SOX9, miR-495, miR-590-3p, and miR-320d were identified as chemoradiotherapy-sensitive genes and miRNAs in colorectal cancer patients based on a microarray dataset

B. DU[1], T. WANG[1], X. YANG[1], J. WANG[2], X. SHI[1], X. WANG[1], D. WU[1], L. FENG[1], L. CHEN[1], W. ZHANG[1,*]

[1]Department of Anorectal Surgery, Gansu Provincial Hospital, Lanzhou, 730000, China; [2]Department of Radiotherapy, Gansu Provincial Hospital, Lanzhou, 730000, China

*Correspondence: zhangweish2016@hotmail.com

We aimed to identify chemoradiotherapy (CRT)-sensitive biomarkers in colorectal cancer (CRC) patients. The GSE15781 dataset used in this study contains 42 samples: 22 CRC tissues (non-CRT: n=13; CRT: n=9) and 20 normal colorectal tissues (non-CRT: n=10; CRT: n=10). Following pretreatment, differentially expressed genes were selected using the limma package. Potential CRT-sensitive genes were identified with Venn analysis and then enriched in function and pathway clusters using the DAVID online tool. Moreover, protein-protein interaction (PPI) network analysis was implemented using the STRING database. The TRRUST database was used to establish a transcription factor (TF)-target transcriptional network. A miRNA–mRNA network was constructed based on relevant databases. miRNA and mRNA expression levels were analyzed using real-time quantitative PCR. A group of 259 candidate CRT-sensitive genes were identified that were mainly enriched in cell cycle regulation, adhesion-associated processes, and the p53 signaling pathway. A PPI network was established that contained striking nodes, including ITGA2, MYC, ESR1, and dihydropyrimidine dehydrogenase (DPYD), among which ESR1 was linked to MYC, and the two nodes were also highlighted in the TF-target regulation network. SRY-box 9 (SOX9) was another key TF. Hsa-miR-590-3p, hsa-miR-495, hsa-miR-320c, and hsa-miR-320d were predominant in the miRNA–mRNA network. Expression levels of SOX9, DPYD mRNA, miR-495, and miR-590-3p were clearly reduced after X-ray treatment in irradiated HT-29 cells, whereas that of miR-320d was notably enhanced. *SOX9* may be a CRT-sensitive gene in CRC patients, and hsa-miR-590-3p, hsa-miR-495, and hsa-miR-320d may be CRT-sensitive microRNAs in CRC patients. Therefore, *SOX9*, hsa-miR-590-3p, hsa-miR-495, and hsa-miR-320d may be used as sensitive biomarkers in CRC patients..

Key words: colorectal cancer, chemoradiotherapy, p53 signaling, protein–protein interaction, transcription factor

Colorectal cancer (CRC) is the fourth leading cause of cancer-related death globally and the third in the USA [1, 2], with an estimated 50,310 deaths and 136,830 new cases in 2014, accounting for 8% and 9% of all cancer cases, respectively [3]. However, the incidence of CRC has recently decreased in several countries because of improvements in early detection methods, such as sigmoidoscopy and colonoscopy [4, 5]. Successful surgery and the use of advanced neoadjuvant therapies also likely contributed to the decreased mortality and improved survival rates.

At present, standardized surgery following preoperative chemoradiotherapy (CRT) with 5-fluorouracil (5-FU) is the standard therapy for advanced CRC [6]. However, several studies have reported that after a follow-up period of 46 months, the survival benefit of preoperative CRT diminishes compared with that of postoperative CRT [7]. Therefore, it is important to elucidate molecular mechanisms that are activated after preoperative CRT administration.

Tumor responses to multiple treatments differ among reports, from complete resistance to complete regression [7, 8]. Thus, the identification of CRT-sensitive molecular biomarkers is important to establish clinical guidelines for CRC treatment. A recent study established an *in vitro* model using 12 CRC cell lines exposed to 5-FU-based CRT and found distinct sensitivities of these cell lines to CRT. Using a linear model analysis, the study identified 4,796 CRT-sensitive genes, including *STAT3*, *DOK3*, and *ERBB2*, which are related to cell cycle regulation or involved in the Wnt

signaling pathway [9]. The use of gene expression profiling (GEP) from 52 rectal cancer patients receiving preoperative radiotherapy identified 33 crucial genes differentially expressed among responders and non-responders, and most were predicted to be associated with cell adhesion, proliferation, and apoptosis [10]. Despite these valuable findings, neither of these studies considered normal controls. A recent GEP study by Snipstad et al. on 21 rectal cancer patients with resectable adenocarcinomas, which included 10 patients who received preoperative CRT, to determine the different effects of CRT on gene expression profiles in CRC tumors and normal controls identified multiple differentially expressed genes (DEGs) that were mainly associated with cell adhesion and leukocyte transendothelial migration [11]. Nevertheless, potential correlations of these genes, especially at the transcriptional and protein levels, were not investigated.

MicroRNAs (miRNAs/miRs) play important roles in responses to CRT for locally advanced rectal cancer [12]. To date, several key miRNAs have been identified to be dysregulated in CRC in response to CRT, including miR-125b, miR-137, miR-21, and miR-143 [13, 14]. Therefore, the GEP GSE15781 was re-analyzed to identify DEGs in CRC patients by comparing tumor and normal tissues, as well as tumor tissues after preoperative CRT and untreated samples, under the more rigorous criteria of $|\log_2$ fold change (FC)$| > 1$ and false discovery rate of <0.05. The main focus of the present study was to identify DEGs that were up-regulated in tumor samples, yet down-regulated in tumor-CRT samples. Enrichment analyses of these DEGs were performed. In addition, protein–protein interaction (PPI) networks and transcription factor (TF)-target regulatory networks were established to explore potential correlations of these genes at the protein and transcriptional levels. Importantly, a miRNA–mRNA regulatory network was established to identify key miRNAs that targeted these DEGs. Using these comprehensive bioinformatics methods, study aimed to provide novel insights into gene alterations and identify preoperative CRT-sensitive biomarkers in CRC patients.

Materials and methods

Data resource. A GEP that was deposited in the public Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo) database under the accession number GSE15781 [11] was downloaded and utilized in this study. Expression profiling was conducted on 42 samples, consisting of 22 CRC tumor tissues (non-CRT: n=13; CRT: n=9) and 20 normal colorectal tissues (non-CRT: n=10; CRT: n=10). All tumor and normal tissues were collected from 21 patients with rectal cancer (10 who received preoperative CRT and 11 who underwent only surgery). The platform for GSE15781 was GPL2986 (ABI Human Genome Survey Microarray version 2). The study protocol was approved by the Regional Ethics Committee of North-Norway, and written consent was obtained from all participants.

Pretreatment of the dataset. Raw data in the series matrix were downloaded to the above platform. A gene expression matrix was obtained according to the probe annotation information. In the conversion process, if more than one probe corresponded to the same gene, probe values were averaged to calculate gene expression. Then, gene expression was normalized with the preprocessCore package in R (http://www.bioconductor.org/packages/3.0/bioc/html/preprocessCore.html) based on the median method.

Selection of DEGs. Following normalization, DEGs between different samples (normal-CRT vs. normal-non-CRT, tumor-CRT vs. tumor-non-CRT, and tumor vs. normal) were assembled according to the Bayesian adjusted t-statistics in the linear models for microarray analysis (limma) package of R (http://www.bioconductor.org/packages/release/bioc/html/limma.html) [15]. Two criteria ($|\log_2$ FC$|>1$ and false discovery rate <0.5) were implemented to select DEGs for each comparison. Since the aim of this study was to identify CRT-sensitive genes in tumor tissues, DEGs in the normal-CRT vs. normal-non-CRT group were filtered from the crucial gene list. Thereafter, Venn analysis using the VennDiagram package, a useful tool to generate high quality Venn diagrams for four sets [16], was used to identify common DEGs in the remaining two groups. Genes that were up-regulated/down-regulated in tumor vs. normal tissues and simultaneously down-regulated/up-regulated in tumor-CRT vs. tumor-non-CRT samples were regarded as potential CRT-sensitive genes. Then, two-way hierarchical clustering analysis was applied for these CRT-sensitive genes using the heatmap package [17].

Enrichment analysis of DEGs. The Gene Ontology database (http://www.geneontology.org/), which stores vast information of gene sets, and the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (http://www.genome.jp/kegg/pathway.html), which provides informative pathways for substantial genes (http://www.genome.jp/kegg/pathway.html), were used to identify potential functions and pathways of three sets of DEGs. Then, these sets were further explored using the Database for Annotation, Visualization and Integrated Discovery online tool (http://david.abcc.Ncifcrf.gov/), which is based on the hypergeometric t-test to calculate significance [18]. A probability (p) value of <0.05 was set as the selection criterion to identify significant function and pathway categories.

