

ORIGINAL ARTICLE

Epigenomics alternations and dynamic transcriptional changes in responses to 5-fluorouracil stimulation reveal mechanisms of acquired drug resistance of colorectal cancer cells

Y Shen^{1,3}, M Tong^{1,3}, Q Liang¹, Y Guo², HQ Sun¹, W Zheng¹, L Ao¹, Z Guo¹ and F She¹

A drug-induced resistant cancer cell is different from its parent cell in transcriptional response to drug treatment. The distinct transcriptional response pattern of a drug-induced resistant cancer cell to drug treatment might be introduced by acquired DNA methylation aberration in the cell exposing to sustained drug stimulation. In this study, we performed both transcriptional and DNA methylation profiles of the HCT-8 wild-type cells (HCT-8/WT) for human colorectal cancer (CRC) and the 5-fluorouracil (5-FU)-induced resistant cells (HCT-8/5-FU) after treatment with 5-FU for 0, 24 and 48 h. Integrated analysis of transcriptional and DNA methylation profiles showed that genes with promoter hypermethylation and concordant expression silencing in the HCT-8/5-FU cells are mainly involved in pathways of pyrimidine metabolism and drug metabolism-cytochrome *P450*. Transcriptional analysis confirmed that genes with transcriptional differences between a drug-induced resistant cell and its parent cell after drug treatment for a certain time, rather than their primary transcriptional differences, are more likely to be involved in drug resistance. Specifically, transcriptional differences between the drug-induced resistant cells and parental cells after drug treatment for 24 h were significantly consistent with the differentially expressed genes (termed as CRG_{5-FU}) between the tissues of nonresponders and responders of CRCs to 5-FU-based therapy and the consistence increased after drug treatment for 48 h (binomial test, P -value = $1.88E - 06$). This study reveals a major epigenetic mechanism inducing the HCT-8/WT cells to acquire resistance to 5-FU and suggests an appropriate time interval (24–48 h) of 5-FU exposure for identifying clinically relevant drug resistance signatures from drug-induced resistant cell models.

The Pharmacogenomics Journal (2018) **18**, 23–28; doi:10.1038/tpj.2016.91; published online 3 January 2017

INTRODUCTION

Colon and rectal cancer (CRC) is a fatal cancer with low overall 5-year survival rate.¹ Until now, 5-fluorouracil (5-FU)-based chemotherapy is the first-line clinical treatment of CRC,² but the response rate is only ~50%,³ and most of patients acquire resistance during chemotherapy.^{4,5,6} The process of a tumor acquiring drug resistance during chemotherapy could be mimicked partially by the common process of establishing drug-resistant cells by exposing parental cells to drug treatment *in vitro* over a period of time. Therefore, it should be helpful to explore drug-induced resistant cell models to understand both innate and acquired drug resistance mechanisms of cancers.

Based on a drug-induced resistant cell model, researchers often first identify differentially expressed genes (DEGs) between parental and resistant cells, termed basally deregulated (BD) genes,^{7–9} and defined these DEGs as drug resistance genes. However, it has been recognized that most of these BD genes might simply represent drug-induced transcriptional changes irrelevant to drug resistance.^{10–12} Recently, taking the DEGs between the nonresponders and responders of CRC patients receiving 5-FU-based therapy (termed CRG_{5-FU}) as reference, we reported that inducible difference (ID) genes, which represent the transcriptional differences between resistant cells and parental

cells synchronously exposing to drug for a certain time, are more likely to be involved in drug resistance. This result suggests a novel strategy to identify clinically relevant drug resistance genes from drug-induced resistant cell models.¹⁰ However, the cells analyzed in our previous study were treated with 5-FU for only 6, 12 and 24 h. Thus, it is necessary to prolong the time of drug treatment to investigate the characteristics of sustained responses.^{13,14} Another problem remains to be addressed is to explain the underlying molecular mechanism for the acquired distinct transcriptional response characteristics of a drug-induced resistant cancer cell to drug treatment. It has been reported that aberrant DNA hypermethylation within gene promoters and consequent gene silencing might play a major role in generating drug-resistant phenotypes.^{15–17} Therefore, we could assume that the distinct transcriptional response characteristics of a drug-induced resistant cancer cell to drug treatment could be introduced by acquired DNA methylation aberration in the cell exposing to sustained drug stimulation.

In this work, we performed both transcriptional and DNA methylational profiles of HCT-8 wild type cells (HCT-8/WT) for human CRC and its 5-FU-induced resistant cell line (HCT-8/5-FU). We first uncovered that the genes with both promoter hypermethylation and expression silencing during 5-FU treatment are mainly involved in pathways of pyrimidine metabolism and drug

¹Department of Bioinformatics, Key Laboratory of Ministry of Education for Gastrointestinal Cancer, Fujian Medical University, Fuzhou, China and ²Department of Preventive Medicine, School of Basic Medicine Sciences, Gannan Medical University, Ganzhou, China. Correspondence: Professor Z Guo or Professor F She, Department of Bioinformatic, Key Laboratory of Ministry of Education for Gastrointestinal Cancer, Fujian Medical University, Fuzhou, Fujian 350108, China.

