Am J Cancer Res 2021;11(4):1522-1539
www.ajcr.us /ISSN:2156-6976/ajcr0123922

*Original Article*

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

Xiaohong Han[1*], Zheng Wang[2*], Lei Zhang[3*], Yinchen Shen[3], Qiaoyun Tan[3], Yongkun Sun[3], Jianfei Wang[3], Xiaoyan Qian[3], Hongying Yang[4], Yuankai Shi[3]

*[1]Clinical Pharmacology Research Center, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing 100032, China; [2]Department of Pathology, Beijing Hospital, Beijing 100730, China; [3]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; [4]Department 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 (folinic 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,

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 stage 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,

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 found 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 oncodriveCLUST [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

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 PLNK 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 clinicopathological 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-

**Table 1.** Characteristics of study subjects at the discovery, validation, and replication stages

| Variable | Discovery stage | | | Validation stage | | | Replication stage (N = 290) |
|---|---|---|---|---|---|---|---|
| | Good (N = 7) | Poor (N = 6) | *P* | Good (N = 30) | Poor (N = 30) | *P* | |
| Sex | | | 1 | | | 1 | |
| Male | 4 | 4 | | 13 | 14 | | 170 |
| Female | 3 | 2 | | 17 | 16 | | 120 |
| Age, years | 58.57±9.96 | 57±9.9 | 0.78 | 61.83±11.14 | 63.27±10.99 | 0.62 | 61.69±12.65 |
| Location[a] | | | 1 | | | 0.43 | |
| Right-sided | 4 | 3 | | 10 | 14 | | 129 |
| Left-sided | 3 | 3 | | 20 | 16 | | 161 |
| Differentiation | | | 0.27 | | | 0.23 | |
| Moderate | 6 | 3 | | 25 | 20 | | 250 |
| Low | 1 | 3 | | 5 | 10 | | 39 |
| Stage | | | NA | | | 0.79 | |
| II | NA | NA | | 14 | 12 | | 107 |
| III | 7 | 6 | | 16 | 18 | | 183 |

Data are shown as number or mean ± standard deviation. [a]Tumors proximal to the splenic flexure are classified as right side of colon cancer, whereas left-sided colon cancer consists of the splenic flexure, descending and sigmoid colon.

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, 59.8 versus 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

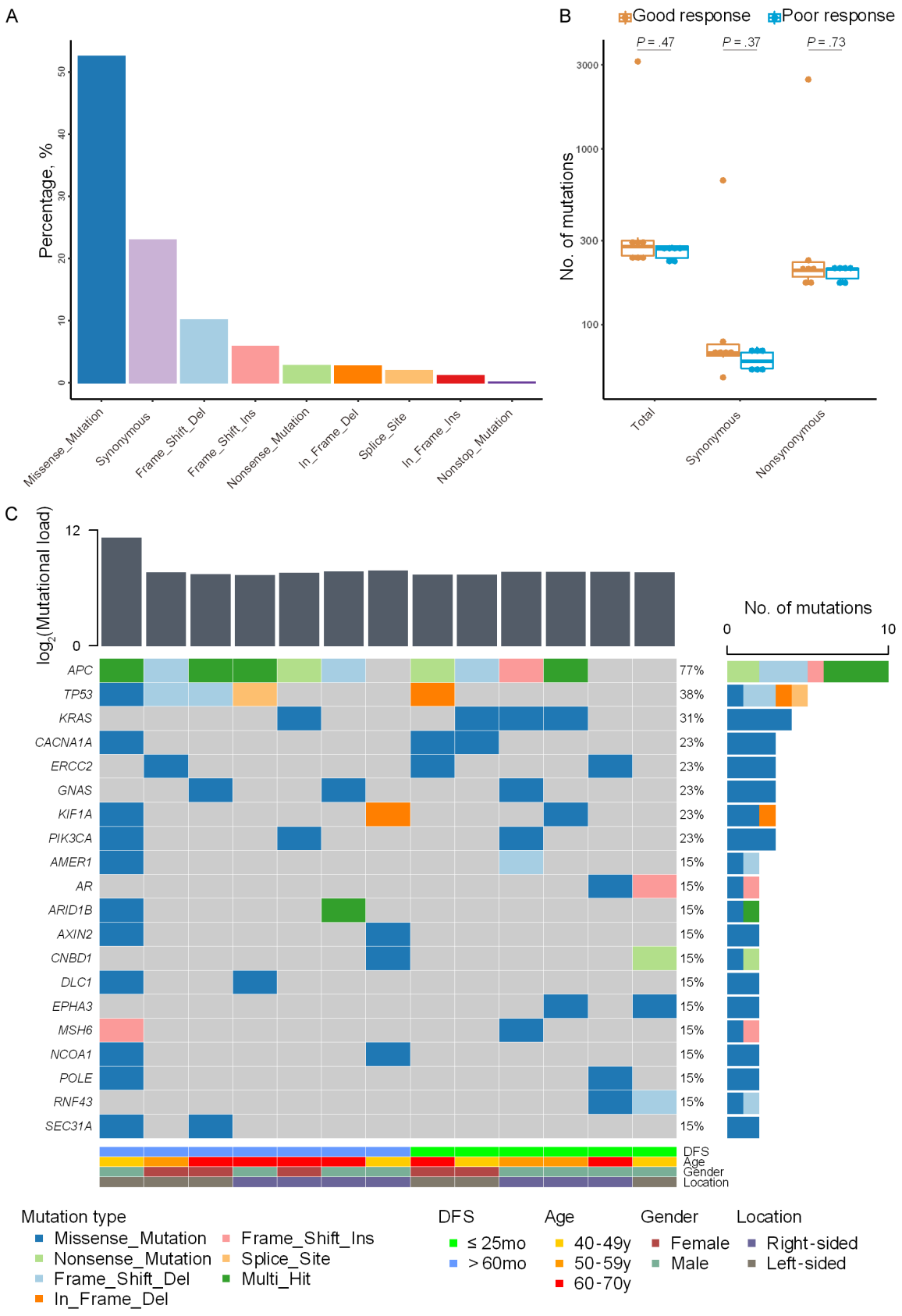

A
Percentage, %
Missense_Mutation
Synonymous
Frame_Shift_Del
Frame_Shift_Ins
Nonsense_Mutation
In_Frame_Del
Splice_Site
In_Frame_Ins
Nonstop_Mutation
B
Good response
Poor response
P = .47
P = .37
P = .73
No. of mutations
Total
Synonymous
Nonsynonymous
C
log2(Mutational load)
No. of mutations
APC 77%
TP53 38%
KRAS 31%
CACNA1A 23%
ERCC2 23%
GNAS 23%
KIF1A 23%
PIK3CA 23%
AMER1 15%
AR 15%
ARID1B 15%
AXIN2 15%
CNBD1 15%
DLC1 15%
EPHA3 15%
MSH6 15%
NCOA1 15%
POLE 15%
RNF43 15%
SEC31A 15%
DFS
Age
Gender
Location
Mutation type
Missense_Mutation
Frame_Shift_Ins
Nonsense_Mutation
Splice_Site
Frame_Shift_Del
Multi_Hit
In_Frame_Del
DFS
≤ 25mo
> 60mo
Age
40-49y
50-59y
60-70y
Gender
Female
Male
Location
Right-sided
Left-sided

**Figure 1.** Somatic mutations landscape across 13 patient tumors at the discovery stage. A. Distribution of somatic mutations across various classification. B. Comparison of mutational load between good and poor responders for total, synonymous, and nonsynonymous mutations. Box plots depict the first, median, and third quartiles, whiskers extend to 1.5 times the interquartile range, and each dot indicates one patient tumor. Differences in the number of somatic mutations for all three comparisons were examined using the Mann-Whitney *U* test. C. Oncoplot for top 20 mutated cancer-associated genes, which are ordered according to decreasing mutational frequency from top to bottom and color-coded for different mutation types. Bars on the right side indicate the number of somatic mutations in each gene. Patients are shown as columns, and top panel indicates nonsynonymous mutational load of each patient. Tile plots in the middle show clinical characteristics of each patient.


**Figure 2.** Comparison of mutated genes between good and poor responders. (A and B) Significantly mutated genes identified in good (A) and poor (B) responders. The tile plot in the middle depicts significantly mutated genes for each patient, with different mutation types having distinct colors and each column indicating one patient. The bar plot in the right shows the number of somatic mutations in each gene, whereas the bar plot in the left denotes statistical significance levels of each gene as -$\log_{10}$ (FDR). (C) A tile plot depicting mutations of three differential mutated genes, with each column indicating one patient.

