86% with a range of 65-95%). Tissue total RNA was isolated with the TRIzol RNA isolation kit (Invitrogen, Carlsbad, CA). The quality of RNA ($A_{260/280}$ >1.8; clear RNA bands for 28S, 18S, and 5S) was confirmed in the Siteman Cancer Center Tissue Procurement Core. Genomic DNA was extracted with Qiagen DNeasy Tissue kit (Qiagen, Hilden, Germany).

Quantitative real-time reverse transcription-PCR. Two-step reverse transcription-PCR was conducted in this study. First, cDNA was synthesized in a 20 μL reaction containing 5 μg of total RNA, 0.5 μg of oligo$(dT)_{20}$VN primer, and 100 units of StrataScript reverse transcriptase (Stratagene, La Jolla, CA). The cDNA samples were then diluted to a concentration of 10 ng total RNA/μL. Ten-microliter reaction mixture for real-time PCR was composed of 5 μL of 2× TaqMan universal PCR master mix (Applied Biosystems, Foster City, CA), 3 μL primer and probe mix (600 nmol/L each forward and reverse primers and 100 nmol/L specific TaqMan probe), and 2 μL cDNA (20 ng). All real-time PCR assays were done in triplicate on an ABI

Table 3. Relative mRNA level, and category and quartile of tumor-normal ratio of the 20 DNA repair pathway genes in 52 colorectal cancers

Gene symbol	Median RNA level in tumors	Median RNA level in normals	Median T/N ratio	Quartile	T/N < 0.8	T/N = 0.8-1.2	T/N > 1.2
*ERCC2**	1.33E+05	5.61E+04	2.85	Top	6	6	40
*POLB**	4.03E+03	1.88E+03	2.72		2	6	44
*ERCC4**	5.02E+03	1.62E+03	2.67		6	7	39
ERCC6†	6.82E+02	2.98E+02	2.28		13	4	35
TP53†	3.11E+05	2.41E+05	2.02		15	5	32
*HMG1**	1.18E+06	6.05E+05	1.95	Second	4	5	43
MSH2†	1.10E+06	4.72E+05	1.92		14	5	33
*ATR**	2.98E+05	1.38E+05	1.88		12	3	37
*ERCC3**	9.95E+04	5.29E+04	1.86		9	4	39
*ERCC1**	2.99E+04	2.19E+04	1.44		4	15	33
XRCC1†	2.50E+05	1.74E+05	1.44	Third	13	9	30
DDB1‡	9.07E+05	6.54E+05	1.36		12	11	29
RAD9	1.87E+05	1.47E+05	1.33		13	12	27
MSH6	1.78E+02	1.50E+02	1.30		18	7	27
XPC	3.28E+04	2.84E+04	1.28		17	8	27
ERCC5	1.80E+04	1.73E+04	1.27	Bottom	19	6	27
POLH	1.42E+05	1.53E+05	1.25		20	6	26
ATM	4.62E+05	4.22E+05	1.06		16	17	19
MLH1	3.26E+05	3.05E+05	1.04		18	10	24
XPA†	4.85E+05	7.13E+05	0.78		27	11	14

Abbreviation: T/N, tumor-normal ratio.
*P < 0.001, compared between tumor and normal tissues by Wilcoxon matched pair test.
†P < 0.01, compared between tumor and normal tissues by Wilcoxon matched pair test.
‡P < 0.05, compared between tumor and normal tissues by Wilcoxon matched pair test.

PRISM 7900HT Sequence Detector System (Applied Biosystems) according to the following program: 50°C for 2 minutes, 95°C for 10 minutes, 40 cycles at 95°C for 20 seconds, and at 60°C for 1 minute. Primers and TaqMan probes used in this study were designed using Primer Express version 1.5 (Applied Biosystems). The sequence of primers and probes specific for each gene are displayed in Table 1. The specificity of each primer/probe set was determined with a pretest showing the specific amplification for a specific gene by gel visualization.

Measurement of relative expression of mRNA. Relative mRNA expression level of each gene was calculated by a curve-fit method (instead of the $\Delta\Delta C_T$ method), which uses nonlinear regression to calculate the transcript abundance in each sample based on the kinetic raw data of fluorescence intensity at each cycle during the PCR reaction (25). Because of a wide range of the standard curve slope (−2.9 to −3.8) for the 21 genes in this study, the $\Delta\Delta C_T$ method was not suitable to analyze relative RNA expression level of these genes. A mathematical regression model for simulation of sigmoid curve was used as follows (25): $R_n = R_{max} / [1 + \exp(-(n - n_{1/2}) / k)]$, where R_n is background-subtracted fluorescence intensity in real-time PCR at cycle n, R_{max} is the maximal fluorescence intensity, $n_{1/2}$ is the cycle number when fluorescence intensity is half of the R_{max}, and k is the slope factor of increase in fluorescence intensity. The regression analysis gives the variables R_{max}, $n_{1/2}$, and k. The initial value of each sample (R_0) equals $R_{max} / [1 + \exp(n_{1/2} / k)]$. An internal reference gene, amyloid β precursor protein, which had nearly identical expression between colon tumor and normal tissues (46:31 copies/cell; ref. 26), was used to control variation in RNA concentration across individual samples. Finally, the value of sample RNA level normalized to the internal reference gene was scaled to a 1× sample so that the minimal level was 1 in the entire data set. Thus, the RNA expression level in this study is a unitless value or called arbitrary unit, relative to both the internal reference gene and the 1× sample.

