

siRNAs with decreased off-target effect facilitate the identification of essential genes in cancer cells

Chunyan Li^{1,*}, Zhenzhen Liu^{1,2,*}, Fang Yang^{1,*}, Wensheng Liu¹, Di Wang¹, Encheng Dong¹, Yu Wang¹, Chung-I Wu¹ and Xuemei Lu¹

¹ CAS Key Laboratory of Genomic and Precision Medicine, Beijing Institute of Genomics, Chinese Academy of Sciences, Chaoyang District, Beijing, P. R. China

² University of Chinese Academy of Sciences, Shijingshan District, Beijing, P. R. China

* These authors have contributed equally to this work

Correspondence to: Xuemei Lu, email: luxm@big.ac.cn

Keywords: essential gene, siRNA, off-target, cell viability, transcriptome

Received: February 12, 2015

Accepted: May 13, 2015

Published: May 25, 2015

This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

ABSTRACT

Since the essential genes are crucial to the proliferation and survival of cancer cells, the interference of these genes is promising to be an option for cancer therapy to overcome heterogeneity. However, the essential genes are highly overestimated by RNA interference (RNAi) screenings, which is mainly caused by the pervasive off-target effect of small interference RNA (siRNA) and short hairpin RNA (shRNA). In the present study, we designed Match-Mismatch paired siRNAs to discriminate the on-target effect from off-target effect of siRNAs on cell viability. Only one of the 7 potential essential genes was validated as essential to cell viability, which demonstrates the high false positive rate in RNAi screenings. We modified the siRNA by introducing random nucleotides (N) into the guide strand to mitigate the off-target effect, without significantly compromising the on-target effect. The whole transcriptome profile analysis of cells transfected with siRNAs with or without N indicates that siRNA-dN (with Ns on both the 2nd and the 18th bases of the guide strand) weakens the off-target effect by decreasing the unintended targets. The optimized siRNAs can be applied in the characterization of essential genes in cancer cells.

INTRODUCTION

Genome-wide sequencing has revolutionized our understanding of cancer genetics, and hundreds to thousands of mutations have been characterized in each cancer type [1-4]. However, it remains difficult to determine the key mutations, which are essential to the progression of cancer cells, from hundreds of genes or functional domains alerted by somatic mutations. Essential genes are those that are indispensable for a certain organism under a certain condition [5]. Identification of essential genes for cell survival and proliferation of cancer cells is an efficient way to identify candidates for drug development and cancer therapy [6, 7]. High-throughput RNA interfering (RNAi) screenings are effective tools that are ubiquitously used to characterize essential genes [8, 9]. Small interference RNA (siRNA) and short hairpin RNA

(shRNA) are central players in RNAi. However, the low validation rate and the lack of overlap between different screenings demonstrate high false positive rate in RNAi screening.

Luo and colleagues systematically identified cell essential genes in 12 cancer cell lines using 45,000 shRNAs targeting ~9,500 human genes. Although 530 genes were supposed to be essential genes (of the top 5%) in at least 5 cell lines, only 2 essential genes were shared by two uncorrelated cell lines [10]. A pooled total of 8203 distinct shRNAs targeting 2924 genes were independently screened in 3 cancer cell lines (DLD-1, HCT 116, and HCC 1954), and only 25 genes were characterized as essential genes among these three cancer cell lines. Only 44 essential genes were shared by the two colorectal cancer cell lines (DLD-1 and HCT 116) (the number of essential genes for DLD-1 and HCT 116 is 88

and 151, respectively) [11]. Since the false positive rate of discovering essential genes will decrease with increase of cell line types in RNAi screening, Cheung et al. and Marcotte et al. assessed the essentiality of 11,194 genes in 102 human cancer cell lines and ~16,000 genes in 72 human cancer cell lines, respectively [6, 12]. However, among the genes that were characterized to be essential in >50% of the cell lines in each study, no more than 60% of the essential genes were shared by these two screenings [6]. The divergence between different screenings demonstrates the high false positive rate in RNAi screenings. The dramatic overestimation on the essential genes is significantly caused by the off-target effects that result from the unintended gene interactions in different cell type context [13, 14].

Two main factors lead to the off-target effect of siRNAs. First, the unintended incorporation of the passenger strand (the strand of the duplex, that is not the guide strand) into the RISC (RNA-induced silencing complex) leads to the off-targets, which could be circumvented by appropriate thermodynamic asymmetry or by the 5'-phosphorylation on the guide strand and the 2'-O-methyl modification on the passenger strand [15-17]. The unintended targets remain widespread even when the guide strand is loaded into the RISC [18]. Second, a large proportion of off-targets showed 3' UTR sequence complementarity to the siRNAs, especially to the 5' end of the guide strand (miRNA-like off-target) [19, 20]. The siRNAs with low seed complement frequencies (SCFs) were experimentally validated to have minimal off-targets [21]. The modification of the seed region (2-8nt of the guide strand), such as 2'-O-methyl ribosyl substitution at position 2 of the guide strand, will enhance the specificity of siRNAs and reduce the off-target effect as well as dramatically increase the cost of siRNAs [22]. Alternatively, pooling siRNAs that target different regions of the target mRNA can diminish the off-target effect [23]. However, the high complexity of the pool makes the number of potentially problematic siRNAs be high. It is better to characterize the siRNAs with good performance and use the pool of these siRNAs for the further study.

In the present study, we developed a combined methodology for discriminating the siRNA on-target effect from the off-target effect using the paired Match-Mismatch siRNA strategy and decreasing the off-target effects by introducing random nucleotides (N) into the guide strand of siRNA. Previous studies reported that the nucleotide(s) replacement at the position 9 to 11 of the siRNA guide strand may result in the loss of most or all on-target effects [24-26]. To avoid the introduction of new off-targets by nucleotide replacement, we developed optimal negative controls (named as Mismatch) for siRNAs by replacing only one of the nucleotides at the position of 9 to 11 of the guide strand. The Mismatch siRNAs with nearly no on-target effect were used in this study. The effects caused by the 'Mismatch' siRNAs with

minimal knockdown efficiency on the gene, therefore, are off-target. To minimize the off-target effects of siRNAs, we designed siRNAs avoiding complementary to the 3' UTR of genes, screened siRNAs with low SCFs, and introduced N (random nucleotides) into the 2nd and the 18th bases of the guide strand to eliminate the off-target effect of siRNAs.