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% in good responders versus 10% in poor responders; $P = 1 \times 10^{-9}$, 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

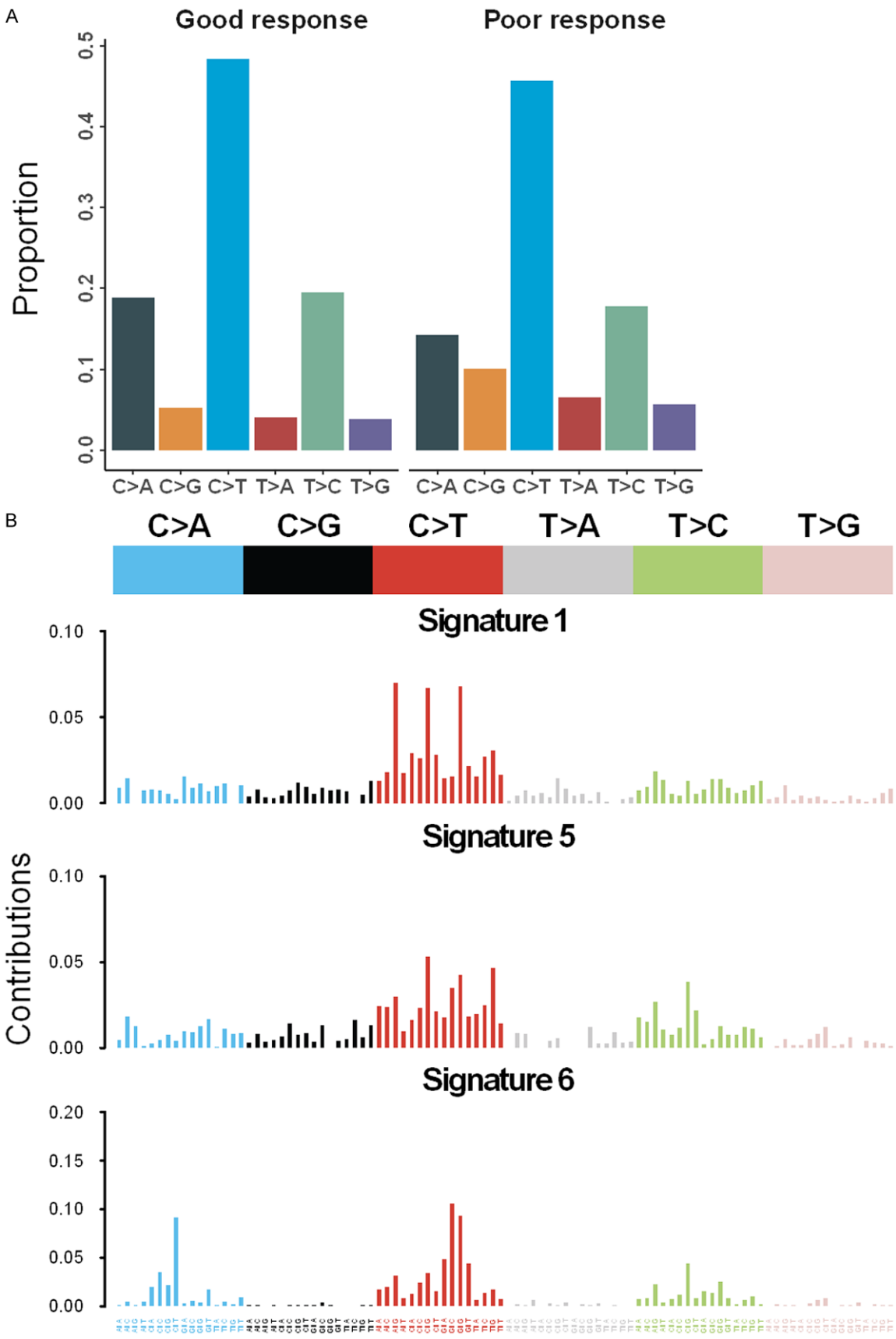

A
Good response
Poor response
Proportion
0.0
0.1
0.2
0.3
0.4
0.5
C>A C>G C>T T>A T>C T>G
C>A C>G C>T T>A T>C T>G
B
C>A
C>G
C>T
T>A
T>C
T>G
Signature 1
0.10
0.05
0.00
Signature 5
Contributions
0.10
0.05
0.00
Signature 6
0.20
0.15
0.10
0.05
0.00

**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

| Gene | Variant | Genotype[a] | Discovery stage | | | | Validation stage | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | No. Poor Responders | No. Good Responders | OR (95% CI) | *P* | No. Poor Responders | No. Good Responders | OR (95% CI) | *P* |
| *APPL2* | rs2272495 | GG | 0 | 6 | 1 | .005 | 18 | 14 | 1 | .44 |
| | | GA + AA | 6 | 1 | 0 (0-0.48) | | 11 | 13 | 0.66 (0.23-1.91) | |
| *SLF1* | rs6891545 | CC | 0 | 6 | 1 | .005 | 17 | 9 | 1 | .048 |
| | | CA + AA | 6 | 1 | 0 (0-0.48) | | 12 | 19 | 0.33 (0.11-0.99) | |
| *WARS2* | rs3790549 | CC | 6 | 1 | 1 | .005 | 19 | 17 | 1 | .56 |
| | | CG + GG | 0 | 6 | NA | | 8 | 10 | 0.72 (0.23-2.23) | |
| *ZNF443* | rs35699767 | CC | 1 | 7 | 1 | .005 | 13 | 13 | 1 | .53 |
| | | CA + AA | 5 | 0 | 0 (0-0.5) | | 13 | 9 | 1.44 (0.46-4.54) | |
| *ADGRE5* | rs2230748 | GG | 6 | 1 | 1 | .005 | | | | |
| | | GA + AA | 0 | 6 | NA | | | | | |
| *IL6R* | rs2228145 | AA | 0 | 6 | 1 | .005 | | | | |
| | | AC + CC | 6 | 1 | 0 (0-0.48) | | | | | |

OR, odds ratio; CI, confidence interval; NA, not available. [a]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

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

| Gene | Variant | Genotype | No. Patients | $HR_{DFS}$ (95% CI) | $P_{DFS}$ | $HR_{OS}$ (95% CI) | $P_{OS}$ |
|---|---|---|---|---|---|---|---|
| *APPL2* | rs2272495 | GG | 140 | 1 [reference] | .27 | 1 [reference] | .23 |
| | | GA + AA | 108 | 0.71 (0.39-1.31) | | 0.63 (0.30-1.34) | |
| *SLF1* | rs6891545 | CC | 119 | 1 [reference] | .007 | 1 [reference] | .07 |
| | | CA + AA | 125 | 0.43 (0.23-0.81) | | 0.50 (0.23-1.08) | |
| *WARS2* | rs3790549 | CC | 171 | 1 [reference] | .75 | 1 [reference] | .59 |
| | | CG + GG | 70 | 1.11 (0.59-2.08) | | 1.23 (0.58-2.63) | |
| *ZNF443* | rs35699767 | CC | 103 | 1 [reference] | .68 | 1 [reference] | .7 |
| | | CA + AA | 105 | 0.86 (0.42-1.76) | | 0.83 (0.32-2.13) | |

$HR_{DFS}$ and $P_{DFS}$ indicate hazard ratio and *P* value for disease-free survival, respectively. $HR_{OS}$ and $P_{OS}$ indicate hazard ratio and *P* value for overall survival, respectively.

adjuvant chemotherapy, we performed survival analyses in a larger replication cohort. In the Kaplan-Meier survival analysis, rs6891545 A allele of *SLF1* was found to be strongly associated with superior DFS (CA + AA versus CC: HR, 0.43; 95% CI, 0.23-0.81; *P* = .007), but exhibited marginally significant improvement in OS (HR, 0.50; 95% CI, 0.23-1.08; *P* = .07) (**Table 3** and **Figure 4A** and **4B**). For other three variants (rs2272495, rs3790549, and rs35699767), no significant associations were observed for either DFS or OS (**Table 3** and Figure S6).

To eliminate the possibility that associations between the four polymorphisms and DFS were biased by confounders, a multivariable Cox proportional hazards regression model was adopted to controlling for sex, age, tumor location, differentiation, and stage. The results revealed that no significant associations with DFS were observed for sex, tumor location, and differentiation, but patients with stage III colon cancer acquired less survival benefit from oxaliplatin-based adjuvant chemotherapy in comparison with ones with stage II colon cancer (HR, 4.69; 95% CI, 1.96-11.24; *P* < .001); and elder patients showed slightly increased hazard of death (HR, 1.03; 95% CI, 1.00-1.06; *P* = 0.04) (**Figure 5**). Notably, we found that the association between rs6891545 and DFS remained statistically significant after adjusting for these confounding factors (HR, 0.42; 95% CI, 0.22-0.80; *P* = .009) (**Figure 5**). However, there were still no statistical significance to be reached for other three polymorphisms (*P* > .05 for both DFS and OS). We also examined the association of tumor location with OS, and likewise detected no association signal (HR, 1.29; 95% CI, 0.63-2.64; *P* = .49).