E-mail: guoz@ems.hrbmu.edu.cn or shefeifei@yeah.net

³These two authors contributed equally to this work.

Received 20 July 2016; revised 6 November 2016; accepted 14 November 2016; published online 3 January 2017

metabolism-cytochrome *P450* that might be a major epigenetic mechanism of a cancer cell to acquire 5FU resistance. Then, we confirmed that ID genes detected at the 24 h time point were significantly consistent with CRG_{5-FU} and the consistency became more significant after drug treatment for 48 h (binomial test, P -value = $1.88E-06$). This result suggested an appropriate time interval (24–48 h) of drug treatment for identifying genes related to drug resistance.

MATERIALS AND METHODS

Cell lines and reagents

HCT-8/WT and HCT-8/5-FU cells were purchased from Keygentec (Nanjing, China). Complete growth medium RPMI-1640 that contains 10% fetal calf serum and 1% penicillin–streptomycin solution was used for cell culture. Both cell lines were maintained at 37 °C in a humidified atmosphere containing 5% CO₂. MTT assay was performed to determine the 5-FU resistance in both of HCT-8/WT and HCT-8/5-FU cell lines. 5-FU and penicillin–streptomycin solution (100 ×) were purchased from Sigma-Aldrich Corp. St. Louis, MO, USA. RPMI-1640 and fetal calf serum were purchased from Thermo Fisher Scientific, 168 Third Avenue, Waltham, MA, USA. Cell culture plates and other consumables were purchased from Corning and Axygen, 836 North Street, Building 300 Suite 3401, Tewksbury, MA 01876, USA.

Gene expression and methylation profiling

HCT-8/WT and HCT-8/5-FU cells were cultured in 5-FU free medium for 48 h, and then both cell lines were transferred to the medium for 0, 24 and 48 h with HCT-8/WT half-maximal inhibitory concentration doses of 5-FU (Figure 1). Untreated cells (0 h) and treated cells (24 and 48 h) were subject to extract RNA to be subjected to RNA microarray analysis. Total RNA was isolated from three independent experiments using TRizol (ThermoFisher) according to the manufacturer's instructions. Before the following experiments, total RNA was quality controlled by formaldehyde denatured agarose gel electrophoresis and ultraviolet–visible spectrophotometer. Total RNA (100 ng) from one experiment was sent to commercial kits for complementary DNA synthesis, complementary RNA synthesis, antisense RNA synthesis and its fragmentation (1792, Ambion, Thermo Fisher Scientific, 168 Third Avenue, Waltham, MA, USA). Fragmented antisense RNAs were hybridized to Affymetrix GeneChip PrimeView Human Gene Expression Array (3420 Central Expressway, Santa Clara, CA 95051) following the manufacturer's protocols.

Methylation profiling of HCT-8/WT and HCT-8/5-FU cells with three independent experiments, respectively, were performed using Infinium HumanMethylation450 BeadChips (Illumina, 5200 Illumina Way, San Diego,

CA 92122, USA). A total amount of 1 mg genomic DNA from one experiment was bisulfite modified using the EZ DNA Methylation Kit (Zymo Research, Irvine, CA, USA). In all, 200 ng of converted DNA was further processed to run BeadArrays according to the manufacturer's instructions. Each locus is represented by fluorescent signals from two bead types corresponding to the methylated (M) and unmethylated (U) alleles, respectively.

Methylation and expression profiling data were available in NCBI (National Center for Biotechnology Information) Gene Expression Omnibus (GEO accession number: GSE81008).

Data processing

Gene expression profiling data were preprocessed with the Robust Microarray Average algorithm. Each probe set ID was mapped to its Entrez gene ID with the corresponding custom CDF files. If multiple probe sets were mapped to the same gene, the expression value for the gene was defined as the arithmetic mean of the values of the multiple probe sets (on the log₂ scale).

The raw signals of unmethylated (*U*) and methylated (*M*) bead types were background corrected and computed into β -values using the GenomeStudio software Methylation Module (5200 Illumina Way, San Diego, CA 92122, USA). A total of 25 943 CpG loci located within the proximal promoter regions of the transcription start sites of 14 109 genes were used as background CpG loci in this study. The β -value was used to estimate the methylation level of the CpG locus using the ratio of intensities between methylated and unmethylated alleles:

$$\beta = \frac{\max(M, 0)}{(\max(M, 0) + \max(U, 0) + 100)}$$

Concordance score

For DEGs from two independent data sets sharing *k* DEGs, of which *s* genes have the same up- or down-regulation directions between the treated cell type and the control cell type, the concordance score was calculated as *s/k*. If *k* genes have both methylation and expression changes, among which *s* genes are hypermethylated (or hypomethylated) and correspondingly downregulated (or upregulated), then the concordance score was calculated as *s/k*.