RESULTS

Match-Mismatch siRNA assays discriminate the siRNA target effect from the off-target effect

Dozens of genes, which were previously investigated to be generally essential for cancer cell viability [6, 7, 11, 27-32], were used as targets for paired Match-Mismatch siRNAs design. The Mismatch siRNAs were designed to have a nucleotide mismatch on the guide strand (as described in Materials and Methods). Using quantitative Real-Time PCR (qRT-PCR) to detect the target gene expression, we screened multiple paired siRNAs to assure that Match siRNAs had an effect on targets and Mismatch siRNAs could be used as optimal negative controls. The criteria were: (i) knock-down efficiency of Match should be greater than 60% compared to the blank control; (ii) knock-down efficiency of Match should be 50% higher than Mismatch.

Off-target effects increase along with the elevation of the siRNA concentration, and are not negligible when the concentration of siRNA is up to 20 nM [18]. Additionally, a high concentration of siRNA leads to toxicity by the competition of RISC with miRNAs *in vivo* [14]. To determine the optimal siRNA concentration without severe toxicity, a gradient of siRNA negative control (NC, without known targets in the transcriptome) was transfected into HCT 116 and the cell viability was analyzed (Figure S1). The results demonstrated that up to 10 nM, siRNA could lead to cell toxicity; therefore, a lower concentration of 5 nM was chosen in this study.

After the filtering, 7 pairs of siRNAs remained (the siRNA sequences are listed in Table S1). Both the target expression level and cell viability were examined in HCT 116 after the siRNA transfection. The qRT-PCR results confirmed that the Match siRNAs had an acceptable knock-down efficiency, while the knock-down efficiency of Mismatch siRNAs was significantly lower than that of the Match (Figure 1; * for $P < 0.05$, and ** for $P < 0.01$). However, only 1 (*CDC45*) of the 7 targets was likely to affect the cell viability in HCT 116. The cell viability dropped to below 0.80 when the *CDC45* was targeted by Match siRNA, while the Mismatch siRNA did not have any effect on the cell viability (approximately 1.0). On the contrary, although the Mismatch siRNA had no effect on the expression of *FZD6*, the cell viability dropped to

Table 1: Design of siRNAs with N

siRNA type	sequences of the guide strand (5'-3')	
	Match	Mismatch
-0N	ACCGUAUGAAG	ACCGUAUGAA A
	UACUUGGCdTdT	UACUUGGCdTdT
-sN	A NCGUAUGAAG	A NCGUAUGAA A
	UACUUGGCdTdT	UACUUGGCdTdT
-dN	A NCGUAUGAAG	A NCGUAUGAA A
	UACUUG N CdTdT	UACUUG N CdTdT

below 0.80. Therefore, the siRNA targeting *FZD6* exerted a strong off-target effect. Only *CDC45*, one of the 7 genes that we screened, was validated as an essential gene, and the validation rate was lower than 15%. To delineate the genuine essential genes in cells, optimization of the shRNAs or siRNAs is necessary.

The on-target effect of siRNAs is not significantly affected by the introduction of N

According to the general discrimination profile depicted by Huang et al., mismatches at the 5' or 3' end of the guide strand have a minor effect on the down-regulation of targets [25]. The introduction of N into these regions may have a negligible effect on the knock-down efficiency, and minimize the off-target effect as well.

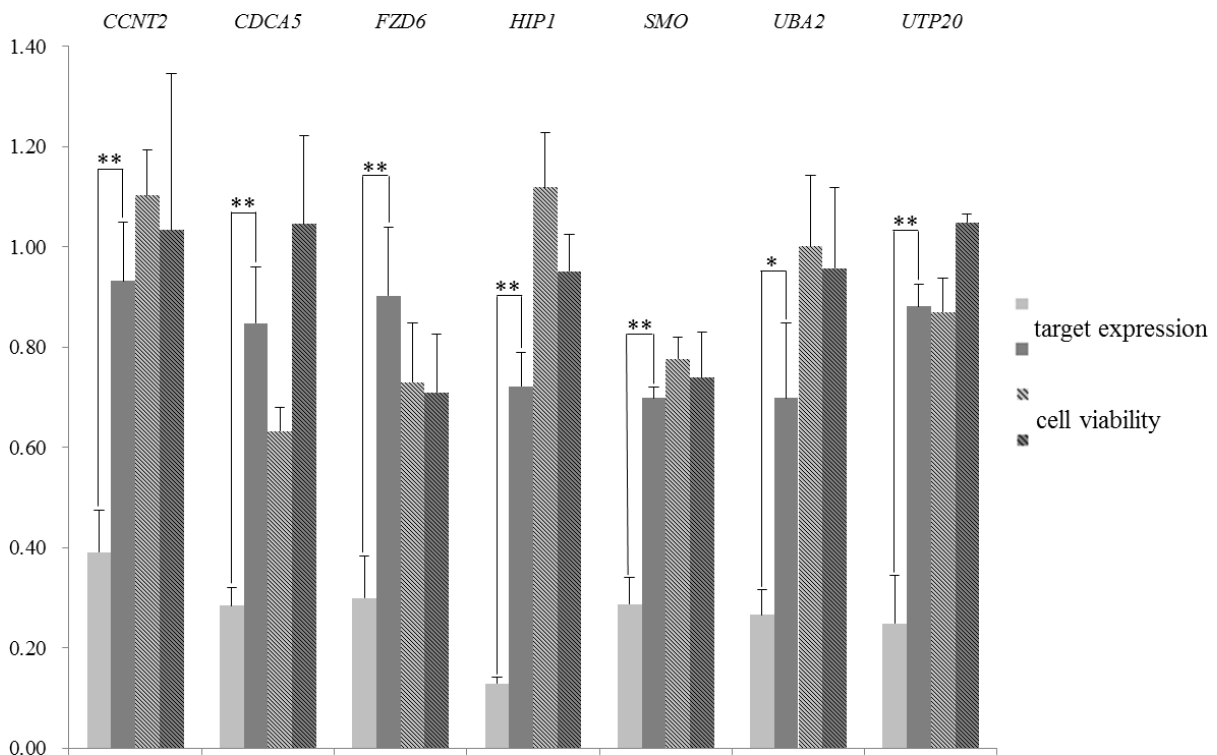


Figure 1: Gene expression of siRNA targets and cell viability after the transfection of Match (light gray bars) or Mismatch (dark gray bars) siRNAs in HCT 116. The target genes of siRNA oligos are labeled at the top of the graph. The y-axis denotes the relative expression level of siRNA target genes (bars without lines) and the cell viability (bars with lines), which were measured against the blank control cells by qRT-PCR and cell viability analysis, respectively (described in Materials and Methods). For the differential expression between cells transfected with Match and Mismatch siRNA oligos, statistically significant *P* values for the Student *t* test are marked by ** (*P* < 0.01), and * (*P* < 0.05) above the bars. Error bars represent the standard deviation.

