Original Article

SLF1 polymorphism predicts response to oxaliplatin-based adjuvant chemotherapy in patients with colon cancer

Xiaohong Han1*, Zheng Wang2*, Lei Zhang3*, Yinchen Shen3, Qiaoyun Tan3, Yongkun Sun3, Jianfei Wang3, Xiaoyan Qian3, Hongying Yang4, Yuankai Shi3

1Clinical Pharmacology Research Center, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing 100032, China; 2Department of Pathology, Beijing Hospital, Beijing 100730, China; 3Department of Medical Oncology, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing Key Laboratory of Clinical Study on Anticancer Molecular Targeted Drugs, Beijing 100021, China; 4Department of Pathology, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing 100021, China. *Equal contributors.

Received October 11, 2020; Accepted February 8, 2021; Epub April 15, 2021; Published April 30, 2021

Abstract: Response to oxaliplatin-based adjuvant chemotherapy varies among patients with stage II and III colon cancer; however, genetic alterations associated with this response remain incompletely characterized. A three-stage analytical framework, including the discovery, validation, and replication stages, was designed to explore genetic alterations modulating response to oxaliplatin-based chemotherapy in adjuvant setting among patients with stage II and III colon cancer receiving complete resection of tumor. Except for several somatic mutated genes, such as ARSD and ACE, showing less definitive associations with response to oxaliplatin-based adjuvant chemotherapy, we found stable associations of rs6891545C > A polymorphism in SLF1 gene, a key component of DNA damage response system, with the response across all three stages. Patients with rs6891545 A allele had significantly lower risk of poor responsiveness to oxaliplatin-based adjuvant chemotherapy at both discovery and validation stages, compared with ones possessing wild homozygous genotype CC (discovery stage: odds ratio, 0; 95% CI, 0-0.48; \( P = .005 \); validation stage: odds ratio, 0.33; 95% CI, 0.11-0.99; \( P = .048 \)). In the replication cohort, rs6891545 A allele was confirmed to be strongly associated with improved DFS (hazard ratio, 0.43; 95% CI, 0.23-0.81; \( P = .007 \)). Notably, the improvement persisted after controlling for sex, age, tumor location, differentiation, and stage (hazard ratio, 0.42; 95% CI, 0.22-0.80; \( P = .009 \)). Moreover, in silico analysis unraveled strong impact of rs6891545 A allele on local secondary structure of SLF1 mRNA, possibly leading to low SLF1 protein expression. We conclude that the rs6891545C > A polymorphism may serve as an independent marker of response to oxaliplatin-based adjuvant chemotherapy in patients with stage II and III colon cancer, with improved clinical benefit observed in patients with the A allele possibly attributable to low expression of SLF1 protein resulting in deficient DNA repair capacity.

Keywords: SLF1, marker, oxaliplatin-based adjuvant chemotherapy, colon cancer

Introduction

Colorectal cancer ranks third and second in terms of incidence and mortality, respectively, posing a major health burden worldwide [1]. Surgical resection remains the mainstay of treatment for stage I-III locoregional cancers [2]. Postoperative adjuvant fluoropyrimidine chemotherapy is first reported to associate with survival advantages of patients with stage III colon cancer by Moertel and colleagues [3]. Previous trials have demonstrated further enhanced survival by fluoropyrimidine-oxaliplatin combination chemotherapy, including FOLFOX (folic acid, fluorouracil, plus oxaliplatin) and XELOX (capecitabine plus oxaliplatin) regimens, in patients receiving curative resection for stage II and stage III colon cancer [4-7]. Nonetheless, adjuvant chemotherapeutic efficacy varies among patients [8, 9]. Accordingly,
Marker of oxaliplatin-based adjuvant chemotherapy response in SLF1

the development of biomarkers to help determine a patient subset with higher likelihood of response has important clinical implications.

Many studies have been conducted to unravel molecular markers of response to adjuvant oxaliplatin-containing chemotherapy in colon cancer. A randomized FOLFOX-based adjuvant trial showed that KRAS-mutated patients had inferior outcome [10], but contrasts with the findings from another large trial, in which data suggested the lack of prognostic value of KRAS mutations in the adjuvant setting [11]. BRAF V600E mutation was also reported to associate with worse overall survival (OS) in FOLFOX-treated patients [11], whereas a nonsignificant impact on OS was observed in the MOSAIC trial [6]. Similarly, in a pooled analysis of two randomized adjuvant trials, favorable prognosis was identified in deficient versus proficient DNA mismatch repair phenotypes [12], but not by other studies [6, 11]. In addition to these biomarkers derived from pre-treatment tumors, circulating tumor DNA detectable after adjuvant chemotherapy also displayed a strong positive association with recurrence risk in stage III colon cancer [13].

Besides, associations between response to oxaliplatin-containing chemotherapy and genetic polymorphisms in genes regarding DNA repair, including ERCC1 and ERCC2 [14-16], drug-metabolizing enzymes, such as GSTP1 [15-18] and GSTM5 [19], immunogenic cell death pathway [20], including CALR and LRP1, as well as other relevant mechanisms, such as SELE [16], have been previously investigated in the context of metastatic colorectal cancer and, less commonly, in the adjuvant setting; however, statistical significances from individual studies remain inconsistent for most pharmacogenetic markers [15, 16]. Of note, the majority of studies adopted the candidate gene approach based on a priori knowledge of the gene’s biological functions, thus hampering the possibility of identifying novel correlates of response. Here, we aim to identify additional predictors of response to oxaliplatin-containing adjuvant chemotherapy in patients with colon cancer, using a three-stage analytical approach which first performs a comprehensive scan across the whole exome followed by further validation in two relatively large clinical cohorts.

