A novel isolation method for cancer prognostic factors via the p53 pathway by a combination of *in vitro* and *in silico* analyses

Yohey Kamijo[1,2,*], Kohichi Kawahara[1,*], Takuma Yoshinaga[3], Hiroyuki Kurata[4], Kazunari Arima[2], and Tatsuhiko Furukawa[1]

[1] Department of Molecular Oncology, Graduate School Medical and Dental Sciences, Kagoshima University, Kagoshima 890-8544, Japan

[2] Department of Chemistry and Bioscience, Faculty of Science, Graduate School of Science and Engineering, Kagoshima University, Kagoshima 890-0065, Japan

[3] Division of Clinical Application, Nanpuh Hospital, Kagoshima 892-8512, Japan

[4] Department of Bioscience and Bioinformatics, Kyushu Institute of Technology, Fukuoka 820-8502, Japan

[*] These authors contributed equally to the manuscript

Correspondence to: Kohichi Kawahara and Tatsuhiko Furukawa, **email**: k-kawahr@m3.kufm.kagoshima-u.ac.jp

Tatsuhiko Furukawa, **email**: furukawa@m3.kufm.kagoshima-u.ac.jp

ABSTRACT

Identifying new therapeutic target genes affecting the survival of patients with cancer is crucial for the development of new cancer therapies. Here, we developed a novel technology combining *in vitro* short hairpin RNA (shRNA) library screening and *in silico* analysis of the tumor transcriptome to identify prognostic factors via the p53 tumor-suppressor pathway. For initial screening, we screened 5,000 genes through selection of shRNAs in p53 wild-type tumor cells that altered sensitivity to the p53 activator actinomycin D (ActD) to identify p53 regulatory genes; shRNAs targeting 322 genes were obtained. Among these 322 genes, seven were prognostic factor candidates whose high expression increased ActD sensitivity while prolonging the survival period in patients with the p53 wild-type genotype. Conversely, we identified 33 genes as prognostic factor candidates among ActD-resistant genes related to a shortened survival period only in p53 wild-type tumors. These 40 genes had biological functions such as apoptosis, drug response, cell cycle checkpoint, and cell proliferation. The 40 genes selected by this method contained many known genes related to the p53 pathway and prognosis in patients with cancer. In summary, we developed an efficient screening method to identify p53-dependent prognostic factors with *in vitro* experimental data and database analysis.

INTRODUCTION

TP53 is one of the most commonly inactivated genes in human cancer; the gene product, p53, acts as a tumor suppressor. P53 promotes the transcription of a variety of genes, including genes encoding BCL2 associated X, apoptosis regulator *(BAX)* [1] and p53 upregulated modulator of apoptosis *(PUMA)* [2], which induce apoptosis, P21 [3] and *14-3-3σ* [4], which arrest the cell cycle; and growth arrest and DNA damage-inducible 45

(GADD45) [5,6], which repairs DNA damage. This variety of target genes confers p53 with tumor-suppressive functions. Under normal conditions, p53 protein levels are kept low by degradation via mouse double minute 2 homolog (MDM2), an E3 ubiquitin ligase for p53 [7,8]. Stresses such as radiation, DNA damage agents, or RNA synthesis inhibitors (e.g., actinomycin D [ActD]) [9-11], as well as oncogenic stress induced by abnormal activation of oncogenes, including Ras [12], and nutrient starvation [13] inhibit MDM2 function, resulting in accumulation

and activation of p53 and p53-dependent cellular responses [14]. P53 is mutated and shows loss of transcription factor activity in about half of all cancers [15]. Moreover, p53 mutational status also greatly affects the survival of patients with cancer [16,17]. For example, the expression of genes such as *PICT1*, which functionally regulates p53 activation, affects survival in patients with colorectal, esophageal, gastric, and lung cancers [18-21]. Moreover, novel p53 regulatory mechanisms, such as modulation of the nucleolar stress response, have also recently been revealed [22]. Thus, from the continuing evolution of our knowledge of p53, it is clear that the molecular mechanisms regulating complex p53 signaling pathways have not yet been fully clarified.

In recent years, high-throughput genomic analysis methods, such as next-generation sequencing (NGS) and genomic-scale microarrays, have been developed. Consequently, large amounts of information regarding the cancer genome and transcriptome have been obtained. These datasets are available from The Cancer Genome Atlas, Gene Expression Omnibus (GEO), and other databases and have greatly contributed to recent developments in tumor biology. Indeed, these analyses have demonstrated the occurrence of previously unreported genetic mutations at an early stage of carcinogenesis in patients with luminal A subtype breast cancer [23]. Furthermore, in other studies, researchers showed that there were no similar genomic mutation patterns in hypermutated colorectal cancer and rectal cancer [24]. Thus, it is becoming clear that each cancer type and subtype have unique expression patterns and/or genomic backgrounds, suggesting that known oncogenes and tumor-suppressor genes are crucial for tumor development. Additionally, bioinformatics analysis is becoming increasingly important because it has enabled us to identify oncogenes and cancer therapeutic targets, which may be difficult to identify using conventional molecular biological methods [23,24]. Accordingly, p53 functional experimental screening combined with bioinformatics technology may be useful for rapid identification of p53 regulatory target genes and therapeutic targets via the p53 pathway, which have been difficult to identify using conventional experimental methods or bioinformatics analysis alone.

In this study, we conducted p53-associated cell death screening combined with bioinformatics analysis to determine gene expression profiles in patients with cancer using the GEO. After sequential screening, we identified candidate 40 prognostic factors associated with p53 function in human cancer.

RESULTS

Combination of *in vitro* and *in silico* analyses for novel genes related to the p53 pathway

First, we performed RNAi screening using a lentiviral shRNA library to identify genes related to p53-dependent apoptosis (Figure 1A). This library comprised 27,500 shRNAs targeting approximately 5,000 genes. Each shRNA had a unique barcode sequence that could be used to identify the inserted shRNA. When infecting cells with lentivirus at a multiplicity of infection (MOI) of 0.1, the possibility that two or more lentiviral particles infected a single cell was less than 5%; thus, lentiviral dual infection events were suppressed [25]. Based on this result, p53 wild-type U2OS human osteosarcoma cells were infected with lentivirus carrying shRNA at an MOI of 0.1 in order to avoid multiple lentiviral infections.

The genes selected by shRNA library screening included genes controlling cell death and cell cycle, regardless of the p53 pathway. Therefore, prognostic factors in colorectal cancer depending on p53 mutational status were selected from the screened genes because only wild-type p53 regulated prognostic factors related to alteration of overall survival (Figure 1B).

We expect to identify prognostic genes related to the p53 pathway via a combination of pooled shRNA library screening to select genes that are sensitive or resistant to p53-activating drugs and bioinformatics analysis to assess the effects of genes on the survival of patients with cancer depending on p53 mutational status (Figure 1C).