In the present study, 2 types of modified siRNAs (sN and dN) were designed through introducing N (an even mixture of A, T, G, and C) to the nucleotides on the guide strand (Table 1). The 0N represents the siRNA without the introduction of N. The sN is the siRNA with N on the 2nd nucleotide of the guide strand. To further increase the complexity of pooling siRNAs and to dilute the off-target effect, the siRNA-dN was designed with N on both the 2nd and the 18th nucleotides of the guide strand.

To confirm that the introduction of N did not affect the knock-down efficiency of the on-targets, 3 (*CCNT2*, *CDCA5*, and *FZD6*) of the 7 previously analyzed targets were examined using 3 pairs of siRNAs (-0N, -sN, and -dN) for each target (Figure 2A). For all 3 targets, the

introduction of N (either sN or dN) did not significantly weaken the on-target effects of Match siRNAs (pairwise Student *t* test, $P > 0.05$, *P* values are listed in Table S2), and the Mismatch siRNAs had little effect on the targets.

The slight difference between knockdown efficiencies of the siRNAs with N and without N could have cascade effects on the downstream genes of siRNA targets, leading to the functional changes of the targeting pathway and consequently, the related cell phenotypes. *FZD6* encodes 7-transmembrane domain proteins, which are receptors for Wnt signaling [33]. According to the “Wnt/ beta-Catenin Signaling” annotation in Cell Signaling Technology, we examined the expression of the direct downstream genes regulated by *FZD6* in Wnt

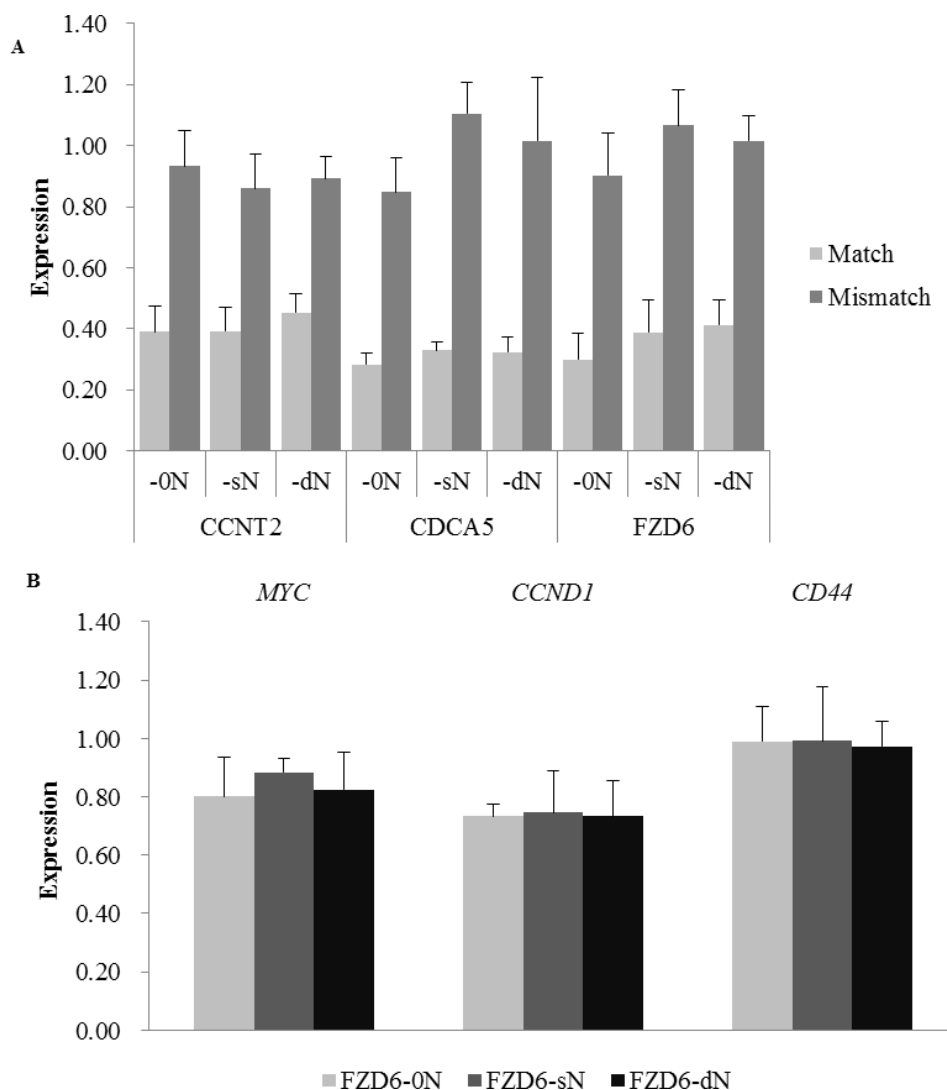


Figure 2: The expression of siRNA target genes and *FZD6* downstream genes after the transfection of siRNAs with N (-sN and -dN) or without N (-0N). **A.** The relative expression levels of three target genes after the transfection of Match (light gray bars) and Mismatch (dark gray bars) siRNAs with or without N were measured by qRT-PCR. The target genes and the type of siRNAs are labeled under the x-axis. **B.** The expression of 3 downstream targets (*MYC*, *CCND1*, and *CD44*) of *FZD6* in the Wnt pathway were examined by qRT-PCR, 24 hours after the transfection of siRNAs with or without N. Bars in light gray, dark gray, and black represent the cells transfected with Match siRNA without N (labeled as FZD6-0N), Match siRNA with single N on the 2nd position of the guide strand (labeled as FZD6-sN), and Match siRNA with double Ns on the 2nd and the 18th position of the guide strand (labeled as FZD6-dN), respectively. Error bars represent the standard deviation.

