Gene signatures associated with drug resistance to irinotecan and oxaliplatin predict a poor prognosis in patients with colorectal cancer

XINRONG SUN[1*], XIANG WANG[1*], WENMING FENG[1], HUIHUI GUO[1],
CHENGWU TANG[1], YONGLIANG LU[2], XIAOBIN XIANG[3] and YING BAO[1]

[1]Department of Gastrointestinal Surgery, The First Affiliated Hospital, Huzhou Teachers College,
The First People's Hospital of Huzhou; [2]Department of Medicine, Huzhou Teachers College, Huzhou, Zhejiang 313000;
[3]Department of Sports Medicine, Guangdong Orthopedic Hospital, Guangzhou, Guangdong 510045, P.R. China

Abstract. The identification of novel survival predictors may help to improve the appropriate management of colorectal cancer (CRC). In the present study, two gene sets associated with irinotecan or oxaliplatin resistance in CRC cell lines were first identified and subsequently applied to the clinical CRC microarray dataset GSE14333. Subsequently, a 60-gene irinotecan resistance-associated signature and a 13-gene oxaliplatin resistance-associated signature were established, which were able to classify CRC patients into high- and low-risk subgroups with varied clinical outcomes [irinotecan-resistance gene signature: hazard ratio (HR)=0.4607, 95% confidence interval (CI)=0.3369-0.6300, P<0.0001; oxaliplatin-resistance gene signature: HR=0.6119, 95% CI=0.4547-0.8233, P=0.0008]. The performance of these two gene expression signatures in predicting outcome risk were also validated in two other independent CRC gene expression microarray datasets, GSE17536 (irinotecan-resistance gene signature: HR=0.5318, 95% CI=0.3359-0.8419, P=0.0079; oxaliplatin-resistance gene signature: HR=0.5383, 95% CI=0.3400-0.8521, P=0.0114) and GSE17537 (irinotecan-resistance gene signature: HR=0.2827, 95% CI=0.1173-0.6813, P=0.0088; oxaliplatin-resistance gene signature: HR=0.2378, 95% CI=0.09773-0.5784, P=0.0023). Furthermore, the combination of these two gene classifiers demonstrated a superior performance in CRC prognosis prediction than either used individually. Therefore, this study proposed novel gene classifier models for CRC prognosis prediction, which may be potentially useful to inform treatment decisions for patients with CRC in clinical settings.

Introduction

Regimens based on the anti-metabolite drug 5-fluorouracil (5-FU) combined with the topoisomerase I inhibitor irinotecan or the DNA-binding agent oxaliplatin, which is also known as oxaliplatin with 5-FU and folinic acid chemotherapy (FOLFOX), are typically used as the initial chemotherapeutic treatment for colorectal cancer (CRC) (1). Although an objective response to chemotherapeutic regimens significantly increases the survival of patients with CRC, a relatively high proportion (50-70%) of innate and acquired resistance is still a major clinical problem faced by these individuals (2). Therefore, predicting the risk of chemotherapy resistance and a poor prognosis is of important clinical significance for patients with CRC.

Common clinicopathological variables, including differentiation and nerve or vessel invasion, demonstrate a relatively weak predictive power for discriminating among CRC patients with varied risks of clinical outcomes (3). Although several molecular characteristics, including chromosomal instability, microsatellite instability and the CpG island methylator phenotype, have been used as prognosis predictors for CRC (4,5), these genetic aberrations are not directly responsible for the responses or resistance to chemotherapeutic drugs.

Based on mRNA expression patterns, a series of gene signatures have been developed for CRC, which are able to divide CRC patients into subclasses that present distinct prognostic profiles (6-8). Considering the association between chemotherapy efficiency and CRC prognosis, the present study hypothesized that drug resistance-associated genes may be used to build gene signature models to predict the clinical outcomes for patients with CRC.

In the present study, using the previously published CRC mRNA microarray expression datasets deposited in the Gene Expression Omnibus (GEO) database, two gene signatures associated with acquired resistance to irinotecan or oxaliplatin

Correspondence to: Dr Ying Bao, Department of Gastrointestinal Surgery, The First Affiliated Hospital, Huzhou Teachers College, The First People's Hospital of Huzhou, 158 Guangchanghou Road, Huzhou, Zhejiang 313000, P.R. China
E-mail: baoying2222@126.com

*Contributed equally

were established. These individual and combined gene signatures demonstrated a predictive power to stratify patients with CRC into good or poor survival groups.

Materials and methods

Datasets. Four gene expression microarray datasets for CRC deposited in the GEO database (https://www.ncbi.nlm.nih.gov/geo/) were used for data mining in the present study. Gene expression data for GSE42387 was obtained from three human colon cancer cell lines (HCT116, HT29 and LoVo) and their sub-cell lines with acquired resistance to irinotecan (active metabolite SN-38) or oxaliplatin (9). GSE42387 was used to identify irinotecan or oxaliplatin resistance-associated genes.

Three clinical CRC tissue gene expression microarray datasets were selected for the gene signature training and validation. Gene expression data of 226 CRC tissues with complete follow-up data from GSE14333 were used as a training dataset to establish a drug-resistance gene signature associated with clinical outcome (10). The gene expression data from two sets of CRC tissues, GSE17536 and GSE17537, were used as validation datasets to test the robustness of the gene signature in predicting survival (11,12). The clinical traits of patients in the three CRC cohorts are summarized in Table I.

