DNA Mismatch Repair Gene Polymorphisms Affect Survival in Pancreatic Cancer

XIAOQUN DONG,^a YANAN LI,^a KENNETH R. HESS,^b JAMES L. ABBRUZZESE,^a DONGHUI LI^a

Departments of ^aGastrointestinal Medical Oncology and ^bBiostatistics, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA

Key Words. DNA mismatch repair • Pancreatic cancer • Single-nucleotide polymorphism • Survival

Disclosures: Xiaoqun Dong: None; Yanan Li: None; Kenneth R. Hess: None; James L. Abbruzzese: None; Donghui Li: None.

The content of this article has been reviewed by independent peer reviewers to ensure that it is balanced, objective, and free from commercial bias. No financial relationships relevant to the content of this article have been disclosed by the authors or independent peer reviewers.

ABSTRACT

Purpose. DNA mismatch repair (MMR) maintains genomic stability and mediates cellular response to DNA damage. We aim to demonstrate whether MMR genetic variants affect overall survival (OS) in pancreatic cancer.

Materials and Methods. Using the Sequenom method in genomic DNA, we retrospectively genotyped 102 single-nucleotide polymorphisms (SNPs) of 13 MMR genes from 706 patients with pancreatic adenocarcinoma seen at The University of Texas MD Anderson Cancer Center. Association between genotype and OS was evaluated using multivariable Cox proportional hazard regression models.

Results. At a false discovery rate of 1% ($p \leq .0015$), 15 SNPs of *EXO1*, *MLH1*, *MSH2*, *MSH3*, *MSH6*, *PMS2*, *PMS2L3*, *TP73*, and *TREX1* in patients with localized disease ($n = 333$) and 6 SNPs of *MSH3*, *MSH6*, and *TP73*

in patients with locally advanced or metastatic disease ($n = 373$) were significantly associated with OS. In multivariable Cox proportional hazard regression models, SNPs of *EXO1*, *MSH2*, *MSH3*, *PMS2L3*, and *TP73* in patients with localized disease, *MSH2*, *MSH3*, *MSH6*, and *TP73* in patients with locally advanced or metastatic disease, and *EXO1*, *MGMT*, *MSH2*, *MSH3*, *MSH6*, *PMS2L3*, and *TP73* in all patients remained significant predictors for OS ($p \leq .0015$) after adjusting for all clinical predictors and all SNPs with $p \leq .0015$ in single-locus analysis. Sixteen haplotypes of *EXO1*, *MLH1*, *MSH2*, *MSH3*, *MSH6*, *PMS2*, *PMS2L3*, *RECQL*, *TP73*, and *TREX1* significantly correlated with OS in all patients ($p \leq .001$).

Conclusion. MMR gene variants may have potential value as prognostic markers for OS in pancreatic cancer patients. *The Oncologist* 2011;16:61–70

INTRODUCTION

Pancreatic cancer is the fourth leading cause of cancer mortality in the United States [1]. Despite advances in the field, pancreatic cancer is one of the most aggressive and therapy-

resistant types of cancer, with the lowest 5-year survival rate of all types of malignancies [1]. There is an urgent need for novel therapeutic strategies and prognostic and predictive biomarkers to better treat this deadly disease.

Correspondence: Donghui Li, Ph.D., Department of Gastrointestinal Medical Oncology, Unit 426, The University of Texas MD Anderson Cancer Center, 1550 Holcombe Boulevard, Houston, Texas 77030-4009. Telephone: 713-834-6690; Fax: 713-834-6153; e-mail: dli@mdanderson.org Received April 27, 2010; accepted for publication November 28, 2010; first published online in *The Oncologist Express* on January 6, 2011. ©AlphaMed Press 1083-7159/2011/\$30.00/0 doi: 10.1634/theoncologist.2010-0127

Figure 1. Selected gene-encoded proteins involved in human mismatch repair (MMR) system. MMR machinery consists of human homologs of MutS α , MutS β , MutL α , MutL β , MutL γ , exonuclease (EXO1, TREX1), DNA polymerase δ (Pol δ), and DNA ligase I [9, 11]. The mismatch-bound MutS α or MutS β recruited MutL α (MutL β /MutL γ). The ternary complex activated EXO1 and TREX1 to commence the degradation of the strand in a 5'→3' direction. Pol δ filled the gap, and DNA ligase I sealed the remaining nick to complete the repair process. MMR eliminated damaged cells via apoptosis network. MutS/MutL/exonuclease can transduce DNA damage signaling to the cell-cycle checkpoint machinery ataxia-telangiectasia mutated and rad3-related (ATR)/checkpoint kinase 1 (CHK1) or ataxia telangiectasia mutated (ATM)/checkpoint kinase 2 (CHK2) directly, which leads to cell-cycle arrest and apoptosis [8, 12]. CHK1 and CHK2 regulate TP73 induction to initiate p53-independent apoptosis [7]. MGMT repairs the O(6)-alkyl-guanine mismatches dependent on MutS α /MutL α /EXO1 [13, 14]. MutS α stimulates RECQL helicase activity; RECQL stimulates EXO1 incision activity and interacts with MLH1 [16].

We have shown previously that genetic variations in drug metabolism and DNA repair correlate with overall survival (OS) of patients with pancreatic cancer [2–5]. However, most of our studies were limited by small sample sizes, small numbers of genes examined, and lack of validation and unknown functional significance of the genetic variants [2–5]. Thus, the potential value of our previous findings in clinical application is not yet known. To confirm and expand on our previous findings, our current study focused on one of the most important DNA repair pathways, mismatch repair (MMR), in a large cohort of 706 patients with pancreatic cancer.

MMR is critical to maintaining genome stability [6]. An MMR deficiency may confer genome-wide instability and contribute to aggressive tumor phenotypes by accumulating genetic alterations [6]. We evaluated 13 genes (Fig. 1), which encode core components involved in the DNA mismatch recognition and removal such as *mutS homolog 2* (MSH2), *MSH3*, *MSH6*, *mutL homolog 1* (MLH1), *MLH3*, *postmeiotic segregation increased 1* (PMS1), *PMS2*, *PMS2-like 3* (PMS2L3), *exonuclease I* (EXO1), *three prime repair exonuclease 1* (TREX1), and *O-6-methylguanine-DNA methyltransferase* (MGMT) and MMR regulators such as *RecQ protein-like/DNA helicase Q1-like* (RECQL) or MMR-induced apoptosis transducer such as *tumor protein 73* (TP73) [7, 8].