MLH1 methylation analysis. Genomic DNAs were first converted with treatment of sodium bisulfite as described previously (24). After treatment, unmethylated cytosine residues are converted to thymine, whereas methylated cytosine residues are retained as cytosine. In addition, a human genomic DNA sample purchased from Promega (Madison, WI) was included either for negative control (unmethylated) when it was converted only by sodium bisulfite, or for positive control (high methylation) when it was treated first by the SssI methylase (New England Biolabs, Beverly, MA) and then converted by sodium bisulfite. The converted DNA samples underwent two rounds of amplification to obtain the gene-specific fragments for the pyrosequencing reaction, which was carried out to quantify the DNA methylation level. The first-round external PCR reaction were done using Amplitaq Gold PCR master mix (ABI, Foster City, CA), 5 pmol of each primer, and 10 ng of the bisulfite-converted genomic DNA in a 20 μL reaction. This external reaction was run for 40 cycles at 94°C, 55°C, and 72°C for 1 minute, whereas the second-round nested reaction was run for 55 cycles at 94°C for 30 seconds, 60°C for 30 seconds, and 72°C for 1 minute. The nested reaction was conducted with 5′-biotinylated reverse or forward primer in a 20 μL reaction containing 2 μL of 1:5 diluted first-round PCR products, and 4 pmol of each primer. Pyrosequencing was done as previously described (27), by using the pyrosequencing PSQ HS96A instrument and software (Biotage, Uppsala, Sweden). To validate specificity of the pyrosequencing reaction, the fragment-only, pyrosequencing-primer-only, and no-primer and no-fragment reactions were included for negative control. The *MLH1* methylation primers were listed in Table 2. Finally, the function of allele quantification in the pyrosequencing PSQ HS96A system is capable of quantifying the DNA methylation level for each CpG locus in the DNA samples. The DNA methylation levels were calculated as ratio of the base C to T peaks at a given CpG site in pyrograms. The average of three experiments was taken for the DNA methylation level of each CpG locus. When the DNA methylation level (expressed as average of multiple sites) was <5%, it is defined as unmethylated. Otherwise, 5% to 20% is denoted as low methylation, 21% to 50% as medium methylation, and >50% as high methylation.

MSI analysis. Analysis of MSI status in genomic DNA samples was done as previously described (28), using the panel of five consensus markers recommended for the detection of MSI in colorectal cancer (29). This panel of markers includes three dinucleotide markers (D2S123, D5S346, and D17S250) and two mononucleotide markers (BAT25 and BAT26). Samples were classified as MSI-high if instability was detected at two or more consensus markers, MSI-low if instability was confined to only one consensus marker, and microsatellite stable if none of the consensus markers revealed instability. The data for the 52 tumors have been previously reported (30).

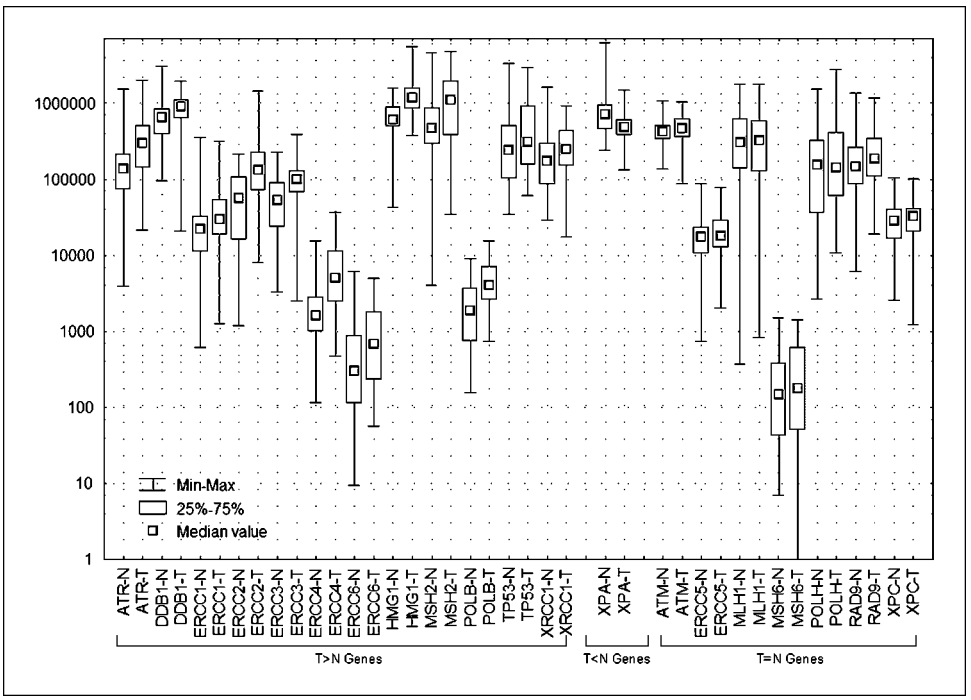

Fig. 1. Box-whisker plot demonstrating the variability in RNA expression level (log scale, relative units to the internal reference gene *APP*) for the 20 DNA repair pathway genes in the colorectal normal (−*N*) and tumor (−*T*) tissues.