*Impact of rs6891545 on 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 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 *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 *SLF1* mRNA secondary structure and found a significant structural difference in local secondary structure of mutant sequence (containing alternative 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 *ARSD* and *ACE*. Specifically, ARSD is a member of the sulfatase family and plays a crucial role in sphingolipid

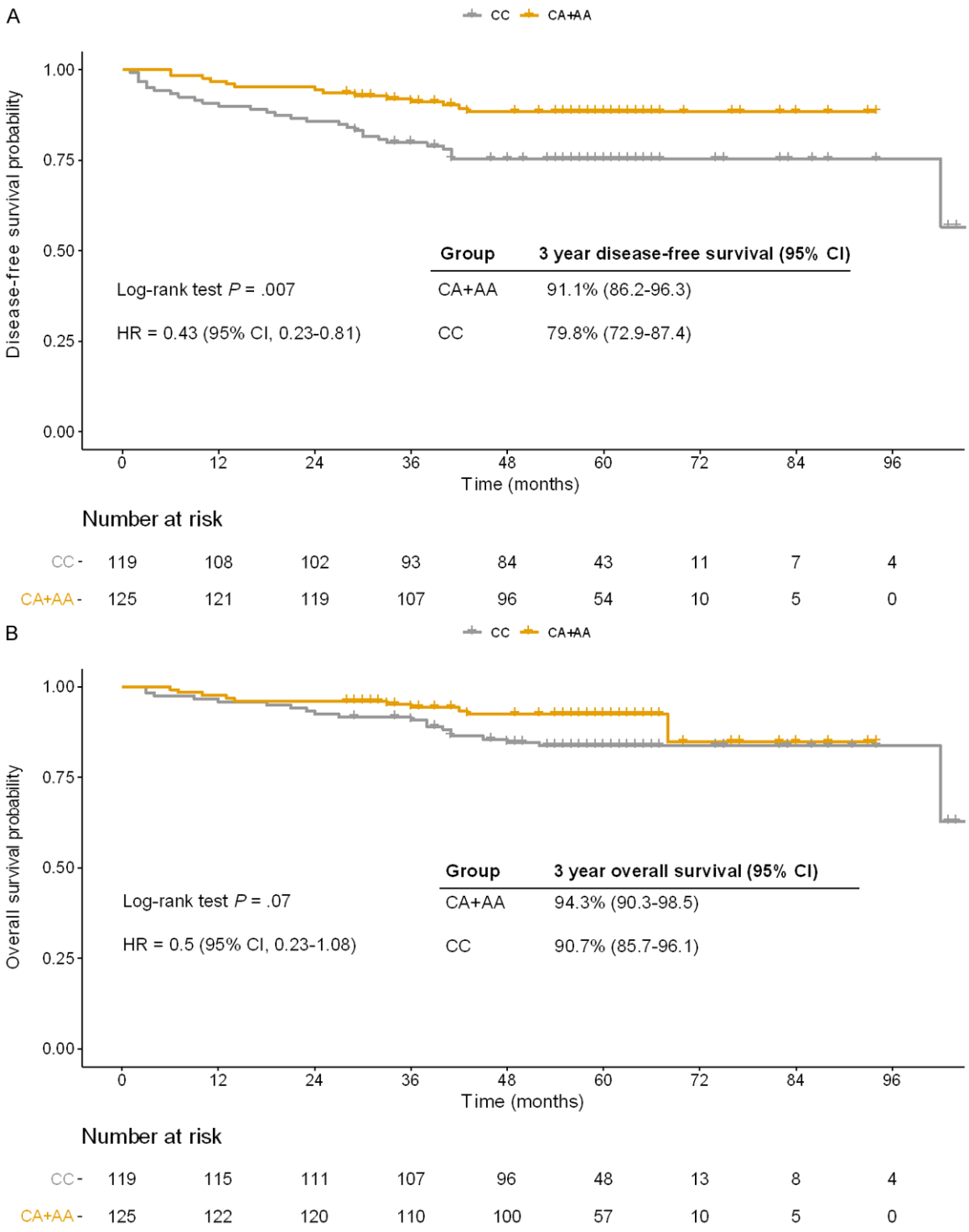



**Figure 4.** Association between rs6891545 and survival outcome in the replication cohort. (A and B) Kaplan-Meier curves of disease-free survival (A) and overall survival (B) in patients with versus without alternative A allele of rs6891545 in *SLF1*. HR indicates hazard ratio.

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-

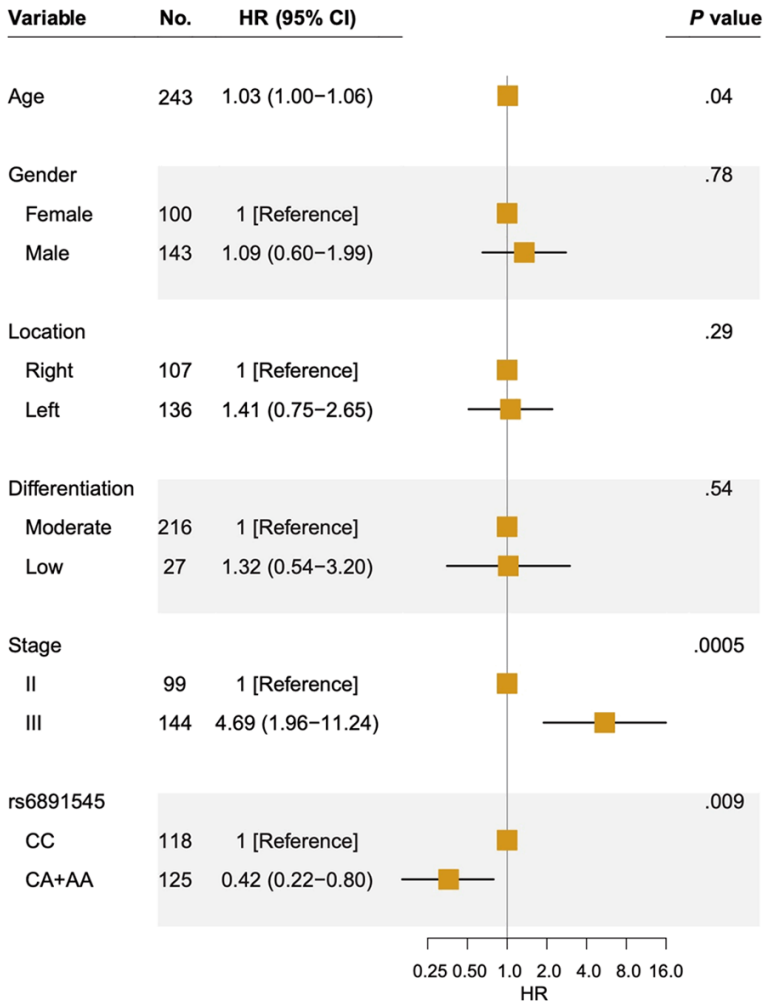



**Figure 5.** Forest plot depicting association between rs6891545 and disease-free survival in the replication cohort. The survival analysis is controlled for sex, age, tumor location, differentiation, and stage by Cox model. The vertical line represents hazard ratio (HR) of 1.0. Square markers show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

strated 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, rs6891545, located in *SLF1* on 5q15, showed consistent association with response to oxaliplatin-based adjuvant chemotherapy across all three stages. Specifically, patients carrying the *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 *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, SLF1 has previously been reported to play a pivotal role in DDR system by suppressing genomic instability *in vitro* as a DNA repair factor [49]. Specifically, 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 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 SLF1 protein expression. Intriguingly, our eQTL analysis