The probability of observing a concordance score (*s/k*) by chance can be evaluated using the cumulative binomial distribution model as following:

$$p = 1 - \sum_{i=0}^{s-1} \binom{k}{i} (p_e)^i (1-p_e)^{k-i}$$

in which p_e is the probability of one gene having the same dysregulation direction in two gene lists by random chance (here, $p_e = 0.5$).

Identification of reproducible DEGs and differentially methylation genes in cell lines

The fold change and average difference metrics are often used to identify DEGs between two types of cell lines with two or three technical replicates.^{10,18} However, both of the methods depend on an arbitrarily determined cutoff value and lack any statistical control.¹⁸ Here, we applied the reproducibility-based pairwise difference (PD) and pairwise fold change (PFC) methods to identify reproducible DEGs and differentially methylation genes between two types of cell lines based on small-scale cell line data.¹⁸ Briefly, the algorithm treats every paired technical replicates of two cell types as an independent experiment and then identifies genes with top-ranked fold change or average difference between the two cell types that are significantly reproducible in the independent experiments. The DEGs and differentially methylation genes between the two cell types were finally identified according to the reproducibility of dysregulation directions in a treated cell type compared with the control cell type. The two parameters of the algorithm, the initial step (that is, the initial number of the top-ranked genes) and the consistency threshold were set as 300 and 90%, respectively, as suggested in our previous study.¹⁸

Human protein–protein interaction (PPI) data

The PPI data including 142 583 distinct interactions and 13 693 human proteins were collected as previously described.¹⁹ The types of interactions between proteins included physical interaction, transcriptional regulation

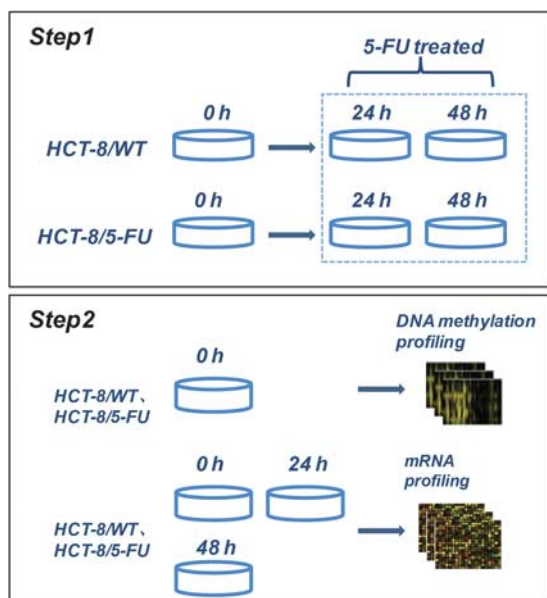


Figure 1. The experimental design in the study.