The proteins encoded by the genes examined in this study play various roles in the MMR process. Human

MutS α homolog (MSH2/MSH6) recognizes base-base mismatches and small insertion/deletion loops (IDLs) [6, 9]. The MutS β homolog MSH2/MSH3 also recognizes IDLs [6, 9]. MutL α homolog MLH1/PMS2 forms a ternary complex with mismatched DNA and MutS α , increases discrimination between heteroduplexes and homoduplexes, and functions in meiotic recombination [6]. Cells deficient in either MutL or MutS homologs, for example, MLH1 and MSH2, exhibit mutator phenotype and microsatellite instability [6]. PMS2L3, a member of the PMS2 family, contains motifs (e.g., KELVEN) that are conserved among MutL proteins [10]. *PMS2L3* encodes a polypeptide similar to PMS2 in exons 2–3 due to frame shift caused by a 10-nucleotide insertion [10]. PMS2L3 codons 180–255 are 97% identical to PMS2 [10]. The MutL β homolog MLH1/PMS1 and the MutL γ homolog MLH1/MLH3 primarily function in meiotic recombination and serve as a backup mechanism for MutL α in repairing base-base mismatches/small IDLs [6]. EXO1 encodes 5'–3' exonuclease which is involved in MMR and recombination by interacting with MSH2 [9, 11]. TREX1 encodes 3'–5' DNA exonuclease that proofreads for DNA polymerase and degrades 3' ends of nicked DNA in granzyme A-mediated cell death [12]. MGMT plays a crucial role in the defense against alkylating agents by interacting with MutS α , MutL α , and EXO1 [13, 14]. MGMT downregulation activates p53 and inhibits pancreatic cancer tumor growth [15]. RECQL is a member of the RecQ DNA helicase family that is required for DNA

replication, recombination, and repair [16, 17]. TP73, a member of the p53 family, is involved in cellular responses to stress and development [18]. MMR-mediated p53-independent apoptosis is stimulated by MLH1/c-Abl/p73 α /GADD45 α retrograde signaling [19].

In the current study, we examined 102 single-nucleotide polymorphisms of the above-mentioned 13 MMR genes in 706 patients with pancreatic cancer and demonstrated their associations with overall survival. Our previous findings on the associations of some of the MMR genes including *RECQL* with the clinical outcome of pancreatic cancer [2–4] are confirmed in this larger data set and new findings on additional genes are also reported.

MATERIALS AND METHODS

Patient Population and Clinical Data

The 706 patients involved in this study include 154 patients with potentially resectable tumors who had been enrolled in phase II clinical trials for preoperative chemoradiation [4] and 552 patients who had been recruited to a case-control study of pancreatic cancer conducted at MD Anderson during February 1999 through May 2007. The patient inclusion criteria are (1) having been diagnosed with a pathologically confirmed pancreatic ductal adenocarcinoma and (2) having a blood or tissue sample for DNA extraction. Each patient had signed an informed consent for medical record review and DNA sample donation. The institutional review board at MD Anderson Cancer Center had approved the previous studies, which had been conducted according to all current ethical guidelines.

We reviewed the 706 patients' medical records and extracted the following types of clinical information: dates of diagnosis and death or last follow-up; tumor stage, resection status, differentiation, size, and site; performance status; and serum markers for liver, kidney, and pancreas functions as well as serum carbohydrate antigen 19-9 (CA19-9) level at diagnosis. Clinical tumor staging at the time of diagnosis had followed the objective computed tomography (CT) criteria: a localized or potentially resectable tumor had been defined as a tumor (1) without evidence of extrapancreatic disease (extensive peripancreatic lymph node involvement), (2) without invasion of the celiac axis or superior mesenteric artery, inferior vena cava, or aorta, or (3) without encasement or occlusion of the superior mesenteric vein–portal vein confluence [20]. Tumor that abutted and encased the superior mesenteric vein, without vessel occlusion or extension to the superior mesenteric artery, had been considered resectable [20]. Locally advanced tumors had been defined as those that were unre-

sectable but without distant metastasis. Tumor margin and lymph node status had been evaluated only in patients with resected tumors. Dates of death were obtained from the MD Anderson Cancer Center Tumor Registry, inpatient electronic medical records, or the United States Social Security Death Index (www.deathindexes.com/ssdi.html).

DNA Extraction, Single-Nucleotide Polymorphism (SNP) Selection, and Genotyping

For this retrospective study, DNA samples of 706 patients for genotyping were available from the original studies [4, 21]. Using Qiagen DNA isolation kits (Qiagen, Valencia, CA), we extracted DNA from archived peripheral blood lymphocytes samples of 679 patients. Blood samples were not collected in 27 patients with resected tumors so archived paraffin sections of normal tissues adjacent to tumors were used for DNA extraction in these patients. Tissue sections were obtained from our Institutional Tissue Bank and normal tissues were identified by certified pathologists. Normal and tumor tissues were expected to have the same genotype for these germline common polymorphic sequence variants. We selected 45 tagging SNPs using the SNPbrowser (Applied Biosystems, Carlsbad, CA) with a cutoff of $r^2 = 0.8$ and a MAF $\geq 10\%$ in White from the HapMap Project database (www.hapmap.org). For more comprehensive coverage of the gene variants, we included 57 SNPs in the coding region (nonsynonymous or synonymous), untranslated region (UTR), promoter region, or splicing sites that have a MAF $\geq 1\%$ in White. The genes, nucleotide substitutions, functions (such as encoding amino acid changes), reference SNP identification numbers, and reported allele frequencies of the 102 SNPs evaluated in this study are summarized in supplemental online Table 1. The protein sequences, structures, homology models, mRNA transcripts, and predicted functions for the SNPs were evaluated using F-SNP software (Queen's University, Kingston, Ontario, Canada) [22]. For genotyping, we used the mass spectroscopy–based MassArray method (Sequenom, Inc., San Diego). Twenty percent of the DNA samples were genotyped in duplicate, with 99.6% concordance. Inconsistent genotyping calling data were excluded from our final analysis.