Identification of individual genes associated with acquired irinotecan or oxaliplatin resistance. Differentially expressed genes between the parent and resistant subclass cell lines from the GSE42387 dataset were analyzed by performing the univariate F-test with a randomized variance model and false discovery rate correction for each gene. A permutation test (n=10,000) was performed for each significant gene. Genes for which $P<0.005$ were selected for further analysis. All the analyses were performed using BRB-ArrayTools software (version 4.5) developed by Dr. Richard Simon and the BRB-ArrayTools Development Team (National Institutes of Health, National Cancer Center, Bethesda, MD, USA) (13).

Construction of irinotecan or oxaliplatin resistance gene signatures associated with the survival of CRC. Significant genes identified as resistance-associated in the GSE42387 dataset were then applied to the GSE14333 training dataset involving clinical CRC samples. The univariate Cox proportional hazards model tool of BRB-ArrayTools was used to test for genes that significantly influenced patient survival (permutation time, 1,000). Genes with a significance threshold of 0.05 were selected. Hierarchical clustering of the CRC samples in the GSE14333 dataset was performed using the acquired resistance signatures.

Subsequently, gene signature models were developed based on a linear combination of the expression levels of survival-associated genes weighted by the regression coefficient derived from the univariate Cox regression analysis for the selected irinotecan- or oxaliplatin-resistance gene signatures. The patient groups were divided into high- or low-risk subgroups based on their risk scores being above or below the median value across all samples, respectively. The prediction accuracy of the gene classifiers was estimated by leave-one-out cross-validation.

Table I. Clinical traits of patients in the three colorectal cancer datasets.

Characteristics	GSE14333	GSE17536	GSE17537
Age, years			
<60	61	59	24
≥60	165	118	31
Gender			
Male	120	96	26
Female	106	80	29
Location			
Left	93	N.A.	N.A.
Right	101	N.A.	N.A.
Unknown	32	N.A.	N.A.
Grade			
1	N.A.	16	1
2	N.A.	134	32
3	N.A.	27	3
Unknown	N.A.	0	19
Stage			
1	41	24	4
2	94	57	15
3	91	57	19
4	0	39	17

External validation of the models using expression data in two independent datasets of clinical CRC samples. The established gene classifiers were validated using datasets from two independent cohorts of patients with CRC (GSE17536 and GSE17537). Gene expression profiles were used to predict survival classification (high- or low-risk) using the same models developed in the training sets.

Effect of combined drug resistance gene signatures on predicting clinical outcomes in CRC datasets. The predictive results of irinotecan- and oxaliplatin-resistance gene signatures, respectively, were combined and the patients of the training and validation datasets were further dichotomized into four sub-classes according to survival risk as follows: High-/high-, low-/high-, high-/low-, and low-/low- subgroups.

Statistical analysis. Categorical variables were compared using the chi-squared test. Except for analyses that were performed using BRB-ArrayTools, all other statistical analyses were performed using MedCalc® software (version 8.1; MedCalc® software, Ostend, Belgium). Survival between groups was compared using the Kaplan-Meier method, and the differences were assessed using the log-rank test. $P<0.05$ was considered to indicate a statistically significant difference.

Results

Determination of irinotecan- and oxaliplatin-resistance gene signatures for CRC using microarray datasets. Using the microarray results of the irinotecan- and oxaliplatin-resistant

Table II. Irinotecan-resistance genes associated with the survival of patients with colorectal cancer in the GSE14333 dataset.

No. identified	Gene symbol	Accession no.	Parametric P-value	FDR	Permutation P-value	Hazard ratio	SD of log intensities
1	LIMS2	NM_001136037	0.0001736	0.0165	0.0002	0.642	0.601
2	TUBA1B	NM_006082	0.0005244	0.0249	0.0003	5.428	0.168
3	KLHDC2	NM_014315	0.0008015	0.0254	0.0008	0.484	0.362
4	PDGFC	NM_016205	0.0018540	0.0412	0.0020	0.740	0.882
5	TGFB1I1	NM_001042454	0.0022273	0.0412	0.0030	0.753	0.836
6	ARHGAP24	NM_001025616	0.0029117	0.0412	0.0040	0.754	0.803
7	PPAP2A	NM_003711	0.0030347	0.0412	0.0029	0.666	0.620
8	DNASE2B	NM_021233	0.0071416	0.0848	0.0064	1.159	1.411
9	KIRREL	NM_001286349	0.0092085	0.0972	0.0088	0.713	0.618
10	EDA2R	NM_001199687	0.0122321	0.1160	0.0132	1.214	0.995
11	VWA5B1	NM_001039500	0.0145063	0.1250	0.0132	1.203	1.002
12	DAPK1	NM_001288729	0.0320929	0.2540	0.0328	0.792	0.752
13	FBXL21	NM_012159	0.0422906	0.3090	0.0416	0.881	1.270

FDR, false discovery rate; SD, standard deviation.

sublines and the parent cell lines from the GSE42387 dataset, 292 and 103 genes were identified to be significantly dysregulated in irinotecan- and oxaliplatin-resistant CRC cell lines, respectively, using a univariate F-test with a randomized variance model and false discovery rate correction in BRB-ArrayTools.