The clusterprofiler package (http://bioconductor.org/packages/2.8/bioc/html/clusterProfiler.html) was used to draw enriched pathways to visually display different pathways of DEGs. Furthermore, to directly identify the functions of potential CRT-sensitive genes, interconnected biological networks of biological process categories and KEGG pathway clusters were integrated using the ClueGO (http://www.ici.upmc.fr/cluego/) and CluePedia (http://www.ici.upmc.fr/cluepedia/) plug-ins included with the Cytoscape software package, with a p-value of <0.05 set as the threshold for the selection of significant items.

Establishment of PPI network. Potential protein interactions of these genes were extracted by integrating CRT-sensitive genes with information in the Search Tool of the Retrieval of Interacting Genes database (http://string-db.org/) [19]. With this selection criterion and combined score of >0.4, a PPI network was constructed and then visualized using the Cytoscape software platform to identify complex networks (http://cytoscape.org/). A node was defined as a gene-encoded protein in the PPI network. The igraph package [20] was used to calculate the connectivity of each node, namely the number of linked proteins in the network, which was deemed as the degree. Hub nodes were considered as those with high degree values.

Construction of a transcriptional regulation network. Based on the TF-target regulatory information in the TRRUST (Transcriptional Regulatory Relationships Unraveled by Sentence-based Text-mining) database of human transcriptional regulatory interactions [21], potential TFs of identified CRT-sensitive genes, as well as target genes, were identified to build a TF-target network, which was then visualized using Cytoscape software.

Construction of a miRNA–mRNA regulatory network. The miRecords (http://c1.accurascience.com/miRecords/), MiRWalk (http://mirnablog.com/mirwalk-2-0-available/), miRanda (http://www.microrna.org/), MirTarget2 [22], PicTar (http://www.pictar.mdc-berlin.de/), PITA (https://genie.weizmann.ac.il/pubs/mir07/mir07_data.html), and TargetScan (http://www.targetscan.org/vert_71/) databases were used to establish a miRNA–mRNA network with the prerequisite that all interactions were recorded in at least three of these databases. Similarly, the network was visualized using Cytoscape software.

Cell culture and treatment. Human colon carcinoma HT-29 cells were provided by Shanghai Iyunbio Co., Ltd, (Shanghai, China) and maintained in 90% Dulbecco's modified Eagle's medium (Gibco, Carlsbad, CA, USA) + 10% fetal bovine serum (Gibco) + 1% penicillin/streptomycin (Gibco) at 37 °C in a humidified chamber supplemented with 5% CO_2. Cells were divided into two groups: a control group (3 samples) and an X-ray treatment group (3 samples). Cells in the X-ray treatment group were irradiated with a single radiation dose of 4 Gy. All cells in both groups were harvested 6 h after irradiation.

Real-time quantitative PCR (qRT-PCR). First-strand cDNA was synthesized using the miRNA First-Strand cDNA Synthesis Kit (EY001; Shanghai Iyunbio Co., Shanghai, China). qRT-PCR was performed using the Power SYBR Green kit (1504490; ABI, Foster City, CA, USA) on an ABI 7500 FAST real-time PCR system kit (ABI) with the following cycling parameters: 50 °C for 3 min, 95 °C for 3 min, 95 °C for 10 s, and 60 °C for 30 s, followed by 40 cycles of 95 °C for 10 s. Melting curve analysis was used to verify the specificity of the primer amplicons. Target mRNA and miRNA expression was quantified using the comparative Ct method and normalized to GAPDH and U6 expres-

Table 1. Primers of each gene and miRNA

Name	Primers (5'–3')
MYC-F	TACAACACCCGAGCAAGGAC
MYC-R	AGCTAACGTTGAGGGGCATC
ITGA2-F	GTGGCTTTCCTGAGAACCGA
ITGA2-R	GATCAAGCCGAGGCTCATGT
SOX9-F	AACTCCAGCTCCTACTACAGCCA
SOX9-R	TCTGCGGGATGGAAGGGA
DPYD-F	TGTTCCACTTCGGCCAAGAA
DPYD-R	CTCACCAAGAGTCGTGTGCT
miR-495-F	TGGTACCTGAAAAGAAGTTGCC
miR-320c-F	AAAAGCTGGGTTGAGAGGGT
miR-320d-F	TCTAAAAGCTGGGTTGAGAGGA
miR-590-3p-F	GCGCGCTAATTTTATGTATAAGCTA
U6-F	CTCGCTTCGGCAGCACA
U6-R	AACGCTTCACGAATTTGCGT
GAPDH-F	TGACAACTTTGGTATCGTGGAAGG
GAPDH-R	AGGCAGGGATGATGTTCTGGAGAG

sions, respectively. Primers used in this study are listed in Table 1.

Statistical analysis. All data are expressed as the mean ± standard deviation and differences between the two groups were compared using the Student's t-test. All statistical analyses were performed using IBM SPSS Statistics for Windows, version 22.0 (IBM Corp., Armonk, NY, USA), and results were mapped using GraphPad Prism 5 (GraphPad Software, San Diego, CA, USA). A p-value of <0.05 was considered statistically significant.

Results

DEGs in different types of samples. Important DEGs among different samples were identified in accordance with the aforementioned selection criteria. There were 398 up- and 501 down-regulated DEGs in tumor samples compared with those in the normal tissue controls, and 1,433 up- and 1,168 down-regulated DEGs were identified in tumor-CRT samples compared with those in tumor-non-CRT samples. The top 10 DEGs in each group are shown in Table 2. A Venn diagram showed that 148 genes were up-regulated in tumor samples and simultaneously down-regulated in tumor-CRT samples, whereas 111 genes were down-regulated in tumor samples and up-regulated in tumor-CRT samples (Figure 1). A heat map of gene clustering analysis of the three sample types (normal, tumor, and tumor-CRT) is shown in Figure 2.

Enriched functions and pathways of DEGs. Results showed that 148 genes were up-regulated in tumor samples and those down-regulated in tumor-CRT samples were mainly enriched in the functions of cell cycle-related regulation (e.g., mitotic cell cycle process, meiotic cell cycle process, and regulation of the cell cycle phase transition), cytokine-associated regulation (e.g., cytoskeleton-dependent

Table 2. Top 10 of DEGs in each group.

Tumor vs. normal up- regulation		Tumor vs. normal down- regulation		Tumor vs. tumor-CRT up-regulation		Tumor vs. tumor-CRT down-regulation	
Gene symbol	Adjusted p-value	Gene symbol	Adjusted p-value	Gene symbol	Adjusted p-value	Gene symbol	Adjusted p-value
ETV4	1.57E-04	OTOP2	1.57E-04	FLJ30435	3.03E-07	MPST	1.36E-06
INHBA	3.32E-04	CARD14	1.57E-04	FGF2	3.03E-07	DKFZP434I216	1.36E-06
CHRNA5	3.32E-04	C9orf100	1.61E-04	SELM	3.03E-07	EPHA1	1.66E-06
IFITM1	3.32E-04	SPIB	1.61E-04	BNC2	4.40E-07	CDS1	1.70E-06
KDELC1	3.32E-04	CA7	1.61E-04	EMILIN1	5.28E-07	EPS8L3	2.71E-06
CXCL2	3.71E-04	PDE6A	3.02E-04	FGF7	1.36E-06	ICA1	2.72E-06
C12orf11	4.16E-04	PYY2	3.02E-04	MEIS1	1.36E-06	CDCA2	3.09E-06
IFITM3	4.74E-04	SST	4.46E-04	TPST1	1.36E-06	BRI3BP	3.55E-06
CXCL3	5.36E-04	SLC4A4	4.74E-04	JAM3	1.36E-06	KRTCAP3	3.55E-06
MYC	5.36E-04	PCSK1N	5.25E-04	GYPC	1.36E-06	BDH	3.55E-06