Materials and methods

Study design and sample collection

To construct a relatively homogenous population, the present study was restricted to stages II and III colon cancer patients receiving complete resection of tumor followed by treatment of FOLFOX or XELOX regimens. Time from surgery to the first confirmed relapse, or alive without recurrence at last contact was defined as disease-free survival (DFS), and time from surgery to last follow-up or death was defined as overall survival (OS). We further defined patients with DFS > 60 months as good responders and ones with DFS ≤ 25 months as poor responders. Our study was designed in three stages. At the discovery stage, 13 patients with stage III colon cancer, which included seven good responders and six poor responders, were recruited at Cancer Hospital (CH), Chinese Academy of Medical Sciences (CAMS) for whole exome sequencing. Next, a total of 60 patients with stage II and III colon cancer from CH of CAMS and Beijing Hospital (BH) were collected as the validation set, which comprised 30 good responders and 30 poor responders. Further, we collected 290 stage II and III colon cancer patients at CH of CAMS and BH as the replication cohort.

All patients were unrelated ethnic Han Chinese recruited between March 2005 and December 2013, and had histopathologically or cytologically confirmed stage II or III colon cancer according to the 7th edition of the International Union against Cancer/American Joint Committee on Cancer staging system, which was reviewed by at least two local pathologists. We also confirmed that no patient had received chemotherapy or radiotherapy previously. All specimens were identified to guarantee tumor cells enrichment through hematoxylin and eosin staining before DNA extraction. Genomic DNA was extracted from formalin-fixed, paraffin-embedded (FFPE) tumor tissues using the QIAamp DNA FFPE Tissue Kit (Qiagen, Hilden, Germany), as per the manufacturer’s protocol. Informed consent was obtained from each patient before sample collection. This study was approved by the Institutional Review Boards of both hospitals.

Whole exome sequencing

Genomic DNA was sheared to 150 to 200 bp using a Covaris system (Covaris, Woburn, MA,
Marker of oxaliplatin-based adjuvant chemotherapy response in SLF1

USA) and quantified on a 2100 Bioanalyzer (Agilent, Santa Clara, CA, USA). Following Agilent’s protocol, DNA libraries were then constructed using Agilent SureSelect Human All Exon V5 Kit and sequenced on HiSeq 2000 platform as paired-end 100 bp reads (Illumina, San Diego, CA, USA). After generation of raw sequencing reads, we removed reads including adapter bases, low quality reads with > 50% bases having base quality ≤ 5, or reads with > 10% unknown bases. The remaining reads were aligned to human reference genome (hg19) using Burrows-Wheeler Aligner (http://bio-bwa.sourceforge.net/). Duplicate reads were marked using Picard (https://broadinstitute.github.io/picard/). Resulting BAM files were feed to the Genome Analysis Toolkit (GATK) to perform local indel realignment, base quality score recalibration, variants calling, and variant quality score recalibration according to GATK best practices [21]. Detected variants were annotated using the ANNOVAR software [22].

We discriminated somatic mutations from germline variants in three steps [23]. First, we retained alterations for further analysis only if they (i) were nonsynonymous alterations located within exons or the intronic 2 base pairs of a splicing junction; (ii) supported the alternative allele by > 3 reads; and (iii) had variant allele fraction (VAF) > 5%. Next, retained alterations were flagged as somatic if they were absent in any of population variants databases below with alternative allele frequency > 0.1%: 1000 Genomes Project phase 3 [24], Exome Aggregation Consortium database (version 0.3) [25], National Heart, Lung, and Blood Institute’s Exome Sequencing Project (ESP-6500SI-V2) [26], and Database of Single Nucleotide Polymorphisms (dbSNP) build 147 [27]; however, although present in dbSNP, retained alterations were flagged as somatic if they were absent in the Catalogue Of Somatic Mutations In Cancer database (COSMIC) (version 70) [28]. Finally, we excluded alterations, which were present in ClinVar archive [29] as benign or likely benign classification, from our putative somatic mutations set.

Analysis of somatic mutations from whole exome sequencing

We first calculated tumor mutational load as the number of somatic mutations for each patient. The proportion of somatic mutations attributable to each category of transitions and transversions (C > A, C > G, C > T, T > A, T > C, and T > G) in each patient was also calculated and then compared between good and poor responders using Pearson’s chi-squared test. In addition, a list of cancer-associated genes was derived from Gillette and colleagues [30].

To identify genes associated with response to oxaliplatin-based adjuvant chemotherapy, significantly mutated genes in good and poor responders were first inferred using oncosdrive-CLUST [31], respectively, with Benjamin-Hochberg corrected P value (false discovery rate [FDR]) < .05 as the threshold of statistical significance. Mutation frequency for each gene with at least one somatic mutation was next compared in good versus poor responders by Fisher’s exact test.

Intratumor heterogeneity of each patient tumor was quantified using the mutant-allele tumor heterogeneity (MATH) algorithm [32], with high MATH value reflecting high intratumor heterogeneity.