Subsequently, these resistance-associated genes were applied to the training cohort of clinical biopsies of CRC tumors (GSE14333). Using the Cox regression analysis, a 60-gene signature for irinotecan resistance and a 13-gene signature for oxaliplatin resistance were generated, which were associated with the poor survival of patients with CRC in the GSE14333 dataset (Tables II and III).

Using these two filtered gene signatures, a survival risk score system was developed by calculating a linear combination of irinotecan- or oxaliplatin-gene signature expression values weighted by their Cox regression coefficients. Based on the risk scores evaluated by the irinotecan- and oxaliplatin-resistance gene signatures, patients with CRC from the GSE14333 dataset were divided into two subclasses (high- and low-risk) according to the median values of the risk scores.

The prognostic significance of irinotecan- and oxaliplatin-resistance gene signatures for CRC. As presented in Fig. 1A and B, the high-risk groups defined by the two resistance gene signatures in the GSE14333 dataset had significantly shorter overall survival times than the low-risk groups [irinotecan-resistance gene signature: hazard ratio (HR)=0.4607, 95% confidence interval (CI)=0.3369-0.6300, log-rank P<0.0001; oxaliplatin-resistance gene signature: HR=0.6119, 95% CI=0.4547-0.8233, P=0.0008]. Furthermore, the irinotecan- or oxaliplatin-resistance gene sets were able to hierarchically cluster the GSE14333 CRC tumors into two subgroups with varying distributions of outcome risk in an unsupervised manner (Fig. 2).

Validation of the prognostic value of irinotecan- and oxaliplatin-resistance gene signatures in independent CRC cohorts. To further investigate the clinical relevance and evaluate the predictive value of the developed gene signature models, a preliminary evaluation in two independent cohorts of patient samples was performed. As presented in Fig. 1C-F, the robustness of the irinotecan- and oxaliplatin resistance-associated gene signatures may be also demonstrated in two large independent cohorts of patients with CRC: GSE17536 (irinotecan-resistance gene signature: HR=0.5318, 95%=0.3359-0.8419, P=0.0079; oxaliplatin-resistance gene signature: HR=0.5383, 95% CI=0.3400-0.8521, P=0.0114) and GSE17537 (irinotecan-resistance gene signature: HR=0.2827, 95% CI=0.1173-0.6813, P=0.0088; oxaliplatin-resistance gene signature: HR=0.2378, 95% CI=0.09773-0.5784, P=0.0023).

Combination of irinotecan- and oxaliplatin-resistance gene signatures for the prognosis of CRC. The combination of the two signatures described above may further add prognostic power for the prediction of the survival of patients with CRC in training and validation datasets. As presented in Fig. 3, high-risk subgroups predicted by the two models had the poorest clinical outcomes in all three CRC cohorts.