DEGs, differentially expressed genes; CRT: chemoradiotherapy.
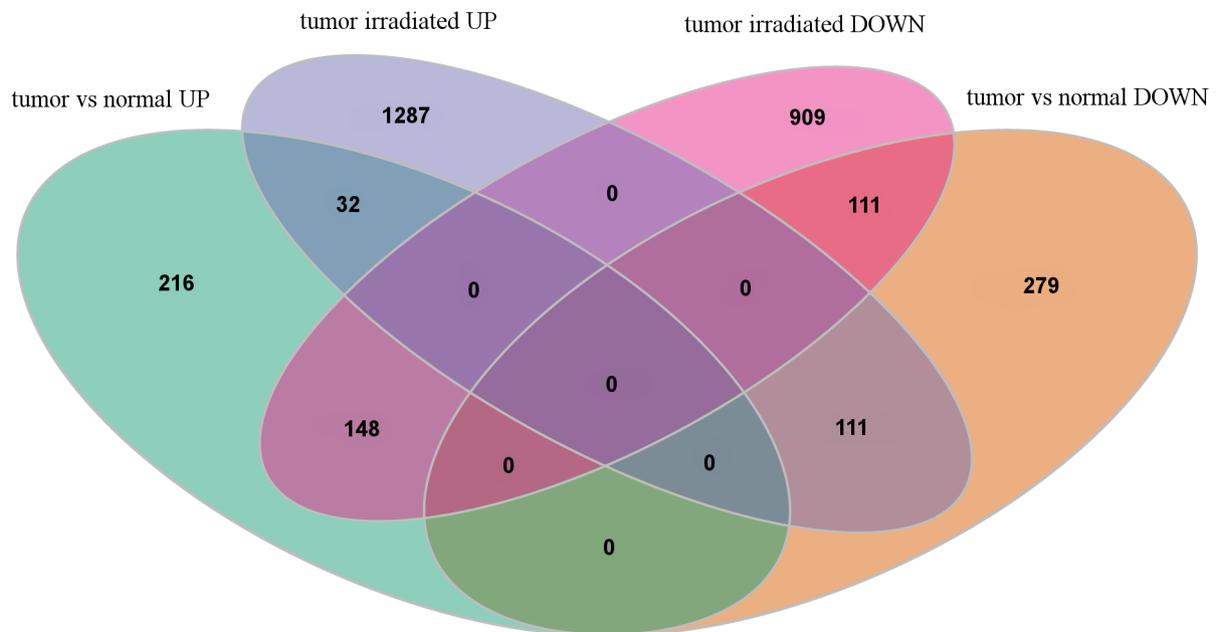

Figure 1. Venn diagram of genes between different comparisons.

cytokinesis and regulation of cytokinesis), and ubiquitin-protein activity (e.g., regulation of ubiquitin-protein ligase activity, regulation of ubiquitin-protein transferase activity, and regulation of ubiquitin-protein ligase activity involved in mitotic cell cycle) (Supplementary Figure 1). In contrast, 111 genes were down-regulated in tumor samples and up-regulated in tumor-CRT samples. These genes were significantly (p<0.05) enriched in the functions of apoptosis-related categories (e.g., positive regulation of extrinsic apoptotic signaling pathway and regulation of epithelial cell apoptotic process) and adhesion-related functions (e.g., positive regulation of focal adhesion assembly and regulation of focal adhesion assembly) (Supplementary Figure 2).

A comparison of the most enriched pathways of DEGs showed that tumor-induced genes were significantly

correlated with cytokine–cytokine receptor interaction (p=1.36E-04; genes: CXCL11, CXCL6, CXCL1), the p53 signaling pathway (p=8.29E-03; genes: CCNB1, CCNB2, SERPINB5), cell cycle regulation (p=6.53E-07; genes: MYC, CCNB1, PTTG1), and DNA replication (p=2.23E-03; genes: PRIM1, RFC3, POLD1), whereas down-regulated genes in tumor-CRT samples were significantly related to metabolic pathways, cell cycle regulation (p=7.49E-07; genes: ORC6L, PTTG1, CDC6), DNA replication (p=3.14E-04, genes: PRIM1, RFC3, LIG1), the PPAR signaling pathway (p=1.19E-03; genes: SCD, PPARG, PCK2), protein digestion and absorption, and the p53 signaling pathway (p=1.02E-02; genes: CCNB1, CCNB2, SERPINB5). Notably, genes involved in the cell cycle regulation, DNA replication, and the p53 signaling pathway were common in these comparisons,
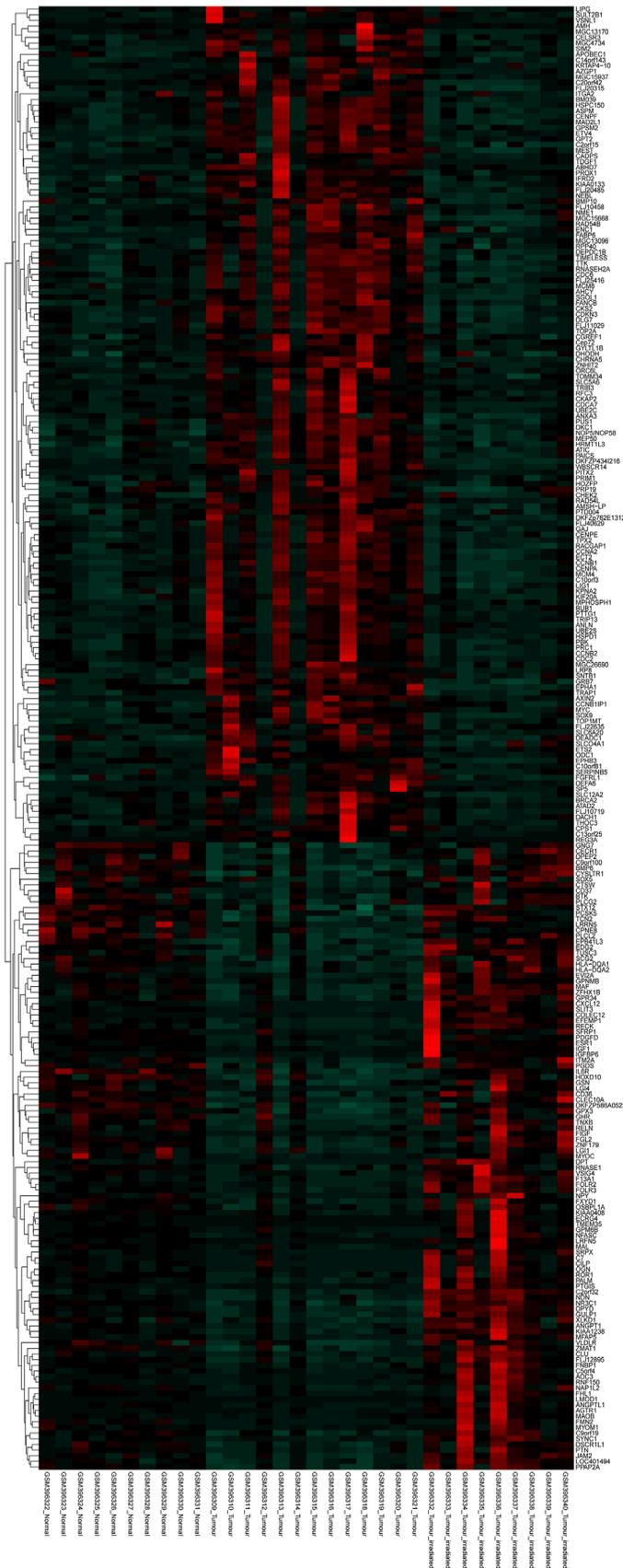
suggesting that genes up-regulated in tumor samples and down-regulated in tumor-CRT samples were tightly associated with these three pathways (Figure 3, Table 3).

PPI network of CRT-sensitive genes. A PPI network of 75 nodes was established based on the predefined parameters (combined score >0.4). Hub nodes with degrees of >15 in the network were KPNA2 (degree=45), ITGA2 (degree=27), RACGAP1 (degree=26), MYC (degree=25), ESR1 (degree=28), dihydropyrimidine dehydrogenase (DPYD; degree=17), and CXCL12 (degree=16), of which the first four were up-regulated in tumor samples and down-regulated in tumor-CRT samples, whereas the expression of the remaining three was opposite (Supplementary Figure 3).

Transcriptional regulation network of CRT-sensitive genes. A TF-target transcriptional regulation network of the 259 CRT-sensitive genes was constructed using the TF-target information in the TRRUST database, in which 15 of the 259 genes were TFs [i.e., MYC, ESR1, NR3C1, ETS2, SRY-box 9 (SOX9), ETV4, PTTG1, PROX1, PITX2, MAF, TRIB3, DACH1, SIM2, MCM4, and SOX5]. As revealed in this transcriptional network, several TFs had numerous targets, including MYC, ETS2, SOX9, ESR1, and NR3C1 (Figure 4).

MiRNA–mRNA network of CRT-sensitive genes. As a result, a set of 295 miRNAs was selected. In the miRNA–mRNA network, miRNAs with the greatest number of targets were hsa-miR-590-3p (targets: MAF, SIM2), hsa-miR-495 (targets: PDGFD, ETS2), hsa-miR-320c (targets: CECR1, CCNA2), and hsa-miR-320d (targets: ESR1, TRAP1) (Figure 5).

Expression of MYC, ITGA2, SOX9, DPYD mRNA, miR-495, miR-590-3p, miR-320d, and miR-320c. qRT-PCR results revealed that expression levels of SOX9, DPYD mRNA, miR-495, miR-590-3p were clearly reduced in irradiated HT-29 cells, whereas that of miR-320d was notably enhanced (Figures 9 and 10). Furthermore, there were no significant differences in the expression patterns of MYC, ITGA2, and miR-320c between the two groups (Figures 6 and 7).