Table 2: Enriched GO terms ($P < 0.01$) by DAVID functional annotation of the differentially expressed genes after the transfection of siRNAs targeting *FZD6*

GO term	number of genes	<i>P</i> value
after the transfection of FZD6-0N		
cell death	15	2.0E-3
death	15	2.2E-3
apoptosis	13	3.6E-3
programmed cell death	13	4.0E-3
NADP metabolic process	3	7.3E-3
after the transfection of FZD6-sN		
positive regulation of anti-apoptosis	4	8.3E-4
regulation of anti-apoptosis	4	1.7E-3
induction of apoptosis by intracellular signals	4	4.6E-3
response to organic substance	12	4.9E-3
response to organic cyclic substance	5	6.8E-3
cofactor metabolic process	6	7.3E-3
response to inorganic substance	6	8.9E-3
after the transfection of FZD6-dN		
response to organic substance	8	6.3E-3

signaling pathway (<http://www.cellsignal.com/contents/science/cst-pathways/science-pathways>). According to the expression of six downstream genes that was detected by RNA-sequencing after the transfection of FZD6-0N or FZD6-sN/dN in the cell line, their expression levels in FZD6-sN/dN transfection did not significantly differ from that in FZD6-0N transfection (Table S4). Three (*MYC*, *CCND1* and *CD44*) of the six downstream targets were reported to be transcriptionally regulated by the Wnt signaling in the colon carcinoma cells from which the HCT 116 (the cell line we used in this study) was derived [34-36]. The expression level of these three downstream genes

with FPRK (fragments per kilobase of exon model per million mapped reads) > 20 in all 4 samples was further validated by qRT-PCR (Figure 2B and Table S4). The expression levels of these three downstream targets in the transfections of siRNAs with N and without N were not different from each other, which was consistent with the RNA-Seq results (Figure S2 and Table S4). The *P* values of analysis of variance (ANOVA) were 0.667, 0.989, and 0.983 for *MYC*, *CCND1* and *CD44*, respectively (Table S4). The results above indicated that the introduction of N did not affect the on-target effect.

The introduction of N into siRNAs minimizes the number of the unintended targets and weakens the off-target effect

Because the Match siRNA targeting *FZD6* had severe off-target effect (Figure 1), whole genome expression profiles of cells with or without the transfection of *FZD6* siRNAs were delineated by RNA-Seq to elucidate the genome wide off-targets (the statistics of the RNA-Seq data is shown in Table S3). The transcriptome of cells transfected with Match siRNA without N (FZD6-0N) had the largest differences from the transcriptome of the blank control cells, compared to those of sN and dN transfected cells (Figure S3). Compared to the blank control, the numbers of the significantly differentially expressed genes in the cells transfected with FZD6-0N, FZD6-sN, and FZD6-dN were 189, 168, and 97, respectively (fold change >1.5 and $P < 0.01$; Figure S4).

The smallest number of genes was affected by FZD6-dN at the transcriptional level. Among the 189 genes, 6 genes with FPKM > 20 in all of the 3 samples were randomly selected to validate their expression levels using qRT-PCR. The expression of these 6 genes in FZD6-sN and FZD6-dN transfected cells was more closed to their expression in 0C than that in the FZD6-0N transfection, which was consistent with the RNA-seq data (Figure 3A and Table S5). It indicates that the unintended effects on the genes' expression were significantly reduced (*, $P < 0.05$; **, $P < 0.01$) by introducing N to the siRNAs.

The function of the differentially expressed genes in cells after the siRNA transfection was analyzed by DAVID (<http://david.abcc.ncifcrf.gov/>). GO (Gene Ontology) terms with $P < 0.01$ are listed in Table 2. In the cells transfected with siRNA *FZD6*, the genes were enriched in GO terms related to cell death and apoptosis. The expression variation of these genes may cause the