Mutational signatures were extracted using standard nonnegative matrix factorization method implemented in the R package maftools [33], based on detected single base substitutions and corresponding trinucleotide context that included immediate bases surrounding the mutated base. Optimal number of signatures was determined as the value where cophenetic correlation dropped significantly in an elbow plot [34]. Extracted signatures were compared with 30 known signatures from COSMIC (https://cancer.sanger.ac.uk/cosmic/signatures_v2) by calculating the cosine similarity.

Germline variants prioritization and validation

To explore associations of germline variants with response to oxaliplatin-based adjuvant chemotherapy, we conducted association analyses for nonsynonymous germline variants in the discovery set using the dominant model (mutational homozygote and heterozygote versus wild homozygote) by Fisher’s exact test. Genotypes of significantly associated variants were next confirmed for the 13 patients from the discovery set. For significant variants passing the genotyping validation, we further validated their associations in the validation set using the dominant model by Fisher’s exact test. Odds ratios (ORs) with 95% confidence
Marker of oxaliplatin-based adjuvant chemotherapy response in SLF1

intervals (CIs) were provided. Genotyping for patients from the discovery and validation sets were both done utilizing iPLEX MassARRAY system (Agena, San Diego, USA) according to the manufacturer’s protocol.

Survival analysis

At the replication stage, Kaplan-Meier survival and multivariable Cox proportional hazards regression analyses were used to examine associations between candidate germline variants and survival outcome. To ruling out the possibility that the association was skewed by confounding factors, we included sex, age, tumor location, differentiation, and stage in our Cox model. The log-rank test was performed to compare survival of patients stratified by either genotype categories or tumor location. Hazard ratios (HRs) with 95% CIs were provided.

Functional analysis of rs6891545 polymorphism

Polyphen-2 [35] and SIFT [36] use structure and sequence homology-based algorithms to assess functional consequences of nonsynonymous polymorphisms, respectively. Polyphen-2 and SIFT scores use the same range, 0.0 to 1.0; a variant with a Polyphen-2 score of 0.0 is predicted to be benign, but a variant with a SIFT score of 1.0 is predicted to be benign. To predict functional impact of the amino acid substitution (p.Ser288Arg) caused by alternative A allele of rs6891545 in SLF1, both programs were used. The Genotype-Tissue Expression portal (https://www.gtexportal.org) was used to identify whether rs6891545 is an expression quantitative trait locus (eQTL), which may affect gene expression in cis and trans. To detect local RNA secondary structure changes induced by rs6891545, the RNAsnp software [37] was used, with the global folding mode (measure = Euclidean distance, minimum length of the sequence interval = 50, cut-off for the base pair probabilities = 0.01) and a folding window size of 501 nt.

Statistical analysis

Fisher’s exact test and Welch t test were used to examine differences in discrete and continuous characteristics between good and poor responders, respectively. The Mann-Whitney U test was used to compare mutational load, signature activities, and intratumor heterogeneity between good and poor responders. Oncoplot was drawn using the R package maftools [33]. Association analyses in the discovery and validation sets were performed with PLINK 1.9 [38]. Survival analyses were performed using the R package survival (https://cran.r-project.org/web/packages/survival/index.html). Survive curves were plotted using the R package survminer (https://cran.r-project.org/web/packages/survminer/index.html). The forest plot was generated using the R package forestplot (https://cran.r-project.org/web/packages/forestplot/index.html). Statistical significance was set at two-tailed $P < .05$ unless otherwise specified.

Results

Characteristics of study subjects

The mean ages (± standard deviation) of patients at the discovery, validation, and replication stages were 57.85±9.55, 62.55±10.99, and 61.69±12.65 years, and 8 (62%), 27 (45%), as well as 179 (58.6%) patients were males, respectively. Detailed demographic and clinico-pathological characteristics of patients at the three stages were summarized in Table 1. We analyzed the differences in these characteristics in the discovery and validation sets, and identified no statistically significant differences (all $P > .05$, Fisher’s exact test and Welch t test) (Table 1). Thus, both association analyses were conducted without characteristics adjustment.

Variants identified from whole exome sequencing

Whole exome sequencing was used to characterize genetic alterations in seven good responders and six poor responders receiving oxaliplatin-based adjuvant chemotherapy. We achieved a median depth of 117.78 × (interquartile range [IQR], 100.63-133.94 ×) and a median coverage of 97.1% (IQR, 96.9%-97.5%) on the target region, with a median 87.8% (IQR, 85.4%-88.9%) of which covered by at least 20 reads (Table S1).

The categorization of genetic alterations into somatic and germline sets per patient led to 6311 somatic mutations (Table S2) and 270003 germline variants. The detailed distri-
bution of somatic mutations across various functional classification is shown in Figure 1A, with missense mutations as the most common classification (52.5%). No significant differences in the mutational load between good and poor responders were observed for total, synonymous, and nonsynonymous somatic mutations (median load, 277 versus 269; 69 versus 62; 203 versus 207, respectively; all \( P > .05 \), Mann-Whitney U test) (Figure 1B). Of 593 cancer-associated genes we collected, 157 (26.5%) were identified in our study, including known colorectal cancer driver genes (\( APC \), \( TP53 \), \( KRAS \), \( PIK3CA \), and \( AXIN2 \)) (Figure 1C).