Pooled shRNA library screening

In order to detect shRNA enrichment or depletion, shRNA barcodes from ActD-stimulated cells and DMSO-treated cells as controls were amplified and counted by NGS. Barcode reads in the ActD-stimulated and control groups were normalized such that the barcode read ratios of both groups could be accurately estimated (Figure 2A). One-hundred sixty-three shRNAs targeting 161 genes showed 1.75-fold change in barcode counts (Figure 2B). We designated these genes as ActD-sensitive genes. Additionally, 164 shRNAs targeting 161 genes showed less than 1/1.75 (0.57)-fold change in barcode counts (Figure 2B). We designated these genes as ActD-resistant genes (Figure 2B).

We conducted functional enrichment analysis with GeneCodis 3 (http://genecodis.cnb.csic.es) to evaluate the biological processes of ActD-sensitive and ActD-resistant genes. ActD-sensitive genes were involved in biological processes such as response to drug, DNA damage

response, and cell cycle (Figure 3A). ActD-resistant genes were involved in biological functions such as drug response, mitotic cell cycle, and DNA repair (Figure 3B). Notably, p53 shRNA was screened as targeting an ActD-sensitive gene (Figure 2B).

Because the screened genes were involved in cell protection, similar to p53 function, and p53 shRNA was selected, the screening selected genes related to the p53 pathway.

Analysis of microarray data

In our shRNA library screening, we assumed that ActD-sensitive and -resistant genes included genes related to the p53 pathway as well as genes that were not related to the p53 pathway but were still involved in cell viability. To select potential target genes related to the p53 pathway based on ActD sensitivity and resistance, we estimated the duration of survival for patients with colorectal cancer with or without *TP53* mutation using GEO datasets.

Microarray raw CEL files of 421 patients with colorectal cancer originated from GEO datasets (*TP53* wild-type: 200; *TP53* mutant: 221). After MAS5.0 normalization and quality control, 418 datasets were obtained (*TP53* wild-type: 197; *TP53* mutant: 221). We estimated colorectal cancer prognoses for each gene using Kaplan-Meier analysis with R to assess patient survival visually and statistically. Patients with colorectal cancer were divided into high and low expression groups for every gene based on the median signal intensity of each probe. As a result, 5974 genes (9258 probes) were found to be related to survival in TP53 wild-type colorectal cancer (Figure 4A).

Apoptosis-related genes involved in the p53 pathway selected by RNAi screening were expected to induce or suppress cell death following ActD stimulation. Similar to p53, high expression of ActD-sensitive genes was found to improve prognosis in patients with *TP53* wild-type cancer. Conversely, survival was worse in the context of high expression of ActD-resistant genes in patients with *TP53* wild-type cancer. Thus, we chose genes that were significantly correlated with prognosis from ActD-sensitive or -resistant genes in patients harboring
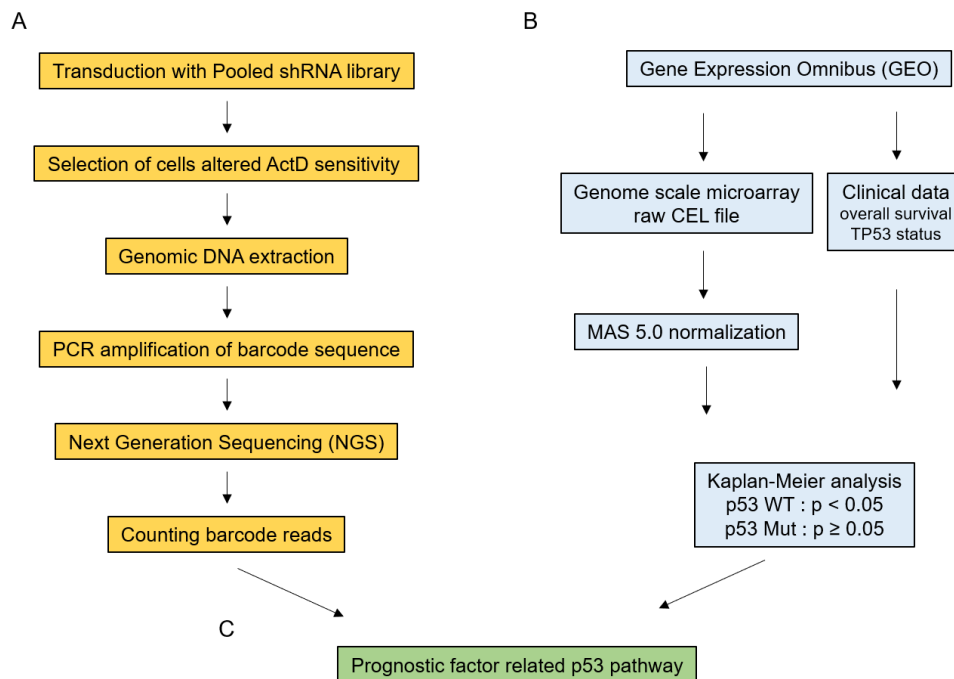

Figure 1: Flowchart of prognosis-related gene identification for targets involved in the p53 pathway by combining pooled shRNA library screening and bioinformatics analysis. (A) Pooled shRNA library screening for identification of ActD-sensitive or -resistant genes. *TP53* wild-type cancer cells were transduced with a pooled shRNA library. Cells showing increased or decreased ActD sensitivity by gene silencing were selected following ActD stimulation. Genomic DNA from selected cells was extracted. The shRNA barcode sequence inserted into genomic DNA was amplified by PCR. Amplified barcode reads were counted using NGS. The same procedures were carried out for DMSO-treated cells as controls. The fold-change (FC) of barcode reads was calculated. **(B)** Survival analysis of patients with cancer based on *TP53* mutational status. The datasets for the human colorectal cancer DNA microarray (GSE39084 and GSE39582) were obtained from the GEO. Each raw CEL file was normalized to MAS 5.0. Probes used to measure gene expression were grouped based on the median level of gene expression in *TP53* wild-type (WT) and mutant (Mut) cancers. The overall survival of patients with cancer was estimated by the Kaplan-Meier method. Significant differences were analyzed using log-rank tests (p < 0.05). Genes with significant difference in *TP53* WT samples but without significant difference in *TP53* Mut samples were selected. **(C)** Genes that showed altered ActD sensitivity in pooled shRNA library screening (A) were combined with genes that affected prognosis in a p53-dependent manner (B). Genes included in both lists were selected as prognostic factors related to the p53 pathway.

wild-type *TP53*. In these patients, 43 ActD-sensitive genes and 47 ActD-resistant genes significantly altered prognosis (Figure 4A). We assumed that these genes did not alter survival outcomes significantly in patients with *TP53* mutant cancer because mutant *TP53* enhanced malignant progression. Therefore, we selected genes that did not alter survival in patients with mutant *TP53*. From this analysis, 7 ActD-sensitive genes and 33 ActD-resistant genes showed a significant prognostic correlation in patients with *TP53* wild-type colorectal cancer, but no significant prognostic correlation in patients with the *TP53* mutant (Figure 4B).