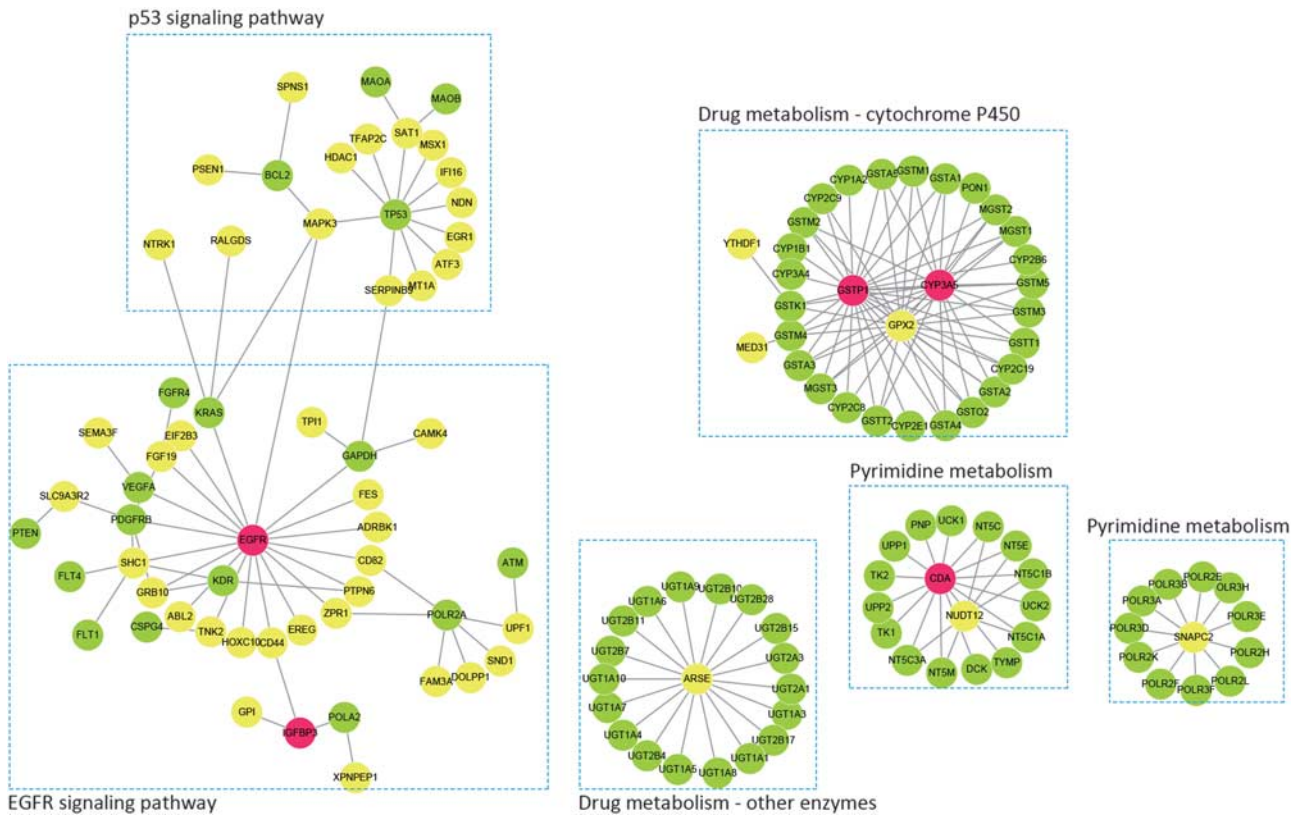


Figure 2. The hypermethylation-mediated downregulations of genes in the HCT-8/5-fluorouracil (5-FU) cells have frequent protein–protein interaction (PPI) links with 5-FU activity-related genes. 5-FU activity-related genes: genes involved in 5-FU transport, metabolism and other downstream effects (such as DNA repair, apoptosis and cell cycle regulation) in the public databases. The green nodes denote 5-FU activity-related genes. The yellow nodes denote the hypermethylation-mediated downregulations of genes in the HCT-8/wild-type (WT) compared with the HCT-8/5-FU cells. The red nodes were overlapped between the two kinds of genes.

and sequential catalysis. Fisher’s exact test was used to test whether the direct PPI links between two gene sets were significant more than what is expected by random chance.

From the Drugbank,²⁰ KEGG (Kyoto Encyclopedia of Genes and Genomes)²¹ and PharmGKB (The Pharmacogenomics Knowledgebase)²² databases, we collected 291 genes associated with the metabolism, transport or activity of 5-FU, of which 254 could be mapped to the PPI network, denoted as 5-FU nodes.

RESULTS

Acquired methylation aberrations in the drug-resistant cells induced by drug treatment

The half-maximal inhibitory concentration of the HCT-8/5-FU cells determined by MTT assay was $4746.31 \mu\text{mol l}^{-1}$, and this was ~45-fold higher than the HCT-8/WT cells ($105.06 \mu\text{mol l}^{-1}$). Thus, the resistance level of HCT-8/5-FU cells was large enough for analyzing molecular changes associated with the mechanism of drug resistance.²³ Then, we explored the potential molecular mechanisms determining the resistance of HCT-8/WT cells to 5-FU stimulation. Using the PD and PFC methods (see Materials and methods), we detected a total of 1658 genes with hypermethylated CpG sites located at the promoter regions in the HCT-8/5-FU compared with the HCT-8/WT cells. Among the 1658 hypermethylated genes, 459 genes were also identified as DEGs by the PFC and PD methods, and 67.76% (311) of these genes showed downregulation in the HCT-8/5-FU cells compared with the HCT-8/WT cells. This concordance score between hypermethylation and downregulation was unlikely to occur by

chance (binomial test, $P\text{-value} = 1.07\text{E}-14$; see Materials and Methods).

For the 311 hypermethylation-mediated downregulated genes in the HCT-8/5-FU cells compared with the HCT-8/WT cells, 242 genes were mapped in PPI network. Among the 242 genes, 24.79% (60) had direct PPI links with at least one of the 254 genes associated with the metabolism, transport or activity of 5-FU, documented in Drugbank, KEGG and PharmGKB databases (see Materials and Methods), significantly higher than the corresponding frequency of 17.91% (2416) for the rest 13 492 genes measured in both the gene expression and methylation platforms in the PPI network (Fisher’s exact test, $P\text{-value} = 8.65\text{E}-04$). As shown in Figure 2, the hypermethylation-mediated downregulated genes participate in pyrimidine metabolism, drug metabolism-cytochrome P450, epidermal growth factor receptor signaling and p53 signaling pathways. The results implied that the hypermethylation-mediated downregulation of genes in the HCT-8/5-FU cells with prolonged time of drug treatment might lead to acquired resistance to 5-FU.

ID₄₈ genes were most significantly consistent with CRG_{5-FU}

The 131 CRG_{5-FU} genes, which represent DEGs between CRC tissue samples of nonresponders and responders for 5-FU-based therapy extracted from three independent data sets,¹⁰ were used to evaluate the clinical relevance of a list of candidate drug resistance genes at the transcriptional level.