Statistical Analysis

We examined the distribution of genotypes for Hardy-Weinberg equilibrium using the goodness-of-fit χ^2 test. We calculated genotype frequency and allele frequency of the SNPs by direct gene counting. We calculated haplotype diversity and linkage disequilibrium index (Lewontin's $|D'|$ and r^2) using SNPalyze (DYNACOM Co., Ltd. Mobara, Japan). Haplotypes were reconstructed by implementing

the EM algorithm using the unphased genotype data in the three major ethnic/racial groups, that is, non-Hispanic white, Hispanics, and blacks. We computed the median follow-up time using censored observations only. We estimated the effect of genotype/haplotype on OS using Cox proportional hazards regression model. We calculated hazard ratios (HR) and 95% confidence interval (CI) in the multivariate Cox proportional hazards regression model with adjustments for sex, age, race, and any significant clinical predictors for OS. OS was calculated from the date of diagnosis to the date of death or last follow-up.

Genotypes were combined in some analyses. The homozygous and heterozygous variants were combined if (1) the homozygous variant was present in ≤ 5 patients and (2) if the genotype effect on survival is dominant, that is, both the heterozygote and homozygote had $p < .05$ in log-rank test compared to the wild type. The heterozygote was combined with the wild type if the genotype effect was recessive; that is, only the homozygote was significantly different from the wild type in survival. We performed statistical testing using SPSS (SPSS Inc., Chicago) and Stata (StataCorp LP, College Station, TX). We calculated the Harrell's C index (concordance) to test the proportional-hazards assumption. We calculated the false discovery rate (FDR) using the beta-uniform mixture method [23]. For 200 comparisons in 102 SNPs for OS in 706 patients, we found that p -values of .0015 and .028 corresponded with FDRs of 1% and 5%, respectively. For the genotype analysis in our study, $p \leq .0015$ was considered statistically significant.

RESULTS

Patients' Characteristics

The patients' characteristics and clinical predictors of OS for the study patients are described in Table 1. The 706 patients included 333 (154 on and 179 off clinical trials) with localized disease, 211 with locally advanced disease, and 162 with metastatic disease. Of the 333 patients with localized tumors, 275 (83%) had undergone tumor resection. Of the 706 patients, 138 (19.5%) were still alive at the end of the original study period, with a median follow-up time of 46.0 months. The median survival duration (MSD) for the entire patient population was 17.2 months (95% CI, 15.8–18.5). Advanced tumor stage, unresected tumor, an elevated CA19-9 serum level or biochemical index, or poor performance status were significant predictors for poor OS in multivariate Cox proportional hazards regression model (data not shown).

Genotype Distribution and Allele Frequencies

The observed minor allele frequencies of the 102 tested SNPs in our study population were comparable with those reported in the general population (supplemental online Table 1). Linkage disequilibrium data of the 102 SNPs in the entire study population and in non-Hispanic whites are summarized in supplemental online Tables 2 and 3, respectively. Most of the 102 tested SNPs followed the Hardy-Weinberg equilibrium (HWE), except for *MGMT* IVS4-44836G>A ($p < .001$), *MLH1* I219V ($p = .003$), *MLH3* P844L ($p = .0054$), *MSH2* IVS1+9A>T ($p = .0035$), IVS9-9A>T ($p < .001$), IVS11-62G>A ($p < .001$), *PMS2* P470S ($p < .001$), N775S ($p < .001$), *PMS2L3* IVS3+9A>G ($p < .001$), *RECQL* IVS11+218T>G ($p = .0013$), IVS8+190A>G ($p < .001$), and IVS4-795C>T ($p = .002$).

Associations of Genotype with Overall Survival

First, we analyzed genotype association with OS in 333 patients with localized disease and 373 patients with locally advanced or metastatic disease separately. Fifteen SNPs of nine genes (i.e., *EXO1*, *MLH1*, *MSH2*, *MSH3*, *MSH6*, *PMS2*, *PMS2L3*, *TP73*, and *TREX1*) were significantly associated with OS at 1% FDR level ($p \leq .0015$) after adjustment for clinical factors in patients with localized disease (Table 2). Six SNPs of *MSH2*, *MSH3*, *MSH6*, and *TP73* were significantly associated with OS in patients with locally advanced or metastatic disease. When data of the 706 patients were pooled to increase power, 16 SNPs were significantly associated with OS. The Harrell's C indexes (concordance) were all >0.61 , indicating the data did not violate the proportionality assumption (data not shown). To reduce population stratification bias, we performed additional data analysis in non-Hispanic only and the results are quite comparable to those in the entire study population. None of the risk estimates had $>10\%$ differences between the two study populations (supplemental online Table 4). Thus, data from the entire study population are presented in this report.

Next, we conducted multivariable Cox proportional hazards regression analysis, including all the SNPs that reached 1% FDR level and all significant clinical predictors. *EXO1*, *MSH2*, *MSH3*, *PMS2L3*, and *TP73* in patients with localized disease; *MSH2*, *MSH3*, *MSH6*, and *TP73* in patients with locally advanced or metastatic disease; and *EXO1*, *MGMT*, *MSH2*, *MSH3*, *MSH6*, *PMS2L3*, and *TP73* in all patients remained significant predictors (Table 3). Results from the same analyses conducted in non-Hispanic whites are summarized in supplemental online Table 5.