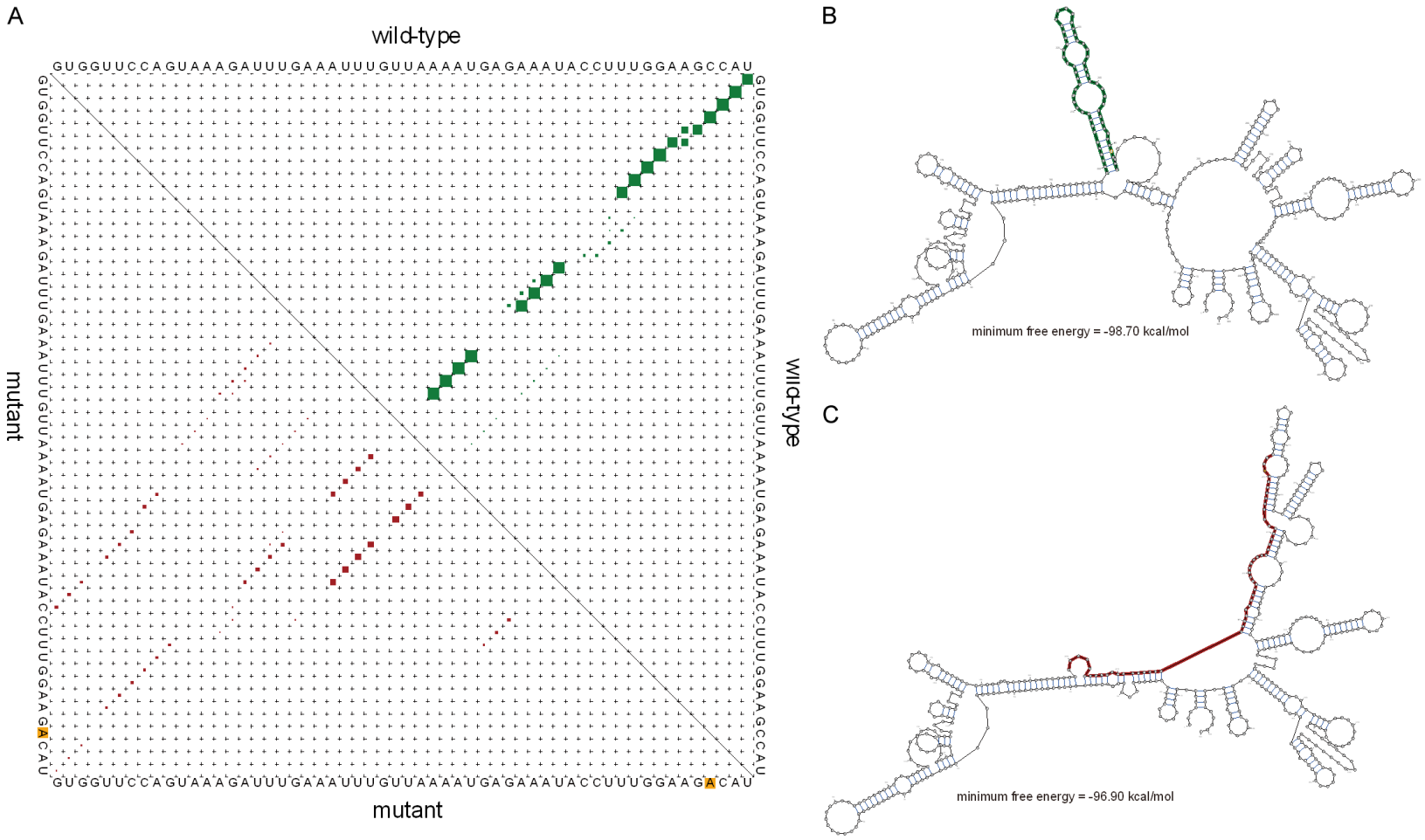



**Figure 6.** Impact of rs6891545 on local secondary structure of *SLF1* mRNA. (A) Base pairing probabilities of the local region containing rs6891545. The dot plot depicts base pairing probabilities of the ensemble structures of wild-type and mutant mRNA sequences corresponds to the local region detected 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.

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

[1] Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA and Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018; 68: 394-424.

[2] Amin MB, Greene FL, Edge SB, Compton CC, Gershenwald JE, Brookland RK, Meyer L, Gress DM, Byrd DR and Winchester DP. The eighth edition AJCC cancer staging manual: continuing to build a bridge from a population-based to a more "personalized" approach to cancer staging. CA Cancer J Clin 2017; 67: 93-99.

[3] Moertel CG, Fleming TR, Macdonald JS, Haller DG, Laurie JA, Goodman PJ, Ungerleider JS, Emerson WA, Tormey DC and Glick JH. Levamisole and fluorouracil for adjuvant therapy of resected colon carcinoma. N Engl J Med 1990; 322: 352-358.

[4] André T, Boni C, Mounedji-Boudiaf L, Navarro M, Tabernero J, Hickish T, Topham C, Zaninelli M, Clingan P, Bridgewater J, Tabah-Fisch I and de Gramont A. Oxaliplatin, fluorouracil, and

leucovorin as adjuvant treatment for colon cancer. N Engl J Med 2004; 350: 2343-2351.

[5] André T, Boni C, Navarro M, Tabernero J, Hickish T, Topham C, Bonetti A, Clingan P, Bridgewater J, Rivera F and de Gramont A. Improved overall survival with oxaliplatin, fluorouracil, and leucovorin as adjuvant treatment in stage II or III colon cancer in the MOSAIC trial. J Clin Oncol 2009; 27: 3109-3116.

[6] André T, de Gramont A, Vernerey D, Chibaudel B, Bonnetain F, Tijeras-Raballand A, Scriva A, Hickish T, Tabernero J, Van Laethem JL, Banzi M, Maartense E, Shmueli E, Carlsson GU, Scheithauer W, Papamichael D, Möehler M, Landolfi S, Demetter P, Colote S, Tournigand C, Louvet C, Duval A, Fléjou JF and de Gramont A. Adjuvant fluorouracil, leucovorin, and oxaliplatin in stage II to III colon cancer: updated 10-year survival and outcomes according to braf mutation and mismatch repair status of the MOSAIC study. J Clin Oncol 2015; 33: 4176-4187.

[7] Yothers G, O'Connell MJ, Allegra CJ, Kuebler JP, Colangelo LH, Petrelli NJ and Wolmark N. Oxaliplatin as adjuvant therapy for colon cancer: updated results of NSABP C-07 trial, including survival and subset analyses. J Clin Oncol 2011; 29: 3768-3774.

[8] Auclin E, Zaanan A, Vernerey D, Douard R, Gallois C, Laurent-Puig P, Bonnetain F and Taieb J. Subgroups and prognostication in stage III colon cancer: future perspectives for adjuvant therapy. Ann Oncol 2017; 28: 958-968.

[9] Fotheringham S, Mozolowski GA, Murray EMA and Kerr DJ. Challenges and solutions in patient treatment strategies for stage II colon cancer. Gastroenterol Rep (Oxf) 2019; 7: 151-161.

[10] Sinicrope FA, Mahoney MR, Smyrk TC, Thibodeau SN, Warren RS, Bertagnolli MM, Nelson GD, Goldberg RM, Sargent DJ and Alberts SR. Prognostic impact of deficient DNA mismatch repair in patients with stage III colon cancer from a randomized trial of FOLFOX-based adjuvant chemotherapy. J Clin Oncol 2013; 31: 3664-3672.

[11] Gavin PG, Colangelo LH, Fumagalli D, Tanaka N, Remillard MY, Yothers G, Kim C, Taniyama Y, Kim S II, Choi HJ, Blackmon NL, Lipchik C, Petrelli NJ, O'Connell MJ, Wolmark N, Paik S and Pogue-Geile KL. Mutation profiling and microsatellite instability in stage II and III colon cancer: an assessment of their prognostic and oxaliplatin predictive value. Clin Cancer Res 2012; 18: 6531-6541.

[12] Zaanan A, Shi Q, Taieb J, Alberts SR, Meyers JP, Smyrk TC, Julie C, Zawadi A, Tabernero J, Mini E, Goldberg RM, Folprecht G, Van Laethem JL, Le Malicot K, Sargent DJ, Laurent-Puig P and Sinicrope FA. Role of deficient DNA mismatch repair status in patients with stage III colon cancer treated with FOLFOX adjuvant chemotherapy: a pooled analysis from 2 randomized clinical trials. JAMA Oncol 2018; 4: 379-383.

[13] Tie J, Cohen JD, Wang Y, Christie M, Simons K, Lee M, Wong R, Kosmider S, Ananda S, McKendrick J, Lee B, Cho JH, Faragher I, Jones IT, Ptak J, Schaeffer MJ, Silliman N, Dobbyn L, Li L, Tomasetti C, Papadopoulos N, Kinzler KW, Vogelstein B and Gibbs P. Circulating tumor DNA analyses as markers of recurrence risk and benefit of adjuvant therapy for stage III colon cancer. JAMA Oncol 2019; 5: 1710-1717.

[14] Yin M, Yan J, Martinez-Balibrea E, Graziano F, Lenz HJ, Kim HJ, Robert J, Im SA, Wang WS, Etienne-Grimaldi MC and Wei Q. ERCC1 and ERCC2 polymorphisms predict clinical outcomes of oxaliplatin-based chemotherapies in gastric and colorectal cancer: a systemic review and meta-analysis. Clin Cancer Res 2011; 17: 1632-1640.

[15] Zaanan A, Dalban C, Emile JF, Blons H, Fléjou JF, Goumard C, Istanbullu M, Calmel C, Alhazmi K, Validire P, Louvet C, de Gramont A, Laurent-Puig P, Taïeb J and Praz F. *ERCC1*, *XRCC1* and *GSTP1* single nucleotide polymorphisms and survival of patients with colon cancer receiving oxaliplatin-based adjuvant chemotherapy. J Cancer 2014; 5: 425-432.

[16] Custodio A, Moreno-Rubio J, Aparicio J, Gallego-Plazas J, Yaya R, Maurel J, Rodríguez-Salas N, Burgos E, Ramos D, Calatrava A, Andrada E, Díaz-López E, Sánchez A, Madero R, Cejas P and Feliu J. Pharmacogenetic predictors of outcome in patients with stage II and III colon cancer treated with oxaliplatin and fluoropyrimidine-based adjuvant chemotherapy. Mol Cancer Ther 2014; 13: 2226-2237.

[17] Stoehlmacher J, Park DJ, Zhang W, Yang D, Groshen S, Zahedy S and Lenz HJ. A multivariate analysis of genomic polymorphisms: prediction of clinical outcome to 5-FU/oxaliplatin combination chemotherapy in refractory colorectal cancer. Br J Cancer 2004; 91: 344-354.