First, we further confirmed the previous report¹⁰ that BD genes, which represent DEGs between parental and resistant cells, mainly represent drug-induced changes. We combined BD genes

Table 1. The consistency scores between BD genes, ID genes and CRG_{5-FU}

Gene set	Overlapped DEG ^a	Consistent DEG ^b	Consistency score (%) ^c	Binomial P-value ^d
BD	54	36	66.67	9.92E-03
ID ₂₄	21	15	71.43	3.92E-02
ID ₄₈	71	55	77.46	1.88E-06

Abbreviations: BD, basally deregulated; DEG, differentially expressed gene; 5-FU, 5-fluorouracil; ID, inducible difference. CRG_{5-FU} indicates DEGs between the pre-chemotherapy clinical specimens of responders and nonresponders receiving 5-FU-based chemotherapy; ID genes are genes detected between drug-treated parental cell line and drug-treated resistant cell line; BD genes are basally deregulated genes detected between parental cell line and resistant cell line. ^aThe number of BD genes or ID genes overlapped with CRG_{5-FU}. ^bThe number of overlapped DEGs with the same deregulation direction. ^cThe consistency score of BD genes or ID genes with CRG_{5-FU}. ^dThe binomial distribution P-value.

Table 2. The consistency scores between BD genes and IP genes

Gene set	Overlapped DEG ^a	Consistent DEG	Consistency score (%) ^b	Binomial P-value
IP ₂₄	1667	1631	97.84	< 1.11E-16
IP ₄₈	2176	2151	98.85	< 1.11E-16

Abbreviations: BD, basally deregulated gene; DEG, differentially expressed gene; IP, inducible parental gene. ^aThe number of BD genes overlapped with IP genes. ^bThe consistency score of BD genes and IP genes.

detected by either the PFC or PD methods and obtained 6405 BD genes in total. The concordance score between BD genes and CRG_{5-FU} was 66.67%, suggesting significant but still weak consistency (Table 1). In addition, the concordance score of BD genes with IP₂₄ and IP₄₈ were 97.84% and 98.85%, respectively (binomial test, all P-values < 1.11E-16, Table 2). Here, the IP₂₄ and IP₄₈ genes represented DEGs between parental and drug-treated parental cells that underwent drug treatment for 24 and 48 h, respectively. The results confirmed that most of the BD genes detected in the HCT-8 cells mainly represent drug-induced changes.¹⁰

The PD and PFC methods were further used to identify ID genes that represent the transcriptional difference between resistant cells and parental cells in response to drug treatment for a certain time. The ID genes were denoted as ID₂₄ and ID₄₈ for the conditions where cells underwent drug treatment for 24 and 48 h, respectively. In total, 1922 ID₂₄ genes and 6177 ID₄₈ genes were identified. The ID₂₄ genes overlapped with 21 genes of the CRG_{5-FU} and the concordance scores of the overlapped genes was 71.43% (binomial test, P-value = 3.92E-02, Table 1), whereas the ID₄₈ genes overlapped with 71 genes of the CRG_{5-FU} and the corresponding score increased to 77.47%. The results showed that the ID₄₈ genes were most significantly consistent with CRG_{5-FU} (binomial test, P-value = 1.88E-06, Table 1). Especially, many of the 55 ID₄₈ genes that were consistent with CRG_{5-FU} (Supplementary Table S1) are known to be correlated with drug resistance or tumor cell proliferation, as exemplified in Supplementary Table S2 for the top 20 ID₄₈ genes ranked by the PFC method. For example, upregulation of CD24 in CRC cell lines exhibits cancer stem cell-like characteristics and enhances 5-FU resistance.²⁴ SLC35A1 is a member of solute carriers and its overexpression can activate the process of absorption and transport of cell inhibitors.²⁵ DLGAP5, DTL, FGFBP1 and CDKN3

promote tumor cell proliferation and their overexpression could stimulate drug resistance.²⁶⁻²⁸

Taken together, the above results suggested an appropriate time point for identifying ID genes from the drug-induced resistant cell model.

DISCUSSION

5-FU-based chemotherapy is widely used in the treatment of CRC and other solid tumors like gastrointestinal, head and neck and breast cancers.²⁹ However, response rates of patients with solid tumors to 5-FU-based chemotherapy are still very low.³⁰ Our analyses suggested that aberrant promoter hypermethylation might be a major epigenetic mechanism contributing to the acquired resistance of cancer cells to 5-FU, supporting the previous reports that CRC patients with a CpG island hypermethylation phenotype could not benefit from 5-FU-based adjuvant chemotherapy.^{31,32}