Table 1. Patients' characteristics and clinical predictors for overall survival

Variable	No. of patients	No. of deaths	MST, mo(s)	$p_{\log\text{-rank}}$
Gender				.48
Female	281	235	17.5	
Male	425	333	16.8	
Age (years)				.22
≤50	90	80	16.4	
51–60	181	142	18.4	
61–70	272	213	17.2	
>70	163	133	16.5	
Race				.63
White	624	502	16.6	
Hispanic	43	35	16.8	
Black	27	21	20.2	
Other	12	10	17.3	
Performance status				<.001
0	148	106	31.0	
1	483	394	16.0	
2–3	75	68	9.5	
Clinical disease				<.001
Localized	333	237	28.5	
Locally advanced	211	188	14.7	
Metastatic	162	143	9.2	
Tumor size (cm)				<.001
≤2	134	94	27.5	
>2	572	474	15.3	
Tumor site				.001
Head	511	411	18.2	
Nonhead	195	157	13.7	
CA19-9 level (U/ml)				<.001
≤47	161	107	32.1	
48–500	291	237	18.0	
500–1,000	73	60	14.5	
>1,000	181	164	10.6	
Tumor differentiation ^a				<.001
Well to moderate	355	261	26.2	
Poorly	139	120	12.0	
Tumor response to therapy ^b				<.001
PR/SD	225	146	35.8	
PD	36	35	9.3	
Tumor resection				<.001
Yes	275	179	35.9	
No	431	389	12.0	

(continued)

Table 1. (Continued)

Variable	No. of patients	No. of deaths	MST, mo(s)	$p_{\log\text{-rank}}$
Biochemical index ^c				.008
0–2	303	230	18.9	
3–6	362	302	16.0	
7–9	41	36	12.9	

^aThis information was unavailable for patients without proper histological samples.^bTumor response to therapy was evaluated in patients who received preoperative chemoradiotherapy only.^cBiochemical index represents the number of serum markers with abnormal value. The markers include aspartate aminotransferase, lactic dehydrogenase, alkaline phosphatase, alanine aminotransferase, amylase, creatinine, hemoglobin, albumin, bilirubin, and fasting glucose.

Abbreviations: CA19-9, carbohydrate antigen 19-9; CI, confidence interval; HR, hazard ratio; MST, median survival time; PD, progressive disease; PR, partial response; SD, stable disease.

Associations of Haplotype Diversity with OS

Haplotype distribution in different ethnic/racial groups is presented in supplemental online Table 6. Because of the large variation in haplotype distribution between ethnic/racial groups and the small number of minorities, haplotype associations with OS were analyzed in non-Hispanic whites as well as in the entire study population with justification for race (Table 4). Sixteen haplotypes of *EXO1*, *MLH1*, *MSH2*, *MSH3*, *MSH6*, *PMS2*, *PMS2L3*, *RECQL*, *TP73*, and *TREX1* genes were significantly associated with OS ($p \leq .0015$). The associations of haplotypes with OS manifested the effects of *EXO1* N279S GA; *MLH1*-92G>A AA; *MSH2* IVS11-62G>A GA/AA and -117T>C CT/CC; *MSH3* P231P AA, T1045A GG, and Q949R GA/AA; *MSH6* D180D CT/CC and P92P GA/GG; *PMS2* IVS1-1121C>T TC/TT; *PMS2L3* IVS3+9A>G GA/GG and IVS1-8C>T CT; *CAATTACCTT* of *RECQL* IVS2+1222T>C, *1236A>G, IVS11+582T>A, IVS11+218T>G, IVS6-1031T>C, IVS5-717C>A, IVS4-795C>T, IVS4+1179T>C, IVS1-92C>T, and IVS1-570T>G; *TP73* Ex2+4G>A AA and *1454C>T CC; *TREX1* Ex14+297A>G AG and Ex14+782T>C CT/CC genotypes. Although none of the individual SNPs reached the 1% FDR significance level for *RECQL*, the haplotype *CAATTACCTT* did reach that level ($p < .001$) when all study participants were included in the analysis (Table 4).

DISCUSSION

Results from the current study confirmed our previous observations of an association between MMR gene variants

Table 2. Association of genotype with overall survival (hazard ratios and 95% confidence interval)

Function	Genotype	Localized		Locally advanced + metastatic		All patients	
		HR (95% CI) ^a	<i>p</i>	HR (95% CI)	<i>p</i>	HR (95% CI)	<i>p</i>
Recognition							
O(6)-alkyl-guanine	<i>MGMT</i> IVS4-44836G>A (GG vs. GA/AA)	1.46 (1.08–1.98)	.007	1.36 (0.92–2.01)	.13	1.42 (1.18–1.71)	<.001
MutSβ	<i>MSH2</i> G322D (GG vs. GA/AA)	1.97 (1.30–2.98)	.001	1.38 (0.96–1.98)	.08	1.57 (1.20–2.06)	<.001
	<i>MSH2</i> IVS12-6T>C (TT/TC vs. CC)	5.68 (2.73–11.8)	<.001	9.94 (4.65–21.2)	<.001	6.97 (4.15–11.7)	<.001
	<i>MSH3</i> Ex24-318G>A (GG vs. GA/AA)	1.82 (1.36–2.42)	<.001	1.28 (1.00–1.49)	.05	1.41 (1.18–1.69)	<.001
	<i>MSH3</i> IVS10-3448A>G (AA/GG vs. GA)	1.71 (1.30–2.25)	<.001	1.00 (0.82–1.19)	.91	1.03 (0.87–1.22)	.74
	<i>MSH3</i> IVS23+437G>A (AA/GA vs. GG)	1.69 (1.27–2.24)	<.001	2.06 (1.43–2.97)	<.001	1.80 (1.38–2.35)	<.001
	<i>MSH3</i> T1045A (AA/GA vs. GG)	0.75 (0.48–1.17)	.20	2.02 (1.38–2.95)	<.001	1.25 (0.94–1.66)	.13
MutSα	<i>MSH6</i> P92P (AA vs. GA/GG)	0.63 (0.47–0.84)	.001	0.67 (0.53–0.86)	.001	0.67 (0.56–0.81)	<.001
	<i>MSH6</i> D180D (TT vs. CT/CC)	0.64 (0.49–0.84)	.001	0.74 (0.59–0.93)	.009	0.71 (0.60–0.84)	<.001
MutLα	<i>MLH1</i> IVS14-19A>G (AA vs. AG/GG)	1.66 (1.23–2.23)	<.001	1.29 (0.99–1.66)	.056	1.40 (1.15–1.70)	<.001
	<i>PMS2</i> N775S (AA/AG vs. GG)	0.44 (0.28–0.70)	<.001	0.83 (0.65–1.05)	.12	0.63 (0.47–0.84)	.001
	<i>PMS2L3</i> IVS3+9A>G (AA vs. GA/GG)	0.50 (0.38–0.67)	<.001	0.83 (0.66–1.05)	.12	0.69 (0.58–0.83)	<.001
Excision							
exonuclease	<i>EXO1</i> N279S (AA vs. GA)	2.36 (1.61–3.45)	<.001	1.36 (0.92–2.01)	.13	1.69 (1.28–2.22)	<.001
	<i>TREX1</i> Ex14+782T>C (TT vs. CC/CT)	0.61 (0.46–0.82)	<.001	0.52 (0.21–1.28)	.15	0.76 (0.64–0.92)	.001
Apoptosis							
TP73	<i>TP73</i> Ex2+3C>T (CC vs. CT)	4.42 (2.11–9.26)	<.001	1.07 (0.64–1.77)	.80	1.37 (0.92–2.07)	.10
	<i>TP73</i> Ex2+4G>A (GG/GA vs. AA)	2.98 (1.89–4.71)	<.001	2.07 (1.29–3.32)	.001	2.49 (1.81–3.41)	<.001
	<i>TP73</i> Ex2+14C>T (CC/TC vs. TT)	1.83 (1.09–3.07)	.02	1.98 (1.22–3.23)	.006	1.90 (1.30–2.77)	.001
	<i>TP73</i> T173T (CC vs. TC/TT)	1.46 (1.08–1.97)	<.001	4.50 (2.37–8.56)	<.001	2.66 (1.45–4.86)	<.001
	<i>TP73</i> A610A (GG vs. GA)	0.65 (0.43–0.99)	.047	0.62 (0.40–0.95)	.03	0.64 (0.47–0.86)	.0015
^a HR and 95% CI were from multivariable Cox regression models including sex, age, race, tumor resection, CA19-9, performance status, biochemical index, and stage, when appropriate. Abbreviations: HR, hazard ratio; CI, confidence interval; CA19-9, carbohydrate antigen 19-9.							