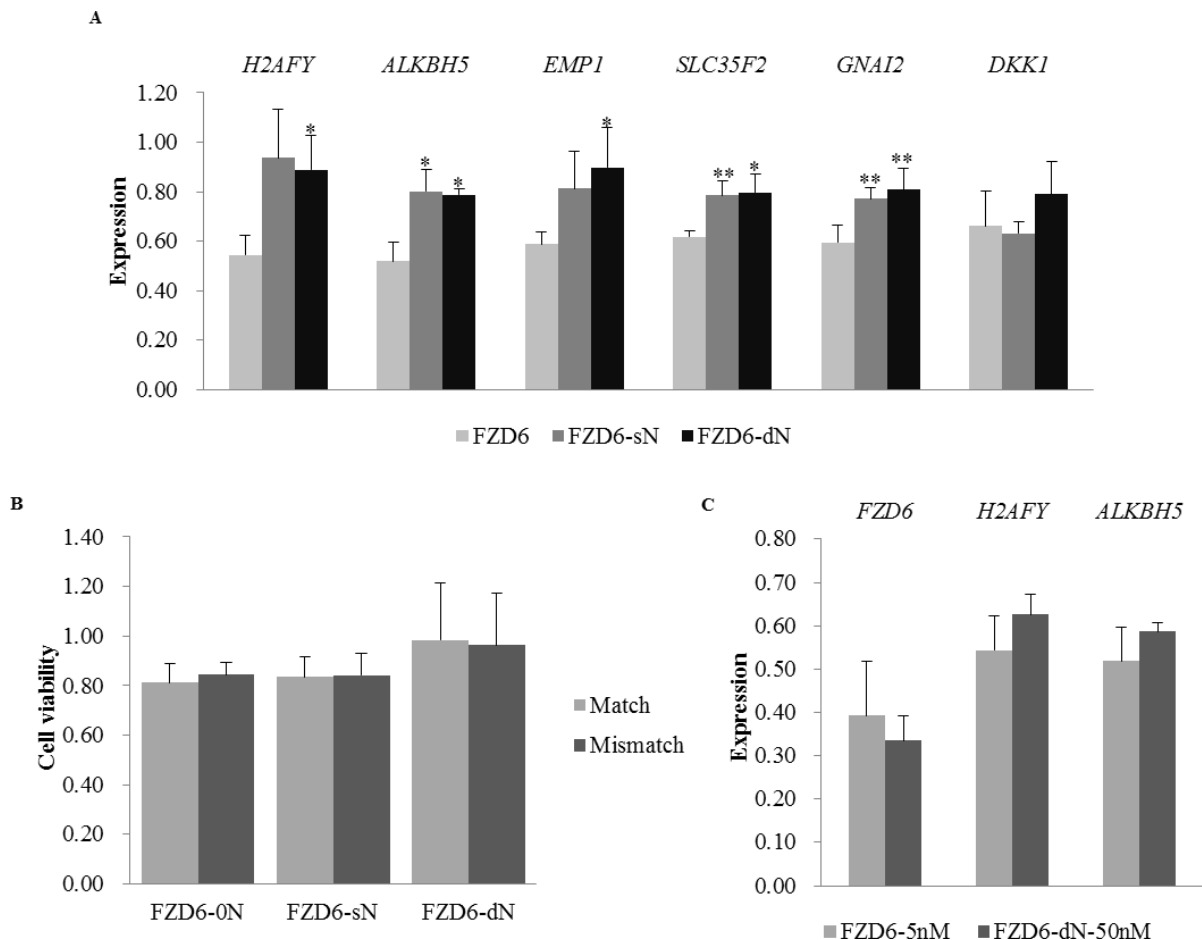


Figure 3: The effects of siRNAs with or without N on the gene expression of *FZD6* and six unintended targets and on cell viability. **A.** The expression of 6 unintended target genes after the transfection of *FZD6* siRNAs with or without N. The bars in different colors represent the cells transfected with siRNAs with or without N, as described in Figure 2B. The off-target effects on the 6 unintended genes' expression were significantly reduced (*, $P < 0.05$; **, $P < 0.01$) by introducing N to the siRNAs. **B.** The cell viability after the transfection of Match (light gray) and Mismatch (dark gray) siRNAs with N (FZD6-sN and FZD6-dN) or without N (FZD6-0N). The drop in cell viability (y-axis) is rescued by the introduction of N into the siRNA oligos (from ~0.80 to ~1.00). **C.** The expression of *FZD6* and 2 unintended targets after the transfection of FZD6-0N (5 nM; light gray) or FZD6-dN (50 nM; dark gray). Error bars represent the standard deviation.

off-target effects on cell viability. Among the 189 genes revealed in FZD6 transfection, 16 genes were annotated to be involved in cell-death. Compared to the blank control, 10 and 5 of the 16 genes were differentially expressed in cells transfected with sN and dN siRNA oligos, respectively, and the difference was even smaller in the FZD6-dN transfected cells than the FZD6-sN transfected cells (Table S6). It suggests that the off-target effect of FZD6-0N on cell viability is largely resulted from the unintended genes, which is reduced by introducing Ns to the siRNA oligos (Figure 3B).

To examine whether the off-target effect of siRNA-dN was increased with the enhancement of the on-target effect by increasing the concentration of siRNA oligos, the expression of *FZD6* and 2 potential off-target genes (*H2AFY* and *ALKBH5*) was examined by qRT-PCR. Although the knock-down efficiency of *FZD6* was slightly stronger after increasing the concentration of the dN siRNA oligo, the off-target effect was still weaker than that exerted by the unmodified siRNA (Figure 3C). The increase of dN siRNA concentration enhances on-target effect rather than off-target effect.

DISCUSSION

Large pools of shRNAs have been used for highly parallel multiplex screening for essential genes in human cell lines [6, 7, 11]. The validation of the essential genes from the screening of human mammary cells revealed that cell viability could drop to below 50% even when the targets have no obvious depletion after the shRNA transfection (the Figure 2B in [7]). The number of essential genes characterized in previous RNAi screenings is listed in Table S7. Although 4 or even more siRNAs/shRNAs were designed for each gene, the essential gene enrichment analysis was frequently done based on the 1-2 siRNAs/shRNAs with strongest effect [6, 12]. Not all siRNAs/shRNAs are effective, whereas the estimation based on 1-2 siRNAs/shRNAs is insufficient to deplete the off-target effect. These studies disclosed that most of the false positive candidates were caused by off-targets, which were prevalent in RNAi screenings. The result in this study shows that the paired Match-Mismatch siRNAs can be applied to discriminate the on-target and off-target effects of the candidate essential genes effectively.

The pooling strategy has been used to diminish the off-target effect of siRNAs [23]. Several siRNAs with a high efficiency of on-target effects have to be screened out, which will elevate the cost. Based on the pooling strategy, introducing N (an even mixture of A, T, G, and C) into the guide strand may be an alternative to dilute the off-target effect. In this study, we designed 2 types of modified siRNAs (sN and dN) to eliminate the off-target effect. The results of transcriptome analysis and cell viability measurement in cancer cell line demonstrate that siRNAs with dN are a promising and easy approach to

characterize essential genes by minimizing the off-target effect. The limitation of this strategy is that to minimize the false positive rate in siRNA screenings for essential genes in cancer cell lines, only siRNAs with strong on-target effect should be used in the introduction of dN to minimize the off-target effect.

Although a limited proportion of genes were essential for each cancer cell line, the essential genes were divergent among different cancer types, even among subtypes [6, 7, 11, 12]. The high level of complexity in gene interaction networks maintains cellular homeostasis, fitness, and survival in both normal and cancer cells [37-39]. Additionally, drug-resistance and relapse are inevitable in single-target therapy because of the heterogeneity found within tumors. Predictably, the large-scale screenings of functional important genes will be required in gene network characterizations and in the development of individualized cancer therapy. To do a large set of RNAi screening in as many cancer cell lines as possible is labor-intensive and makes the ranking algorithms for essential genes even harder. It is convenient to knock-down multiple essential genes at the same time using combined siRNAs with dN, developed herein.

The single nucleotide variations (SNVs) are divergent among different cancer types and different cancer patients, even among the cancer cells within the tumor [40]. The number of SNVs varies from hundreds to millions in exome [41]. The somatic SNVs in the coding region of tumors can be used in the siRNA screening to discriminate cancer cells from normal cells, since tumor-specific siRNAs are mismatched and hence may not effectively bind to the gene region with a single nucleotide difference in normal cells [25, 26]. Based on the somatic SNVs detected in tumors, the optimized Match-siRNAs in this study can be utilized in developing personalized cancer therapy by specifically targeting essential genes in cancer cells and eliminating the toxicity of siRNAs in normal cells in the precision medicine.