Furthermore, the PPI network analysis showed that many downregulated genes mediated by promoter hypermethylation had intensive direct interactions with genes participating in pyrimidine metabolism, drug metabolism-cytochrome P450, epidermal growth factor receptor signaling and p53 signaling pathways. Especially, three hub nodes (CYP3A5, GSTP1 and GPX2) in the PPI network are in the drug metabolism-cytochrome P450 pathway (Figure 2) and their suppression by promoter hypermethylation could induce drug inactivation within CRC cells^{33,34} and/or promote CRC cell survival and migration.³⁵ Notably, the suppression of some of these genes, such as GSTP1,³⁴ might mediate the multidrug resistance in CRC cells. More generally, some aberrant pathways induced by the epigenetically altered genes might increase multidrug resistance levels regardless of canonical drug mechanisms.³⁶ Further study is needed to investigate whether the acquired epigenetic aberrations in a resistant cancer cell induced by a particular drug could also increase the cell's resistance to other drugs that might be an epigenetic mechanism of clinical multidrug resistance.

The DNA methylation aberrations acquired in the process of inducing drug-resistant cells make the drug-induced resistant cancer cells different from the parental cells in transcriptional response to drug treatment. Our analysis confirmed that genes with transcriptional differences between a drug-induced resistant cell and its parent cell after drug treatment for 24-48 h are more likely to be involved in drug resistance. Compared with the ID₂₄ genes,¹⁰ the ID₄₈ genes became more consistent with the clinically relevant drug resistance genes (CRG_{5-FU}). Obviously, if the drug treatment time is too short, the cancer cells could not receive the maximal benefit from drug dosing²³ and the transient changes in expression levels might be unstable.^{10,37} On the other hand, if the drug treatment time is too long, the expression changes of genes might reflect secondary and tertiary responses³⁷ and could not mirror the cycles of chemotherapy that patients receive in clinical treatment.²³ In fact, we have treated both HCT-8/WT and HCT-8/5-FU cells with 5-FU for 72 hours. The log2FC values and AD values of the 55 ID₄₈ genes which were consistent with CRG_{5-FU} were shown in Figure 3a-b. We found that the transcriptional changes of these genes at 72 h time point were different from those at 24-48 h. Moreover, we found the cellular morphology feature and density changed largely at 72 h, which would be unsuitable for experimental operations to obtain signatures related to mechanisms of drug action.³⁷

In summary, this study reveals a major epigenetic mechanism of the HCT-8/WT cells to acquire drug resistance to 5-FU. Furthermore, the results suggest an appropriate time interval (24-48 h) to extract clinically relevant drug resistance signatures from drug-induced resistant cell models. Further study should be done to explore whether we could extract predictive signatures from the drug resistance genes identified from drug-induced resistant cell

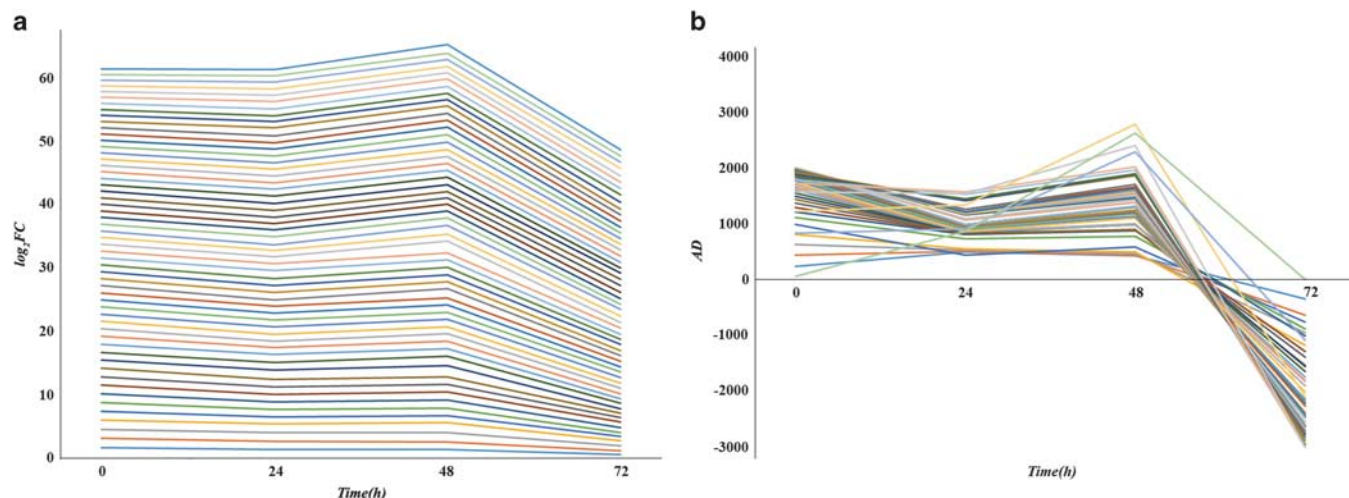


Figure 3. The log₂ fold change (FC) values (a) and average difference (AD) values (b) of 55 genes of 5-fluorouracil (5-FU) resistance.

models to predict responses of patients to 5-FU-based chemotherapy.^{10,38,39}

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

This work was supported by Natural Science Foundation of China (Grant Nos. 81372213, 81572935 and 81271784).