[18] Boige V, Mendiboure J, Pignon JP, Loriot MA, Castaing M, Barrois M, Malka D, Trégouët DA, Bouché O, Le Corre D, Miran I, Mulot C, Ducreux M, Beaune P and Laurent-Puig P. Pharmacogenetic assessment of toxicity and outcome in patients with metastatic colorectal cancer treated with LV5FU2, FOLFOX, and FOLFIRI: FFCD 2000-05. J Clin Oncol 2010; 28: 2556-2564.

[19] Kap EJ, Seibold P, Scherer D, Habermann N, Balavarca Y, Jansen L, Zucknick M, Becker N, Hoffmeister M, Ulrich A, Benner A, Ulrich CM, Burwinkel B, Brenner H and Chang-Claude J.

SNPs in transporter and metabolizing genes as predictive markers for oxaliplatin treatment in colorectal cancer patients. Int J Cancer 2016; 138: 2993-3001.

[20] Arai H, Xiao Y, Loupakis F, Kawanishi N, Wang J, Battaglin F, Soni S, Zhang W, Mancao C, Salhia B, Mumenthaler SM, Weisenberger DJ, Liang G, Cremolini C, Falcone A, Millstein J and Lenz HJ. Immunogenic cell death pathway polymorphisms for predicting oxaliplatin efficacy in metastatic colorectal cancer. J Immunother Cancer 2020; 8: e001714.

[21] DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, Philippakis AA, del Angel G, Rivas MA, Hanna M, McKenna A, Fennell TJ, Kernytsky AM, Sivachenko AY, Cibulskis K, Gabriel SB, Altshuler D and Daly MJ. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet 2011; 43: 491-498.

[22] Wang K, Li M and Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res 2010; 38: e164.

[23] Sukhai MA, Misyura M, Thomas M, Garg S, Zhang T, Stickle N, Virtanen C, Bedard PL, Siu LL, Smets T, Thijs G, Van Vooren S, Kamel-Reid S and Stockley TL. Somatic tumor variant filtration strategies to optimize tumor-only molecular profiling using targeted next-generation sequencing panels. J Mol Diagn 2019; 21: 261-273.

[24] Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, Korbel JO, Marchini JL, McCarthy S, McVean GA and Abecasis GR. A global reference for human genetic variation. Nature 2015; 526: 68-74.

[25] Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, O'Donnell-Luria AH, Ware JS, Hill AJ, Cummings BB, Tukiainen T, Birnbaum DP, Kosmicki JA, Duncan LE, Estrada K, Zhao F, Zou J, Pierce-Hoffman E, Berghout J, Cooper DN, Deflaux N, DePristo M, Do R, Flannick J, Fromer M, Gauthier L, Goldstein J, Gupta N, Howrigan D, Kiezun A, Kurki MI, Moonshine AL, Natarajan P, Orozco L, Peloso GM, Poplin R, Rivas MA, Ruano-Rubio V, Rose SA, Ruderfer DM, Shakir K, Stenson PD, Stevens C, Thomas BP, Tiao G, Tusie-Luna MT, Weisburd B, Won HH, Yu D, Altshuler DM, Ardissino D, Boehnke M, Danesh J, Donnelly S, Elosua R, Florez JC, Gabriel SB, Getz G, Glatt SJ, Hultman CM, Kathiresan S, Laakso M, McCarroll S, McCarthy MI, McGovern D, McPherson R, Neale BM, Palotie A, Purcell SM, Saleheen D, Scharf JM, Sklar P, Sullivan PF, Tuomilehto J, Tsuang MT, Watkins HC, Wilson JG, Daly MJ and MacArthur DG; Exome Aggregation Consortium. Analysis of protein-coding genetic variation in 60,706 humans. Nature 2016; 536: 285-291.

[26] Fu W, O'Connor TD, Jun G, Kang HM, Abecasis G, Leal SM, Gabriel S, Rieder MJ, Altshuler D, Shendure J, Nickerson DA, Bamshad MJ and Akey JM. Analysis of 6,515 exomes reveals the recent origin of most human protein-coding variants. Nature 2013; 493: 216-220.

[27] Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM and Sirotkin K. dbSNP: the NCBI database of genetic variation. Nucleic Acids Res 2001; 29: 308-311.

[28] Tate JG, Bamford S, Jubb HC, Sondka Z, Beare DM, Bindal N, Boutselakis H, Cole CG, Creatore C, Dawson E, Fish P, Harsha B, Hathaway C, Jupe SC, Kok CY, Noble K, Ponting L, Ramshaw CC, Rye CE, Speedy HE, Stefancsik R, Thompson SL, Wang S, Ward S, Campbell PJ and Forbes SA. COSMIC: the catalogue of somatic mutations in cancer. Nucleic Acids Res 2018; 47: D941-D947.

[29] Landrum MJ, Lee JM, Benson M, Brown G, Chao C, Chitipiralla S, Gu B, Hart J, Hoffman D, Hoover J, Jang W, Katz K, Ovetsky M, Riley G, Sethi A, Tully R, Villamarin-Salomon R, Rubinstein W and Maglott DR. ClinVar: public archive of interpretations of clinically relevant variants. Nucleic Acids Res 2016; 44: D862-8.

[30] Gillette MA, Satpathy S, Cao S, Dhanasekaran SM, Vasaikar SV, Krug K, Petralia F, Li Y, Liang WW, Reva B, Krek A, Ji J, Song X, Liu W, Hong R, Yao L, Blumenberg L, Savage SR, Wendl MC, Wen B, Li K, Tang LC, MacMullan MA, Avanessian SC, Kane MH, Newton CJ, Cornwell M, Kothadia RB, Ma W, Yoo S, Mannan R, Vats P, Kumar-Sinha C, Kawaler EA, Omelchenko T, Colaprico A, Geffen Y, Maruvka YE, da Veiga Leprevost F, Wiznerowicz M, Gümüş ZH, Veluswamy RR, Hostetter G, Heiman DI, Wyczalkowski MA, Hiltke T, Mesri M, Kinsinger CR, Boja ES, Omenn GS, Chinnaiyan AM, Rodriguez H, Li QK, Jewell SD, Thiagarajan M, Getz G, Zhang B, Fenyö D, Ruggles KV, Cieslik MP, Robles AI, Clauser KR, Govindan R, Wang P, Nesvizhskii AI, Ding L, Mani DR and Carr SA. Proteogenomic characterization reveals therapeutic vulnerabilities in lung adenocarcinoma. Cell 2020; 182: 200-225, e35.

[31] Tamborero D, Gonzalez-Perez A and Lopez-Bigas N. OncodriveCLUST: exploiting the positional clustering of somatic mutations to identify cancer genes. Bioinformatics 2013; 29: 2238-2244.

[32] Mroz EA, Tward AM, Hammon RJ, Ren Y and Rocco JW. Intra-tumor genetic heterogeneity and mortality in head and neck cancer: analysis of data from the cancer genome atlas. PLoS Med 2015; 12: e1001786.