Discussion

In the present study, two sets of genes that may be used clinically to predict the prognosis in three independent CRC cohorts were identified from a series of experimental data on chemotherapy resistance. One gene set included 60 genes associated with irinotecan-resistance, while another gene set included 13 genes associated with oxaliplatin resistance. These gene lists were compared with the previously identified gene signature for CRC, revealing that cyclin K (CCNK) in the oxaliplatin-resistance signature in the present study was also identified in a 13-gene prognostic signature (ColoGuideEx) for stage II CRC
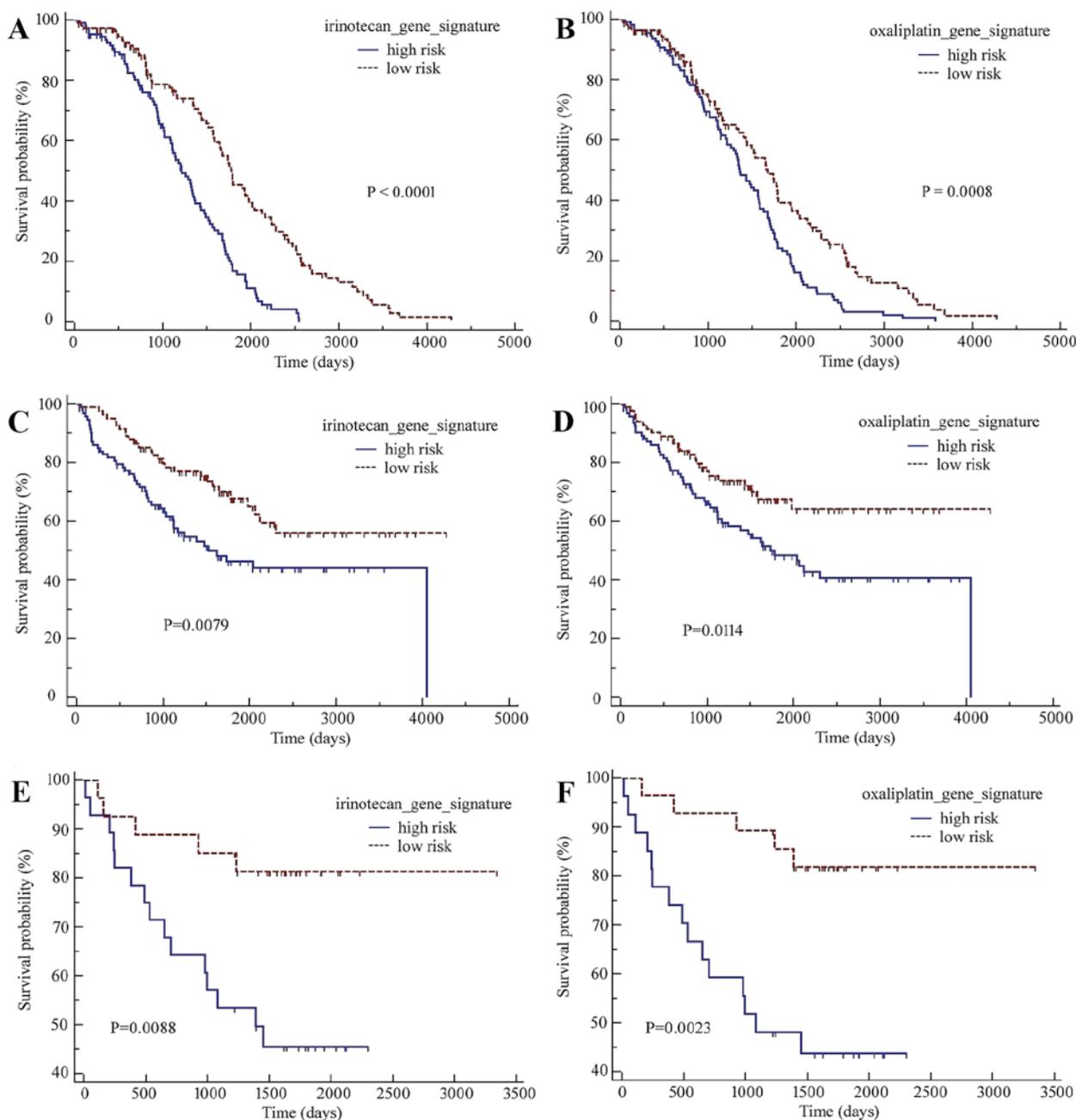
Figure 1. Kaplan-Meier curves of the low-risk and high-risk subgroups defined by the irinotecan-resistance and oxaliplatin-resistance gene signatures. (A) Irinotecan-resistance signature. (B) Oxaliplatin-resistance signature. The GSE17536 dataset's (C) irinotecan- and (D) oxaliplatin-resistance signatures. The GSE17537 dataset's (E) irinotecan- and (F) oxaliplatin-resistance signatures. These two gene signatures were able to stratify patients with colorectal cancer into low- and high-risk subgroups with significant differences in terms of the prognosis in the GSE14333 training dataset.

prognosis (14), while CCNK in the irinotecan-resistance signature was included in a 42-gene signature predictive for sensitivity to radiochemotherapy in patients with locally advanced rectal cancer (15). Using Gene Set Enrichment Analysis, the present study identified that the 60-gene irinotecan-resistance signature significantly overlapped with genes that are dysregulated in numerous epithelial cancer cell lines overexpressing an oncogenic form of the KRAS GTPase proto-oncogene. It was additionally identified that these gene signatures were independent of other factors that affect the outcome of patients with CRC. Furthermore, the combination of these two novel signatures was shown to further improve their robustness in predicting the survival of patients with CRC.

Except for clinicopathological characteristics, responses to standard chemotherapy regimens, particularly standard first-line therapies such as irinotecan and oxaliplatin, are an important determinant affecting the clinical outcome of patients with CRC. Antitumor activity in irinotecan- and oxaliplatin-sensitive or refractory CRC is determined by the existing diverse gene expression patterns (9). Therefore, the results of the present study support the hypothesis that the patterns of expression of numerous drug-resistance genes may be successful in distinguishing between improved and poor outcomes for patients with CRC. However, different from other resistance-associated signatures for CRCs, these two gene classifiers are currently used to predict overall survival

Table III. Oxaliplatin resistance genes associated with the survival of patients with colorectal cancer in the GSE14333 dataset.