Finally, 40 genes were analyzed for function using Gene Ontology analysis. These genes were found to be associated with p53-related functions, including apoptotic process, drug response, cell cycle checkpoint, and cell proliferation (Figure 5).

DISCUSSION

In the present study, we selected 40 genes that altered survival prognosis in patients with colorectal cancer in a p53 function-dependent manner by combining shRNA library screening and database analysis of gene expression. Many genes selected by this combination method were involved in processes relevant to the p53 pathway, including apoptosis, drug response, cell cycle checkpoint, and cell proliferation. In fact, these genes included known p53 pathway genes, such as *CCNB1* [26], *ANKHD1* [27], nuclear receptor coactivator 3 (*NCOA3*) [28], *SIPA1* [29], *IKBKB* [30], and *LITAF* [31].

Consequently, this method combining *in vitro* shRNA library screening with *in silico* survival prognostic analysis was found to be effective for selection of genes related to the p53 pathway. We also examined whether this method could be used to identify prognostic factors related to the p53 pathway in patients with cancer. Importantly, *TP53* status has been shown to affect prognosis in patients with colon cancer because patients with p53 mutant colorectal cancer have poorer survival outcomes than patients with p53 wild-type cancer [32,33]. In addition, the gene expression level of the p53 pathway regulator *PICT1* has been correlated with prognosis in patients with p53 wild-type cancer [18-21]. Thus, a method based on analysis of p53 mutational status-dependent survival in patients with cancer may be useful for effective identification of survival factors via the p53 pathway.

Of the ActD-sensitive genes selected by shRNA library screening, high expression of seven genes showed a better prognosis in patients with p53 wild-type colorectal cancer. Moreover, high expression of 33 ActD-resistant genes was associated with poor prognosis in patients with *TP53* wild-type cancer (Figure 6A, Supplementary Figure 1). All 40 genes, however, did not affect prognosis in patients with *TP53* mutant cancer (Figure 6B, Supplementary Figure 1), demonstrating a dependence on *TP53* mutational status. The genes selected using this method included *NCOA3* [34,35], *HOXA1* [36], *FOLR1* [37], *SOCS1* [38], and *PIK4CA* [39], which showed similar expression patterns related to survival in patients with cancer. This result suggested that our method was a relatively simple method for identification of genes
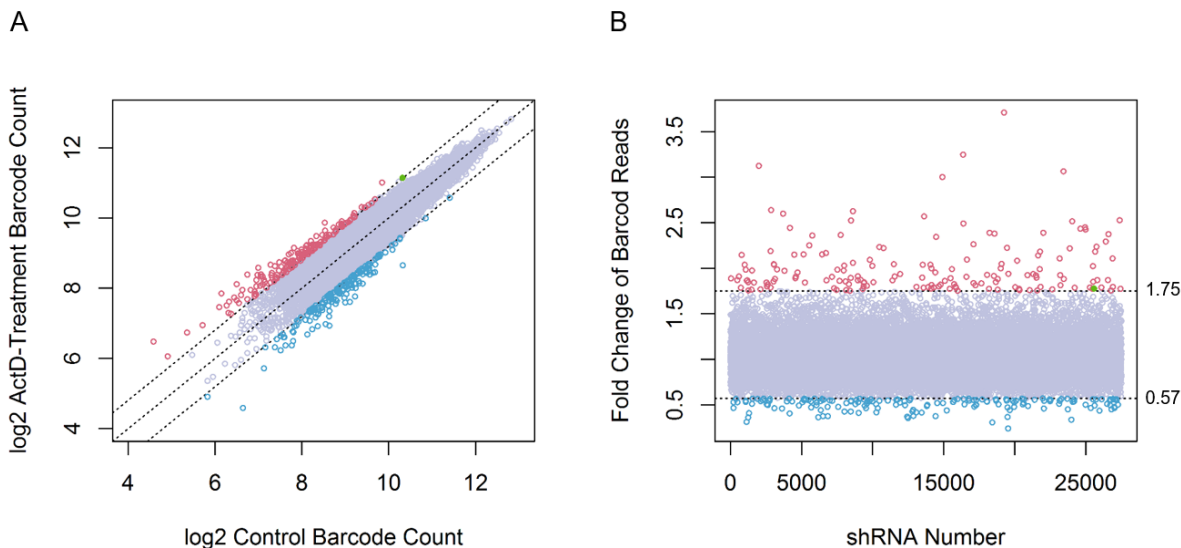
A

B

Figure 2: Selection of candidate p53-regulatory genes or p53 target genes by pooled shRNA library screening. (A) *TP53* wild-type U2OS cells were transduced with pooled shRNA library by lentivirus infection. After puromycin selection, infected cells were treated with ActD or DMSO as a control for 3 days. Genomic DNA was extracted from ActD-treated U2OS cells. The barcode sequence of the inserted shRNA was analyzed by NGS to count the barcode reads. Barcode read counts were log2 transformed. The plot indicates correlations of barcode reads in ActD- or DMSO-treated samples. Red circles, shRNAs targeting ActD-sensitive genes; blue circles, shRNAs targeting ActD-resistant genes; filled green circle, shRNAs targeting *TP53*. **(B)** Fold-changes in barcode reads for each shRNA. Red circles indicate 1.75-fold changes in shRNA barcode reads. Blue circles indicate less than 0.57-fold changes in shRNA barcode reads. Filled green circle indicates *TP53* shRNA.

functionally associated with the p53 pathway.

Notably, the *NCOA3* gene identified by our combination method was shown to be involved in both regulation of the p53 pathway and determination of survival in patients with cancer. NCOA3 protein promotes the transcription of the *TRAF4* gene, which encodes tumor necrosis factor receptor-associated factor 4 (TRAF4). High TRAF4 expression induces downregulation of p53 because TRAF4 destabilizes p53 by inhibiting the p53 deubiquitinase herpesvirus-associated ubiquitin-specific protease [28]. Furthermore, patients with breast cancer with high *NCOA3* expression have a poorer prognosis than patients with low *NCOA3* expression [34,35]. In our *in silico* analysis of a colorectal cancer gene expression database, the expression detected by two independent microarray probes for *NCOA3* indicated a significant correlation with survival in patients with colorectal cancer depending on *TP53* genotype status; thus, these results suggest that *NCOA3* was a prognostic factor via the p53 pathway. These findings strongly support that our method was an excellent approach for selecting prognostic factors regulated by the p53 pathway in patients with cancer (Figure 6A, 6B, Supplementary Figure 1).