[33] Mayakonda A, Lin DC, Assenov Y, Plass C and Koeffler HP. Maftools: efficient and comprehensive analysis of somatic variants in cancer. Genome Res 2018; 28: 1747-1756.
[34] Gaujoux R and Seoighe C. A flexible R package for nonnegative matrix factorization. BMC Bioinformatics 2010; 11: 367.
[35] Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS and Sunyaev SR. A method and server for predicting damaging missense mutations. Nat Methods 2010; 7: 248-249.
[36] Sim NL, Kumar P, Hu J, Henikoff S, Schneider G and Ng PC. SIFT web server: predicting effects of amino acid substitutions on proteins. Nucleic Acids Res 2012; 40: W452-7.
[37] Sabarinathan R, Tafer H, Seemann SE, Hofacker IL, Stadler PF and Gorodkin J. The RNAsnp web server: predicting SNP effects on local RNA secondary structure. Nucleic Acids Res 2013; 41: W475-9.
[38] Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM and Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. GigaScience 2015; 4: 7.
[39] Yang Q, Zhu C, Zhang Y, Wang Y, Wang Y, Zhu L, Yang X, Li J, Nie H, Jiang S, Zhang X, Cao X, Li Q, Zhang X, Tian G, Hu L, Zhu L, Zhao G and Zhang Z. Molecular analysis of gastric cancer identifies genomic markers of drug sensitivity in Asian gastric cancer. J Cancer 2018; 9: 2973-2980.
[40] Al Amri WS, Baxter DE, Hanby AM, Stead LF, Verghese ET, Thorne JL and Hughes TA. Identification of candidate mediators of chemoresponse in breast cancer through therapy-driven selection of somatic variants. Breast Cancer Res Treat 2020; 183: 607-616.
[41] Trojani A, Di Camillo B, Tedeschi A, Lodola M, Montesano S, Ricci F, Vismara E, Greco A, Veronese S, Orlacchio A, Martino S, Colombo C, Mura M, Nichelatti M, Colosimo A, Scarpati B, Montillo M and Morra E. Gene expression profiling identifies ARSD as a new marker of disease progression and the sphingolipid metabolism as a potential novel metabolism in chronic lymphocytic leukemia. Cancer Biomark 2012; 11: 15-28.
[42] Ogretmen B. Sphingolipid metabolism in cancer signalling and therapy. Nat Rev Cancer 2018; 18: 33-50.
[43] Pinter M and Jain RK. Targeting the renin-angiotensin system to improve cancer treatment: implications for immunotherapy. Sci Transl Med 2017; 9: eaan5616.
[44] Yang Y, Ma L, Xu Y, Liu Y, Li W, Cai J and Zhang Y. Enalapril overcomes chemoresistance and potentiates antitumor efficacy of 5-FU in colorectal cancer by suppressing proliferation, angiogenesis, and NF-κB/STAT3-regulated proteins. Cell Death Dis 2020; 11: 477.
[45] Marusyk A, Janiszewska M and Polyak K. Intratumor heterogeneity: the rosetta stone of therapy resistance. Cancer Cell 2020; 37: 471-484.
[46] Wilop S, von Hobe S, Crysandt M, Esser A, Osieka R and Jost E. Impact of angiotensin I converting enzyme inhibitors and angiotensin II type 1 receptor blockers on survival in patients with advanced non-small-cell lung cancer undergoing first-line platinum-based chemotherapy. J Cancer Res Clin Oncol 2009; 135: 1429-1435.
[47] Woynarowski JM, Faivre S, Herzig MC, Arnett B, Chapman WG, Trevino AV, Raymond E, Chaney SG, Vaisman A, Varchenko M and Juniewicz PE. Oxaliplatin-induced damage of cellular DNA. Mol Pharmacol 2000; 58: 920-927.
[48] Bouwman P and Jonkers J. The effects of deregulated DNA damage signalling on cancer chemotherapy response and resistance. Nat Rev Cancer 2012; 12: 587-598.
[49] Räschle M, Smeenk G, Hansen RK, Temu T, Oka Y, Hein MY, Nagaraj N, Long DT, Walter JC, Hofmann K, Storchova Z, Cox J, Bekker-Jensen S, Mailand N and Mann M. Proteomics reveals dynamic assembly of repair complexes during bypass of DNA cross-links. Science 2015; 348: 1253671.
[50] Chen YJ, Roumeliotis TI, Chang YH, Chen CT, Han CL, Lin MH, Chen HW, Chang GC, Chang YL, Wu CT, Lin MW, Hsieh MS, Wang YT, Chen YR, Jonassen I, Ghavidel FZ, Lin ZS, Lin KT, Chen CW, Sheu PY, Hung CT, Huang KC, Yang HC, Lin PY, Yen TC, Lin YW, Wang JH, Raghav L, Lin CY, Chen YS, Wu PS, Lai CT, Weng SH, Su KY, Chang WH, Tsai PY, Robles AI, Rodriguez H, Hsiao YJ, Chang WH, Sung TY, Chen JS, Yu SL, Choudhary JS, Chen HY, Yang PC and Chen YJ. Proteogenomics of non-smoking lung cancer in east asia delineates molecular signatures of pathogenesis and progression. Cell 2020; 182: 226-244, e17.
[51] Xu JY, Zhang C, Wang X, Zhai L, Ma Y, Mao Y, Qian K, Sun C, Liu Z, Jiang S, Wang M, Feng L, Zhao L, Liu P, Wang B, Zhao X, Xie H, Yang X, Zhao L, Chang Y, Jia J, Wang X, Zhang Y, Wang Y, Yang Y, Wu Z, Yang L, Liu B, Zhao T, Ren S, Sun A, Zhao Y, Ying W, Wang F, Wang G, Zhang Y, Cheng S, Qin J, Qian X, Wang Y, Li J, He F, Xiao T and Tan M. Integrative proteomic characterization of human lung adenocarcinoma. Cell 2020; 182: 245-261, e17.
[52] Cashman R, Zilberberg A, Priel A, Philip H, Varvak A, Jacob A, Shoval I and Efroni S. A single nucleotide variant of human PARP1 determines response to PARP inhibitors. NPJ Precis Oncol 2020; 4: 10.

[53] Meguid RA, Slidell MB, Wolfgang CL, Chang DC and Ahuja N. Is there a difference in survival between right- versus left-sided colon cancers? Ann Surg Oncol 2008; 15: 2388-2394.

[54] Weiss JM, Pfau PR, O'Connor ES, King J, LoConte N, Kennedy G and Smith MA. Mortality by stage for right-versus left-sided colon cancer: analysis of surveillance, epidemiology, and end results-medicare data. J Clin Oncol 2011; 29: 4401-4409.

[55] Warschkow R, Sulz MC, Marti L, Tarantino I, Schmied BM, Cerny T and Güller U. Better survival in right-sided versus left-sided stage I-III colon cancer patients. BMC Cancer 2016; 16: 554.

[56] Lee L, Erkan A, Alhassan N, Kelly JJ, Nassif GJ, Albert MR and Rt Monson J. Lower survival after right-sided versus left-sided colon cancers: is an extended lymphadenectomy the answer? Surg Oncol 2018; 27: 449-455.

[57] Petrelli F, Tomasello G, Borgonovo K, Ghidini M, Turati L, Dallera P, Passalacqua R, Sgroi G and Barni S. Prognostic survival associated with left-sided vs right-sided colon cancer: a systematic review and meta-analysis. JAMA Oncol 2017; 3: 211-219.

[58] Karim S, Brennan K, Nanji S, Berry SR and Booth CM. Association between prognosis and tumor laterality in early-stage colon cancer. JAMA Oncol 2017; 3: 1386-1392.

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

| Sample ID | Total effective reads | Total effective yield (Mb) | Number of reads uniquely mapped to genome | Number of reads uniquely mapped to target | Fraction of uniquely mapped on target | Average sequencing depth on target | Coverage of target region | Fraction of target covered with at least 20x | Fraction of target covered with at least 10x | Fraction of target covered with at least 4x |
|---|---|---|---|---|---|---|---|---|---|---|
| 277706 | 122133465 | 10670.67 | 111518642 | 80034844 | 71.80% | 124.32 | 97.10% | 88.60% | 92.80% | 95.20% |
| 277998 | 134493901 | 11843.52 | 125255596 | 88962597 | 71.00% | 136.71 | 97.80% | 90.10% | 93.80% | 95.90% |
| 279098 | 106302581 | 9300.22 | 97087232 | 67051920 | 69.10% | 104.04 | 97.30% | 85.40% | 91.30% | 94.80% |
| 283332 | 45172463 | 3825.64 | 37858443 | 22730742 | 60.00% | 35.98 | 95.90% | 66.00% | 81.20% | 90.50% |
| 287863 | 75223569 | 6536.31 | 67540573 | 44658639 | 66.10% | 69.42 | 96.40% | 80.30% | 88.30% | 93.30% |
| 289158 | 121322280 | 10725.12 | 113970047 | 87147126 | 76.50% | 133.94 | 97.10% | 88.90% | 93.10% | 95.40% |
| 291839 | 113146100 | 9916.84 | 104022205 | 75806215 | 72.90% | 117.78 | 96.90% | 87.40% | 92.10% | 94.90% |
| 297152 | 92485037 | 8039.34 | 82826991 | 55205956 | 66.70% | 85.76 | 96.90% | 83.50% | 90.10% | 94.20% |
| 297830 | 98390735 | 8596.67 | 90011771 | 64926776 | 72.10% | 100.63 | 96.60% | 85.80% | 91.20% | 94.40% |
| 321025 | 130064595 | 11413.78 | 120230719 | 90999914 | 75.70% | 141.58 | 97.10% | 87.90% | 92.30% | 95.10% |
| 329616 | 143906275 | 12725.25 | 135522215 | 102558594 | 75.70% | 156.89 | 97.50% | 90.80% | 94.10% | 95.90% |
| 333072 | 132996279 | 11765.38 | 125437112 | 86870470 | 69.30% | 132.27 | 98.20% | 89.10% | 93.70% | 96.30% |
| 336428 | 116714769 | 10304.44 | 109283096 | 76233595 | 69.80% | 116.47 | 97.90% | 87.80% | 93.00% | 95.80% |