^aHR and 95% CI were from multivariable Cox regression models including sex, age, race, tumor resection, CA19-9, performance status, biochemical index, and stage, when appropriate. Abbreviations: HR, hazard ratio; CI, confidence interval; CA19-9, carbohydrate antigen 19-9.

and OS in pancreatic cancer. At the FDR level of 1%, 10 of the 102 tested SNPs from 7 of the 13 MMR genes examined (i.e., *EXO1*, *MGMT*, *MSH2*, *MSH3*, *MSH6*, *PMS2L3*, and *TP73*) were significant independent predictors for OS in multi-SNP Cox proportional hazards regression models with adjustment for clinical predictors. These findings support a role for MMR polymorphisms in modifying the tumor phenotype and clinical outcome of pancreatic cancer.

MMR is a complicated network with multiple functions, one being to maintain genome stability by removing mismatched or distorted DNA structures [6]. However, MMR also plays an important role in apoptosis signaling when cells are overwhelmed with genotoxic insults [6]. Thus, MMR deficiency contributes to both genomic instability and altered DNA damage response [6]. As MMR machinery is crucial in maintaining genomic stability, our finding indicates that dysfunctional MMR may contribute to rapid

tumor progression via accumulated DNA damage and failed activation of apoptosis signaling pathways.

Although MMR deficiency has not been recognized as a major genetic alteration in pancreatic tumors [24], individuals with hereditary Lynch syndrome (an autosomal dominant genetic disorder caused by germline mutations in MMR genes) also have increased risk of pancreatic cancer [25]. The strong associations between MMR genetic variation and OS observed in our study add supporting evidence to the role of DNA MMR deficiency in pancreatic cancer progression.

We found that genetic variations in *EXO1*, *MGMT*, *MLH1*, *MSH2*, *MSH3*, *PMS2*, *PMS2L3*, and *TREX1* correlated with OS mainly in patients with localized disease; *MSH2*, *MSH3*, *MSH6*, and *TP73* correlated with OS mainly in those with locally advanced or metastatic disease; and *MSH2*, *MSH3*, *MSH6*, and *TP73* correlated with OS in all

Table 3. Predictors of overall survival in multivariate Cox regression models (hazard ratio and 95% confidence interval)

		Localized (<i>n</i> = 333)		Locally advanced + metastatic (<i>n</i> = 373)		All patients (<i>n</i> = 706)	
Function	Variable, single SNP	HR (95% CI) ^a	<i>p</i>	HR (95% CI)	<i>p</i>	HR (95% CI)	<i>p</i>
Recognition							
O(6)-alkyl-guanine	<i>MGMT</i> IVS4-44836G>A	1.40 (0.96–1.81)	.08	1.28 (0.93–1.96)	.17	1.42 (1.16–1.74)	<.001
MutSβ	<i>MSH2</i> IVS12-6T>C	5.70 (2.60–13.0)	<.001	9.94 (3.00–33.0)	<.001	4.06 (2.12–7.79)	<.001
	<i>MSH3</i> T1045A	0.89 (0.54–1.44)	.62	2.08 (1.32–3.29)	<.001	1.23 (0.89–1.69)	.21
	<i>MSH3</i> IVS10-3448A>G	1.85 (1.37–2.51)	<.001	1.01 (0.75–1.35)	.94	1.01 (0.84–1.21)	.94
	<i>MSH3</i> IVS23+437G>A	1.38 (1.01–1.89)	.04	1.94 (1.28–2.93)	.001	1.74 (1.30–2.33)	<.001
MutSα	<i>MSH6</i> P92P	0.71 (0.52–0.98)	.038	0.68 (0.55–0.85)	<.001	0.73 (0.61–0.89)	.001
MutLα	<i>PMS2</i> N775S	0.48 (0.29–0.82)	.006	0.83 (0.56–1.23)	.36	0.69 (0.51–0.93)	.015
	<i>PMS2L3</i> IVS3+9A>G	0.50 (0.37–0.69)	<.001	0.80 (0.62–1.03)	.087	0.65 (0.54–0.78)	<.001
Excision							
Exonuclease	<i>EXO1</i> N279S	2.02 (1.34–3.04)	<.001	1.27 (0.80–1.84)	.28	1.61 (1.20–2.15)	.001
	<i>TREX1</i> Ex14+782T>C	0.79 (0.57–1.10)	.17	0.84 (0.66–1.08)	.18	0.77 (0.66–0.93)	.006
TP73	Apoptosis						
	<i>TP73</i> Ex2+3C>T	5.10 (2.37–11.0)	<.001	1.04 (0.57–1.91)	.90	1.54 (0.95–2.47)	.10
	<i>TP73</i> Ex2+4G>A	4.00 (1.86–8.57)	<.001	1.93 (1.20–3.12)	.007	1.81 (1.47–2.31)	.004
	<i>TP73</i> Ex2+14C>T	2.13 (0.91–4.98)	.08	2.10 (1.19–3.71)	.01	2.21 (1.66–3.67)	.02
	<i>TP73</i> T173T	3.13 (1.38–7.11)	.006	6.41 (3.10–12.8)	<.001	4.24 (2.39–7.54)	<.001
Clinical predictors	Tumor resection (yes vs. no)	5.38 (3.68–7.86)	<.001	4.90 (1.78–13.5)	.001	3.53 (2.68–4.63)	<.001
	CA19-9 level (<1000 vs. ≥1000)	1.22 (1.06–1.41)	.006	1.17 (1.05–1.31)	.005	1.19 (1.09–1.29)	<.001
	Performance status (0 vs. 1–3)	1.80 (1.19–2.73)	<.001	1.91 (1.53–2.39)	<.001	1.86 (1.57–2.20)	<.001