In this study, we provided a convenient method to identify prognostic factors functionally associated with p53 from a large number of candidate genes. In many studies, several genes involved in the p53 pathway have been identified *in vitro* [40], *in vivo* [41], or both. By conventional methods, genes of interest have been used to investigate phenotypic changes in cells and model organisms, such as knockout mice. However, these methods are expensive and time consuming owing to the processes of genetic engineering and construction of analytical systems. Accordingly, it has been difficult to identify genes involved in the p53 pathway from many candidate genes within a short period of time. Moreover, functional analyses of individual genes are still required in order to identify prognostic factors associated with the tumor-suppressor function of p53. These analyses are also limited by the high cost and time-consuming procedures and are therefore not convenient to perform. Furthermore, false-positive genes in candidates selected through large-scale functional screening with library expression analysis can make it difficult to identify true-positives.

In this study, we provided a simple and highly accurate technology to estimate candidate gene functionally related to the p53 pathway without interference of false-positive genes and without the need to perform validation of individual genes. We used osteosarcoma U2OS cells *in vitro* and patients with colorectal cancer *in silico* in our screening because U2OS cells are suitable for the study of p53 signaling and there is sufficient data for a number of patients associated with p53 mutational status in colorectal cancer available in the GEO database for technical issues. We believe our method using two different types of cancer is suitable for isolation of p53-signaling genes for the following two reasons: 1) to a considerable extent, p53 signaling is common
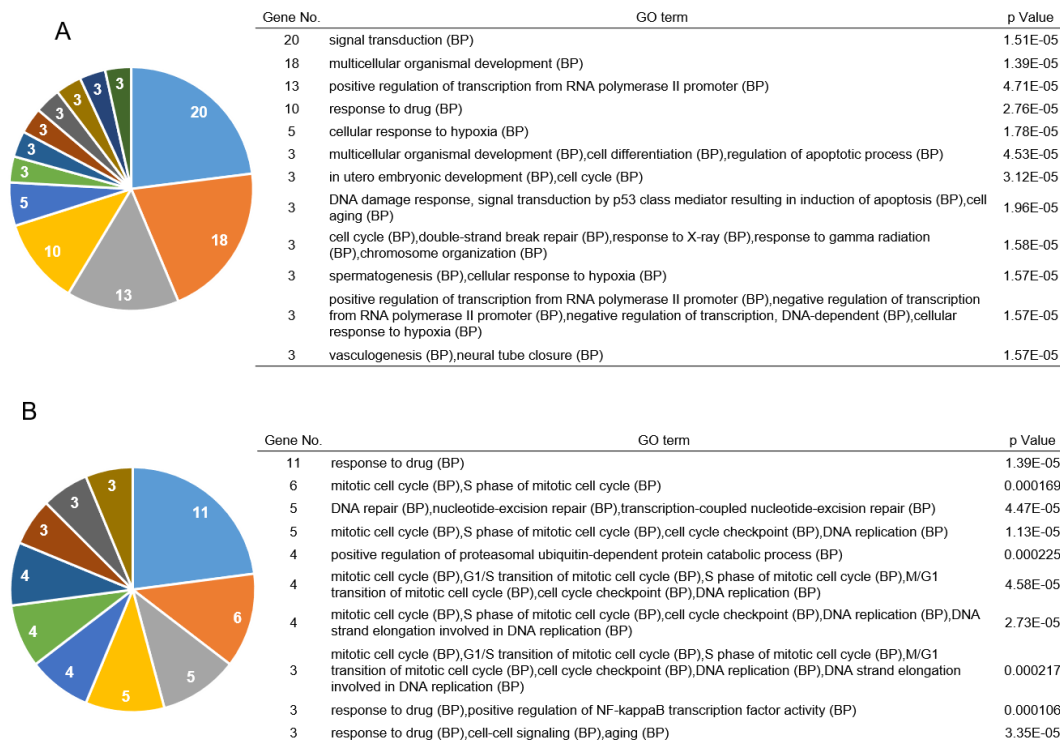

A

Gene No.	GO term	p Value
20	signal transduction (BP)	1.51E-05
18	multicellular organismal development (BP)	1.39E-05
13	positive regulation of transcription from RNA polymerase II promoter (BP)	4.71E-05
10	response to drug (BP)	2.76E-05
5	cellular response to hypoxia (BP)	1.78E-05
3	multicellular organismal development (BP),cell differentiation (BP),regulation of apoptotic process (BP)	4.53E-05
3	in utero embryonic development (BP),cell cycle (BP)	3.12E-05
3	DNA damage response, signal transduction by p53 class mediator resulting in induction of apoptosis (BP),cell aging (BP)	1.96E-05
3	cell cycle (BP),double-strand break repair (BP),response to X-ray (BP),response to gamma radiation (BP),chromosome organization (BP)	1.58E-05
3	spermatogenesis (BP),cellular response to hypoxia (BP)	1.57E-05
3	positive regulation of transcription from RNA polymerase II promoter (BP),negative regulation of transcription from RNA polymerase II promoter (BP),negative regulation of transcription, DNA-dependent (BP),cellular response to hypoxia (BP)	1.57E-05
3	vasculogenesis (BP),neural tube closure (BP)	1.57E-05

B

Gene No.	GO term	p Value
11	response to drug (BP)	1.39E-05
6	mitotic cell cycle (BP),S phase of mitotic cell cycle (BP)	0.000169
5	DNA repair (BP),nucleotide-excision repair (BP),transcription-coupled nucleotide-excision repair (BP)	4.47E-05
5	mitotic cell cycle (BP),S phase of mitotic cell cycle (BP),cell cycle checkpoint (BP),DNA replication (BP)	1.13E-05
4	positive regulation of proteasomal ubiquitin-dependent protein catabolic process (BP)	0.000225
4	mitotic cell cycle (BP),G1/S transition of mitotic cell cycle (BP),S phase of mitotic cell cycle (BP),M/G1 transition of mitotic cell cycle (BP),cell cycle checkpoint (BP),DNA replication (BP)	4.58E-05
4	mitotic cell cycle (BP),S phase of mitotic cell cycle (BP),cell cycle checkpoint (BP),DNA replication (BP),DNA strand elongation involved in DNA replication (BP)	2.73E-05
3	mitotic cell cycle (BP),G1/S transition of mitotic cell cycle (BP),S phase of mitotic cell cycle (BP),M/G1 transition of mitotic cell cycle (BP),cell cycle checkpoint (BP),DNA replication (BP),DNA strand elongation involved in DNA replication (BP)	0.000217
3	response to drug (BP),positive regulation of NF-kappaB transcription factor activity (BP)	0.000106
3	response to drug (BP),cell-cell signaling (BP),aging (BP)	3.35E-05

Figure 3: Functional analysis of genes identified in the shRNA library screening. Results of Gene Ontology (GO) enrichment analysis. **(A)** ActD-sensitive and **(B)** ActD-resistant genes identified by shRNA screening were analyzed using GeneCodis3 to identify the related biological processes. Significantly enriched genes are shown. (A) $p \leq 4.70713 \times 10^{-5}$, (B) $p \leq 2.24936 \times 10^{-4}$.

between osteosarcoma and colorectal cancer [40,42]; and 2) using different types of cancer may generally reduce the screening biases generated by a specific single cancer cell type. Furthermore, our method may be applicable to other genes, such as *HER2* amplification in patients with breast cancer, because patients with breast cancer and amplified *HER2* have poor survival compared to those patients with normal *HER2* [43]. However, some genes selected in this study have not been reported to be involved in p53 regulation or survival in patients with cancer. Thus, further functional validation of these genes is needed. Identification of prognostic factors via the p53 pathway may lead to the identification of novel therapeutic target genes in patients with cancer.