**Comparative analysis of mutated genes and intratumor heterogeneity in good and poor responders**

Nine and five significantly mutated genes were detected in good and poor responders, respectively. (All FDR < .05) (Figure 2A and 2B). Of these genes, five genes, including \( RBMXL3 \), \( ELP4 \), \( SSC5D \), \( TCHH \), and \( ZNF343 \), were significant exclusively in the good responders' group, whereas \( LFNG \) was observed to be significantly mutated only in the group of poor responders. Of these, \( TCHH \) was previously linked to responsiveness to chemotherapy in gastric cancer [39] and breast cancer [40]. We further compared somatic mutational frequencies in good responders with those in poor responders across all genes. Notably, none of those significantly mutated genes above exhibited differential mutations between two groups (Table S3). But we found that two genes (\( ARSD \) and \( ACE \)), albeit marginally significant, had been implicated in regulation of cancer growth, metastasis, and response to platinum-based chemotherapy [41-44]. Specifically, \( ARSD \) mutations were observed in 4 of 7 good (57%) versus 0 of 6 poor (0%) responders, whereas 0 of 7 good (0%) and 3 of 6 poor (50%) responders had \( ACE \) mutations (both \( P = .07 \), Fisher's exact test) (Figure 2C and Table S3). However, there were no definite biological roles reported for another marginally significant gene, \( GOLGA6L2 \).

Given that intratumor heterogeneity has been hypothesized to associate with inferior outcome and therapeutic resistance in multiple cancer types [45], we compared the intratumor heterogeneity in good versus poor responders. Consistent with the hypothesis, poor responders to adjuvant chemotherapy evidenced a trend to higher intratumor heterogeneity than good responders (median MATH, 46.6; \( P = .07 \), one-tailed Mann-Whitney U test) (Figure S1).

**Mutational signatures in good and poor responders**

To further investigate whether patterns of somatic mutations associated with response to
Marker of oxaliplatin-based adjuvant chemotherapy response in SLF1
Marker of oxaliplatin-based adjuvant chemotherapy response in SLF1

oxaliplatin-based adjuvant chemotherapy, we first examined the distribution of six different types of single base conversions and found that C > A transversion was significantly enriched in good responders (18.9% in good responders versus 14.3% in poor responders; \(P = 1.4 \times 10^{-4}\), chi-squared test) (Figure 3A). In addition, significant predominance in poor responders were observed for another three types of transversions, including C > G (5.2% versus 10%; \(P = 1 \times 10^{-9}\), chi-squared test), T > A (4.1% versus 6.5%; \(P = 4.7 \times 10^{-4}\), chi-squared test), and T > G (3.9% vs 5.7%; \(P = .004\), chi-squared test).
Marker of oxaliplatin-based adjuvant chemotherapy response in \textit{SLF1}

![Graph showing marker of oxaliplatin-based adjuvant chemotherapy response in SLF1](image)

**A**

Good response vs. Poor response

**B**

Signature 1, Signature 5, Signature 6

Contributions
Marker of oxaliplatin-based adjuvant chemotherapy response in SLF1

Figure 3. Comparison of somatic mutational patterns between good and poor responders. A. A bar plot depicting proportions of somatic mutations attributable to each single base conversion type stratified by the response status. B. Patterns of three mutational signatures, which are displayed as per six substitution subtypes and sequence context immediately 5' and 3' to the mutated base. The vertical axis denotes the proportion of somatic mutations contributed by each of 96 possible mutation types in the signature. The contribution bars are depicted in different colors according to the six substitution subtypes.

Table 2. Candidate germline polymorphisms associated with response to oxaliplatin-based adjuvant chemotherapy

<table>
<thead>
<tr>
<th>Gene</th>
<th>Variant</th>
<th>Genotype</th>
<th>Discovery stage</th>
<th>Validation stage</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>No. Poor Responders</td>
<td>No. Good Responders</td>
</tr>
<tr>
<td>APPL2</td>
<td>rs2272495</td>
<td>GG</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GA + AA</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>SLF1</td>
<td>rs6891545</td>
<td>CC</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CA + AA</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>WARS2</td>
<td>rs3790549</td>
<td>CC</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CG + GG</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>ZNF443</td>
<td>rs35699767</td>
<td>CC</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CA + AA</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>ADGRE5</td>
<td>rs2230748</td>
<td>GG</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GA + AA</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>IL6R</td>
<td>rs2228145</td>
<td>AA</td>
<td>0</td>
<td>6</td>
</tr>
</tbody>
</table>

OR, odds ratio; CI, confidence interval; NA, not available. *The wild homozygote of each polymorphism is set to the reference.

test). We next used the nonnegative matrix factorization technique to decompose somatic mutations into three mutational signatures (Figures S2 and 3B). The cosine similarity against COSMIC signatures revealed proposed etiologies of these signatures (Figure S3), including spontaneous deamination of 5-methylcytosine (signature 1), unknown environmental exposure (signature 5), and defective DNA mismatch repair (signature 6). Given that the proportions of somatic mutations attributable to each signature varied between patients (Figure S4), we asked whether mutational signatures could associate with response to oxaliplatin-based adjuvant chemotherapy. However, we found that there were no significant differences in mutational activities between good and poor responders for all three signatures (median activity, signature 1, 0.49 versus 0.63; signature 5, 0.46 versus 0.35; signature 6, 0.02 versus 0.02; all P > .05, Mann-Whitney U test) (Figure S5).