No. identified	Gene symbol	Accession no.	Parametric P-value	FDR	Permutation P-value	Hazard ratio	SD of log intensities
1	MMP16	NM_005941	0.0000037	0.000951	<1e-07	0.659	0.814
2	GPHA2	NM_130769	0.0000248	0.00284	<1e-07	1.425	0.894
3	IRX1	NM_024337	0.0000331	0.00284	0.0001	0.729	1.120
4	FKBP6	NM_001135211	0.000192	0.01230	0.0001	0.710	0.829
5	RAB17	NM_022449	0.0002755	0.01420	0.0005	1.752	0.570
6	MYH7B	NM_020884	0.0004024	0.01720	0.0003	0.855	1.728
7	TLX2	NM_016170	0.000472	0.01730	0.0003	0.832	1.44
8	HTR4	NM_000870	0.000621	0.01990	0.0009	0.737	0.700
9	PIK3R5	NM_001142633	0.0007926	0.02260	0.0007	2.057	0.345
10	SCGB1D1	NM_006552	0.000989	0.02540	0.0014	0.805	1.177
11	P2RX3	NM_002559	0.0013767	0.03070	0.0017	0.825	1.326
12	ID4	NM_001546	0.0014566	0.03070	0.0012	0.788	1.090
13	A2M	NM_000014	0.0017583	0.03070	0.0010	0.681	0.671
14	IPMK	NM_152230	0.0017827	0.03070	0.0020	1.565	0.559
15	TEX13A	NM_031274	0.0017903	0.03070	0.0022	0.817	1.120
16	CRYBB3	NM_004076	0.0020744	0.03330	0.0021	1.961	0.361
17	GPR112	NM_153834	0.002965	0.04480	0.0033	0.797	1.030
18	C12orf49	NM_024738	0.0031541	0.04500	0.0038	2.279	0.321
19	ECEL1	NM_004826	0.004649	0.06290	0.0056	0.736	0.76
20	CHADL	NM_138481	0.0055728	0.06590	0.0063	0.778	0.782
21	CCNK	NM_001099402	0.0056307	0.06590	0.0056	0.576	0.389
22	GAP43	NM_001130064	0.005916	0.06590	0.0075	0.767	0.809
23	TSPAN7	NM_004615	0.0060566	0.06590	0.0058	0.823	1.058
24	NOB1	NM_014062	0.0061516	0.06590	0.0059	1.691	0.403
25	SYT12	NM_001177880	0.0081639	0.08390	0.0089	1.203	1.114
26	NTM	NM_001048209	0.0087813	0.08680	0.0091	0.854	1.245
27	SYCE2	NM_001105578	0.0095563	0.09100	0.0093	0.830	1.102
28	PRKACG	NM_002732	0.0105488	0.09680	0.0119	0.834	1.127
29	RNF146	NM_001242844	0.0109557	0.09710	0.0124	0.593	0.384
30	KCNMA1	NM_001014797	0.0114415	0.09720	0.0121	0.855	1.245
31	ETS1	NM_001143820	0.0120369	0.09720	0.0112	0.707	0.59
32	EDA2R	NM_001199687	0.0122321	0.09720	0.0132	1.214	0.995
33	ADARB2-AS1	NM_001098830	0.012478	0.09720	0.0121	0.815	0.858
34	RCBTB1	NM_018191	0.0136522	0.10200	0.0132	1.459	0.502
35	SNAI2	NM_003068	0.0144012	0.10200	0.0139	0.781	0.826
36	VWA5B1	NM_001039500	0.0145063	0.10200	0.0132	1.203	1.002
37	PXDN	NM_012293	0.0147868	0.10200	0.0140	0.746	0.715
38	DUOX2	NM_014080	0.0155175	0.10200	0.0150	0.904	1.973
39	ADAMTS20	NM_025003	0.0155347	0.10200	0.0160	1.268	0.844
40	FGF11	NM_004112	0.0183579	0.11800	0.0171	0.791	0.759
41	ATXN1L	NM_001137675	0.0191998	0.12000	0.0179	0.552	0.312
42	MUC16	NM_024690	0.0216876	0.12700	0.0231	1.165	1.141
43	APOBEC3B	NM_001270411	0.0220595	0.12700	0.0228	1.226	0.917
44	RGS2	NM_002923	0.0222933	0.12700	0.0217	0.838	1.032
45	DUSP1	NM_004417	0.0225976	0.12700	0.0215	0.810	0.863
46	CHRNA1	NM_000079	0.0228582	0.12700	0.0240	0.806	0.774
47	MFGE8	NM_001114614	0.0232019	0.12700	0.0233	0.819	0.878
48	APOL5	NM_030642	0.026395	0.14100	0.0255	0.746	0.536
49	HAS2	NM_005328	0.0326012	0.17100	0.0324	0.803	0.737
50	AUH	NM_001698	0.0354341	0.18200	0.0335	1.703	0.317
51	HAUS4	NM_001166269	0.0369736	0.18600	0.0360	1.445	0.400

Table III. Continued.

No. identified	Gene symbol	Accession no.	Parametric P-value	FDR	Permutation P-value	Hazard ratio	SD of log intensities
52	ZFP57	NM_001109809	0.0377616	0.18700	0.0359	1.203	0.882
53	BOLL	NM_001284358	0.040459	0.19600	0.0423	1.355	0.529
54	XPR1	NM_001135669	0.0415493	0.19800	0.0422	1.449	0.413
55	MLXIP	NM_014938	0.0423385	0.19800	0.0430	0.770	0.575
56	CEACAM19	NM_001127893	0.0462696	0.21200	0.0461	1.118	1.423
57	KLF12	NM_007249	0.0481573	0.21400	0.0482	0.784	0.665
58	BLVRB	NM_000713	0.0482109	0.21400	0.0474	1.286	0.588
59	GLTPD1	NM_001029885	0.0496136	0.214	0.0478	1.325	0.553
60	EID2	NM_153232	0.0498994	0.214	0.0528	0.605	0.309