In conclusion, we established a combination method with large-scale shRNA library screening *in vitro* and *in silico* analysis of gene expression databases for patients with cancer to identify prognostic factors based on a functional biological pathway. Our method established a new technology for the identification of prognostic factors associated with the p53 pathway in patients with cancer, enabling analysis of thousands of genes at one time. The genes selected by this method affected the survival of patients with cancer. Thus, these genes may be novel targets of anticancer drugs.

MATERIALS AND METHODS

Cell culture and treatments

U2OS and HEK293FT cells were obtained from the American Type Culture Collection (Manassas, VA, USA) and Thermo Fisher Scientific (Waltham, MA, USA), respectively. Cells were cultured in Dulbecco's modified Eagle's medium (DMEM; Nissui, Tokyo, Japan) containing 10% fetal bovine serum in an atmosphere containing 5% CO_2 at 37°C with 100% humidity.

Preparation of lentivirus

HEK293FT cells (6.0×10^6 cells) were cultured in 10-cm dishes and cultured for 1 day. The plasmids pMDLg/pRRE (6.5 μg), pRSV-Rev (2.5 μg), pMD2.G (3.5 μg), and 2.5 μg of Human DECIPHER Module 1 (Cellecta, Mountain View, CA, USA), i.e., lentivirus vectors expressing 27,500 shRNAs targeting approximately 5,000 genes involved in cell signaling, were mixed in 500 μL serum-free DMEM. Separately, 20 μL of
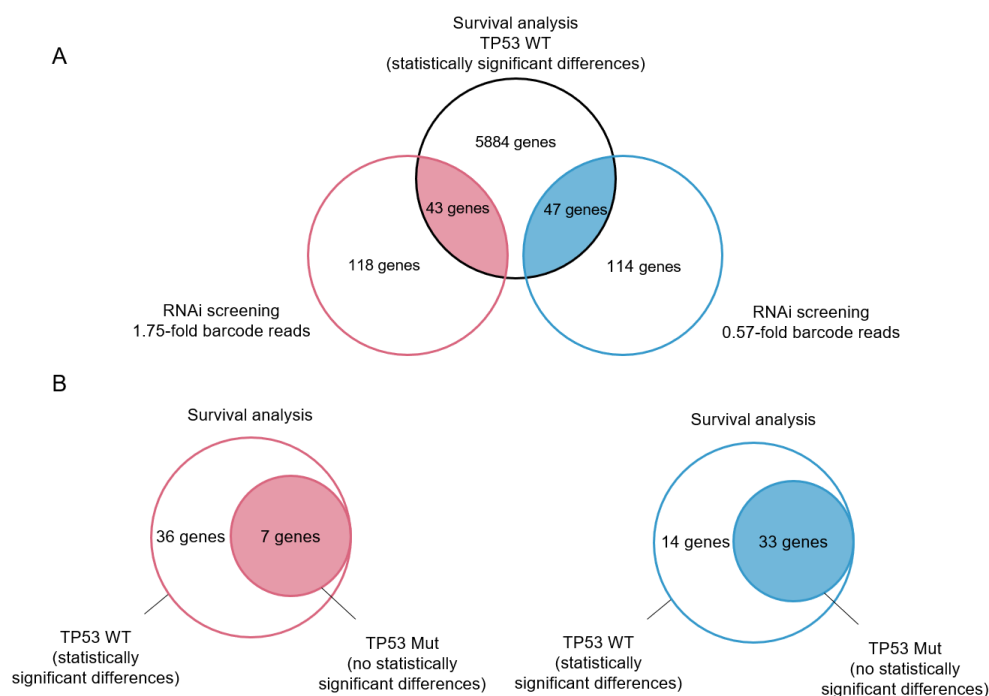

Figure 4: Identification of genes associated with survival in patients with cancer with regard to p53 mutation status using bioinformatics. (A) Venn diagram indicating the overlap of genes related to survival in patients with *TP53* WT colorectal cancer and genes selected by shRNA library screening. Patient survival was estimated using the Kaplan-Meier method. Statistical significance was determined by log-rank tests ($p < 0.05$). Red, 1.75-fold changes in barcode reads indicating ActD-sensitive genes; blue, less than 0.57-fold changes in barcode reads indicating ActD-resistant genes. **(B)** Venn diagram indicating overlap of genes not related to prognoses in colorectal cancer specimens with *TP53* Mut and genes in (A). Red, ActD-sensitive genes; blue, ActD-resistant genes. Survival analysis and statistical tests were the same as in (A).

Lipofectamine 2000 (Thermo Fisher Scientific) was mixed with 500 µL serum-free DMEM. These two solutions were mixed. Then, HEK293FT cells were transfected with the plasmid mixture and cultured for 2.5 days. The cultured medium was centrifuged at 3,500 rpm for 5 min, and the supernatant containing lentiviral particles was then harvested. The supernatants were stored at -80°C.

Because Human DECIPHER Module 1 contained the red fluorescent protein (RFP) coding sequence, cells infected with lentivirus expressed RFP. To determine the lentiviral titer, the numbers of U2OS cells expressing RFP and of 4',6-diamidino-2-phenylindole dihydrochloride (DAPI; Dojindo, Kumamoto, Japan)-positive U2OS cells were counted after lentiviral infection. U2OS cells were plated in 6-well plates at 5.0×10^5 cells/well. The prepared lentivirus stock solution was serially diluted twice in DMEM to a final volume of 500 µL. The solutions were then added to the cells. After 2 days, lentivirus-infected cells were stained with DAPI diluted 4,000 times for 10 min. The cells were observed with a confocal laser scanning microscope (LSM 700; Carl Zeiss, Oberkochen, Germany). Cells stained with DAPI and expressing RFP were counted as the total number of cells and the number of cells harboring shRNA, respectively. A dilution ratio close to 10% was defined as an MOI of 0.1.