Candidate germline polymorphisms associated with response to oxaliplatin-based adjuvant chemotherapy at the discovery and validation stages

As a further step toward understanding genetic factors that modulate response to oxaliplatin-based adjuvant chemotherapy, we next investigated nonsynonymous germline variants. Association analyses for nonsynonymous germline variants in the discovery cohort identified six candidate missense polymorphisms across six genes (Table 2), including rs2272495 in APPL2, rs6891545 in SLF1, rs3790549 in WARS2, rs35699767 in ZNF443, rs2230748 in ADGRE5, and rs2228145 in IL6R. Of the six polymorphisms, four (rs3790549, rs6891545, rs2272495, and rs35699767) were confirmed with the iPLEX MassARRAY system. We next selected these four polymorphisms to perform the validation in an independent cohort with 60 patients with stage II and III colon cancer receiving oxaliplatin-based adjuvant chemotherapy. As shown in Table 2, alternative A allele of rs6891545 in the SLF1 gene was identified to significantly reduce the risk of poor responsiveness to oxaliplatin-based adjuvant chemotherapy in the dominant model (CA + AA versus CC: OR, 0.33; 95% CI, 0.11-0.99; P = .048). However, other three variants failed to reach statistical significance at this validation stage.

Survival analysis in the replication cohort

To further confirm associations of the four polymorphisms with response to oxaliplatin-based
Marker of oxaliplatin-based adjuvant chemotherapy response in \textit{SLF1}

**Table 3. Association between the 4 germline polymorphisms and survival outcome in the replication cohort**

<table>
<thead>
<tr>
<th>Gene</th>
<th>Variant</th>
<th>Genotype</th>
<th>No. Patients</th>
<th>HR\textsubscript{DFS} (95% CI)</th>
<th>P\textsubscript{DFS}</th>
<th>HR\textsubscript{OS} (95% CI)</th>
<th>P\textsubscript{OS}</th>
</tr>
</thead>
<tbody>
<tr>
<td>\textit{APPL2}</td>
<td>rs2272495</td>
<td>GG</td>
<td>140</td>
<td>1 [reference]</td>
<td>.27</td>
<td>1 [reference]</td>
<td>.23</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GA + AA</td>
<td>108</td>
<td>0.71 (0.39-1.31)</td>
<td>.007</td>
<td>0.63 (0.30-1.34)</td>
<td>.07</td>
</tr>
<tr>
<td>\textit{SLF1}</td>
<td>rs6891545</td>
<td>CC</td>
<td>119</td>
<td>1 [reference]</td>
<td>.007</td>
<td>1 [reference]</td>
<td>.007</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CA + AA</td>
<td>125</td>
<td>0.43 (0.23-0.81)</td>
<td>.50</td>
<td>0.50 (0.23-1.08)</td>
<td>.50</td>
</tr>
<tr>
<td>\textit{WARS2}</td>
<td>rs3790549</td>
<td>CC</td>
<td>171</td>
<td>1 [reference]</td>
<td>.75</td>
<td>1 [reference]</td>
<td>.59</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CG + GG</td>
<td>70</td>
<td>1.11 (0.59-2.08)</td>
<td>1.23</td>
<td>1.23 (0.58-2.63)</td>
<td>.75</td>
</tr>
<tr>
<td>\textit{ZNF443}</td>
<td>rs35699767</td>
<td>CC</td>
<td>103</td>
<td>1 [reference]</td>
<td>.68</td>
<td>1 [reference]</td>
<td>.75</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CA + AA</td>
<td>105</td>
<td>0.86 (0.42-1.76)</td>
<td>0.83</td>
<td>0.83 (0.32-2.13)</td>
<td>.83</td>
</tr>
</tbody>
</table>

\(HR\textsubscript{DFS}\) and \(P\textsubscript{DFS}\) indicate hazard ratio and \(P\) value for disease-free survival, respectively. \(HR\textsubscript{OS}\) and \(P\textsubscript{OS}\) indicate hazard ratio and \(P\) value for overall survival, respectively.

Impact of rs6891545 on \textit{SLF1} mRNA expression and local secondary structure

We used Polyphen-2 and SIFT to predict whether amino acid change (p.Ser288Arg) attributable to rs6891545C > A polymorphism is deleterious to \textit{SLF1} protein function, and found that both softwares classified it as tolerated (Polyphen-2 score = 0; SIFT score = 1). We next explored the eQTL role of rs6891545 and observed a positive association of rs6891545 A allele with \textit{SLF1} mRNA expression in peripheral blood (\(P = .05\)). Single nucleotide polymorphisms in both coding and non-coding regions of mRNA are known to induce mRNA secondary structure changes [37]. Thus, we used RNAsnp to check the impact of rs6891545 on \textit{SLF1} mRNA secondary structure and found a significant structural difference in local secondary structure of mutant sequence (containing alternate A allele) versus that of wild-type sequence (containing reference C allele) (\(P = .05\)) (Figure 6A). The optimal secondary structures of global mutant and wild-type sequences around rs6891545 are presented in Figure 6B and 6C.