REFERENCES

- 1 Siegel R, Ma J, Zou Z, Jemal A. Cancer statistics, 2014. *CA Cancer J Clin* 2014; **64**: 9–29.
- 2 Graham JS, Cassidy J. Adjuvant therapy in colon cancer. *Expert Rev Anticancer Ther* 2012; **12**: 99–109.
- 3 Adlard JW, Richman SD, Seymour MT, Quirke P. Prediction of the response of colorectal cancer to systemic therapy. *Lancet Oncol* 2002; **3**: 75–82.
- 4 Vinod BS, Antony J, Nair HH, Puliappadamba VT, Saikia M, Narayanan SS *et al*. Mechanistic evaluation of the signaling events regulating curcumin-mediated chemosensitization of breast cancer cells to 5-fluorouracil. *Cell Death Dis* 2013; **4**: e505.
- 5 Ando T, Ishiguro H, Kuwabara Y, Kimura M, Mitsui A, Sugito N *et al*. Relationship between expression of 5-fluorouracil metabolic enzymes and 5-fluorouracil sensitivity in esophageal carcinoma cell lines. *Dis Esophagus* 2008; **21**: 15–20.
- 6 Holohan C, Van Schaeybroeck S, Longley DB, Johnston PG. Cancer drug resistance: an evolving paradigm. *Nat Rev Cancer* 2013; **13**: 714–726.
- 7 Hansen SN, Westergaard D, Thomsen MB, Vistesen M, Do KN, Fogh L *et al*. Acquisition of docetaxel resistance in breast cancer cells reveals upregulation of ABCB1 expression as a key mediator of resistance accompanied by discrete upregulation of other specific genes and pathways. *Tumour Biol* 2015; **36**: 4327–4338.
- 8 Chen Z, Zhang L, Xia L, Jin Y, Wu Q, Guo H *et al*. Genomic analysis of drug resistant gastric cancer cell lines by combining mRNA and microRNA expression profiling. *Cancer Lett* 2014; **350**: 43–51.
- 9 Nakamura A, Nakajima G, Okuyama R, Kuramochi H, Kondoh Y, Kanemura T *et al*. Enhancement of 5-fluorouracil-induced cytotoxicity by leucovorin in 5-fluorouracil-resistant gastric cancer cells with upregulated expression of thymidylate synthase. *Gastric Cancer* 2014; **17**: 188–195.
- 10 Tong M, Zheng W, Lu X, Ao L, Li X, Guan Q *et al*. Identifying clinically relevant drug resistance genes in drug-induced resistant cancer cell lines and post-chemotherapy tissues. *Oncotarget* 2015; **6**: 41216–41227.
- 11 Gillet JP, Calcagno AM, Varma S, Marino M, Green LJ, Vora MI *et al*. Redefining the relevance of established cancer cell lines to the study of mechanisms of clinical anti-cancer drug resistance. *Proc Natl Acad Sci USA* 2011; **108**: 18708–18713.
- 12 Boyer J, Allen WL, McLean EG, Wilson PM, McCulla A, Moore S *et al*. Pharmacogenomic identification of novel determinants of response to chemotherapy in colon cancer. *Cancer Res* 2006; **66**: 2765–2777.
- 13 Qu XA, Rajpal DK. Applications of Connectivity Map in drug discovery and development. *Drug Discov Today* 2012; **17**: 1289–1298.
- 14 Bar-Joseph Z, Gitter A, Simon I. Studying and modelling dynamic biological processes using time-series gene expression data. *Nat Rev Genet* 2012; **13**: 552–564.
- 15 Housman G, Byler S, Heerboth S, Lapinska K, Longacre M, Snyder N *et al*. Drug resistance in cancer: an overview. *Cancers (Basel)* 2014; **6**: 1769–1792.
- 16 Meyer UA, Zanger UM, Schwab M. Omics and drug response. *Annu Rev Pharmacol Toxicol* 2013; **53**: 475–502.
- 17 Shen DW, Pouliot LM, Hall MD, Gottesman MM. Cisplatin resistance: a cellular self-defense mechanism resulting from multiple epigenetic and genetic changes. *Pharmacol Rev* 2012; **64**: 706–721.
- 18 Ao L, Yan H, Zheng T, Wang H, Tong M, Guan Q *et al*. Identification of reproducible drug-resistance-related dysregulated genes in small-scale cancer cell line experiments. *Sci Rep* 2015; **5**: 11895.
- 19 Shen X, Li S, Zhang L, Li H, Hong G, Zhou X *et al*. An integrated approach to uncover driver genes in breast cancer methylation genomes. *PLoS One* 2013; **8**: e61214.
- 20 Law W, Knox C, Djoumbou Y, Jewison T, Guo AC, Liu Y *et al*. DrugBank 4.0: shedding new light on drug metabolism. *Nucleic Acids Res* 2014; **42** (Database issue): D1091–D1097.
- 21 Kanehisa M, Goto S, Sato Y, Furumichi M, Tanabe M. KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res* 2012; **40** (Database issue): D109–D114.
- 22 Thorn CF, Klein TE, Altman RB. Pharmacogenomics and bioinformatics: PharmGKB. *Pharmacogenomics* 2010; **11**: 501–505.
- 23 McDermott M, Eustace AJ, Busschots S, Breen L, Crown J, Clynes M *et al*. In vitro development of chemotherapy and targeted therapy drug-resistant cancer cell lines: a practical guide with case studies. *Front Oncol* 2014; **4**: 40.
- 24 Ke J, Wu X, He X, Lian L, Zou Y, Wang H *et al*. A subpopulation of CD24(+) cells in colon cancer cell lines possess stem cell characteristics. *Neoplasia* 2012; **59**: 282–288.
- 25 Huang Y, Sadee W. Membrane transporters and channels in chemoresistance and -sensitivity of tumor cells. *Cancer Lett* 2006; **239**: 168–182.
- 26 Liao W, Liu W, Yuan Q, Liu X, Ou Y, He S *et al*. Silencing of DLGAP5 by siRNA significantly inhibits the proliferation and invasion of hepatocellular carcinoma cells. *PLoS One* 2013; **8**: e80789.
- 27 Baraniskin A, Birkenkamp-Demtroder K, Maghnoouj A, Zollner H, Munding J, Klein-Scory S *et al*. MiR-30a-5p suppresses tumor growth in colon carcinoma by targeting DTL. *Carcinogenesis* 2012; **33**: 732–739.
- 28 Tassi E, Wellstein A. The angiogenic switch molecule, secreted FGF-binding protein, an indicator of early stages of pancreatic and colorectal adenocarcinoma. *Semin Oncol* 2006; **33** (6 Suppl 11): S50–S56.
- 29 Longley DB, Harkin DP, Johnston PG. 5-Fluorouracil: mechanisms of action and clinical strategies. *Nat Rev Cancer* 2003; **3**: 330–338.
- 30 Johnston PG, Kaye S. Capecitabine: a novel agent for the treatment of solid tumors. *Anticancer Drugs* 2001; **12**: 639–646.