Discussion

The results of bioinformatics analyses showed that 259 preoperative CRT-sensitive critical genes were identified in CRC patients, and most of them were highly correlated with cell cycle regulation, adhesion-associated processes, and the p53 signaling pathway. ITGA2, MYC,

Figure 2. Heat map of gene expression levels in normal, tumor, and tumor-CRT samples. Expression profiles of samples are shown on the X-axis, whereas gene names are shown on the Y-axis. Genes in red were up-regulated and those in green were down-regulated. CRT: chemoradiotherapy.
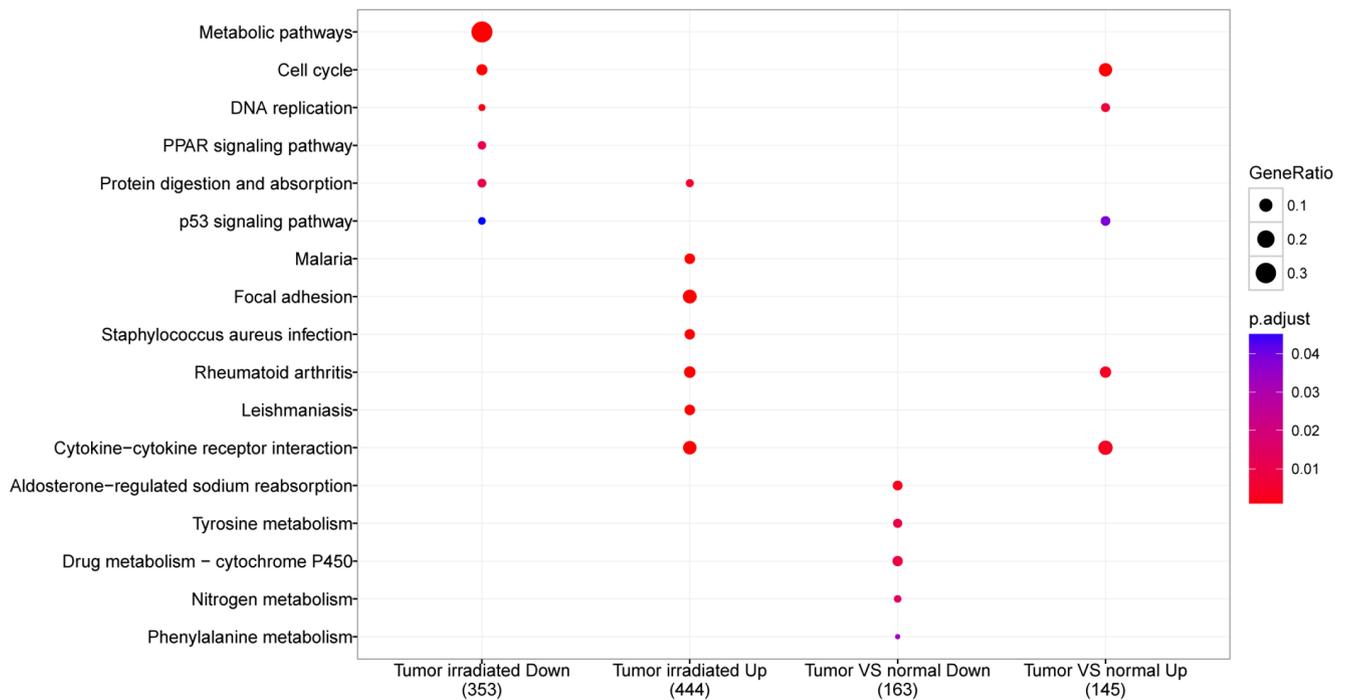
Figure 3. Pathway enrichment analysis of DEGs in different samples.

Table 3. Enriched pathways of DEGs.

	Pathway term	Count	Genes	p-value
SET1	hsa04110:Cell cycle	16	CDC7, CDC6, TTK, ORC6L, MYC et al.	6.53E-07
	hsa04060:Cytokine–cytokine receptor interaction	19	CXCL1, IL6, FLT1, IL8, CXCL5 et al.	1.36E-04
	hsa03030:DNA replication	6	PRIM1, RFC3, POLD1, LIG1, RNASEH2A et al.	2.23E-03
	hsa04115:p53 signaling pathway	7	CCNB1, CCNB2, SERPINB5, SERPINE1, CHEK2 et al.	8.30E-03
	hsa04512:ECM-receptor interaction	7	COL4A2, COL4A1, CD44, ITGA2, COL1A1, THBS2, SPP1	2.20E-02
	hsa04621:NOD-like receptor signaling pathway	6	CXCL1, IL6, IL8, CXCL2, IL1B et al.	2.23E-02
	hsa04914:Progesterone-mediated oocyte maturation	7	CCNB1, MAD2L1, CCNB2, BUB1, ANAPC7 et al.	2.44E-02
	hsa03440:Homologous recombination	4	POLD1, BRCA2, RAD54B, RAD54L	3.56E-02
	hsa00250:Alanine, aspartate and glutamate metabolism	4	CAD, CPS1, GPT2, PPAT	4.61E-02
	hsa00330:Arginine and proline metabolism	5	ODC1, ALDH1B1, ARG2, NOS3, CPS1	4.89E-02
SET2	hsa04110:Cell cycle	25	CDC6, E2F5, TP53, ORC6L, MYC et al.	7.49E-07
	hsa03030:DNA replication	10	PRIM1, RFC3, POLE2, LIG1, POLD2 et al.	3.14E-04
	hsa03320:PPAR signaling pathway	13	SCD, EHHADH, PPARG, PCK2, DBI et al.	1.19E-03
	hsa00100:Steroid biosynthesis	6	EBP, CYP51A1, SQLE, DHCR24, FDFT1 et al.	3.33E-03
	hsa00480:Glutathione metabolism	10	GSS, GPX2, ODC1, GSR, OPLAH et al.	3.79E-03
	hsa00983:Drug metabolism	9	XDH, UGT1A10, CYP3A5, CES2, UGT1A8 et al.	5.01E-03
	hsa00330:Arginine and proline metabolism	10	ODC1, PYCR1, ABP1, GOT1, ACY1 et al.	5.67E-03
	hsa00520:Amino sugar and nucleotide sugar metabolism	9	NANS, GMDS, GFPT1, GNPNAT1, HK2 et al.	5.86E-03
	hsa04115:p53 signaling pathway	11	CCNB1, CASP3, CCNB2, SERPINB5, RRM2 et al.	1.02E-02
	hsa00512:O-Glycan biosynthesis	7	GALNT3, GCNT3, GALNT7, GALNT6, GALNT5 et al.	1.03E-02

SET1: Genes up-regulated in tumor samples and down-regulated in tumor-CRT samples; SET2: Genes down-regulated in tumor samples and up-regulated in tumor-CRT samples. CRT: chemoradiotherapy. Count: Gene numbers enriched in a specific pathway category.