Discussion

In this work, we adopted a three-stage analytical approach to investigate genetic alterations modulating response to oxaliplatin-based adjuvant chemotherapy in stage II and III colon cancer patients. Our analysis of somatic mutations from whole exome sequencing unraveled several mutated genes which may associate with the response, including \textit{ARSD} and \textit{ACE}. Specifically, \textit{ARSD} is a member of the sulfatase family and plays a crucial role in sphingolipid
Marker of oxaliplatin-based adjuvant chemotherapy response in SLF1

Figure 4. Association between rs6891545 and survival outcome in the replication cohort. (A and B) Kaplan-Meier curves of disease-free survival (A) and overall survival (B) in patients with versus without alternative A allele of rs6891545 in SLF1. HR indicates hazard ratio.

metabolism [41], whose aberrations have been reported to contribute to chemotherapy resistance in various cancers [42]. As a receptor of renin-angiotensin system, ACE catalyzes the conversion of angiotensin I into angiotensin II, and its inhibitors have previously been demon-
Marker of oxaliplatin-based adjuvant chemotherapy response in \textit{SLF1}

<table>
<thead>
<tr>
<th>Variable</th>
<th>No.</th>
<th>HR (95% CI)</th>
<th>( P ) value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>243</td>
<td>1.03 (1.00–1.06)</td>
<td>.04</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td>.78</td>
</tr>
<tr>
<td>Female</td>
<td>100</td>
<td>1 [Reference]</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>143</td>
<td>1.09 (0.60–1.99)</td>
<td></td>
</tr>
<tr>
<td>Location</td>
<td></td>
<td></td>
<td>.29</td>
</tr>
<tr>
<td>Right</td>
<td>107</td>
<td>1 [Reference]</td>
<td></td>
</tr>
<tr>
<td>Left</td>
<td>136</td>
<td>1.41 (0.75–2.65)</td>
<td></td>
</tr>
<tr>
<td>Differentiation</td>
<td></td>
<td></td>
<td>.54</td>
</tr>
<tr>
<td>Moderate</td>
<td>216</td>
<td>1 [Reference]</td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>27</td>
<td>1.32 (0.54–3.20)</td>
<td></td>
</tr>
<tr>
<td>Stage</td>
<td></td>
<td></td>
<td>.0005</td>
</tr>
<tr>
<td>II</td>
<td>99</td>
<td>1 [Reference]</td>
<td></td>
</tr>
<tr>
<td>III</td>
<td>144</td>
<td>4.69 (1.96–11.24)</td>
<td></td>
</tr>
<tr>
<td>rs6891545</td>
<td></td>
<td></td>
<td>.009</td>
</tr>
<tr>
<td>CC</td>
<td>118</td>
<td>1 [Reference]</td>
<td></td>
</tr>
<tr>
<td>CA+AA</td>
<td>125</td>
<td>0.42 (0.22–0.80)</td>
<td></td>
</tr>
</tbody>
</table>

\textbf{Figure 5.} Forest plot depicting association between rs6891545 and disease-free survival in the replication cohort. The survival analysis is controlled for sex, age, tumor location, differentiation, and stage by Cox model. The vertical line represents hazard ratio (HR) of 1.0. Square markers show estimated HRs. Error bars indicate 95% confidence intervals (CIs).

rs6891545, located in \textit{SLF1} on 5q15, showed consistent association with response to oxaliplatin-based adjuvant chemotherapy across all three stages. Specifically, patients carrying the \textit{SLF1} rs6891545 A allele showed superior DFS in comparison with ones possessing wild CC homozygous genotype (3-year DFS, 91.1% versus 79.8%), regardless of sex, age, tumor location, differentiation, and stage.

Our results could be reasonably explained by biological significance of rs6891545 and its associated gene \textit{SLF1}. As a third generation platinum agent, oxaliplatin induces lethal DNA lesions such as DNA interstrand cross-links, which could block replication fork progression by covalently linking both strands of DNA [47]. Increased damage to cellular DNA enhances antitumor efficacy of platinum agents such as oxaliplatin, whereas DNA damage response (DDR) system has been demonstrated to attenuate the cytotoxicity of chemotherapeutic agents, thus driving resistance and tumor relapse [48]. Of note, \textit{SLF1} has previously been reported to play a pivotal role in DDR system by suppressing genomic instability \textit{in vitro} as a DNA repair factor [49]. Specifically, \textit{SLF1} links RAD18 to SLF2 for the formation of RAD18-SLF1-SLF2 complex, then this complex recruits the SMC5-SMC6 cohesion complex to damaged DNA in response to interstrand cross-links [49]. This raises the possibility that, colon cancer cells overexpressing \textit{SLF1} protein may be more resistant to DNA lesions derived from oxaliplatin-based adjuvant chemotherapy by positive regulation of the DDR system, suggesting that this adjuvant chemotherapy would be more beneficial in patients with low \textit{SLF1} protein expression. Intriguingly, our eQTL analysis

stratified to suppress tumorigenesis and angiogenesis in several cancer models such as colorectal cancer [44]. We also noted that a retrospective analysis of non-small cell lung cancer patients treated with first-line platinum-based chemotherapy indicated a superior survival in patients receiving ACE inhibitors compared with non-recipients [46].