Table 4. Spearman rank correlation matrix of the 20 DNA repair pathway genes

DNA repair pathway	Nucleotide excision repair								
Gene symbol	***DDB1***	***ERCC1***	***ERCC2***	***ERCC3***	***ERCC4***	***ERCC5***	***ERCC6***	***XPA***	***XPC***
DDB1	(1.00)								
ERCC1	0.05	(1.00)							
ERCC2	0.10	0.56	(1.00)						
ERCC3	0.58	0.16	0.23	(1.00)					
ERCC4	0.35	0.48	0.20	**0.62**	(1.00)				
ERCC5	0.57	0.12	0.13	0.53	0.47	(1.00)			
ERCC6	0.21	0.54	0.33	0.43	**0.67**	0.38	(1.00)		
XPA	0.19	0.28	−0.16	0.16	0.12	0.16	0.10	(1.00)	
XPC	0.50	0.21	0.17	**0.66**	0.43	0.54	0.34	0.30	(1.00)
XRCC1	0.19	−0.10	−0.09	0.39	0.10	0.31	−0.02	0.17	0.34
MLH1	0.09	0.43	0.28	0.19	0.27	0.20	0.45	0.09	0.14
MSH2	**0.61**	0.19	0.33	0.54	0.40	**0.64**	0.32	0.13	0.37
MSH6	0.27	−0.24	−0.07	0.28	0.03	0.29	−0.12	−0.04	0.26
POLB	0.22	**0.67**	0.63	0.07	0.21	0.17	0.37	0.19	0.11
POLH	0.13	−0.45	−0.19	0.03	−0.22	0.13	−0.36	−0.14	0.20
HMG1	0.14	**0.61**	0.52	0.20	0.35	0.24	**0.61**	0.03	0.02
ATM	0.03	−0.02	0.01	0.02	−0.08	0.15	−0.01	0.33	0.05
ATR	0.43	0.47	0.28	0.46	0.54	0.59	**0.60**	0.22	0.40
RAD9	0.16	0.09	0.30	0.17	−0.05	0.15	0.03	−0.03	0.34
TP53	0.15	0.30	0.19	0.33	0.56	0.26	0.48	−0.10	0.18

NOTE: All Spearman correlation scores >0.60 ($P < 0.001$) are displayed in bold font.

Labeling of apoptotic cells by immunohistochemistry. Terminal deoxyribonucleotide transferase–mediated nick-end labeling (TUNEL) staining allows the *in situ* detection of apoptotic cells using an immune peroxidase detection system. Fixed tissue sections (5 μmol/L) were rehydrated and immunostained using the *In situ* Cell Death Detection kit, POD TUNEL assay (Roche Diagnostics Corporation, Indianapolis, IN), according to the specifications of the manufacturer. Nonspecific antifluorescein antibody binding was blocked by washing the slides in 1% bovine serum albumin in PBS thrice for 10 minutes each, and the slides were rinsed with PBS. The slides were finally counterstained with Harris hematoxylin for 20 seconds to reveal the nuclei, mounted with xylene-based medium, and covered with glass coverslips. The result was scored semiquantitatively, 0 as negative, 1 as positive cells <5% per high power field, 2 as positive cells 6% to 10%, and 3 as >10% positive cells. The final score was averaged from three continuous sections of each tumor.

Statistical analysis. Descriptive statistical analyses were done using the software STATISTICA from StatSoft, Inc. (Tulsa, OK). The difference in RNA expression between paired tumor and normal samples was evaluated using Wilcoxon matched-pair test. The association of patient's phenotypes (e.g., age, gender, tumor location, DNA methylation, or MSI status) to RNA expression level was evaluated with either the Mann-Whitney test or the Kruskal-Wallis test. Spearman rank correlations were used to compare continuous variables.

Results

Differential expression of the pathway genes. Many of the pathway genes (65%) were differentially expressed between tumor and normal tissues. One gene (*XPA*) had a significantly lower mRNA expression level in the 52 colorectal tumor tissues than in the matched adjacent normal tissues (median 0.78-fold lower, $P = 0.004$); but 12 of 20 (60%) DNA repair pathway genes studied had significantly higher mRNA level in the tumors than in the normal tissues (median range 1.36- to 2.85-fold higher, all $P < 0.01$). There were no significant differences between paired tumor and normal samples in 7 of 20 genes (35%, $P = 0.061$-0.682; Table 3). Furthermore, patients were classified according to a tumor-normal ratio of <0.8, 0.8-1.2, or >1.2 for each gene. A category list is presented in Table 3 to further address the differential expression of the 20 pathway genes.