FDR, false discovery rate; SD, standard deviation.
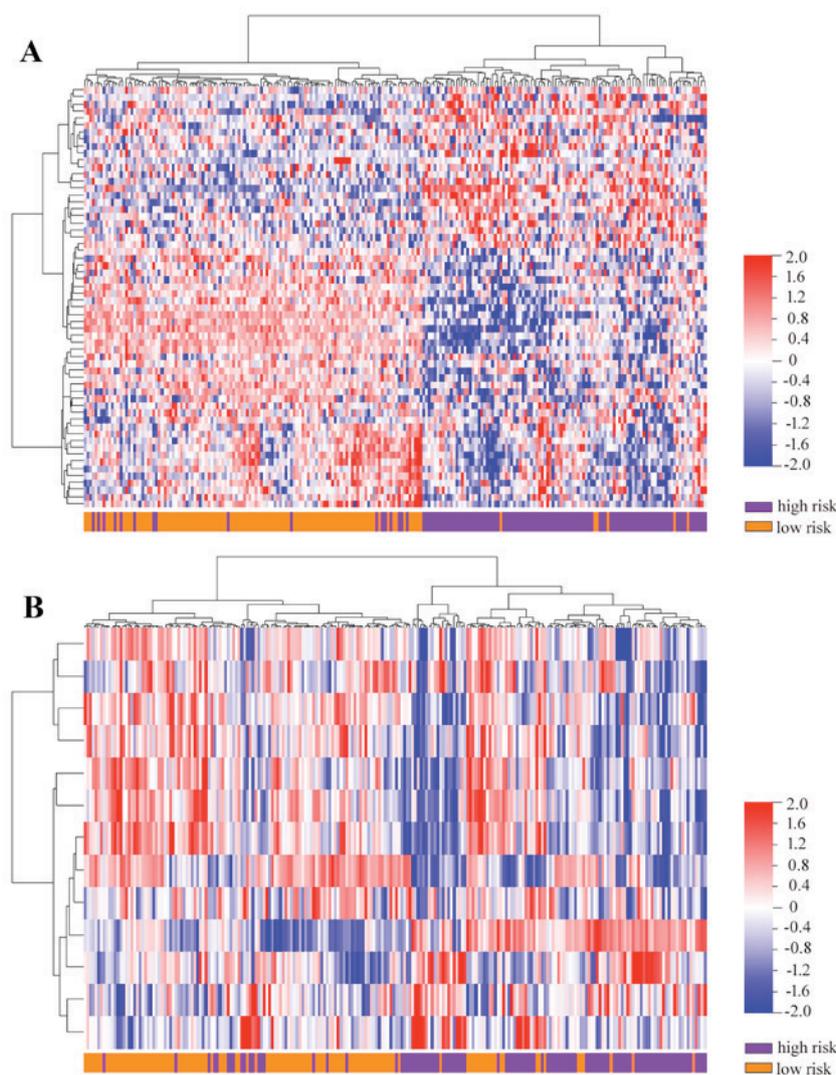
Figure 2. Unsupervised hierarchical clustering of the colorectal cancer samples from the GSE14333 dataset using drug resistance-associated gene signatures. The (A) irinotecan-resistance and (B) oxaliplatin-resistance gene signatures were able to hierarchically cluster the samples into two subgroups with varied distributions of outcome risk.

for CRC patients, but not to distinguish between sensitive and resistant tumors (9). The utility of the signatures identified in the present study for predicting chemotherapy activity remain to be clarified.
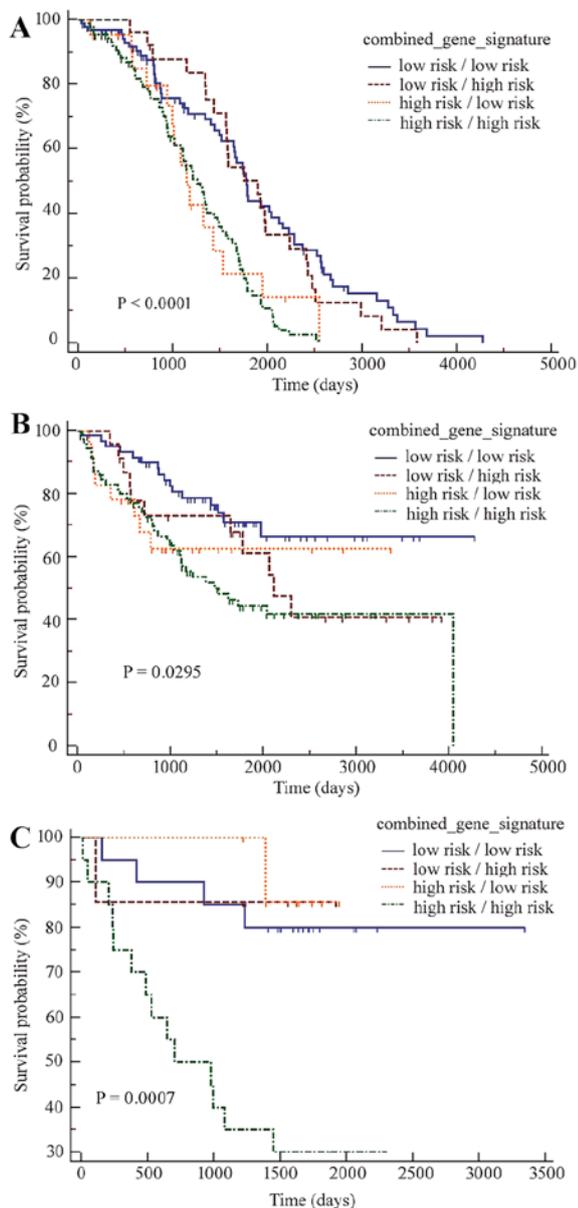
Figure 3. Kaplan-Meier curves based on the combined drug resistance gene signatures for patients with colorectal cancer in the three cohorts. In all three datasets, (A) GSE14333, (B) GSE17536 and (C) GSE17537, patients predicted as high-risk by the two gene signatures had the poorest clinical outcomes.