ESR1, and DPYD were the most striking nodes in the PPI network. Notably, MYC and ESR1 were also highlighted in the TF-target regulation network, whereas SOX9 was a crucial TF in the transcriptional network. Interestingly, MYC was enriched in the cell cycle-related pathways and ITGA2 in the extracellular matrix (ECM)-receptor interaction pathway. SERPINB5 was dramatically enriched in the p53 signaling pathway and was indirectly linked to the crucial node DPYD.
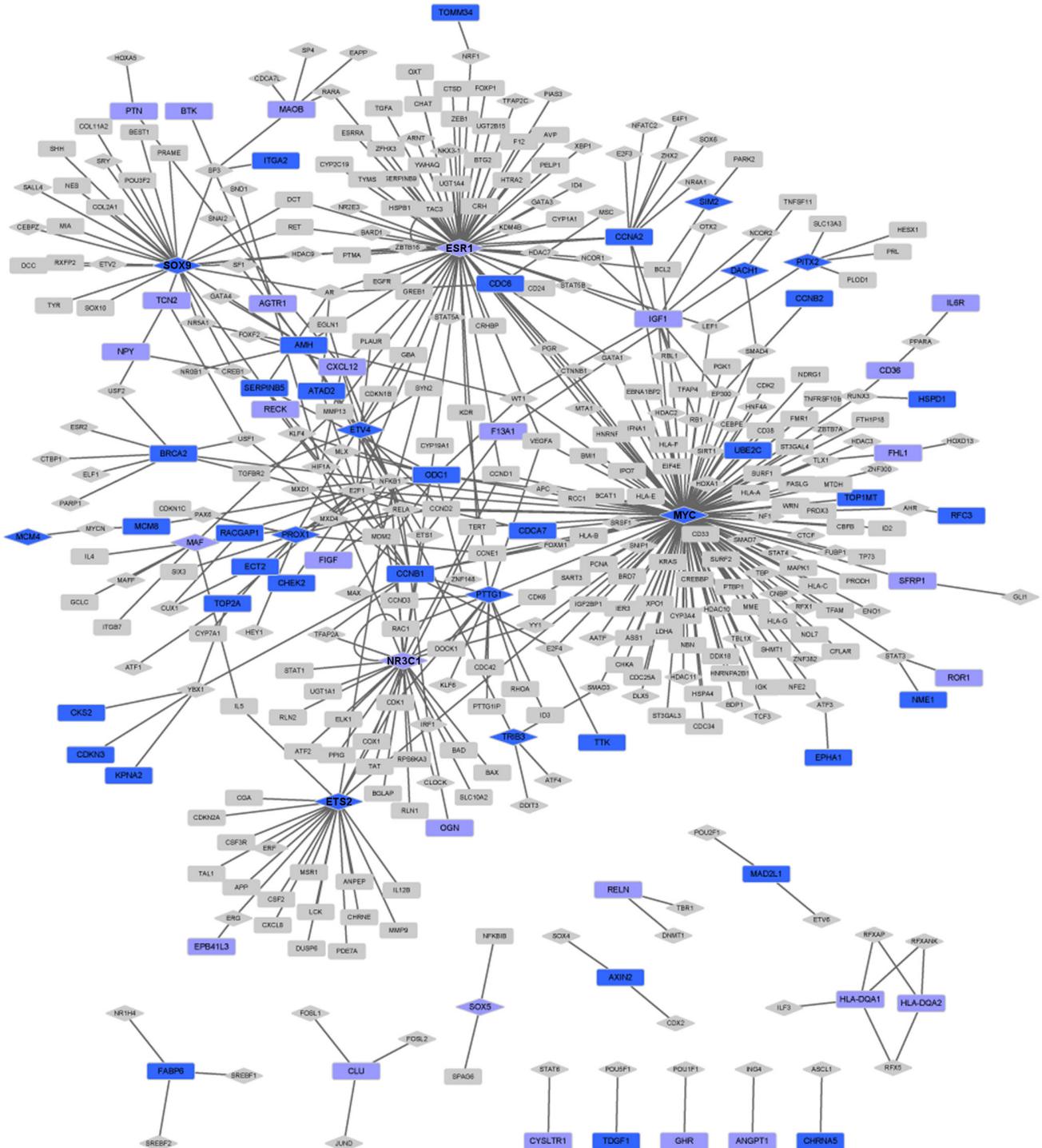
Figure 4. The TF–target network of 259 CRT-sensitive genes. Diamond represents TFs and rectangle denotes targets. Blue represents genes/TFs up-regulated in tumor samples and down-regulated in tumor-CRT samples, and purple represents genes/TFs down-regulated in tumor samples and up-regulated in tumor-CRT samples. CRT: chemoradiotherapy, TF: transcription factor.

The *MYC* oncogene encodes the TF c-MYC, which regulates the cell cycle in tumor development. c-MYC is associated with metabolic alterations during tumorigenesis and induces energy metabolism in cancer cells [23]. In addition, the deregulated expression of c-MYC could result in abnormal tumor cell proliferation and immortalization [24]. MYC is induced by the activation of the Wnt/β-catenin pathway, which plays an important role in CRC progres-
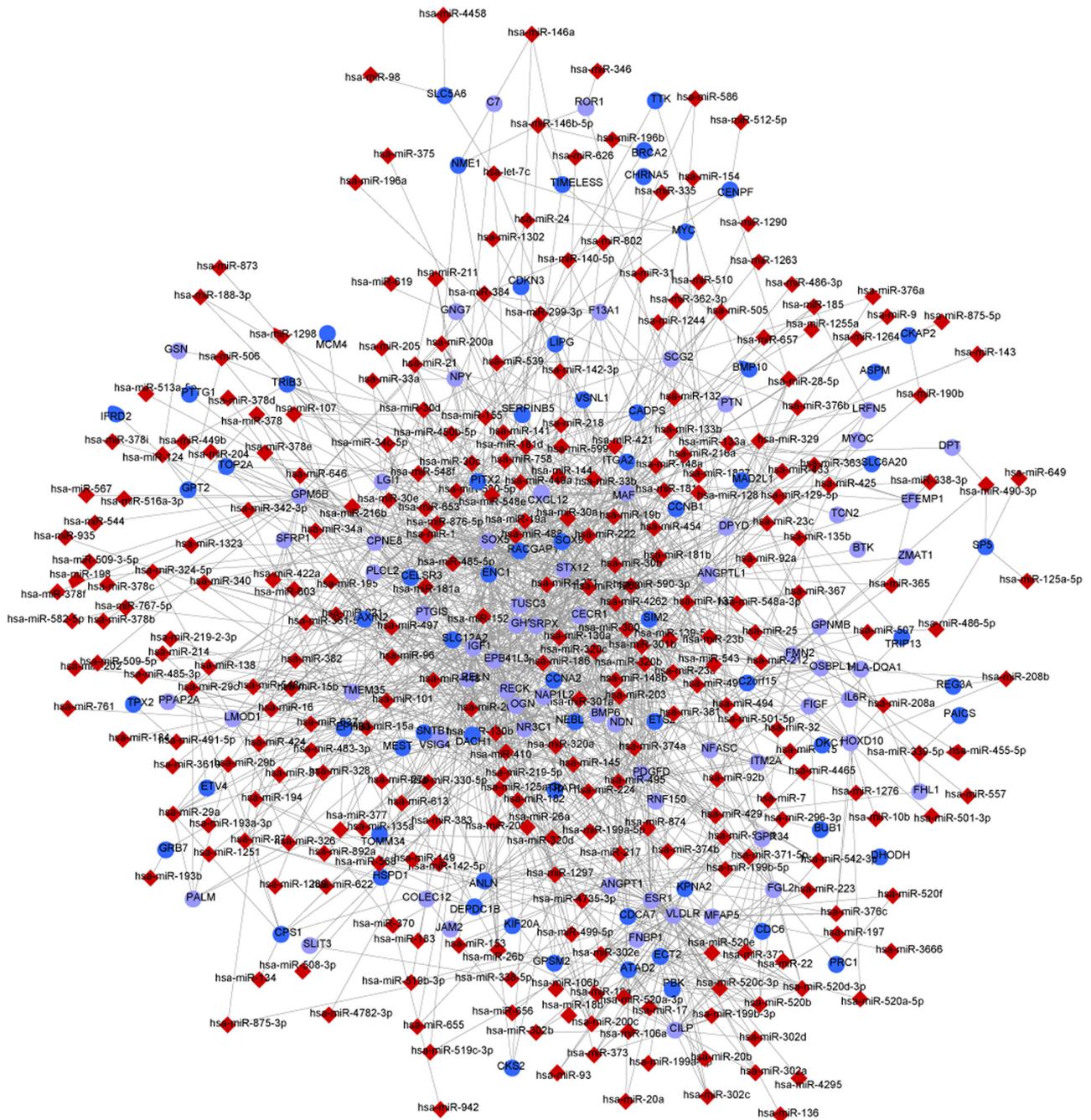
Figure 5. A miRNA–mRNA regulatory network of 259 CRT-sensitive genes. Diamonds represent miRNAs and circles represent genes. Blue represents genes up-regulated in tumor samples and down-regulated in tumor-CRT samples, and purple represents genes down-regulated in tumor samples and up-regulated in tumor-CRT samples. CRT: chemoradiotherapy.

sion [25]. More recently, the Desmoglein 3 (DSG3) gene was identified as a prognostic target in rectal adenocarcinoma patients receiving preoperative CRT [26]. Reportedly, plakoglobin serves as an inhibitor of the Wnt/β-catenin signaling pathway, and DSG3 knockdown reportedly facilitates the inhibitory effect of plakoglobin [27]. Furthermore,