MATERIALS AND METHODS

siRNA design and synthesis

To minimize the off-target effect of siRNAs, each oligo targets less than 2 genes in which more than 12 consecutive basepairs are complementary to the guide strand of the siRNA oligos. The 3'UTR is avoided for the target region of siRNAs. The siRNAs that are completely complementary to the targets are named 'Match', while the siRNAs with a non-complementary nucleotide to the targets are named 'Mismatch'. The Mismatch siRNAs are designed according to the previously mentioned rule [25]. The sN and dN oligos are 2 types of modified siRNAs: the sN is the siRNA with a random nucleotide (N) on the

2nd base of the guide strand, and dN is the siRNA with Ns on both the 2nd and the 18th bases of the guide strand. All of the siRNAs were synthesized by Guangzhou RiboBio Co., Ltd.

Cell culture

The HCT 116 cell line was obtained from the cell bank of Shanghai Institutes for Biological Sciences, CAS. The cells were cultured in McCoy's 5A (Sigma) plus 10% FBS (Hyclone) with penicillin (100 U/ml) plus streptomycin (100 µg/ml) (Sigma) at 37 °C with 5% CO₂ in a humidified incubator (Thermo Scientific), and were expanded on a 6-cm dish until the cells were seeded on 96/24-well plates. The cells were counted using a hemocytometer.

siRNA transfection

Cells were seeded into the wells one day before transfection. For cell viability analysis, 2×10³ HCT 116 cells in 50 µl of medium without antibiotics were seeded in each of the internal 60 wells, and 200 µl of PBS were added to each of the outside 36 wells of the 96-well plate. For knock-down efficiency detection, 3×10⁴ HCT 116 cells in 400 µl of medium without antibiotics were seeded in each well of the 24-well plates.

The transfection was mediated by Lipofectamine™ 2000 (Invitrogen) according to the manufacturer's instructions. The medium was replaced by 200 µl or 500 µl of fresh medium without antibiotics for 96-well and 24-well plates separately, 6 hours after transfection. The final concentration of siRNA was 5 nM unless noted otherwise.

Cell viability analysis

Three days after transfection, cell viability was measured using the CellTiter 96® AQueous One Solution Cell Proliferation Assay (Promega). A volume of 20 µl CellTiter 96® AQueous One Solution Reagent was pipetted into each well of the 96-well plates containing samples. After incubation at 37°C for 2 hours, the absorbance at 490 nm was recorded. More than 3 independent experiments were carried out for each treatment, and triplicates were conducted in the same plate. The cell viability was the ratio of the absorbance at 490 nm divided by the blank control (0C), which was transfected with the transfection reagents without siRNA oligo.

qRT-PCR (quantitative real-time PCR)

qRT-PCR was carried out 24 hours after the siRNA transfection to quantitate the on-target and the off-target

effects. Total RNA was extracted using RNA prep Pure Cell/Bacteria Kit (Tiangen). A total of 1 µg of extracted total RNA from each sample was reverse-transcribed by Reverse Transcription System (Promega) according to the manufacturer's protocol. qRT-PCR experiments were performed using Maxima SYBR Green/ROX qPCR Master Mix (Thermo) in triplicate 20 µl reactions according to manufacturer's protocol on an AB 7500. Thermal cycling was organized in 2 steps: an initial denaturation step of 10 min at 95°C, followed by 40 repeated cycles of 95°C for 10 sec and 60°C for 30 sec. Melting curves were obtained by increasing the temperature from 60°C to 95°C with a plate reading every 0.2°C. The housekeeping gene *GAPDH* was used as an endogenous control. Primer sequences of candidate genes were designed on the basis of sequence data obtained from the NCBI database (<http://www.ncbi.nlm.nih.org>). The expression level of each sample was normalized according to its *GAPDH* content. The $\Delta\Delta C_t$ method was used to measure the expression level of each candidate. $\Delta C_t = C_t(\text{target}) - C_t(\text{GAPDH})$. $\Delta\Delta C_t = \Delta C_t(\text{sample}) - \Delta C_t(0C)$. The fold change of each gene was calculated by the equation $2^{-\Delta\Delta C_t}$, which represented each gene's relative expression level against the blank control cells (0C) [42].

RNA-Seq library construction

Total RNA was extracted using a RNA prep Pure Cell/Bacteria Kit (Tiangen). Three milligrams of total RNA from each of the 3 independent experiments was mixed for the whole transcriptome library construction. The library preparation was performed according to the TruSeq™ RNA Sample Preparation Guide standard protocol. Sequencing was conducted on a Hi-Seq 2000 (Illumina).

Transcriptome sequencing, mapping and expression analysis

Sequencing reads of 101 bp in length were aligned to the reference genome (GRCh37.73/hg19, downloaded from Ensembl database, <ftp.ensembl.org>) with the Tophat alignment software tools [43]. The uniquely aligned reads were used for calculation by Cufflinks software and the gene expression levels were normalized using the FPKM (fragments per kilobase of exon model per million mapped reads) [44]. The expression level of each gene is listed in Table S5. The results of differentially expressed genes were analyzed and visualized by the Bioconductor function "CummeRbund" in the R language [45]. Gene expression RNA-Seq data have been deposited in the Gene Expression Omnibus (GEO) archive with the accession number of GSE62968.

Statistical analysis

Data are expressed as the mean \pm standard deviation from at least 3 independent experiments performed in at least duplicate. The differences were considered significantly different at $P < 0.05$ (denoted by *) and $P < 0.01$ (denoted by **) using a double-sided Student *t* test.

ACKNOWLEDGMENTS

We thank Changqing Zeng, Jian Lu, and Yang Shen for useful discussions and suggestions. We thank Minjie Zhang for the help in the data analysis, and Min Sun, Weijie Li and Bing Zhang at the Genomics Core Facility in the Beijing Institute of Genomics for their technical support of sequencing.

CONFLICTS OF INTEREST

The authors declare that they have no conflict of interest.