It may be hypothesized that individual genes in the two resistance signatures in the present study have functional significance in the process of developing irinotecan or oxaliplatin resistance. Among the two gene lists, transforming growth factor-β-induced transcript 1 has been validated as a potential biomarker to predict the effects of FOLFOX4 chemotherapy in patients with CRC (16). Tubulin α-1B chain (TUBA1B) in the oxaliplatin-resistance signature and inhibitor of DNA binding 4 (ID4), Ets-1, mucin 16, regulator of G-protein signaling 2, dual specificity protein phosphatase 4 and milk fat globule-EGF factor 8 in the irinotecan-resistance signature, have been identified as prognostic factors for CRC survival in previous studies (17-22). Functionally, TUBA1B was revealed to be associated with resistance to paclitaxel in patients with hepatocellular carcinoma (23). ID4 may partici-pate in the chemoresistance associated with gain-of-function

mutations in p53 (24). In particular, the two genes ectodys-plasin A2 Receptor (EDA2R) and Von Willebrand factor A domain-containing 5B1 (VWA5B1) overlap between the irinotecan resistance- and oxaliplatin resistance-gene signa-tures. EDA2R, also termed XEDAR and tumor necrosis factor receptor superfamily member 27, is a member of the tumor necrosis factor receptor super-family, and is a novel p53 target with an important role in colorectal carcinogenesis (25-27). The function of the VWA5B1 gene remains unknown and its association with chemotherapy resistance has not yet been investigated. The results of the present study indicate that the roles of EDA2R, VWA5B1 and other gene signatures in the development of irinotecan or oxaliplatin resistance warrant further experimental investigation.

In conclusion, the present study developed two sets of gene signatures from chemoresistance-associated experimental microarray data that were able to effectively and reproducibly classify CRC tumors according to poor or improved prognosis. These findings highlight the importance of chemotherapy resistance in determining the prognosis of patients with CRC, and led the present study to propose that chemotherapy resistance-associated genes may be a novel source for the establishment of gene classifiers to categorize CRC into subgroups with varied clinical outcomes. Furthermore, the functional significance of these novel gene sets in developing drug resistance in CRC also requires further investigation.

Acknowledgements

References

1. Gustavsson B, Carlsson G, Machover D, Petrelli N, Roth A, Schmoll HJ, Tveit KM and Gibson F: A review of the evolution of systemic chemotherapy in the management of colorectal cancer. Clin Colorectal Cancer 14: 1-10, 2015.
2. Wang F, Liu R, Lee SW, Sloss CM, Couget J and Cusack JC: Heparin-binding EGF-like growth factor is an early response gene to chemotherapy and contributes to chemotherapy resis-tance. Oncogene 26: 2006-2016, 2007.
3. Bianchi P, Laghi L, Delconte G and Malesci A: Prognostic value of colorectal cancer biomarkers. Cancers (Basel) 3: 2080-2105, 2011.
4. Lao VV and Grady WM: Epigenetics and colorectal cancer. Nat Rev Gastroenterol Hepatol 8: 686-700, 2011.
5. Simons CC, Hughes LA, Smits KM, Khalid-de Bakker CA, de Bruïne AP, Carvalho B, Meijer GA, Schouten LJ, van den Brandt PA, Weijenberg MP and van Engeland M: A novel classification of colorectal tumors based on microsatellite instability, the CpG island methylator phenotype and chromo-somal instability: Implications for prognosis. Ann Oncol 24: 2048-2056, 2013.
6. Wang L, Shen X, Wang Z, Xiao X, Wei P, Wang Q, Ren F, Wang Y, Liu Z, Sheng W, et al: A molecular signature for the prediction of recurrence in colorectal cancer. Mol Cancer 14: 22, 2015.
7. Hur K, Toiyama Y, Schetter AJ, Okugawa Y, Harris CC, Boland CR and Goel A: Identification of a metastasis-specific MicroRNA signature in human colorectal cancer. J Natl Cancer Inst 107: pii:dju492, 2015.
8. Hu Y, Chen HY, Yu CY, Xu J, Wang JL, Qian J, Zhang X and Fang JY: A long non-coding RNA signature to improve prognosis prediction of colorectal cancer. Oncotarget 5: 2230-2242, 2014.

9. Jensen NF, Stenvang J, Beck MK, Hanáková B, Belling KC, Do KN, Viuff B, Nygård SB, Gupta R, Rasmussen MH, *et al*: Establishment and characterization of models of chemotherapy resistance in colorectal cancer: Towards a predictive signature of chemoresistance. Mol Oncol 9: 1169-1185, 2015.