DSG3 contributes to cancer cell growth via mediation of the DSG3-plakoglobin-TCF/LEF-MYC/CCND1/MMP signaling pathway [28]. These findings collectively imply that MYC is linked to DSG3, which may also be sensitive to preoperative CRT. The results of the present study showed that *MYC* was up-regulated in tumor samples, but down-regulated in
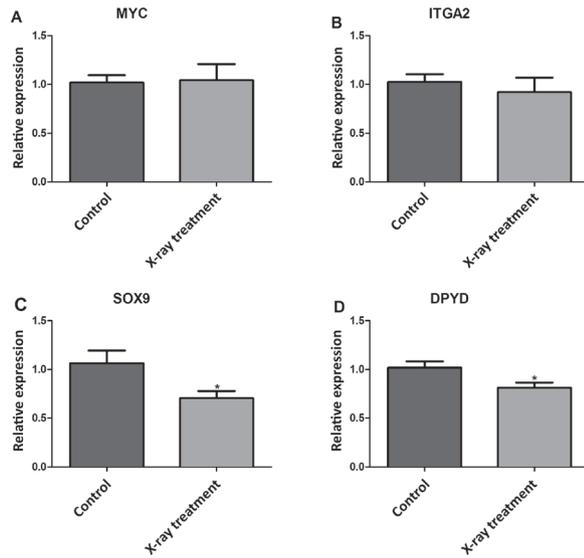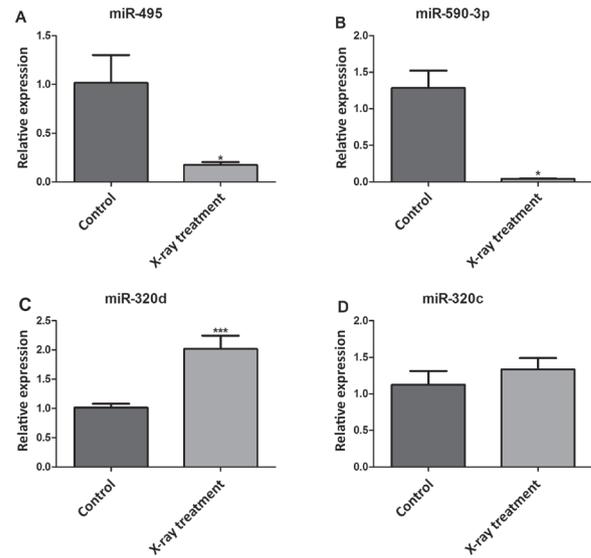
Figure 6. Expression levels of (A) *MYC*, (B) *ITGA2*, (C) *SOX9*, and (D) *DPYD* in HT-29 cells with and without X-ray treatment, as assessed using qRT-PCR. *p<0.05. qRT-PCR: real-time quantitative polymerase chain reaction.

Figure 7. Expression levels of (A) miR-495, (B) miR-590-3p, (C) miR-320d, and (D) miR-320c in HT-29 cells with and without X-ray treatment, as assessed using qRT-PCR. *p<0.05, ***p<0.01. qRT-PCR: real-time quantitative polymerase chain reaction.

tumor-CRT samples and significantly enriched in cell cycle-related pathways. Additionally, MYC was a hub node in both PPI and TF-target networks.

The integrin subunit alpha 2 (*ITGA2*) gene is associated with cell adhesion, and ECM and plays an important role in CRC development [35]. The protein is abundant in the extracellular vesicles of primary CRC SW480 cells compared with those of metastatic SW620 cells [36]. More importantly, a recent study found that *ITGA2* is a candidate gene for the prognosis of locally advanced rectal cancer in response to preoperative CRT [37]. These findings strongly supported our prediction that *ITGA2* is a CRT-sensitive gene that is up-regulated in tumor samples and down-regulated in tumor-CRT samples. Additionally, activation of *ITGA1*, a paralog of *ITGA2*, is enhanced by c-MYC in CRC [38]. In accordance with this result, *ITGA2* was linked to *MYC* in the predicted PPI network. However, mRNA levels of *MYC* and *ITGA2* in HT-29 cells were not obviously affected by X-ray treatment, possibly because the cells used in the present study were only irradiated with a single radiation dose of 4 Gy and harvested 6 h after irradiation, whereas patients (GSE15781) received 50 Gy delivered in 25 2-Gy fractions over a 5-week period. Therefore, we speculated that the genes associated with cell cycle regulation and cell adhesion were not affected by the single radiation dose or the relatively short cell culture duration. Thus, further studies are warranted to determine whether *MYC* and *ITGA2* can be used as therapeutic targets for CRC patients after preoperative CRT.

The Wnt pathway controls the proliferation of crypt cells and plays a key role in CRC onset and progression [39]. *SOX9* is a member of this pathway that maintains epithelial homeostasis in the crypts. Hence, *SOX9* deregulation could alter cellular activities that contribute to colorectal carcinogenesis [40], as *SOX9* mutation has been reported in CRC samples [41]. However, to our knowledge, no study has yet investigated the response of *SOX9* to CRT. Based on these findings, SOX9 was predicted as a crucial CRT-sensitive TF that was up-regulated in tumor samples and down-regulated in tumor-CRT samples, whereas the expression of this gene was clearly reduced in irradiated HT-29 cells, suggesting that *SOX9* is a possible novel preoperative CRT-sensitive gene in CRC patients.

DPYD is involved in various pathways, including purine metabolism and drug metabolism [42]. In advanced CRC, haplotype polymorphisms of *DPYD* are predictive of severe toxicity of the drug capecitabine [43], suggesting that altered gene expression of *DPYD* is potentially CRT-sensitive. In the present study, this gene was a crucial node in the PPI network that was down-regulated in tumor samples and up-regulated in tumor-CRT samples, whereas DPYD mRNA levels were down-regulated in irradiated HT-29 cells. As a possible explanation of this finding, the patients in the study of Snipstad et al. received capecitabine (Xeloda; Roche, Basel, Switzerland) at 825 mg/m² two times per day during the entire radiation period. Thus, we speculated that DPYD was a CRT-sensitive gene in CRC patients who received capecitabine treatment via a drug metabolism pathway. Nonetheless, further studies are needed to validate these predictions.

Regarding miRNAs, hsa-miR-590-3p is reportedly up-regulated in CRC [47]. The microarray analysis results showed that hsa-miR-495 was also increased in CRC stromal tissue compared with normal stromal tissue [48]. In the present

study, the expression levels of miR-590-3p and miR-495 were down-regulated in irradiated HT-29 cells. However, neither has been identified as a CRT-related miRNA. Hsa-miR-320 was predicted as a regulator of *CDC42* expression, which is involved in G1 arrest and inhibition of CRC cell invasion [49]. Notably, the results of an in vitro experiment indicated that miR-320 enhanced the sensitivity of human CRC cells in response to CRT via targeting FOXM1, an important regulator of the cell cycle, tumor cell invasion, and apoptosis [50]. Consistent with these observations, the results of the present study showed that the miR-320d expression level was enhanced in irradiated HT-29 cells and hsa-miR-590-3p, hsa-miR-495, and hsa-miR-320d were major miRNAs in the regulation of CRT-sensitive DEGs, suggesting that they might also be CRT-sensitive in CRC patients.

There was a major limitation to this study, as the sensitive genes and miRNAs were only identified in human colon carcinoma cell lines. Therefore, further studies with larger sample sizes are needed to verify these results.

In conclusion, several sensitive genes, such as *SOX9* and *DPYD*, in response to preoperative CRT were identified in CRC patients that may function via the mediation of pathways related to cell cycle regulation, the Wnt/β-catenin and p53 signaling pathways, and drug metabolism. Hsa-miR-590-3p, hsa-miR-495, and hsa-miR-320d may be sensitive to preoperative CRT in CRC patients. Therefore, *SOX9*, hsa-miR-590-3p, hsa-miR-495, and hsa-miR-320d may be useful as sensitive biomarkers of CRC.

Supplementary information is available in the online version of the paper.

References

[1]   HAGGAR FA, BOUSHEY RP. Colorectal cancer epidemiology: incidence, mortality, survival, and risk factors. Clin Colon Rectal Surg 2009; 22: 191–197. https://doi.org/10.1055/s-0029-1242458

[2]   SIEGEL R, DESANTIS C, JEMAL A. Colorectal cancer statistics, 2014. CA Cancer J Clin 2014; 64: 104–117. https://doi.org/10.3322/caac.21220

[3]   SIEGEL R, MA J, ZOU Z, JEMAL A. Cancer statistics, 2014. CA Cancer J Clin 2014; 64: 9–29. https://doi.org/10.3322/caac.21208

[4]   CHENG L, ENG C, NIEMAN LZ, KAPADIA AS, DU XL. Trends in Colorectal Cancer Incidence by Anatomic Site and Disease Stage in the United States From 1976 to 2005. Am J Clin Oncol 2011; 34: 573–580. https://doi.org/10.1097/COC.0b013e3181fe41ed

[5]   MEZA R, JIHYOUN J, RENEHAN AG, LUEBECK EG. Colorectal cancer incidence trends in the United States and United Kingdom: evidence of right- to left-sided biological gradients with implications for screening. Can Res 2010; 70: 5419–5429. https://doi.org/10.1158/0008-5472.CAN-09-4417

[6]   MIGNANELLI ED, LFSTOCCHI CL. Downstaging after chemoradiotherapy for locally advanced rectal cancer: is there more (tumor) than meets the eye? Dis Colon Rectum 2010; 53: 251–256. https://doi.org/10.1007/DCR.0b013e3181bcd3cc

[7]   SAUER R, BECKER H, HOHENBERGER W, RODEL C, WITTEKIND C et al. Preoperative versus postoperative chemoradiotherapy for rectal cancer. N Engl J Med 2004; 351: 1731–1740. https://doi.org/10.1056/NEJMoa040694

[8]   MINSKY BD. Is preoperative chemoradiotherapy still the treatment of choice for rectal cancer? J Clin Oncol 2009; 27: 5115–5116. https://doi.org/10.1200/JCO.2009.22.9112