Given marginal statistical significance of these mutated genes and limited findings derived from somatic mutations, we asked whether germline variants could provide additional information in predicting the response. Indeed, our analysis of germline variants demonstrated that a missense germline polymorphism,
Marker of oxaliplatin-based adjuvant chemotherapy response in SLF1

Figure 6. Impact of rs6891545 on local secondary structure of SLF1 mRNA. (A) Base pairing probabilities of the local region containing rs6891545. The dot plot depicts base pairing probabilities of the ensemble structures of wild-type and mutant mRNA sequences corresponds to the local region detected with maximum structural change. The alternative allele is highlighted with a yellow box. The indices (i, j) of the matrix show a dot if the bases at position i and j form a base pair. The size of the dots is proportional to the base pairing probability where small dots indicate low probability to form a base pair and large dots indicate high. The upper triangle shows base pairing probabilities for the wild-type sequence (green) and the lower triangle for the mutant sequence (red). (B and C) Optimal secondary structures of wild-type sequence and mutant sequence in the folding window. The secondary structures in planar graphic representation indicate the minimum free energy structures of global wild-type (B) and mutant sequences (C). The global sequence is from the folding window, which flanks rs6891545 with either side 250 nt.
Marker of oxaliplatin-based adjuvant chemotherapy response in SLF1

revealed a positive impact of rs6891545 A allele on mRNA expression of SLF1 in cis. Notably, the discrepancy between expressions of protein and the cognate mRNA has been shown in the growing body of proteogenomic literature [30, 50, 51]. A possible mechanistic explanation regarding this discrepancy may be the influence the polymorphism exerts on mRNA secondary structure, which could decrease ribosome translation rate and thus lead to consequent low protein level. This hypothesis has been experimentally validated in a prior work focusing on a polymorphism in PARP1 gene [52]. Indeed, using RNA folding algorithm RNAsnp, we found that rs6891545 A allele exhibited a significant local structural effect, thus contributing to reduced SLF1 protein expression. Taken together, this finding indicates that, the clinical observation that patients with the SLF1 rs6891545 A allele had superior prognosis in the adjuvant setting is possibly due to deficient DNA repair capacity caused by low SLF1 protein expression, which in turn boosts the antitumor efficacy of oxaliplatin-based chemotherapy in these group of patients.

In addition, we explored the association of tumor location with survival outcome in the replication cohort. Although with contradictory results, accumulating studies have investigated the prognostic role of primary tumor location in colorectal cancer [53-56]. A meta-analysis described that left-sided colon cancer was associated with greatly reduced death risk [57]. Nonetheless, a population-based cohort study reported no link between tumor location and OS among patients with stage I-III colon cancer [58]. In our study, we observed no association of tumor location with survival benefit from oxaliplatin-based chemotherapy utilized in the adjuvant setting, but larger clinical studies are warranted, considering our limited sample size.

In conclusion, to our knowledge, the current study is the first to unravel clinical significance of SLF1 in the context of adjuvant oxaliplatin-based treatment of patients with stage II and III colon cancer. Our findings suggest that SLF1 rs6891545C > A polymorphism may be a helpful prognostic factor in the adjuvant setting, with improved clinical benefit as a consequence of deficient DNA repair capacity which is attributable to the A allele leading to low expression of SLF1 protein. This polymorphism may hold promise as marker for aiding clinicians in making optimal treatment decisions, although additional prospective studies and experimental work are required to confirm these findings.

Acknowledgements

This work was supported by the International Science & Technology Cooperation Program of China (2013DFE33110) and China National Major Project for New Drug Innovation (2017ZX09304015). We thank BGI Inc. and DNAlead Inc. for providing whole exome sequencing and iPLEX MassARRAY platforms, respectively.

Disclosure of conflict of interest

None.

Address correspondence to: Yuankai Shi, Department of Medical Oncology, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing Key Laboratory of Clinical Study on Anticancer Molecular Targeted Drugs, Beijing 100021, China. Tel: +86-10-87788293; Fax: +86-10-87788781; E-mail: syuankai@cicams.ac.cn