Variability of mRNA expression of the pathway genes. Variability for each gene was large (Fig. 1); the coefficient of variance ranged from 43.4% to 155.0% (median 88.4%) in the tumor tissues and from 41.6% to 166.6% (median 110.4%) in the normal tissues. Similarly, the fold change of gene expression in the 52 colorectal cancer patients was wide for most of the 20 DNA repair pathway genes; it ranged from *ATM* 11.1 to *MLH1* 2,209.3 (median 85.4) in the tumor and also from *ATM* 8.0 to *MLH1* 4,761.5 (median 127.2) in the normal tissue. Although some genes had a significantly different mRNA level in the tumors, not all patients fell into a single category of either tumor-normal ratio of <0.8 or >1.2 (Table 3). Like the RNA level itself, the tumor-normal ratio of the 52 patients for each gene also varied widely, with the coefficient of variance from 63.7% to 274.0% (median 142.9%) and the fold change from 28.2 to 4,825.3 (median 246.5).

Relative expression of the pathway genes. By scaling entire data set to a 1× sample, we were able to look at relative expression across multiple genes. As shown in Fig. 1 and Table 3, the relative expression level across the 20 DNA repair genes varied largely. The median mRNA expression level in the tumor samples ranged from the lowest *MSH6* (1.78×10^2 units) to the highest *HMG1* (1.18×10^6 units), representing a 6,653-fold change; and in the normal samples, again *MSH6* (1.50×10^2 units) was the lowest one, and the highest one was *XPA* (7.13×10^5 units), representing a 4,768-fold change (Table 3). The median tumor-normal ratios for the 20 genes ranged from 0.78 to 2.85.

Correlation of the pathway genes. Coordinated expression was noted among the 20 DNA repair pathway genes. For instance, *ERCC6*, *HMG1*, *MSH2*, and *POLB* each had a Spearman rank correlation score of $R_S \geq 0.60$ with three other

Table 4. Spearman rank correlation matrix of the 20 DNA repair pathway genes (Cont'd)

Base excision repair	Mismatch repair			Translesional synthesis		Damage recognition	Pathway regulation			
XRCC1	*MLH1*	*MSH2*	*MSH6*	*POLB*	*POLH*	*HMG1*	*ATM*	*ATR*	*RAD9*	*TP53*
(1.00)										
0.25	(1.00)									
0.17	0.21	(1.00)								
0.31	0.07	0.12	(1.00)							
−0.19	0.36	0.27	−0.08	(1.00)						
0.36	−0.17	−0.13	**0.61**	−0.26	(1.00)					
−0.08	0.46	0.28	−0.02	0.67	−0.29	(1.00)				
−0.08	−0.22	0.10	−0.12	0.12	−0.04	0.10	(1.00)			
0.22	0.36	**0.64**	0.01	0.32	−0.30	0.49	0.12	(1.00)		
−0.05	0.16	0.12	0.14	0.20	0.20	0.09	0.13	−0.03	(1.00)	
0.18	0.22	0.28	0.04	0.16	−0.27	0.40	−0.06	0.51	−0.02	(1.00)

genes in the colon tumor tissues (all $P < 0.001$; Table 4). Moreover, among the nine NER genes evaluated, each pair of *ERCC3* and *ERCC4*, *ERCC4* and *ERCC6*, and *XPC* and *ERCC3* had a closer correlation in the RNA expression than other NER genes ($R_S > 0.60$, $P < 0.001$). The three MMR genes in this study had no significant correlations in the RNA expression.

MSI status, MLH1 methylation and apoptosis index. Ten of the 52 tumors (19%) were classified as MSI-high. Four tumors (8%) showed instability at only one locus (classified as MSI-L), and 38 (73%) showed no instability at any of the five loci examined (classified as microsatellite stable). The median tumor RNA expressions of *MLH1* and *XRCC1* were significantly lower in patients with MSI-high than those with microsatellite stable (*MLH1*: 177,332 versus 362,339 units, *XRCC1*: 138,754 versus 310,419 units; Mann-Whitney *U* test, $P = 0.04$ and 0.02 for *MLH1* and *XRCC1*, respectively), but another two MMR genes (*MSH2* and *MSH6*) and other DNA repair genes had no statistical difference (Fig. 2). In the DNA methylation experiment, five CpG dinucleotide loci in *MLH1* promoter region (−248 to −222 from the transcriptional start site) were evaluated and a total of 20 tumors were defined as *MLH1* methylated, of which nine, five, and six tumors had high methylation (64.9-90.6%, median 80.0%), medium methylation (21.1-44.0%, median 31.4%), and low methylation (5.0-16.1%, median 7.2%), respectively. A reciprocal relationship between the average DNA methylation degree and the RNA level was observed for *MLH1* in the 20 methylated tumors ($R_S = -0.624$, $P = 0.003$; Fig. 3). Furthermore, there was concordance between the *MLH1* high-methylation tumors (89%; eight of nine) and the MSI-high tumors (80%; 8 of 10) in colorectal cancer patients (Table 5). Forty-six tumors had available sections for the TUNEL staining (used as apoptosis index), of which 25 showed low apoptosis (average apoptosis index score was <1.5) and 21 high apoptosis (apoptosis index score equal to or >1.5). All 20 pathway genes had no obvious difference in their RNA level between the high and low apoptosis tumors (all $P > 0.1$, Mann-Whitney *U* test); and there was no difference in apoptosis index between MSI-high and microsatellite stable tumors.