^aHR (95% CI) were from multivariable Cox regression models including sex, age, race, tumor resection, CA19-9, performance status, biochemical index, and stage, when appropriate. All SNPs with *p* ≤ .0015 in single-locus analysis for each patient group were included in the model.

Abbreviations: CA19-9, carbohydrate antigen 19-9; CI, confidence interval; HR, hazard ratio; SNP, single-nucleotide polymorphism.

patients. Although some of the inconsistent observations between patients with different stages of diseases may suggest false positive findings, our previous observation on the association between *EXO1*, *MSH2*, *MSH3*, and *TP73* SNPs with OS in patients with resectable tumors [4] were confirmed in this larger patient population that was under stringent FDR control. In addition, we also observed a significant association between *PMS2L3* IVS3+9A>G and OS in the current study. This intronic SNP is in linkage with a 3'UTR SNP Ex7-32G>A. Bioinformatics has predicted that the latter SNP may confer a protein with altered solvent accessibility [22]. *PMS2L3* has a similar function to that of *PMS2* in MutLα, which is to play a role in maintaining genome integrity via MMR and DNA damage-induced apoptosis [26].

Loss of *MSH3* protein expression has been related to tumor progression in colorectal cancer [27]. Our finding that the *MSH3* genetic variants contributed to shorter OS in pa-

tients with locally advanced or metastatic pancreatic cancer suggests that *MSH3* deficiency may contribute to rapid tumor progression, possibly through accumulation of larger IDLs in patients with late-stage disease.

P73 is a homolog of the tumor suppressor gene *p53*, but its function in tumor development has not been established [28]. *P73* plays an important role in apoptosis and chemosensitivity; thus, it has been recognized as a potential therapeutic target in cancer treatment [29]. Our findings on the association between *P73* SNPs and OS support a role for this protein in pancreatic cancer progression.

We found *RECQL* haplotype was related to shorter OS in the 706 patients with pancreatic cancer. *RECQL* is a member of the RECQ DNA helicase family that is required for DNA replication, recombination, and repair [30]. Germ-line mutations of three of the five members of this protein family have been associated with hereditary syndromes characterized with chromosome instability, premature ag-

Table 4. Association of haplotype with overall survival (hazard ratio and 95% confidence interval)

Haplotype	All patients				Non-Hispanic white
	Frequency	MST	HR (95% CI) ^a	<i>p</i>	HR (95% CI) ^a
<i>EXO1</i>					
CCTTATGG	0.2141	17.6	1.0		1.0
CCTT <u>G</u> TGG	0.024	14.8	1.88 (1.26–2.79)	.001	2.01 (1.36–2.97)
CCTC <u>G</u> G <u>G</u> G	0.0136	7.2	3.40 (2.05–5.64)	<.001	2.41 (1.48–3.92)
<i>MLH1</i>					
TGCG	0.322	16.5	1.0		1.0
T <u>A</u> TG	0.0108	10.5	2.14 (1.31–3.49)	.001	1.96 (1.16–3.32)
<i>MSH2</i>					
TACGT	0.4236	18.2	1.0		1.0
TAC <u>A</u> C	0.0065	10.8	4.34 (1.91–9.82)	<.001	3.57 (1.47–8.70)
<i>MSH3</i>					
GAGGGTAG	0.1779	17.6	1.0		1.0
<u>A</u> AGGATAG	0.004	9.3	8.19 (2.59–25.9)	<.001	8.27 (2.45–26.3)
G <u>G</u> AGATAG	0.003	7.8	5.15 (1.89–14.0)	.001	6.20 (1.93–19.9)
<i>MSH6</i>					
GTCCA	0.1908	17.6	1.0		1.0
G <u>C</u> CC <u>G</u>	0.1658	18.7	0.68 (0.54–0.86)	<.001	0.69 (0.56–0.85)
<i>PMS2</i>					
TGCCGGTA	0.1382	14.4	1.0		1.0
TGCT <u>T</u> GGTA	0.0348	30.4	0.50 (0.31–0.80)	.001	0.47 (0.30–0.75)
<i>PMS2L3</i>					
GAGCAC	0.3663	14.7	1.0		1.0
G <u>G</u> GCAC	0.1923	20.4	0.72 (0.61–0.86)	<.001	0.71 (0.60–0.84)
G <u>G</u> GC <u>A</u> T	0.0191	34.9	0.49 (0.31–0.77)	.001	0.30 (0.15–0.66)
<i>RECQL</i>					
TGTGCCTTCGAACATT	0.2263	18.1	1.0		1.0
<u>CA</u> ATTACCTTGGCATT	0.0141	10.7	4.15 (2.03–8.45)	<.001	3.29 (1.51–1.17) ^b
<u>CA</u> ATTACCTTAA <u>A</u> ATT	0.005	9.0	21.3 (8.24–54.8)	<.001	5.73 (1.39–23.7) ^b
<i>TP73</i>					
GCAGTCCG	0.2724	18.7	1.0		1.0
<u>A</u> CAGTCCG	0.0144	8.3	2.13 (1.30–3.51)	.0015	2.91 (1.45–5.92)
<u>A</u> CAG <u>C</u> CCG	0.006	6.2	4.44 (1.97–10.0)	<.001	3.88 (1.92–9.40)
<i>TREX1</i>					
CAAT	0.3764	15.0	1.0		1.0
CA <u>A</u> C	0.1638	19.8	0.71 (0.59–0.85)	<.001	0.70 (0.57–0.85)
C <u>A</u> GT	0.025	29.9	0.46 (0.28–0.73)	.001	0.47 (0.30–0.74)