10. Oh SC, Park YY, Park ES, Lim JY, Kim SM, Kim SB, Kim J, Kim SC, Chu IS, Smith JJ, *et al*: Prognostic gene expression signature associated with two molecularly distinct subtypes of colorectal cancer. Gut 61: 1291-1298, 2012.

11. Smith JJ, Deane NG, Wu F, Merchant NB, Zhang B, Jiang A, Lu P, Johnson JC, Schmidt C, Bailey CE, *et al*: Experimentally derived metastasis gene expression profile predicts recurrence and death in patients with colon cancer. Gastroenterology 138: 958-968, 2010.

12. Freeman TJ, Smith JJ, Chen X, Washington MK, Roland JT, Means AL, Eschrich SA, Yeatman TJ, Deane NG and Beauchamp RD: Smad4-mediated signaling inhibits intestinal neoplasia by inhibiting expression of β-catenin. Gastroenterology 142: 562-571.e2, 2012.

13. Simon R, Lam A, Li MC, Ngan M, Menenzes S and Zhao Y: Analysis of gene expression data using BRB-Array Tools. Cancer Inform 3: 11-17, 2007.

14. Agesen TH, Sveen A, Merok MA, Lind GE, Nesbakken A, Skotheim RI and Lothe RA: ColoGuideEx: A robust gene classifier specific for stage II colorectal cancer prognosis. Gut 61: 1560-1567, 2012.

15. Rimkus C, Friederichs J, Boulesteix AL, Theisen J, Mages J, Becker K, Nekarda H, Rosenberg R, Janssen KP and Siewert JR: Microarray-based prediction of tumor response to neoadjuvant radiochemotherapy of patients with locally advanced rectal cancer. Clin Gastroenterol Hepatol 6: 53-61, 2008.

16. Li S, Lu X, Chi P and Pan J: Identification of Nkx2-3 and TGFB1I1 expression levels as potential biomarkers to predict the effects of FOLFOX4 chemotherapy. Cancer Biol Ther 13: 443-449, 2012.

17. Umetani N, Takeuchi H, Fujimoto A, Shinozaki M, Bilchik AJ and Hoon DS: Epigenetic inactivation of ID4 in colorectal carcinomas correlates with poor differentiation and unfavorable prognosis. Clin Cancer Res 10: 7475-7483, 2004.

18. Tokuhara K, Ogata Y, Nakagawa M and Shirouzu K: Ets-1 expression in vascular endothelial cells as an angiogenic and prognostic factor in colorectal carcinoma. Int Surg 88: 25-33, 2003.

19. Streppel MM, Vincent A, Mukherjee R, Campbell NR, Chen SH, Konstantopoulos K, Goggins MG, Van Seuningen I, Maitra A and Montgomery EA: Mucin 16 (cancer antigen 125) expression in human tissues and cell lines and correlation with clinical outcome in adenocarcinomas of the pancreas, esophagus, stomach, and colon. Hum Pathol 43: 1755-1763, 2012.

20. Jiang Z, Wang Z, Xu Y, Wang B, Huang W and Cai S: Analysis of RGS2 expression and prognostic significance in stage II and III colorectal cancer. Biosci Rep 30: 383-390, 2010.

21. De Vriendt V, De Roock W, Di Narzo AF, Tian S, Biesmans B, Jacobs B, Budinska E, Sagaert X, Rossi S, D'Ario G, *et al*: DUSP 4 expression identifies a subset of colorectal cancer tumors that differ in MAPK activation, regardless of the genotype. Biomarkers 18: 516-524, 2013.

22. Jinushi T, Shibayama Y, Kinoshita I, Oizumi S, Jinushi M, Aota T, Takahashi T, Horita S, Dosaka-Akita H and Iseki K: Low expression levels of microRNA-124-5p correlated with poor prognosis in colorectal cancer via targeting of SMC4. Cancer Med 3: 1544-1552, 2014.

23. Lu C, Zhang J, He S, Wan C, Shan A, Wang Y, Yu L, Liu G, Chen K, Shi J, *et al*: Increased α-tubulin1b expression indicates poor prognosis and resistance to chemotherapy in hepatocellular carcinoma. Dig Dis Sci 58: 2713-2720, 2013.

24. Dell'Orso S, Ganci F, Strano S, Blandino G and Fontemaggi G: ID4: A new player in the cancer arena. Oncotarget 1: 48-58, 2010.

25. Tanikawa C, Ri C, Kumar V, Nakamura Y and Matsuda K: Crosstalk of EDA-A2/XEDAR in the p53 signaling pathway. Mol Cancer Res 8: 855-863, 2010.

26. Tanikawa C, Furukawa Y, Yoshida N, Arakawa H, Nakamura Y and Matsuda K: XEDAR as a putative colorectal tumor suppressor that mediates p53-regulated anoikis pathway. Oncogene 28: 3081-3092, 2009.

27. Wong KK, Izaguirre DI, Kwan SY, King ER, Deavers MT, Sood AK, Mok SC and Gershenson DM: Poor survival with wild-type TP53 ovarian cancer? Gynecol Oncol 130: 565-569, 2013.