Genome-wide RNAi screening

For RNAi screening, 3.0×10^6 U2OS cells were plated in 8.9 mL DMEM in 20 dishes (10-cm) and infected with lentiviral particles at an MOI of 0.1. Since lentivirus-infected cells exhibited resistance to puromycin (Invivogen, San Diego, CA, USA), the cells were selected with 2.5 µg/mL puromycin. After 2 days, puromycin-selected cells were stimulated with 7.5 nM ActD (Wako, Osaka, Japan; the half-maximal inhibitory concentration [IC_{50}]) for 3 days in order to induce p53 hyperactivation-related apoptosis. Genomic DNA was isolated from stimulated cells, and the barcode sequence was amplified by PCR. Amplified fragments were sequenced by NGS (Cosmo Bio, Tokyo, Japan), and changes in ActD sensitivity were evaluated by counting shRNA barcode reads. Target gene silencing by shRNA was found to promote apoptosis through p53 when the number of barcode reads in the ActD-stimulated group was greater than that in the control group. Conversely, target gene silencing by shRNA was found to suppress apoptosis through p53 when the number of barcode reads in the ActD-stimulated group was lesser than that in the control group.

Collection and analysis of human colon and colorectal cancer microarray data

The datasets from GSE 39084 and GSE 39582 (*TP53* wild-type: 200 specimens, *TP53* mutant: 221 specimens, *TP53* mutation unknown: 215 specimen) were downloaded from GEO. The raw CEL files were normalized by MAS 5.0 using an Expression Console (Affymetrix, Thermo Fisher Scientific) with an Affymetrix mask file. Quality control (background < 120, RawQ < 10, percent present calls < 55, scaling factor < 3, GAPDH 3′ to 5′ ratio < 5, beta-actin 3′ to 5′ ratio < 5) was performed. Probe signal intensities were converted to log2 values with R. After log2 conversion, probes with zero values were excluded.

Survival analysis

Patients with *TP53* wild-type or mutant colorectal cancer were divided into high and low expression groups according to the median signal intensity of each microarray probe. The overall survival in the two groups was compared by the Kaplan-Meier method using the R-3.1.2 and "survival" packages. Significant differences
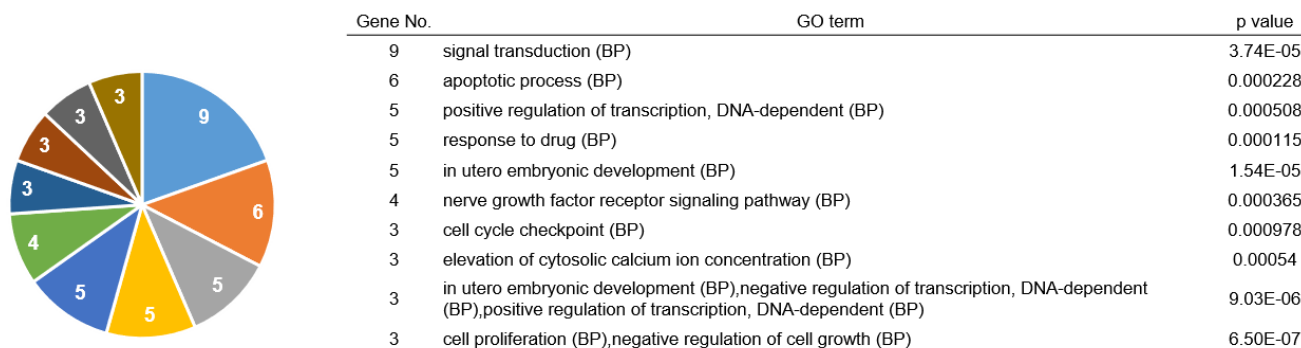

Gene No.	GO term	p value
9	signal transduction (BP)	3.74E-05
6	apoptotic process (BP)	0.000228
5	positive regulation of transcription, DNA-dependent (BP)	0.000508
5	response to drug (BP)	0.000115
5	in utero embryonic development (BP)	1.54E-05
4	nerve growth factor receptor signaling pathway (BP)	0.000365
3	cell cycle checkpoint (BP)	0.000978
3	elevation of cytosolic calcium ion concentration (BP)	0.00054
3	in utero embryonic development (BP),negative regulation of transcription, DNA-dependent (BP),positive regulation of transcription, DNA-dependent (BP)	9.03E-06
3	cell proliferation (BP),negative regulation of cell growth (BP)	6.50E-07