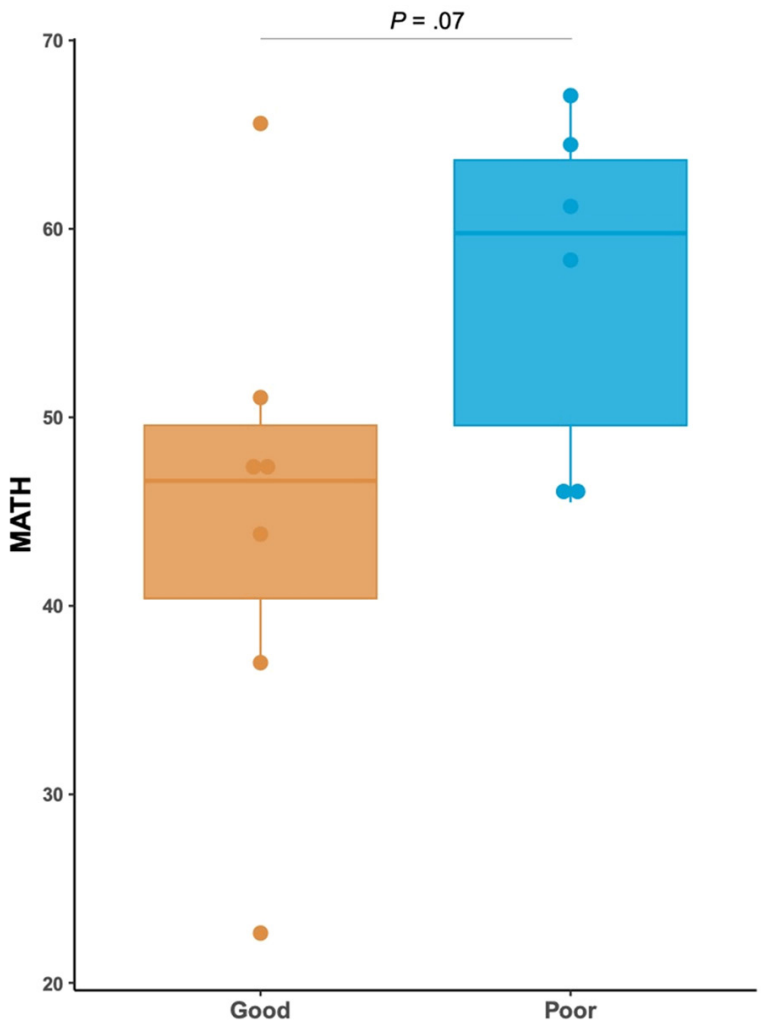



**Figure S1.** MATH values are compared between good and poor responders using one-tailed Mann-Whitney U test. Box plots show the first, median, and third quartiles, whiskers extend to 1.5 times the interquartile range, and each dot indicates one patient tumor.

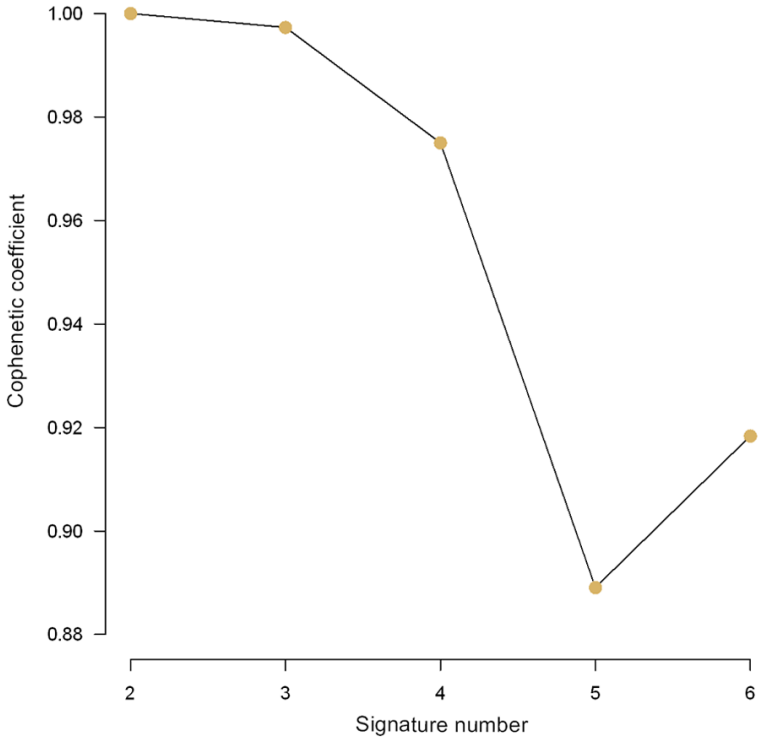



**Figure S2.** The elbow plot depicts the cophenetic correlation on the Y axis for varying signature numbers (X axis), where the correlation shows the first substantial decrease on signature number 3.

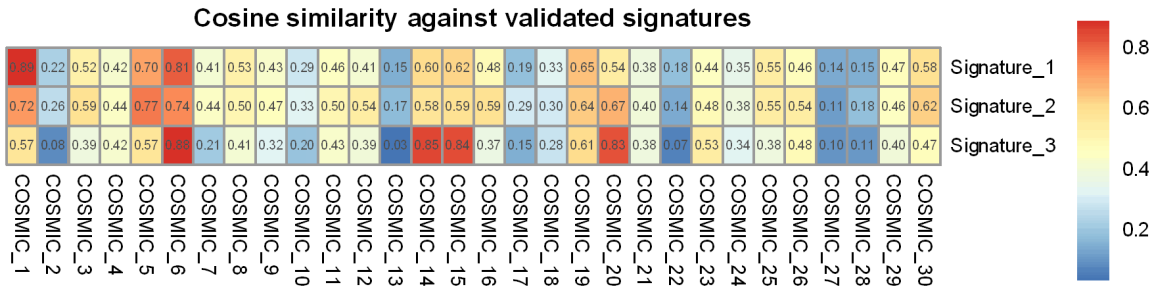



**Figure S3.** The row depicts cosine similarity between each mutational signature identified in our study and 30 COSMIC signatures. Shown in the cell is the specific similarity value, filled with colors from dark blue (lower similarity), to dark red (higher similarity).

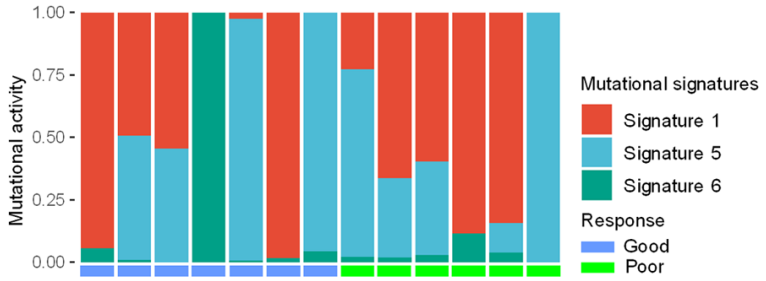



**Figure S4.** Filled bars depict the mutational signature activities of 13 patients stratified by response status, and the vertical axis denotes the proportion of mutations attributable to each mutational signature.