Association between RNA level and clinicopathologic features. No significant association between RNA level was found by patient age, gender, tumor location, cellularity, histologic type, or grade of tumor differentiation for any of the 20 pathway genes. In addition, no association between RNA level and the apoptosis index was observed for all genes analyzed in this study, including the four coordinated genes (*ERCC6*, *HMG1*, *MSH2*, and *POLB*). No patient outcome data are available for further analysis of correlation with the RNA level or methylation level.

Discussion

The DNA repair system is important for maintaining the stability of normal genomic function of cells. In human cancer, one of the most promising molecular phenotypes investigated to date is MSI and the high MSI phenotype has been suggested to be involved in development of colorectal cancer through DNA repair defects, particularly in MMR genes such as *MLH1*.

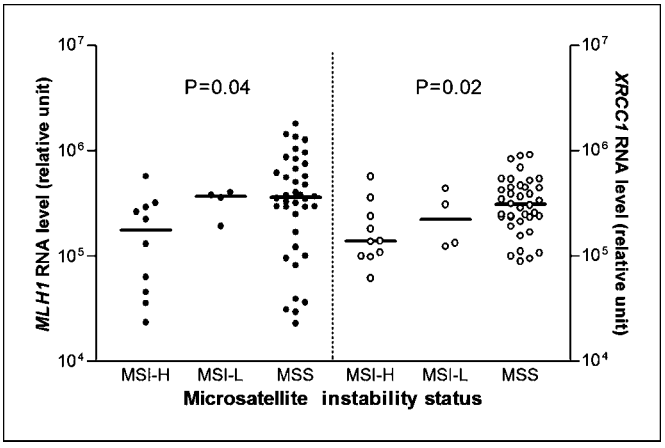

Fig. 2. Tumor RNA expression plots (log scale) by the MSI status for *MLH1* (●) and *XRCC1* (○). Mann-Whitney *U* test, $P = 0.04$ (*MLH1*) and 0.02 (*XRCC1*) when the MSI-high (*MSI-H*) group was compared with the microsatellite stable (*MSS*) group.

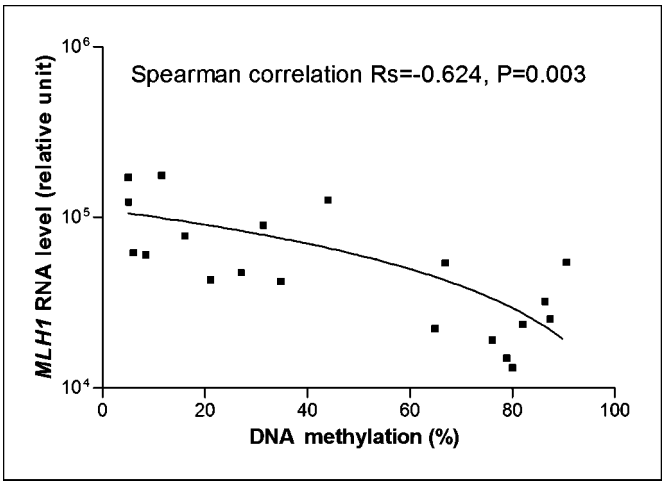

Fig. 3. Correlation between the *MLH1* RNA level (log scale) and DNA methylation in the *MLH1*-methylated (>5%) tumors.

Approximately 15% of colorectal cancers are characterized by genomic MSI, reflecting inactivation of the MMR genes (31, 32). Our study showed an inverse correlation between the *MLH1* DNA methylation and RNA expression (Fig. 3), and between the MSI status and the RNA expression level for the MMR gene *MLH1* (Fig. 2), which is consistent with the concept of inactivation of cellular DNA repair function (e.g., through DNA methylation), leading to genomic MSI. This association is further confirmed by concordance between *MLH1* methylation and MSI-high in the colorectal tumors. Our data did not reveal a clear RNA-MSI relationship for other MMR genes (*MSH2* and *MSH6*), likely reflecting additional functional influences of these genes. Recent studies have shown that the MMR proteins MSH2 and MSH6 functionally interact and MSH6 is degraded in the absence of its binding partner (33). Thus, the mechanisms that cause genomic MSI may not be only through transcriptional regulation of the MMR genes. Also, our data did not show any associations between tumor apoptosis, MSI status, and the RNA level of DNA repair pathway genes, suggesting that tumor apoptosis activity, and similarly the response of tumor cells to DNA-damaging agents, may not rely solely on gene activity of the DNA repair pathway.