- 31 Jover R, Nguyen TP, Perez-Carbonell L, Zapater P, Paya A, Alenda C *et al*. 5-Fluorouracil adjuvant chemotherapy does not increase survival in patients with CpG island methylator phenotype colorectal cancer. *Gastroenterology* 2011; **140**: 1174–1181.
- 32 Ogino S, Meyerhardt JA, Kawasaki T, Clark JW, Ryan DP, Kulke MH *et al*. CpG island methylation, response to combination chemotherapy, and patient survival in advanced microsatellite stable colorectal carcinoma. *Virchows Arch* 2007; **450**: 529–537.
- 33 Martinez C, Garcia-Martin E, Pizarro RM, Garcia-Gamito FJ, Agundez JA. Expression of paclitaxel-inactivating CYP3A activity in human colorectal cancer: implications for drug therapy. *Br J Cancer* 2002; **87**: 681–686.
- 34 Li H, Li L. Relationship of GSTP1 lower expression and multidrug resistance reversing of curcumin on human colon carcinoma cells. *Zhonghua Yi Xue Za Zhi* 2015; **95**: 2478–2482.
- 35 Emmink BL, Laoukili J, Kipp AP, Koster J, Govaert KM, Fatrai S *et al*. GPx2 suppression of H2O2 stress links the formation of differentiated tumor mass to metastatic capacity in colorectal cancer. *Cancer Res* 2014; **74**: 6717–6730.
- 36 Geeleher P, Cox NJ, Huang RS. Cancer biomarker discovery is improved by accounting for variability in general levels of drug sensitivity in pre-clinical models. *Genome Biol* 2016; **17**: 190.
- 37 Lamb J, Crawford ED, Peck D, Modell JW, Blat IC, Wrobel MJ *et al*. The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. *Science* 2006; **313**: 1929–1935.
- 38 Basik M, Aguilar-Mahecha A, Rousseau C, Diaz Z, Tejpar S, Spatz A *et al*. Biopsies: next-generation biospecimens for tailoring therapy. *Nat Rev Clin Oncol* 2013; **10**: 437–450.
- 39 Loi S, Symmans WF, Bartlett JM, Fumagalli D, Van't Veer L, Forbes JF *et al*. Proposals for uniform collection of biospecimens from neoadjuvant breast cancer clinical trials: timing and specimen types. *Lancet Oncol* 2011; **12**: 1162–1168.



This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-sa/4.0/>

© The Author(s) 2018

Supplementary Information accompanies the paper on the The Pharmacogenomics Journal website (<http://www.nature.com/tpj>)