^aHR value was from multivariable Cox regression model including sex, age, race, clinical stage, tumor resection, CA19-9, performance status, and biochemical index. Haplotype with the highest frequency was used as the reference group, and those with *p*-values >.0015 in Cox regression models were not listed in the table.

^b*p* > .0015 Cox regression model.

Abbreviations: MST, median survival time (months); HR: hazard ratio; CI, confidence interval; CA19-9, carbohydrate antigen 19-9.

ing, and cancer susceptibility [30]. Although RECQL has not been associated with any hereditary syndromes, accu-

mulating evidence suggests that RECQL is a regulating gene not only for the MMR network but also for cell cycle

checkpoints [31]. RECQL might be a potential therapeutic target in cancer treatment [32].

In our current study, we found several nonsynonymous SNPs (e.g., *MLH1* I219V, *MSH3* T1045A, *PMS2* N775S, *EXO1* N279S, and *TREX1* K125Q) and synonymous SNPs (e.g., *MSH6* P92P and D180D; *TP73* T173T and A610A) to be significantly associated with OS. Although the synonymous SNPs do not lead to amino acid change, some experimental evidence suggests that they could affect protein function [33]. In addition, some 5'UTR SNPs (e.g., *TP73* Ex2+3C>T, Ex2+4G>A, and Ex2+14C>T) have been implicated in protein function modulation. These SNPs may cause frame-shift coding and alternative splicing, or they may affect transcription factor binding [34].

Furthermore, several 3'UTR SNPs (e.g., *MSH3* Ex24-318G>A, *TREX1* Ex14+297A>G, and Ex14+782T>C) in our study were also associated with OS. The 3'UTR SNPs may contain sequence motifs crucial for regulating transcription, mRNA stability, and cellular localization of the mRNA or microRNA binding [35]. A few tagging SNPs in our study (e.g. *MLH1* IVS14-19A>G, *MSH3* IVS10-3448A>G, and IVS23+437G>A) were associated with OS, and they are linked with the nonsynonymous SNP *MLH1* I219V and *MSH3* T1045A. As expected, the minor allele frequencies of the functional SNPs were relatively low, and thus a large sample size was required to demonstrate their phenotypic significance in a human study. Although some tagging SNPs might not be functional, they might be in linkage disequilibrium with other functional variants that are not measured in the current study. Larger studies covering functional SNPs with MAF <5% would be required to fully identify the responsible genetic variants.

The strengths of our study include a large sample size, detailed clinical information, a focused DNA repair pathway in gene selection, and comprehensive SNP coverage. The limitations of our study include the heterogeneous patient population, the potential false-positive findings related to multiple comparisons, and the lack of consideration of treatment effect. To overcome some of these limitations,

we conducted data analyses stratified by disease stage and then applied a stringent FDR control *p*-value. Since it is well known that most treatment regimens do not make a significant impact on OS of patients with pancreatic cancer, we adjusted for almost all known clinical predictors in the Cox proportional hazards regression models.

The SNPs that significantly deviated from the HWE might indicate the following: (1) There are genotyping errors; (2) there are selection bias in patient recruitment; and (3) the variant allele of these SNPs might be associated with risk of pancreatic cancer, so they are either over- or under-represented in the cancer patients compared with the general population. Three SNPs, that is, *MGMT* IVS4-44836G>A, *PMS2* N775S, and *PMS2L3* IVS3+9A>G, that are not in HWE were significantly associated with OS. Although these associations could be false positive, the major results on other SNPs that followed the HWE are unlikely biased by the three SNPs.

Our results identifying the strong association between MMR gene variants and OS support a role for DNA MMR deficiency in pancreatic cancer progression. Because tissue samples are extremely hard to obtain from most patients with pancreatic cancer, genetic markers that are measurable in peripheral blood DNA are clinically preferred and valuable. Additional studies are warranted to establish and validate a genetic risk-prediction model for pancreatic cancer. Such studies could identify some genes as potential therapeutic targets in the treatment of this deadly disease.

ACKNOWLEDGMENTS

We are grateful for support from the National Institutes of Health (NIH) R01 grant CA098380 (to D.L.), SPORE P20 grant CA101936 (to J.L.A.), NIH Cancer Center Core grant CA16672, and a research grant from the Lockton Research Funds (to D.L.).

AUTHOR CONTRIBUTIONS

Conception/Design: Donghui Li, Yanan Li
Provision of study material or patients: Donghui Li, James L. Abbruzzese
Collection and/or assembly of data: Xiaqun Dong, Yanan Li
Data analysis and interpretation: Donghui Li, Xiaqun Dong, Kenneth R. Hess
Manuscript writing: Donghui Li, Xiaqun Dong
Final approval of manuscript: Donghui Li, James L. Abbruzzese