[9]   SPITZNER M, EMONS G, KRAMER F, GAEDCKE J, RAVE-FRANK M et al. A Gene Expression Signature for Chemoradiosensitivity of Colorectal Cancer Cells. Int J Radiat Oncol Biol Phys 2010; 78: 1184–1192. https://doi.org/10.1016/j.ijrobp.2010.06.023

[10] WATANABE T, KOMURO Y, KIYOMATSU T, KANAZAWA T, KAZAMA Y et al. Prediction of sensitivity of rectal cancer cells in response to preoperative radiotherapy by DNA microarray analysis of gene expression profiles. Can Res 2006; 66: 3370–3374. https://doi.org/10.1158/0008-5472.CAN-05-3834

[11] SNIPSTAD K, FENTON CG, KJAEVE J, CUI G, ANDERSSEN E et al. New specific molecular targets for radio-chemotherapy of rectal cancer. Mol Oncol 2010; 4: 52–64. https://doi.org/10.1016/j.molonc.2009.11.002

[12] DELLA VITTORIA SCARPATI G, FALCETTA F, CARLOMAGNO C, UBEZIO P, MARCHINI S et al. A Specific miRNA Signature Correlates With Complete Pathological Response to Neoadjuvant Chemoradiotherapy in Locally Advanced Rectal Cancer. Int J Radiat Oncol Biol Phys 2012; 83: 1113–1119. https://doi.org/10.1016/j.ijrobp.2011.09.030

[13] SVOBODA M, IZAKOVICOVA HOLLA L, SEFR R, VRTKOVA I, KOCAKOVA I et al. Micro-RNAs miR125b and miR137 are frequently upregulated in response to capecitabine chemoradiotherapy of rectal cancer. Int J Oncol 2008; 33: 541–547.

[14] DREBBER U, LAY M, WEDEMEYER I, VALLBOHMER D, BOLLSCHWEILER E et al. Altered levels of the oncomicroRNA 21 and the tumor-supressor microRNAs 143 and 145 in advanced rectal cancer indicate successful neoadjuvant chemoradiotherapy. Int J Oncol 2011; 39: 409–415. https://doi.org/10.3892/ijo.2011.1036

[15] DIBOUN I, WERNISCH L, ORENGO CA, KOLTZENBURG M. Microarray analysis after RNA amplification can detect pronounced differences in gene expression using limma. BMC Genomics 2006; 7: 252. https://doi.org/10.1186/1471-2164-7-252

[16] CHEN H, BOUTROS PC. VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R. BMC Bioinformatics 2011; 12: 35. https://doi.org/10.1186/1471-2105-12-35

[17] AN F, ZHANG Z, XIA M. Functional analysis of the nasopharyngeal carcinoma primary tumor-associated gene interaction network. Mol Med Report 2015; 12: 4975–4980. https://doi.org/10.3892/mmr.2015.4090

[18] HUANG DW, SHERMAN BT, TAN Q, KIR J, LIU D et al. DAVID Bioinformatics Resources: expanded annotation database and novel algorithms to better extract biology from large gene lists. Nucleic Acids Res 2007; 35: 169–175. https://doi.org/10.1093/nar/gkm415

[19] SZKLARCZYK D, FRANCESCHINI A, KUHN M, SIMONOVIC M, ROTH A et al. The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. Nucleic Acids Res 2011; 39: D561–D568. https://doi.org/10.1093/nar/gkq973

[20] CSARDI G, NEPUSZ T. The Igraph Software Package for Complex Network Research. Inter J Complex Sys 2006; 1695. Accessed at http://www.interjournal.org/manuscript_abstract.php?361100992

[21] HAN H, SHIM H, SHIN D, SHIM JE, KO Y et al. TRRUST: a reference database of human transcriptional regulatory interactions. Sci Rep 2015; 5: 11432. https://doi.org/10.1038/srep11432

[22] SAITO T, SAETROM P. A two-step site and mRNA-level model for predicting microRNA targets. BMC Bioinformatics 2010; 11: 612. https://doi.org/10.1186/1471-2105-11-612

[23] DANG CV, LE A, GAO P. MYC-induced cancer cell energy metabolism and therapeutic opportunities. Clin Cancer Res 2009; 15: 6479–6483. https://doi.org/10.1158/1078-0432.CCR-09-0889

[24] SINGH AM, DALTON S. The cell cycle and Myc intersect with mechanisms that regulate pluripotency and reprogramming. Cell Stem Cell 2009; 5: 141–149. https://doi.org/10.1016/j.stem.2009.07.003

[25] NAN H, MORIKAWA T, SUURINIEMI M, IMAMURA Y, WERNER L et al. Aspirin use, 8q24 single nucleotide polymorphism rs6983267, and colorectal cancer according to CTNNB1 alterations. J Natl Cancer Inst 2013; 105: 1852–1861. https://doi.org/10.1093/jnci/djt331

[26] CHAO TB, LI CF, LIN CY, TIAN YF, CHANG IW et al. Prognostic significance of DSG3 in rectal adenocarcinoma treated with preoperative chemoradiotherapy. Future Oncol 2016; 12: 1457–1467. https://doi.org/10.2217/fon-2016-0071

[27] KWON OC, WOO KS, KIM TM, KIM DJ, HONG J-T et al. [Physicochemical characteristics of garlic (Allium sativum L.) on the high temperature and pressure treatment]. Korean J Food Sci Tech 2006; 38: 331–336.

[28] CHEN YJ, LEE LY, CHAO YK, CHANG JT, LU YC et al. DSG3 facilitates cancer cell growth and invasion through the DSG3-plakoglobin-TCF/LEF-Myc/cyclin D1/MMP signaling pathway. PLoS One 2013; 8: 64088. https://doi.org/10.1371/journal.pone.0064088

[29] HISHIDA M, NOMOTO S, INOKAWA Y, HAYASHI M, KANDA M et al. Estrogen receptor 1 gene as a tumor suppressor gene in hepatocellular carcinoma detected by triple-combination array analysis. Int J Oncol 2013; 43: 88–94. https://doi.org/10.3892/ijo.2013.1951

[30] COPPEDE F, LOPOMO A, SPISNI R, MIGLIORE L. Genetic and epigenetic biomarkers for diagnosis, prognosis and treatment of colorectal cancer. World J Gastroenterol 2014; 20: 943–956. https://doi.org/10.3748/wjg.v20.i4.943

[31] WORTHLEY DL, WHITEHALL VL, BUTTENSHAW RL, IRAHARA N, GRECO SA et al. DNA methylation within the normal colorectal mucosa is associated with pathway-specific predisposition to cancer. Oncogene 2010; 29: 1653–1662. https://doi.org/10.1038/onc.2009.449

[32] SUGA Y, MIYAJIMA K, OIKAWA T, MAEDA J, USUDA J et al. Quantitative p16 and ESR1 methylation in the peripheral blood of patients with non-small cell lung cancer. Oncol Rep 2008; 20: 1137–1142.

[33] MOLINARI C, CASADIO V, FOCA F, ZINGARETTI C, GIANNINI M et al. Gene methylation in rectal cancer: Predictive marker of response to chemoradiotherapy? J Cell Physiol 2013; 228: 2343–2349. https://doi.org/10.1002/jcp.24405

[34] STAUB E, GROENE J, HEINZE M, MENNERICH D, ROEPCKE S et al. An expression module of WIPF1-coexpressed genes identifies patients with favorable prognosis in three tumor types. J Mol Med (Berl) 2009; 87: 633–644. https://doi.org/10.1007/s00109-009-0467-y

[35] GERGER A, HOFMANN G, LANGSENLEHNER U, RENNER W, WEITZER W et al. Integrin alpha-2 and beta-3 gene polymorphisms and colorectal cancer risk. Int J Colorectal Dis 2009; 24: 159–163. https://doi.org/10.1007/s00384-008-0587-9

[36] CHOI DS, CHOI DY, HONG BS, JANG SC, KIM DK et al. Quantitative proteomics of extracellular vesicles derived from human primary and metastatic colorectal cancer cells. Journal of Extracellular Vesicles 2012; 11: 1. https://doi.org/10.3402/jev.v1i0.18704

[37] MILLINO C, MARETTO I, PACCHIONI B, DIGITO M, DE PAOLI A et al. Gene and microRNA Expression Are Predictive of Tumor Response in Rectal Adenocarcinoma Patients Treated with Preoperative Chemoradiotherapy. J Cell Physiol 2017; 232: 426–435. https://doi.org/10.1002/jcp.25441

[38] BOUDJADI S, BEAULIEU JF. MYC and integrins interplay in colorectal cancer. Oncoscience 2016; 3: 50–51.