GRANT SUPPORT

This work was supported by the Strategic Priority Research Program of the Chinese Academy of Sciences [XDB13040300 to X.M.L and C.I.W.], the National High-tech R&D Program (863 Program) of China [2012AA022502 to X.M.L.], the National Basic Research Program (973 Program) of China [2014CB542006 to C.I.W.], the Natural Science Foundation of China [91131903 to X.M.L.], and the Chinese Academy of Sciences [KJZD-EW-L06-01 to C.I.W.].

REFERENCES

1. Bea S, Valdes-Mas R, Navarro A, Salaverria I, Martin-Garcia D, Jares P, Gine E, Pinyol M, Royo C, Nadeu F, Conde L, Juan M, Clot G, Vizan P, Di Croce L, Puente DA, et al. Landscape of somatic mutations and clonal evolution in mantle cell lymphoma. *Proceedings of the National Academy of Sciences of the United States of America*. 2013; 110: 18250-18255.
2. Pleasance ED, Cheetham RK, Stephens PJ, McBride DJ, Humphray SJ, Greenman CD, Varela I, Lin ML, Ordonez GR, Bignell GR, Ye K, Alipaz J, Bauer MJ, Beare D, Butler A, Carter RJ, et al. A comprehensive catalogue of somatic mutations from a human cancer genome. *Nature*. 2010; 463: 191-196.
3. Song Y, Li L, Ou Y, Gao Z, Li E, Li X, Zhang W, Wang J, Xu L, Zhou Y, Ma X, Liu L, Zhao Z, Huang X, Fan J, Dong L, et al. Identification of genomic alterations in oesophageal squamous cell cancer. *Nature*. 2014; 509: 91-95.
4. Vogelstein B, Papadopoulos N, Velculescu VE, Zhou SB,

- Diaz LA and Kinzler KW. *Cancer Genome Landscapes*. *Science*. 2013; 339: 1546-1558.
5. Zhang R and Lin Y. DEG 5.0, a database of essential genes in both prokaryotes and eukaryotes. *Nucleic Acids Research*. 2009; 37: D455-D458.
6. Marcotte R, Brown KR, Suarez F, Sayad A, Karamboulas K, Krzyzanowski PM, Sircoulomb F, Medrano M, Fedyshyn Y, Koh JL, van Dyk D, Fedyshyn B, Luhova M, Brito GC, Vizeacoumar FJ, Vizeacoumar FS, et al. Essential gene profiles in breast, pancreatic, and ovarian cancer cells. *Cancer Discovery*. 2012; 2: 172-189.
7. Silva JM, Marran K, Parker JS, Silva J, Golding M, Schlabach MR, Elledge SJ, Hannon GJ and Chang K. Profiling essential genes in human mammary cells by multiplex RNAi screening. *Science*. 2008; 319: 617-620.
8. Tuschl T and Borkhardt A. Small interfering RNAs: a revolutionary tool for the analysis of gene function and gene therapy. *Molecular Interventions*. 2002; 2: 158-167.
9. Kiefer J, Yin HH, Que QQ and Mousses S. High-throughput siRNA screening as a method of perturbation of biological systems and identification of targeted pathways coupled with compound screening. *Methods in Molecular Biology*. 2009; 563: 275-287.
10. Luo B, Cheung HW, Subramanian A, Sharifnia T, Okamoto M, Yang X, Hinkle G, Boehm JS, Beroukhi R, Weir BA, Mermel C, Barbie DA, Awad T, Zhou X, Nguyen T, Piquani B, et al. Highly parallel identification of essential genes in cancer cells. *Proceedings of the National Academy of Sciences of the United States of America*. 2008; 105: 20380-20385.
11. Schlabach MR, Luo J, Solimini NL, Hu G, Xu QK, Li MZ, Zhao ZM, Smogorzewska A, Sowa ME, Ang XLL, Westbrook TF, Liang AC, Chang K, Hackett JA, Harper JW, Hannon GJ, et al. Cancer proliferation gene discovery through functional Genomics. *Science*. 2008; 319: 620-624.
12. Cheung HW, Cowley GS, Weir BA, Boehm JS, Rusin S, Scott JA, East A, Ali LD, Lizotte PH, Wong TC, Jiang GZ, Hsiao J, Mermel CH, Getz G, Barretina J, Gopal S, et al. Systematic investigation of genetic vulnerabilities across cancer cell lines reveals lineage-specific dependencies in ovarian cancer. *Proceedings of the National Academy of Sciences of the United States of America*. 2011; 108: 12372-12377.
13. Kaelin WG. Use and Abuse of RNAi to Study Mammalian Gene Function. *Science*. 2012; 337: 421-422.
14. Sigoillot FD and King RW. Vigilance and validation: Keys to success in RNAi screening. *ACS Chem Biol*. 2011; 6: 47-60.
15. Chen PY, Weinmann L, Gaidatzis D, Pei Y, Zavolan M, Tuschl T and Meister G. Strand-specific 5'-O-methylation of siRNA duplexes controls guide strand selection and targeting specificity. *RNA*. 2008; 14: 263-274.
16. Reynolds A, Leake D, Boese Q, Scaringe S, Marshall WS and Khvorovova A. Rational siRNA design for RNA