Figure 5: Functional analysis of genes selected by pooled shRNA library screening and survival analysis. Results of Gene Ontology (GO) enrichment analysis for 40 genes related to the survival of patients with *TP53* wild-type colorectal cancer and genes not related to prognosis in patients with *TP53* mutant cancer selected by shRNA library screening. Enrichment analysis of biological processes was performed using GeneCodis 3. Significantly enriched genes are shown ($p \leq 0.00097773$).
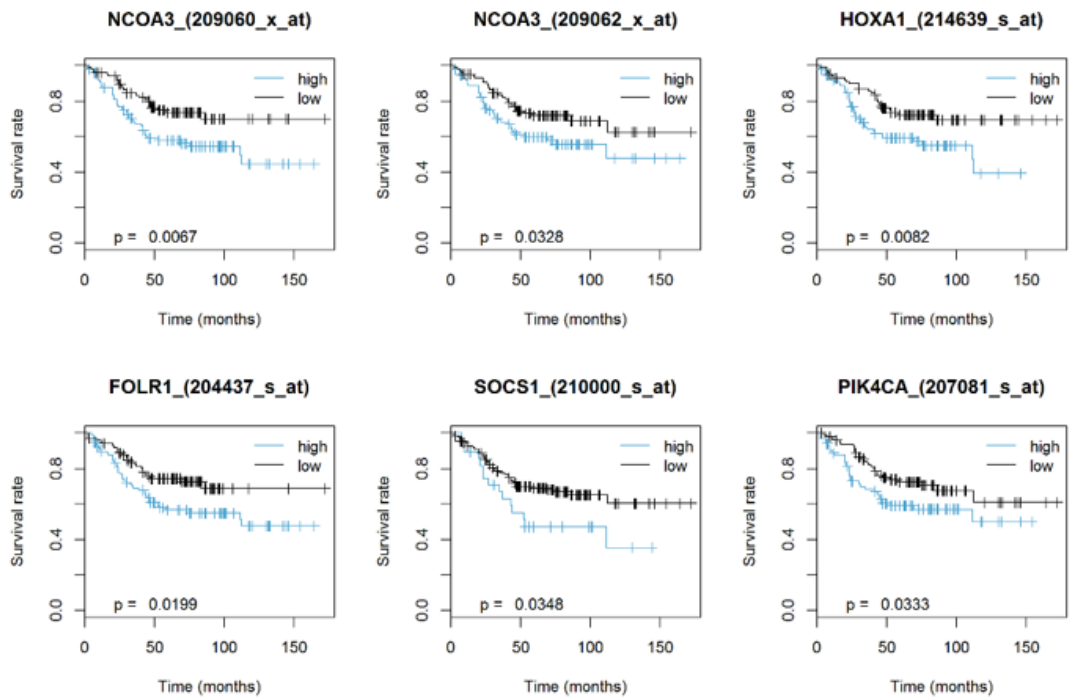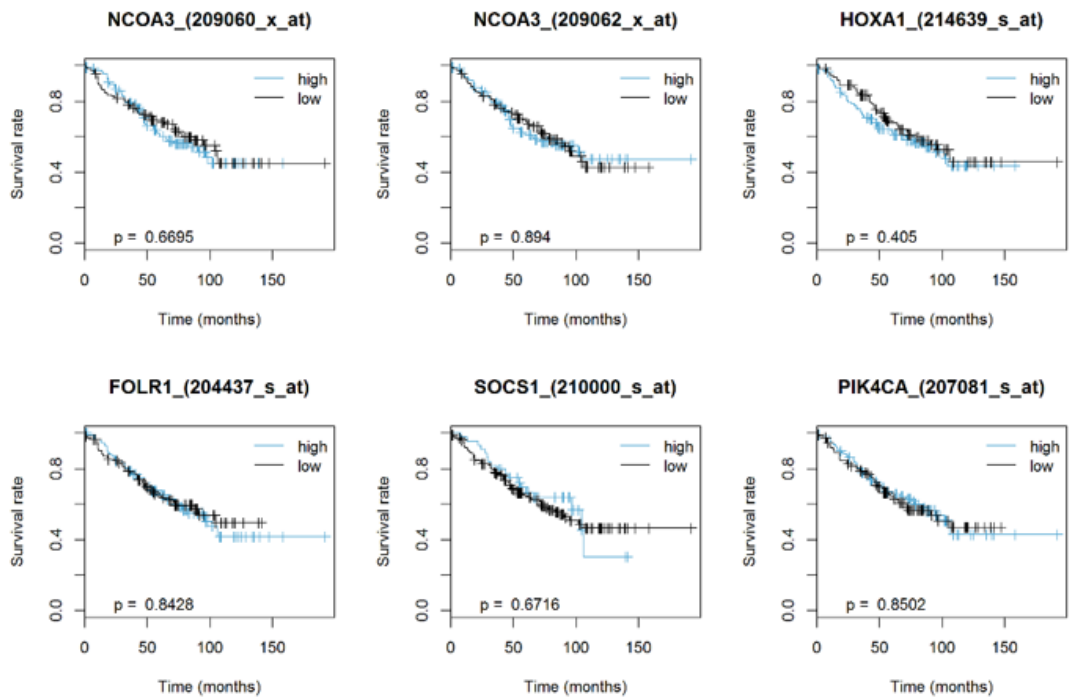
Figure 6: Expression of selected candidate genes influenced the survival of patients with colorectal cancer in a p53-dependent manner. Kaplan-Meier survival plots for patients with (A) *TP53* WT and (B) Mut colorectal cancer to determine the effects of selected genes (*NCOA3*, *HOXA1*, *FOLR1*, *SOCS1*, and *PIK4CA*) selected by shRNA library screening and database analysis on survival. Log rank tests were used for statistical analysis ($p < 0.05$).

in survival were calculated by log-rank tests and defined as having a *P* value of less than 0.05. This analysis was performed in patients with *TP53* wild-type and mutant cancer.

Functional analysis of prognostic factors

Gene Ontology analysis was performed using GeneCodis 3 (http://genecodis.cnb.csic.es/analysis). Annotations were conducted with GO biological processes. *HILPDA* (also called *HIG2*), *RPS17P16* (also called *LOC402057*), and *CD31* (also called *PECAM1*) were not registered in GeneCodis 3 and were therefore excluded from this analysis.

Author Contributions

Conception and design, K.K., and T.F.; experimental analysis and interpretation, Y.K., and K.K.; bioinformatical analysis and interpretation, Y.K., and H.K.; writing and revisions, Y.K., K.K., T.Y., H.K., K.A., and T.F.

ACKNOWLEDGEMENTS

CONFLICT OF INTEREST

The authors have no conflicts of interest to declare.

REFERENCES

1. Miyashita T, Reed JC. Tumor suppressor p53 is a direct transcriptional activator of the human bax gene. Cell. 1995; 80:293–99.

2. Nakano K, Vousden KH, Ashcroft M, Taya Y, Ashkenazi A, Dixit VM, Bates S, Bennett M, Macdonald K, Chan SW, Luzio JP, Simari R, Weissberg P, et al. PUMA, a novel proapoptotic gene, is induced by p53. Mol Cell. 2001; 7:683–94.

3. Bouvard V, Zaitchouk T, Vacher M, Duthu A, Canivet M, Choisy-Rossi C, Nieruchalski M, May E. Tissue and cell-specific expression of the p53-target genes: bax, fas, mdm2 and waf1/p21, before and following ionising irradiation in mice. Oncogene. 2000; 19:649–60.

4. Hermeking H, Lengauer C, Polyak K, He TC, Zhang L, Thiagalingam S, Kinzler KW, Vogelstein B. 14-3-3σ is a p53-regulated inhibitor of G2/M progression. Mol Cell. 1997; 1:3–11.

5. Kastan MB, Zhan Q, el-Deiry WS, Carrier F, Jacks T, Walsh WV, Plunkett BS, Vogelstein B, Fornace AJ Jr. A mammalian cell cycle checkpoint pathway utilizing p53 and GADD45 is defective in ataxia-telangiectasia. Cell. 1992; 71:587–97.

6. Smith ML, Chen IT, Zhan Q, Bae I, Chen CY, Gilmer TM, Kastan MB, O'Connor PM, Fornace AJ Jr. Interaction of the p53-regulated protein Gadd45 with proliferating cell nuclear antigen. Science. 1994; 266:1376–80.

7. Kubbutat MH, Jones SN, Vousden KH. Regulation of p53 stability by Mdm2. Nature. 1997; 387:299–303.

8. Honda R, Tanaka H, Yasuda H. Oncoprotein MDM2 is a ubiquitin ligase E3 for tumor suppressor p53. FEBS Lett. 1997; 420:25–27.

9. Kastan MB, Onyekwere O, Sidransky D, Vogelstein B, Craig RW. Participation of p53 protein in the cellular response to DNA damage. Cancer Res. 1991; 51:6304–11.

10. Perry RP, Kelley DE. Inhibition of RNA synthesis by actinomycin D: characteristic dose-response of different RNA species. J Cell Physiol. 1970; 76:127–39.

11. Perry RP, Kelley DE. Persistent synthesis of 5S RNA when production of 28S and 18S ribosomal RNA is inhibited by low doses of actinomycin D. J Cell Physiol. 1968; 72:235–46.