REFERENCES

- Jemal A, Siegel R, Ward E et al. Cancer statistics, 2009. *CA Cancer J Clin* 2009;59:225–249.
- Li D, Frazier M, Evans DB et al. Single nucleotide polymorphisms of RecQ1, RAD54L, and ATM genes are associated with reduced survival of pancreatic cancer. *J Clin Oncol* 2006;24:1720–1728.
- Li D, Liu H, Jiao L et al. Significant effect of homologous recombination DNA repair gene polymorphisms on pancreatic cancer survival. *Cancer Res* 2006;66:3323–3330.
- Dong X, Jiao L, Li Y et al. Significant associations of mismatch repair gene polymorphisms with clinical outcome of pancreatic cancer. *J Clin Oncol* 2009;27:1592–1529.
- Okazaki T, Javle M, Tanaka M et al. Single nucleotide polymorphisms of gemcitabine metabolic genes and pancreatic cancer survival and drug toxicity. *Clin Cancer Res* 2010;16:320–329.
- Jiricny J. The multifaceted mismatch-repair system. *Nat Rev Mol Cell Biol* 2006;7:335–346.
- Urist M, Tanaka T, Poyurovsky MV et al. p73 induction after DNA damage

- is regulated by checkpoint kinases Chk1 and Chk2. *Genes Dev* 2004;18:3041–3054.
- 8 Morgan MA, Parsels LA, Parsels JD et al. Role of checkpoint kinase 1 in preventing premature mitosis in response to gemcitabine. *Cancer Res* 2005;65:6835–6842.
 - 9 Constantin N, Dzantiev L, Kadyrov FA et al. Human mismatch repair: reconstitution of a nick-directed bidirectional reaction. *J Biol Chem* 2005;280:39752–39761.
 - 10 Nicolaides NC, Carter KC, Shell BK et al. Genomic organization of the human PMS2 gene family. *Genomics* 1995;30:195–206.
 - 11 Zhang Y, Yuan F, Presnell SR et al. Reconstitution of 5'-directed human mismatch repair in a purified system. *Cell* 2005;122:693–705.
 - 12 Yang YG, Lindahl T, Barnes DE. Trex1 exonuclease degrades ssDNA to prevent chronic checkpoint activation and autoimmune disease. *Cell* 2007;131:873–886.
 - 13 Klapacz J, Meira LB, Luchetti DG et al. O6-methylguanine-induced cell death involves exonuclease 1 as well as DNA mismatch recognition in vivo. *Proc Natl Acad Sci U S A* 2009;106:576–581.
 - 14 Kaina B, Christmann M, Naumann S et al. MGMT: key node in the battle against genotoxicity, carcinogenicity and apoptosis induced by alkylating agents. *DNA Repair* 2007;6:1079–1099.
 - 15 Konduri SD, Ticku J, Bobustuc GC et al. Blockade of MGMT expression by O6 benzyl guanine leads to inhibition of pancreatic cancer growth and induction of apoptosis. *Clin Cancer Res* 2009;15:6087–6095.
 - 16 Doherty KM, Sharma S, Uzdilla LA et al. RECQ1 helicase interacts with human mismatch repair factors that regulate genetic recombination. *J Biol Chem* 2005;280:28085–28094.
 - 17 Sharma S, Stumpo DJ, Balajee AS et al. RECQL, a member of the RecQ family of DNA helicases, suppresses chromosomal instability. *Mol Cell Biol* 2007;27:1784–1794.
 - 18 Stiewe T. The p53 family in differentiation and tumorigenesis. *Nat Rev Cancer* 2007;7:165–168.
 - 19 Li LS, Morales JC, Hwang A et al. DNA mismatch repair-dependent activation of c-Abl/p73alpha/GADD45alpha-mediated apoptosis. *J Biol Chem* 2008;283:21394–21403.
 - 20 Evans DB, Varadhachary GR, Crane CH et al. Preoperative gemcitabine-based chemoradiation for patients with resectable adenocarcinoma of the pancreatic head. *J Clin Oncol* 2008;26:3496–3502.
 - 21 Li D, Suzuki H, Liu B et al. DNA repair gene polymorphisms and risk of pancreatic cancer. *Clin Cancer Res* 2009;15:740–746.
 - 22 Lee PH, Shatkay H. F-SNP: computationally predicted functional SNPs for disease association studies. *Nucleic Acids Res* 2008;36:D820–D824.
 - 23 Pounds S, Cheng C. Improving false discovery rate estimation. *Bioinformatics* 2004;20:1737–1745.
 - 24 Jones S, Zhang X, Parsons DW et al. Core signaling pathways in human pancreatic cancers revealed by global genomic analyses. *Science* 2008;321:1801–1806.
 - 25 Kastrinos F, Mukherjee B, Tayob N et al. Risk of pancreatic cancer in families with Lynch syndrome. *JAMA* 2009;302:1790–1795.
 - 26 Luo Y, Lin FT, Lin WC. ATM-mediated stabilization of hMutL DNA mismatch repair proteins augments p53 activation during DNA damage. *Mol Cell Biol* 2004;24:6430–6444.
 - 27 Plaschke J, Kruger S, Jeske B et al. Loss of MSH3 protein expression is frequent in MLH1-deficient colorectal cancer and is associated with disease progression. *Cancer Res* 2004;64:864–870.
 - 28 Melino G, De Laurenzi V, Vousden KH. p73: Friend or foe in tumorigenesis. *Nat Rev Cancer* 2002;2:605–615.
 - 29 Lunghi P, Costanzo A, Mazzera L et al. The p53 family protein p73 provides new insights into cancer chemosensitivity and targeting. *Clin Cancer Res* 2009;15:6495–6502.
 - 30 Khakhar RR, Cobb JA, Bjergbaek L et al. RecQ helicases: multiple roles in genome maintenance. *Trends Cell Biol* 2003;13:493–501.
 - 31 Wu Y, Brosh RM Jr. Distinct roles of RECQ1 in the maintenance of genomic stability. *DNA Repair* 2010;9:315–324.
 - 32 Gupta R, Brosh RM Jr. DNA repair helicases as targets for anti-cancer therapy. *Curr Med Chem* 2007;14:503–517.
 - 33 Kimchi-Sarfaty C, Oh JM, Kim IW et al. A “silent” polymorphism in the MDR1 gene changes substrate specificity. *Science* 2007;315:525–528.
 - 34 Pickering BM, Willis AE. The implications of structured 5' untranslated regions on translation and disease. *Semin Cell Dev Biol* 2005;16:39–47.
 - 35 Mazumder B, Seshadri V, Fox PL. Translational control by the 3'-UTR: the ends specify the means. *Trends Biochem Sci* 2003;28:91–98.