Previous studies have shown that *TP53* is required for efficient NER processing of UV-induced DNA lesions (12–14, 34) and in the present study *TP53* was overexpressed in the colon cancer tissues. Although no correlation was observed between *TP53* and the other 19 DNA repair pathway genes in our study ($P > 0.05$), functional TP53 sure has critical interaction with many of the pathway genes and MDM2 can largely affect protein level of the functional TP53 (35). Besides, our data showed a large interpatient variability of the RNA expression of 20 DNA repair pathway genes, suggesting a substantial source of variable chemosensitivity in colorectal cancer. With the large variability of the pathway genes, individualized cancer therapy will be needed for clinical practice where cancer patients have a wide range of response to certain compounds, from no response at all to fatal outcome (36).

DNA repair subpathways are a multistep, multielement process to remove the damaged DNA section and to resynthesize that part of the DNA strand, except for reversion repair pathways. Thus, interplay involved in multiple genes is critical and more informative for identifying genes responsible for chemosensitivity. In the past 10 years, high-throughput approaches for gene expression profiling have defined key determinants that may play a role in prediction of cancer patient outcome, as well as response to chemotherapy (37–39). For instance, Wang et al. (37), using microarray technology, have identified a 23-gene signature that predicts recurrence in Dukes' B patients with 78% overall accuracy. Murakami et al. (39) used cDNA microarray technology to examine the effect of UVB irradiation on 588 cancer-related genes in human keratinocytes and identified several UV-reactive genes, including *ERCC1* and *XRCC1*. In our study, with the Spearman rank correlation test, a variety of correlations were observed among the 20 DNA repair pathway genes in the colon cancer tissues. Those genes with a higher Spearman correlation score (>0.60; Fig. 1) may have coregulation or coordinated expression in a certain mode. Although all highly correlated genes are not on the same chromosome, they may share regulatory domains or other transacting mechanisms. For example, *ERCC4*, located in chromosome 16, is required for incision of the damaged DNA backbone at the 5′ side of the lesion in the NER process. The 3′ side incision by *ERCC5* requires *ERCC6*, which is located on chromosome 10 and had a Spearman score of 0.67 with *ERCC4*. Thus, these two genes could have a certain common transacting mechanism for DNA repair. Moreover, *HMG1* is located on chromosome 13 and *POLB* on chromosome 8, and they had a Spearman correlation score of 0.67. Those two genes participate in the DNA damage recognition and the translesional synthesis process, respectively. A positive correlation for these two genes may be suggestive of a synergetic association between them to fulfill DNA repair, but is contradictive to the fact that HMG1 inhibits the translesional synthesis by POLB and POLH. These results from clustering or correlation analysis will need to be further validated with functional studies and/or clinical trails designed to use expression profiles of these genes in tumors to predict response to DNA-damaging drugs.

Defects in DNA repair give rise to hypersensitivity of tumors to DNA damage, and overexpression of DNA repair genes render tumor resistance to DNA-damaging agents. A recent study has reported that low levels of the NER enzymes ERCC1, ERCC4, and XPA are related to the favorable response of testis tumors to cisplatin-based chemotherapy (40). Ferry et al. (6) have also shown that several NER genes, including *ERCC1*, are constitutively overexpressed in the highly cisplatin-resistant cell line C200 compared with the sensitive parental A2780 cells. In

Table 5. Relationship between MLH1 methylation and MSI status

MLH1 methylation	MSI-H (n = 10)	MSI-L (n = 4)	MSS (n = 38)
HM (n = 9)	8	0	1
MM (n = 5)	0	2	3
LM (n = 6)	0	0	6
UM (n = 29)	2	2	25

NOTE: Three samples failed in the DNA methylation experiment. Abbreviations: MSI-H, MSI-high; MSI-L, MSI-low; MSS, microsatellite stable; HM, high methylation; MM, medium methylation; LM, low methylation; UM, unmethylated.

our study, differential expression analysis of the DNA repair pathway genes indicates that 12 of the 20 pathway genes had increased expression in the tumor tissues, whereas only one (*XPA*) was higher in the nonmalignant tissues (Table 2). It is noted that a number of the DNA repair pathway genes were overexpressed, which may contribute at least in part to low response rates from single platinum agent treatment of colorectal cancer. On the other hand, the resistance of colorectal cancer to platinum drugs may not likely come from *XPA* in most patients with colorectal cancers (27 of 52, 52%) because of its lower tumor RNA expression level. Meanwhile, *ATR* involvement in platinum resistance has been suggested (15, 17), and in our study the *ATR* mRNA expression was significantly overexpressed in the colorectal tumor tissues (1.88-fold, $P < 0.001$). Thus, *ATR* may become a molecular marker for prediction of response in platinum-based chemotherapies.