[39] SEGDITSAS S, TOMLINSON I. Colorectal cancer and genetic alterations in the Wnt pathway. Oncogene 2006; 25: 7531–7537. https://doi.org/10.1038/sj.onc.1210059

[40] PANZA A, VOTINO C, GENTILE A, VALVANO MR, COLANGELO T et al. Peroxisome proliferator-activated receptor γ-mediated induction of microRNA-145 opposes tumor phenotype in colorectal cancer. Biochim Biophys Acta 2014; 1843: 1225–1236. https://doi.org/10.1016/j.bbamcr.2014.03.003

[41] KATSIOS C, ZIOGAS DE, ROUKOS DH, BALTOGIANNIS G. Targeted therapy for colorectal cancer resistance to EGF receptor antibodies and new trends. Expert Rev Gastroenterol Hepatol 2013; 7: 5–8. https://doi.org/10.1586/egh.12.60

[42] L'ESPERANCE S, BACHVAROVA M, TETU B, MES-MASSON AM, BACHVAROV D. Global gene expression analysis of early response to chemotherapy treatment in ovarian cancer spheroids. BMC Genomics 2008; 9: 99. https://doi.org/10.1186/1471-2164-9-99

[43] DEENEN MJ, TOL J, BURYLO AM, DOODEMAN VD, DE BOER A et al. Relationship between single nucleotide polymorphisms and haplotypes in DPYD and toxicity and efficacy of capecitabine in advanced colorectal cancer. Clin Cancer Res 2011; 17: 3455–3468. https://doi.org/10.1158/1078-0432.CCR-10-2209

[44]   STEGH AH. Targeting the p53 signaling pathway in cancer therapy - the promises, challenges and perils. Expert Opin Ther Targets 2012; 16: 67–83. https://doi.org/10.1517/14728 222.2011.643299

[45]   ZHANG N, LI X, WU CW, DONG Y, CAI M et al. microRNA-7 is a novel inhibitor of YY1 contributing to colorectal tumorigenesis. Oncogene 2013; 32: 5078–5088. https://doi.org/10.1038/onc.2012.526

[46]   LIN XC, ZHU Y, CHEN WB, LIN LW, CHEN DH et al. Integrated analysis of long non-coding RNAs and mRNA expression profiles reveals the potential role of lncRNAs in gastric cancer pathogenesis. Int J Oncol 2014; 45: 619–628. https://doi.org/10.3892/ijo.2014.2431

[47]   CUMMINS JM, HE Y, LEARY RJ, PAGLIARINI R, DIAZ LA JR et al. The colorectal microRNAome. Proc Natl Acad Sci U S A 2006; 103: 3687–3692. https://doi.org/10.1073/pnas.0511155103

[48]   NISHIDA N, NAGAHARA M, SATO T, MIMORI K, SUDO T et al. Microarray analysis of colorectal cancer stromal tissue reveals upregulation of two oncogenic miRNA clusters. Clin Cancer Res 2012; 18: 3054–3070. https://doi.org/10.1158/1078-0432.CCR-11-1078

[49]   LIU M, LANG N, QIU M, XU F, LI Q et al. miR-137 targets Cdc42 expression, induces cell cycle G1 arrest and inhibits invasion in colorectal cancer cells. Int J Cancer 2011; 128: 1269–1279. https://doi.org/10.1002/ijc.25452

[50]   WAN LY, DENG J, XIANG XJ, ZHANG L, YU F et al. miR-320 enhances the sensitivity of human colon cancer cells to chemoradiotherapy in vitro by targeting FOXM1. Biochem Biophys Res Commun 2015; 457: 125–132. https://doi.org/10.1016/j.bbrc.2014.11.039

SOX9, miR-495, miR-590-3p, and miR-320d were identified as chemoradiotherapy-sensitive genes and miRNAs in colorectal cancer patients based on a microarray dataset

B. DU[1], T. WANG[1], X. YANG[1], J. WANG[2], X. SHI[1], X. WANG[1], D. WU[1], L. FENG[1], L. CHEN[1], W. ZHANG[1,*]

Supplemental Material
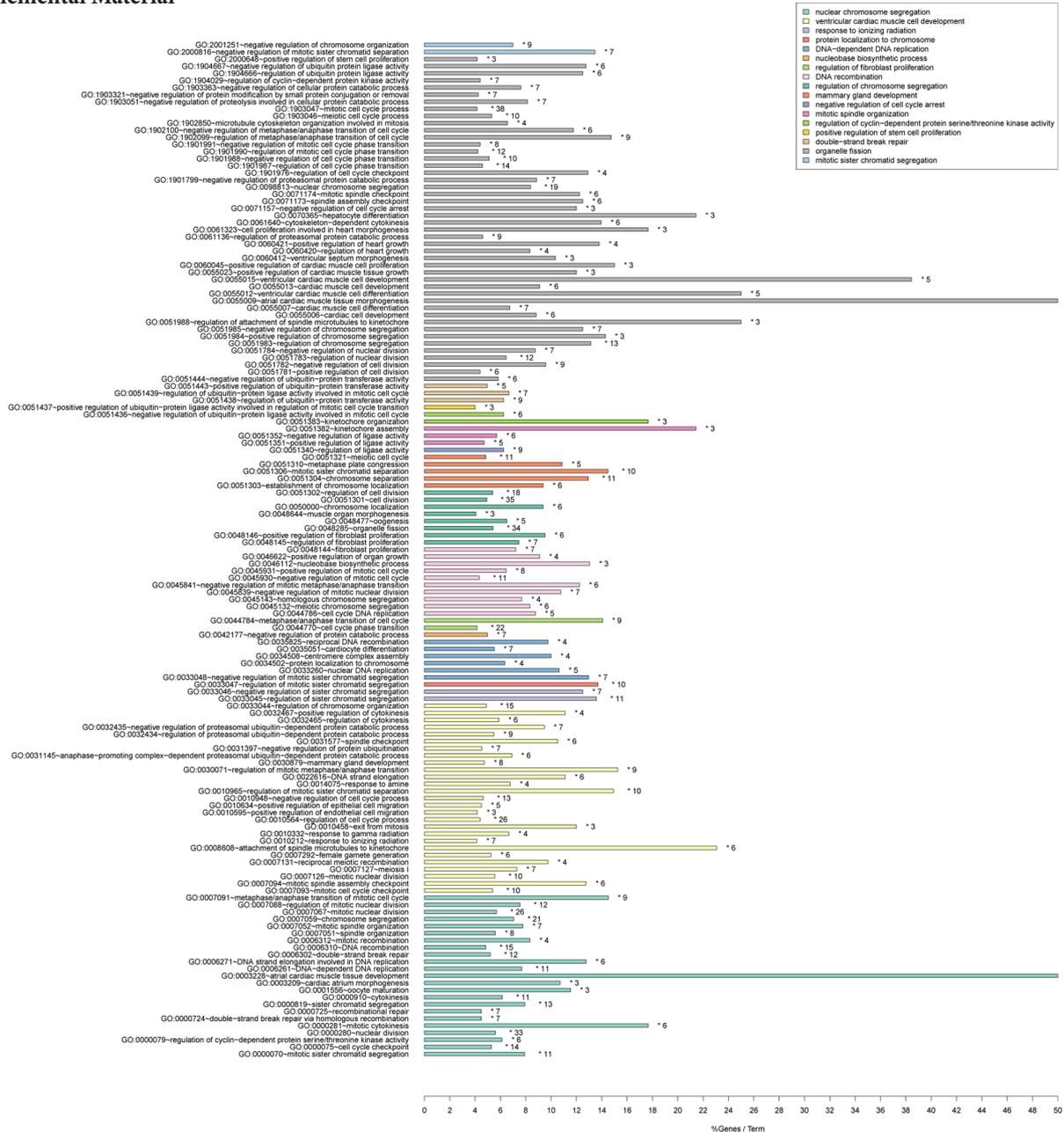
Supplementary Figure 1. Functional enrichment analysis of 148 genes up-regulated in tumor samples and down-regulated in tumor-CRT samples. The abscissa represents the ratio of the numbers of up-regulated genes enriched in the GO term to the total number of genes in that GO term. The Y-axis represents the name of GO-term and the numbers in the figure indicate the up-regulated gene numbers enriched in the GO term. CRT: chemoradiotherapy.

Supplementary Figure 2. Functional enrichment analysis of 111 genes down-regulated in tumor samples and up-regulated in tumor-CRT samples. CRT: chemoradiotherapy. The abscissa represents the ratio of the numbers of down-regulated genes enriched in the GO term to the total number of genes in that GO term. The Y-axis represents the name of GO-term and the numbers in the figure indicate the down-regulated gene numbers enriched in the GO term.

Supplementary Figure 3. Protein–protein interaction network of 259 CRT-sensitive genes. Blue circles represent genes up-regulated in tumor samples and down-regulated in tumor-CRT samples, whereas purple circles represent genes down-regulated in tumor samples and up-regulated in tumor-CRT samples. CRT: chemoradiotherapy.