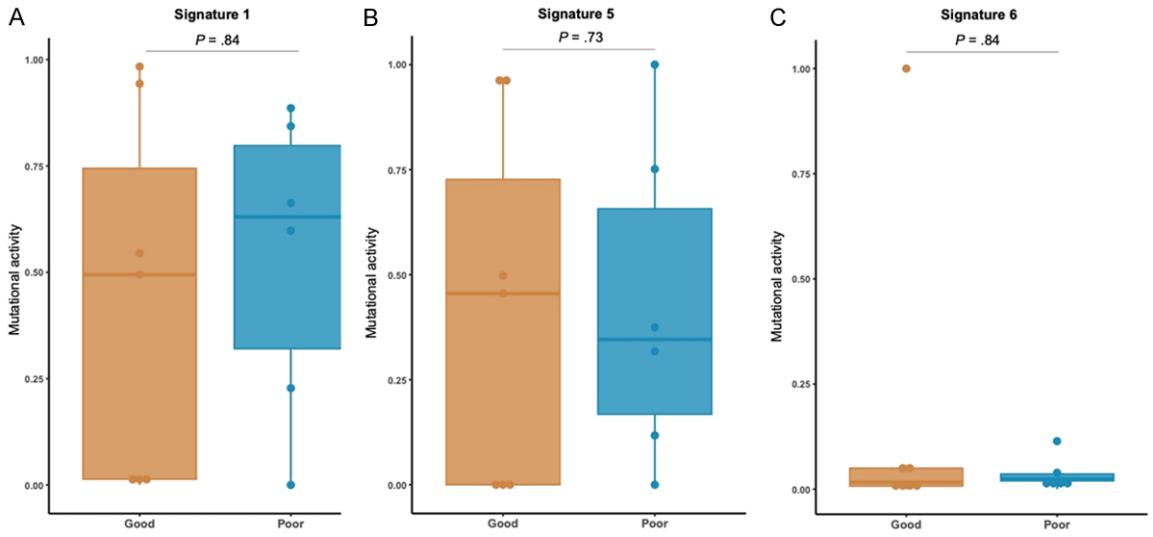
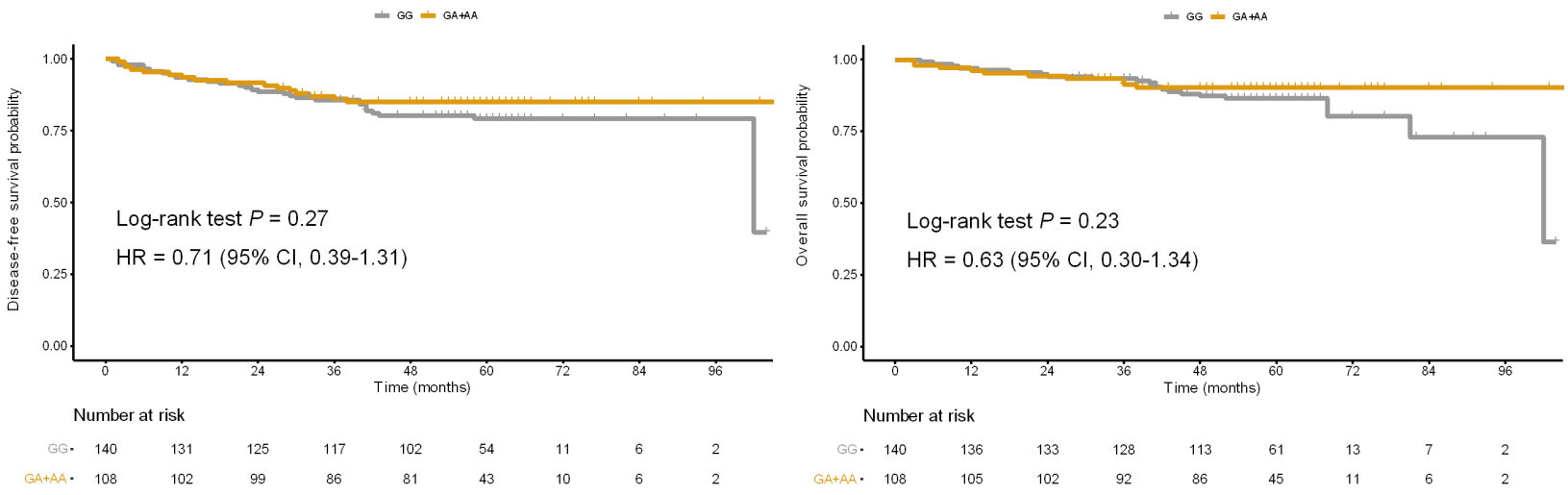
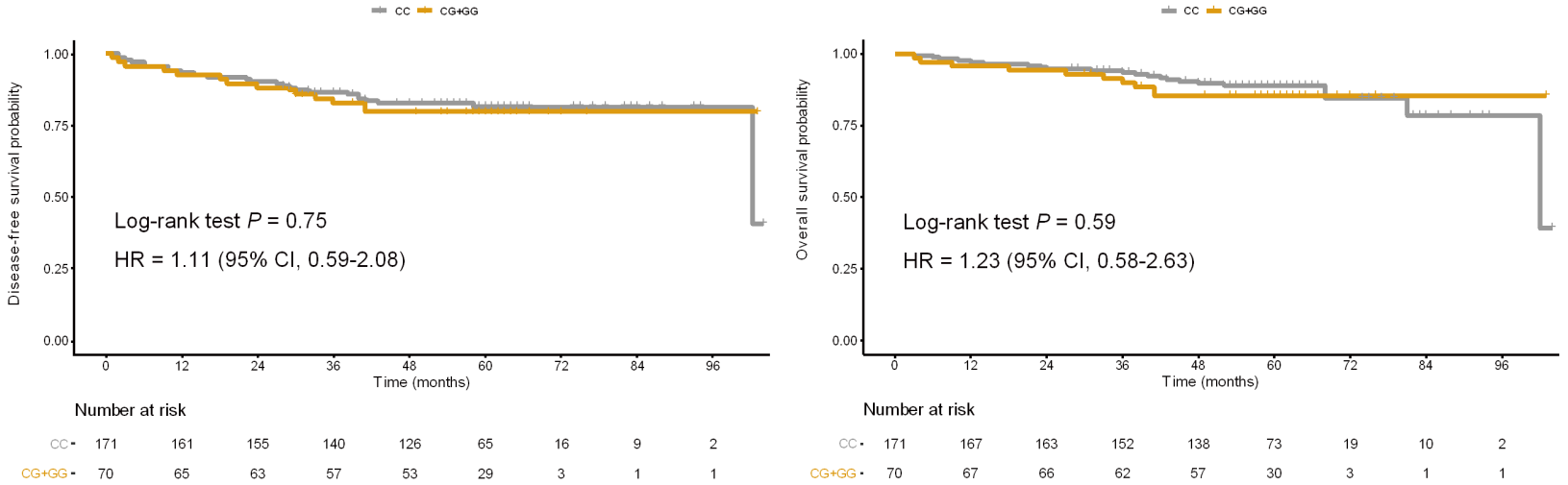



**Figure S5.** Comparison of mutational activities attributable to signature 1 associated with spontaneous deamination of 5-methylcytosine (A), signature 5 associated with unknown environmental exposure (B), and signature 6 associated with defective DNA mismatch repair (C) between good and poor responders. 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.

A


rs2272495
GG GA+AA
Disease-free survival probability
1.00
0.75
0.50
0.25
0.00
Log-rank test P = 0.27
HR = 0.71 (95% CI, 0.39-1.31)
0 12 24 36 48 60 72 84 96
Time (months)
Number at risk
GG 140 131 125 117 102 54 11 6 2
GA+AA 108 102 99 86 81 43 10 6 2
GG GA+AA
Overall survival probability
1.00
0.75
0.50
0.25
0.00
Log-rank test P = 0.23
HR = 0.63 (95% CI, 0.30-1.34)
0 12 24 36 48 60 72 84 96
Time (months)
Number at risk
GG 140 136 133 128 113 61 13 7 2
GA+AA 108 105 102 92 86 45 11 6 2

B


rs3790549
CC CG+GG
Disease-free survival probability
1.00
0.75
0.50
0.25
0.00
Log-rank test P = 0.75
HR = 1.11 (95% CI, 0.59-2.08)
0 12 24 36 48 60 72 84 96
Time (months)
Number at risk
CC 171 161 155 140 126 65 16 9 2
CG+GG 70 65 63 57 53 29 3 1 1
CC CG+GG
Overall survival probability
1.00
0.75
0.50
0.25
0.00
Log-rank test P = 0.59
HR = 1.23 (95% CI, 0.58-2.63)
0 12 24 36 48 60 72 84 96
Time (months)
Number at risk
CC 171 167 163 152 138 73 19 10 2
CG+GG 70 67 66 62 57 30 3 1 1

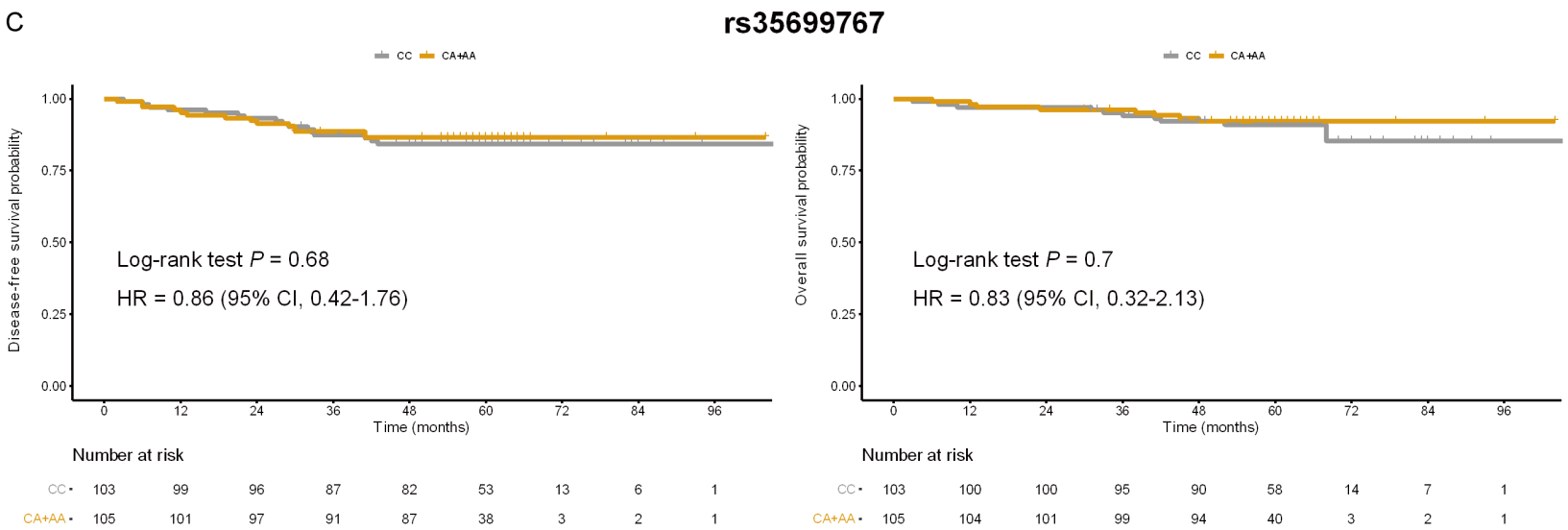



**Figure S6.** A. Kaplan-Meier curves of disease-free survival and overall survival in patients with versus without alternative A allele of rs2272495 in APPL2. B. Kaplan-Meier curves of disease-free survival and overall survival in patients with versus without alternative G allele of rs3790549 in WARS2. 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.