Development of intrinsic or acquired resistance to DNA-damaging drugs may rely on several elements, such as the gene activity of the DNA repair pathway, defects in whole cell uptake and adduct formation, and the so-called secondary lesions (DNA strand breaks and apoptosis; refs. 2, 41, 42). For instance, it has been shown that oxaliplatin forms fewer DNA adducts than cisplatin, and oxaliplatin adducts are poorly recognized by the MMR system (2). Moreover, the ability of oxaliplatin adducts to activate signal transduction pathways, ultimately leading to apoptotic DNA fragmentation, differs from cisplatin (2, 3, 41, 42). In answer to the issue of chemosensitivity to DNA-damaging agents, by analyzing the differential expression, individual variability, and coexpression, our RNA expression profiling study has provided an initial characterization of the DNA repair pathways to aid in further understanding of the cellular DNA damage/repair system in human colorectal cancer.

Acknowledgments

We thank the staff of the Siteman Cancer Center Tissue Procurement Core and the Washington University Genome Sequencing Center.

References

1. Christmann M, Tomicic MT, Roos WP, Kaina B. Mechanisms of human DNA repair: an update. Toxicology 2003;193:3–34.
2. Brabec V, Kasparkova J. Molecular aspects of resistance to antitumor platinum drugs. Drug Resist Updat 2002;5:147–61.
3. Rosell R, Taron M, Barnadas A, Scagliotti G, Sarries C, Roig B. Nucleotide excision repair pathways involved in cisplatin resistance in non-small-cell lung cancer. Cancer Control 2003;10:297–305.
4. Kunkel TA, Erie DA. DNA mismatch repair. Annu Rev Biochem 2005;74:681–710.
5. Furuta T, Ueda T, Aune G, Sarasin A, Kraemer KH, Pommier Y. Transcription-coupled nucleotide excision repair as a determinant of cisplatin sensitivity of human cells. Cancer Res 2002;62:4899–902.
6. Ferry KV, Hamilton TC, Johnson SW. Increased nucleotide excision repair in cisplatin-resistant ovarian cancer cells: role of ERCC1-XPF. Biochem Pharmacol 2000;60:1305–13.
7. Beljanski V, Marzilli LG, Doetsch PW. DNA damage-processing pathways involved in the eukaryotic cellular response to anticancer DNA cross-linking drugs. Mol Pharmacol 2004;65:1496–506.
8. Zamble DB, Mikata Y, Eng CH, Sandman KE, Lippard SJ. Testis-specific HMG-domain protein alters the responses of cells to cisplatin. J Inorg Biochem 2002; 91:451–62.
9. Mitkova E, Ugrinova I, Pashev IG, Pasheva EA. The inhibitory effect of HMGB-1 protein on the repair of cisplatin-damaged DNA is accomplished through the acidic domain. Biochemistry 2005;44:5893–8.
10. Sergent C, Franco N, Chapusot C, et al. Human colon cancer cells surviving high doses of cisplatin or oxaliplatin *in vitro* are not defective in DNA mismatch repair proteins. Cancer Chemother Pharmacol 2002; 49:445–52.
11. Lawes DA, Pearson T, Sengupta S, Boulos PB. The role of MLH1, MSH2 and MSH6 in the development of multiple colorectal cancers. Br J Cancer 2005;93: 472–7.
12. Adimoolam S, Ford JM. p53 and regulation of DNA damage recognition during nucleotide excision repair. DNA Repair (Amst) 2003;2:947–54.
13. Lin X, Ramamurthi K, Mishima M, Kondo A, Christen RD, Howell SB. P53 modulates the effect of loss of DNA mismatch repair on the sensitivity of human colon cancer cells to the cytotoxic and mutagenic effects of cisplatin. Cancer Res 2001;61:1508–16.
14. Zhou B, Liu X, Mo X, et al. The human ribonucleotide reductase subunit hRRM2 complements p53R2 in response to UV-induced DNA repair in cells with mutant p53. Cancer Res 2003;63:6583–94.
15. Shechter D, Costanzo V, Gautier J. Regulation of DNA replication by ATR: signaling in response to DNA intermediates. DNA Repair (Amst) 2004;3:901–8.
16. Yin Y, Zhu A, Jin YJ, et al. Human RAD9 checkpoint control/proapoptotic protein can activate transcription of p21. Proc Natl Acad Sci U S A 2004;101:8864–9.
17. Yazlovitskaya EM, Persons DL. Inhibition of cisplatin-induced ATR activity and enhanced sensitivity to cisplatin. Anticancer Res 2003;23:2275–9.
18. Ponz de Leon M, Benatti P, Di Gregorio C, et al. Genetic testing among high-risk individuals in families with hereditary nonpolyposis colorectal cancer. Br J Cancer 2004;90:882–7.
19. Raut CP, Pawlik TM, Rodriguez-Bigas MA. Clinicopathologic features in colorectal cancer patients with microsatellite instability. Mutat Res 2004;568:275–82.