References

Marker of oxaliplatin-based adjuvant chemotherapy response in SLF1


Marker of oxaliplatin-based adjuvant chemotherapy response in SLF1


Marker of oxaliplatin-based adjuvant chemotherapy response in SLF1


Marker of oxaliplatin-based adjuvant chemotherapy response in *SLF1*

Table S1. Sequencing metrics for whole exome sequencing across 13 patient tumors

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>Total effective reads</th>
<th>Total effective yield (Mb)</th>
<th>Number of reads uniquely mapped to genome</th>
<th>Number of reads uniquely mapped to target</th>
<th>Fraction of uniquely mapped on target</th>
<th>Average sequencing depth on target</th>
<th>Coverage of target region</th>
<th>Fraction of target covered with at least 20x</th>
<th>Fraction of target covered with at least 10x</th>
<th>Fraction of target covered with at least 4x</th>
</tr>
</thead>
<tbody>
<tr>
<td>277706</td>
<td>122133465</td>
<td>10670.67</td>
<td>111518642</td>
<td>80034844</td>
<td>71.80%</td>
<td>124.32</td>
<td>97.10%</td>
<td>88.60%</td>
<td>92.80%</td>
<td>95.20%</td>
</tr>
<tr>
<td>277998</td>
<td>134493901</td>
<td>11843.52</td>
<td>12525596</td>
<td>88962597</td>
<td>71.00%</td>
<td>136.71</td>
<td>97.80%</td>
<td>90.10%</td>
<td>93.80%</td>
<td>95.90%</td>
</tr>
<tr>
<td>279098</td>
<td>106302581</td>
<td>9300.22</td>
<td>97087232</td>
<td>67051920</td>
<td>69.10%</td>
<td>104.04</td>
<td>97.30%</td>
<td>85.40%</td>
<td>91.30%</td>
<td>94.80%</td>
</tr>
<tr>
<td>283332</td>
<td>45172463</td>
<td>3825.64</td>
<td>37858443</td>
<td>22730742</td>
<td>60.00%</td>
<td>35.98</td>
<td>95.90%</td>
<td>66.00%</td>
<td>81.20%</td>
<td>90.50%</td>
</tr>
<tr>
<td>287863</td>
<td>75223569</td>
<td>6536.31</td>
<td>67540573</td>
<td>44658639</td>
<td>66.10%</td>
<td>69.42</td>
<td>96.40%</td>
<td>80.30%</td>
<td>88.30%</td>
<td>93.30%</td>
</tr>
<tr>
<td>289158</td>
<td>121322280</td>
<td>10725.12</td>
<td>113970047</td>
<td>87147126</td>
<td>76.50%</td>
<td>133.94</td>
<td>97.10%</td>
<td>88.90%</td>
<td>93.10%</td>
<td>95.40%</td>
</tr>
<tr>
<td>291839</td>
<td>113146100</td>
<td>9916.84</td>
<td>104022206</td>
<td>75806215</td>
<td>72.90%</td>
<td>117.78</td>
<td>96.90%</td>
<td>87.40%</td>
<td>92.10%</td>
<td>94.90%</td>
</tr>
<tr>
<td>297152</td>
<td>92485037</td>
<td>8039.34</td>
<td>82826991</td>
<td>55205956</td>
<td>66.70%</td>
<td>85.76</td>
<td>96.90%</td>
<td>83.50%</td>
<td>90.10%</td>
<td>94.20%</td>
</tr>
<tr>
<td>297830</td>
<td>98390735</td>
<td>8596.67</td>
<td>90011771</td>
<td>64926776</td>
<td>72.10%</td>
<td>100.63</td>
<td>96.60%</td>
<td>85.80%</td>
<td>91.20%</td>
<td>94.40%</td>
</tr>
<tr>
<td>321025</td>
<td>130064595</td>
<td>11413.78</td>
<td>120230719</td>
<td>90999914</td>
<td>75.70%</td>
<td>141.58</td>
<td>97.10%</td>
<td>87.90%</td>
<td>92.30%</td>
<td>95.10%</td>
</tr>
<tr>
<td>329616</td>
<td>143906275</td>
<td>12725.25</td>
<td>135522215</td>
<td>102558594</td>
<td>75.70%</td>
<td>156.89</td>
<td>97.50%</td>
<td>90.80%</td>
<td>94.10%</td>
<td>95.90%</td>
</tr>
<tr>
<td>333072</td>
<td>132996279</td>
<td>11765.38</td>
<td>125437112</td>
<td>86870470</td>
<td>69.30%</td>
<td>132.27</td>
<td>98.20%</td>
<td>89.10%</td>
<td>93.70%</td>
<td>96.30%</td>
</tr>
<tr>
<td>336428</td>
<td>116714769</td>
<td>10304.44</td>
<td>109283096</td>
<td>76233595</td>
<td>69.80%</td>
<td>116.47</td>
<td>97.90%</td>
<td>87.80%</td>
<td>93.00%</td>
<td>95.80%</td>
</tr>
</tbody>
</table>
Figure S1. MATH values are compared between good and poor responders using one-tailed Mann-Whitney U test. Box plots show the first, median, and third quartiles, whiskers extend to 1.5 times the interquartile range, and each dot indicates one patient tumor.
Marker of oxaliplatin-based adjuvant chemotherapy response in \textit{SLF1}

\textbf{Figure S2.} The elbow plot depicts the cophenetic correlation on the Y axis for varying signature numbers (X axis), where the correlation shows the first substantial decrease on signature number 3.

\textbf{Figure S3.} The row depicts cosine similarity between each mutational signature identified in our study and 30 COSMIC signatures. Shown in the cell is the specific similarity value, filled with colors from dark blue (lower similarity), to dark red (higher similarity).
Marker of oxaliplatin-based adjuvant chemotherapy response in \textit{SLF1}

\textbf{Figure S4.} Filled bars depict the mutational signature activities of 13 patients stratified by response status, and the vertical axis denotes the proportion of mutations attributable to each mutational signature.

\textbf{Figure S5.} Comparison of mutational activities attributable to signature 1 associated with spontaneous deamination of 5-methylcytosine (A), signature 5 associated with unknown environmental exposure (B), and signature 6 associated with defective DNA mismatch repair (C) between good and poor responders. Mann-Whitney U test was used in (A-C). Box plots show the first, median, and third quartiles, whiskers extend to 1.5 times the interquartile range, and each dot indicates one patient tumor.
Marker of oxaliplatin-based adjuvant chemotherapy response in *SLF1*

**A**

**rs2272495**

Log-rank test $P = 0.27$

HR = 0.71 (95% CI, 0.39-1.31)

**B**

**rs3790549**

Log-rank test $P = 0.75$

HR = 1.11 (95% CI, 0.59-2.08)
Marker of oxaliplatin-based adjuvant chemotherapy response in SLF1

Figure S6. A. Kaplan-Meier curves of disease-free survival and overall survival in patients with versus without alternative A allele of rs2272495 in APPL2. B. Kaplan-Meier curves of disease-free survival and overall survival in patients with versus without alternative G allele of rs3790549 in WARS2. C. Kaplan-Meier curves of disease-free survival and overall survival in patients with versus without alternative A allele of rs35699767 in ZNF443. HR indicates hazard ratio.