- interference. *Nature Biotechnology*. 2004; 22: 326-330.
17. Tomari Y, Matranga C, Haley B, Martinez N and Zamore PD. A protein sensor for siRNA asymmetry. *Science*. 2004; 306: 1377-1380.
 18. Jackson AL, Bartz SR, Schelter J, Kobayashi SV, Burchard J, Mao M, Li B, Cavet G and Linsley PS. Expression profiling reveals off-target gene regulation by RNAi. *Nature Biotechnology*. 2003; 21: 635-637.
 19. Birmingham A, Anderson EM, Reynolds A, Ilsley-Tyree D, Leake D, Fedorov Y, Baskerville S, Maksimova E, Robinson K, Karpilow J, Marshall WS and Khvorova A. 3' UTR seed matches, but not overall identity, are associated with RNAi off-targets. *Nature Methods*. 2006; 3: 199-204.
 20. Jackson AL, Burchard J, Schelter J, Chau BN, Cleary M, Lim L and Linsley PS. Widespread siRNA "off-target" transcript silencing mediated by seed region sequence complementarity. *RNA*. 2006; 12: 1179-1187.
 21. Anderson EM, Birmingham A, Baskerville S, Reynolds A, Maksimova E, Leake D, Fedorov Y, Karpilow J and Khvorova A. Experimental validation of the importance of seed complement frequency to siRNA specificity. *RNA*. 2008; 14: 853-861.
 22. Jackson AL, Burchard J, Leake D, Reynolds A, Schelter J, Guo J, Johnson JM, Lim L, Karpilow J, Nichols K, Marshall W, Khvorova A and Linsley PS. Position-specific chemical modification of siRNAs reduces "off-target" transcript silencing. *RNA*. 2006; 12: 1197-1205.
 23. Kittler R, Surendranath V, Heninger AK, Slabicki M, Theis M, Putz G, Franke K, Caldarelli A, Grabner H, Kozak K, Wagner J, Rees E, Korn B, Frenzel C, Sachse C, Sonnichsen B, et al. Genome-wide resources of endoribonuclease-prepared short interfering RNAs for specific loss-of-function studies. *Nature Methods*. 2007; 4: 337-344.
 24. Buehler E, Chen YC and Martin S. C911: A Bench-Level Control for Sequence Specific siRNA Off-Target Effects. *Plos One*. 2012; 7: e51942.
 25. Huang H, Qiao RP, Zhao DY, Zhang T, Li YX, Yi F, Lai FF, Hong JM, Ding XF, Yang ZJ, Zhang LH, Du Q and Liang ZC. Profiling of mismatch discrimination in RNAi enabled rational design of allele-specific siRNAs. *Nucleic Acids Research*. 2009; 37: 7560-7569.
 26. Schwarz DS, Ding HL, Kennington L, Moore JT, Schelter J, Burchard J, Linsley PS, Aronin N, Xu ZS and Zamore PD. Designing siRNA that distinguish between genes that differ by a single nucleotide. *Plos Genetics*. 2006; 2: 1307-1318.
 27. Chuang PT, Kawcak T and McMahon AP. Feedback control of mammalian Hedgehog signaling by the Hedgehog-binding protein, Hip1, modulates Fgf signaling during branching morphogenesis of the lung. *Genes & Development*. 2003; 17: 342-347.
 28. Pasca di Magliano M and Hebrok M. Hedgehog signalling in cancer formation and maintenance. *Nature Reviews Cancer*. 2003; 3: 903-911.
 29. Sicklick JK, Li YX, Jayaraman A, Kannangai R, Qi Y, Vivekanandan P, Ludlow JW, Owzar K, Chen W, Torbenson MS and Diehl AM. Dysregulation of the Hedgehog pathway in human hepatocarcinogenesis. *Carcinogenesis*. 2006; 27: 748-757.
 30. Tanigaki K, Han H, Yamamoto N, Tashiro K, Ikegawa M, Kuroda K, Suzuki A, Nakano T and Honjo T. Notch-RBP-J signaling is involved in cell fate determination of marginal zone B cells. *Nature Immunology*. 2002; 3: 443-450.
 31. Wu QL, Zierold C and Ranheim EA. Dysregulation of Frizzled 6 is a critical component of B-cell leukemogenesis in a mouse model of chronic lymphocytic leukemia. *Blood*. 2009; 113: 3031-3039.
 32. Yauch RL, Gould SE, Scales SJ, Tang T, Tian H, Ahn CP, Marshall D, Fu L, Januario T, Kallop D, Nannini-Pepe M, Kotkow K, Marsters JC, Rubin LL and de Sauvage FJ. A paracrine requirement for hedgehog signalling in cancer. *Nature*. 2008; 455: 406-410.
 33. Lyons JP, Mueller UW, Ji H, Everett C, Fang X, Hsieh JC, Barth AIM and McCrea PD. Wnt-4 activates the canonical beta-catenin-mediated Wnt pathway and binds frizzled-6 CRD: functional implications of Wnt/beta-catenin activity in kidney epithelial cells. *Experimental Cell Research*. 2004; 298: 369-387.
 34. He TC, Sparks AB, Rago C, Hermeking H, Zawel L, da Costa LT, Morin PJ, Vogelstein B and Kinzler KW. Identification of c-MYC as a target of the APC pathway. *Science*. 1998; 281: 1509-1512.
 35. Tetsu O and McCormick F. beta-catenin regulates expression of cyclin D1 in colon carcinoma cells. *Nature*. 1999; 398: 422-426.
 36. Wielenga VJM, Smits R, Korinek V, Smit L, Kielman M, Fodde R, Clevers H and Pals ST. Expression of CD44 in Apc and Tcf mutant mice implies regulation by the WNT pathway. *American Journal of Pathology*. 1999; 154: 515-523.
 37. Ashworth A, Lord CJ and Reis JS. Genetic Interactions in Cancer Progression and Treatment. *Cell*. 2011; 145: 30-38.
 38. Deutscher D, Meilijson I, Kupiec M and Ruppin E. Multiple knockout analysis of genetic robustness in the yeast metabolic network. *Nature Genetics*. 2006; 38: 993-998.
 39. Horn T, Sandmann T, Fischer B, Axelsson E, Huber W and Boutros M. Mapping of signaling networks through synthetic genetic interaction analysis by RNAi. *Nature Methods*. 2011; 8: 341-346.
 40. Burrell RA, McGranahan N, Bartek J and Swanton C. The causes and consequences of genetic heterogeneity in cancer evolution. *Nature*. 2013; 501: 338-345.
 41. Lawrence MS, Stojanov P, Polak P, Kryukov GV, Cibulskis K, Sivachenko A, Carter SL, Stewart C, Mermel CH, Roberts SA, Kiezun A, Hammerman PS, McKenna A, Drier Y, Zou L, Ramos AH, et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature*. 2013; 499: 214-218.
 42. Livak KJ and Schmittgen TD. Analysis of relative gene

expression data using real-time quantitative PCR and the $2^{-(\Delta\Delta C_T)}$ method. *Methods*. 2001; 25: 402-408.

43. Trapnell C, Pachter L and Salzberg SL. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics*. 2009; 25: 1105-1111.
44. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ and Pachter L. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature Biotechnology*. 2010; 28: 511-515.
45. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL and Pachter L. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature Protocols*. 2012; 7: 562-578.