20. Wang Y, Friedl W, Lamberti C, et al. Hereditary nonpolyposis colorectal cancer: frequent occurrence of large genomic deletions in MSH2 and MLH1 genes. Int J Cancer 2003;103:636–41.
21. Ricciardiello L, Goel A, Mantovani V, et al. Frequent loss of hMLH1 by promoter hypermethylation leads to microsatellite instability in adenomatous polyps of patients with a single first-degree member affected by colon cancer. Cancer Res 2003;63:787–92.
22. Teodoridis JM, Hall J, Marsh S, et al. CpG island methylation of DNA damage response genes in advanced ovarian cancer. Cancer Res 2005;65:8961–7.
23. Stefansson I, Akslen LA, MacDonald N, et al. Loss of hMSH2 and hMSH6 expression is frequent in sporadic endometrial carcinomas with microsatellite instability: a population-based study. Clin Cancer Res 2002;8:138–43.
24. Whelan AJ, Babb S, Mutch DG, et al. MSI in endometrial carcinoma: absence of MLH1 promoter methylation is associated with increased familial risk for cancers. Int J Cancer 2002;99:697–704.
25. Liu W, Saint DA. Validation of a quantitative method for real time PCR kinetics. Biochem Biophys Res Commun 2002;294:347–53.
26. Yu J, Shannon WD, Watson MA, McLeod HL. Gene expression profiling of the irinotecan pathway in colorectal cancer. Clin Cancer Res 2005;11:2053–62.
27. Marsh S, King CR, Garsa AA, McLeod HL. Pyrosequencing of clinically relevant polymorphisms. Methods Mol Biol 2005;311:97–114.
28. Goodfellow PJ, Buttin BM, Herzog TJ, et al. Prevalence of defective DNA mismatch repair and MSH6 mutation in an unselected series of endometrial cancers. Proc Natl Acad Sci U S A 2003;100:5908–13.
29. Lipton LR, Johnson V, Cummings C, et al. Refining the Amsterdam Criteria and Bethesda Guidelines: testing algorithms for the prediction of mismatch repair mutation status in the familial cancer clinic. J Clin Oncol 2004;22:4934–43.
30. Marsh S, Mallon MA, Goodfellow P, McLeod HL. Concordance of pharmacogenetic markers in germline and colorectal tumor DNA. Pharmacogenomics 2005; 6:873–7.
31. Goel A, Arnold CN, Niedzwiecki D, et al. Characterization of sporadic colon cancer by patterns of genomic instability. Cancer Res 2003;63:1608–14.
32. Hatch SB, Lightfoot HM, Jr., Garwacki CP, et al. Microsatellite instability testing in colorectal carcinoma: choice of markers affects sensitivity of detection of mismatch repair-deficient tumors. Clin Cancer Res 2005;11:2180–7.
33. Caldes T, Godino J, Sanchez A, et al. Immunohistochemistry and microsatellite instability testing for selecting MLH1, MSH2 and MSH6 mutation carriers in hereditary non-polyposis colorectal cancer. Oncol Rep 2004;12:621–9.
34. Adimoolam S, Ford JM. p53 and DNA damage-inducible expression of the xeroderma pigmentosum group C gene. Proc Natl Acad Sci U S A 2002;99: 12985–90.
35. Bond GL, Hu W, Bond EE, et al. A single nucleotide polymorphism in the MDM2 promoter attenuates the p53 tumor suppressor pathway and accelerates tumor formation in humans. Cell 2004;119:591–602.
36. Watters JW, McLeod HL. Cancer pharmacogenomics: current and future applications. Biochim Biophys Acta 2003;1603:99–111.
37. Wang Y, Jatkoe T, Zhang Y, et al. Gene expression profiles and molecular markers to predict recurrence of Dukes' B colon cancer. J Clin Oncol 2004;22: 1564–71.
38. van 't Veer LJ, Dai H, van de Vijver MJ, et al. Gene expression profiling predicts clinical outcome of breast cancer. Nature 2002;415:530–6.
39. Murakami T, Fujimoto M, Ohtsuki M, Nakagawa H. Expression profiling of cancer-related genes in human keratinocytes following non-lethal ultraviolet B irradiation. J Dermatol Sci 2001;27:121–9.
40. Welsh C, Day R, McGurk C, Masters JR, Wood RD, Koberle B. Reduced levels of XPA, ERCC1 and XPF DNA repair proteins in testis tumor cell lines. Int J Cancer 2004;110:352–61.
41. Raymond E, Faivre S, Chaney S, Woynarowski J, Cvitkovic E. Cellular and molecular pharmacology of oxaliplatin. Mol Cancer Ther 2002;1:227–35.
42. Faivre S, Chan D, Salinas R, Woynarowska B, Woynarowski JM. DNA strand breaks and apoptosis induced by oxaliplatin in cancer cells. Biochem Pharmacol 2003;66:225–37.

Clinical Cancer Research

DNA Repair Pathway Profiling and Microsatellite Instability in Colorectal Cancer

Jinsheng Yu, Mary A. Mallon, Wanghai Zhang, et al.

Clin Cancer Res 2006;12:5104-5111.