12. Bos JL. ras Oncogenes in Human Cancer : A Review. Cncaer Research. 1989: 4682-4689.

13. Lee CH, Inoki K, Karbowniczek M, Petroulakis E, Sonenberg N, Henske EP, Guan KL. Constitutive mTOR activation in TSC mutants sensitizes cells to energy starvation and genomic damage via p53. EMBO J. 2007; 26:4812–23.

14. Kruse JP, Gu W. Modes of p53 regulation. Cell. 2009; 137:609–22.

15. Vousden KH, Lu X. Live or let die: the cell's response to p53. Nat Rev Cancer. 2002; 2:594–604.

16. Young KH, Leroy K, Møller MB, Colleoni GW, Sánchez-Beato M, Kerbauy FR, Haioun C, Eickhoff JC, Young AH, Gaulard P, Piris MA, Oberley TD, Rehrauer WM, et al. Structural profiles of TP53 gene mutations predict clinical outcome in diffuse large B-cell lymphoma: an international collaborative study. Blood. 2008; 112:3088–98.

17. O'Shea D, O'Riain C, Taylor C, Waters R, Carlotti E, Macdougall F, Gribben J, Rosenwald A, Ott G, Rimsza

LM, Smeland EB, Johnson N, Campo E, et al. The presence of TP53 mutation at diagnosis of follicular lymphoma identifies a high-risk group of patients with shortened time to disease progression and poorer overall survival. Blood. 2008; 112:3126–29.

18. Sasaki M, Kawahara K, Nishio M, Mimori K, Kogo R, Hamada K, Itoh B, Wang J, Komatsu Y, Yang YR, Hikasa H, Horie Y, Yamashita T, et al. Regulation of the MDM2-P53 pathway and tumor growth by PICT1 via nucleolar RPL11. Nat Med. 2011; 17:944–51.

19. Uchi R, Kogo R, Kawahara K, Sudo T, Yokobori T, Eguchi H, Sugimachi K, Maehama T, Mori M, Suzuki A, Komune S, Mimori K. PICT1 regulates TP53 via RPL11 and is involved in gastric cancer progression. Br J Cancer. 2013; 109:2199–206.

20. Okamura K, Takayama K, Kawahara K, Harada T, Nishio M, Otsubo K, Ijichi K, Kohno M, Iwama E, Fujii A, Ota K, Koga T, Okamoto T, et al. PICT1 expression is a poor prognostic factor in non-small cell lung cancer. Oncoscience. 2014; 1:375–82.

21. Suzuki A, Kogo R, Kawahara K, Sasaki M, Nishio M, Maehama T, Sasaki T, Mimori K, Mori M. A new PICTure of nucleolar stress. Cancer Sci. 2012; 103:632–37.

22. Golomb L, Volarevic S, Oren M. p53 and ribosome biogenesis stress: the essentials. FEBS Lett. 2014; 588:2571–79.

23. Network TC, Cancer T, Atlas G, Muzny DM, Bainbridge MN, Chang K, Dinh HH, and Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. Nature. 2012; 490:61–70.

24. Muzny DM, Bainbridge MN, Chang K, Dinh HH, Drummond JA, Fowler G, Kovar CL, Lewis LR, Morgan MB, Newsham IF, Reid JG, Santibanez J, Shinbrot E, et al, and Cancer Genome Atlas Network. Comprehensive molecular characterization of human colon and rectal cancer. Nature. 2012; 487:330–37.

25. Nolan-Stevaux O, Tedesco D, Ragan S, Makhanov M, Chenchik A, Ruefli-Brasse A, Quon K, Kassner PD. Measurement of Cancer Cell Growth Heterogeneity through Lentiviral Barcoding Identifies Clonal Dominance as a Characteristic of In Vivo Tumor Engraftment. PLoS One. 2013; 8:e67316.

26. Fang Y, Yu H, Liang X, Xu J, Cai X. Chk1-induced CCNB1 overexpression promotes cell proliferation and tumor growth in human colorectal cancer. Cancer Biol Ther. 2014; 15:1268–79.

27. Dhyani A, Duarte AS, Machado-Neto JA, Favaro P, Ortega MM, Olalla Saad ST. ANKHD1 regulates cell cycle progression and proliferation in multiple myeloma cells. FEBS Lett. 2012; 586:4311–18.

28. Yi P, Xia W, Wu RC, Lonard DM, Hung MC, O'Malley BW. SRC-3 coactivator regulates cell resistance to cytotoxic stress via TRAF4-mediated p53 destabilization. Genes Dev. 2013; 27:274–87.

29. Tanaka H, Tamura A, Sekai M, Hamazaki Y, Minato N. Increased c-Myc activity and DNA damage in hematopoietic progenitors precede myeloproliferative disease in Spa-1-deficiency. Cancer Sci. 2011; 102:784–91.

30. Yang PM, Huang WC, Lin YC, Huang WY, Wu HA, Chen WL, Chang YF, Chou CW, Tzeng CC, Chen YL, Chen CC. Loss of IKKbeta activity increases p53 stability and p21 expression leading to cell cycle arrest and apoptosis. J Cell Mol Med. 2010; 14:687–98.

31. Zhu J, Jiang J, Zhou W, Zhu K, Chen X. Differential regulation of cellular target genes by p53 devoid of the PXXP motifs with impaired apoptotic activity. Oncogene. 1999; 18:2149–55.

32. Hamelin R, Laurent-Puig P, Olschwang S, Jego N, Asselain B, Remvikos Y, Girodet J, Salmon RJ, Thomas G. Association of p53 mutations with short survival in colorectal cancer. Gastroenterology. 1994; 106:42–48.

33. Tortola S, Marcuello E, González I, Reyes G, Arribas R, Aiza G, Sancho FJ, Peinado MA, Capella G. p53 and K-ras gene mutations correlate with tumor aggressiveness but are not of routine prognostic value in colorectal cancer. J Clin Oncol. 1999; 17:1375–81.

34. Burandt E, Jens G, Holst F, Jänicke F, Müller V, Quaas A, Choschzick M, Wilczak W, Terracciano L, Simon R, Sauter G, Lebeau A. Prognostic relevance of AIB1 (NCoA3) amplification and overexpression in breast cancer. Breast Cancer Res Treat. 2013; 137:745–53.

35. Zhao C, Yasui K, Lee CJ, Kurioka H, Hosokawa Y, Oka T, Inazawa J. Elevated expression levels of NCOA3, TOP1, and TFAP2C in breast tumors as predictors of poor prognosis. Cancer. 2003; 98:18–23.

36. Bitu CC, Destro MF, Carrera M, da Silva SD, Graner E, Kowalski LP, Soares FA, Coletta RD. HOXA1 is overexpressed in oral squamous cell carcinomas and its expression is correlated with poor prognosis. BMC Cancer. 2012; 12:146.

37. Notaro S, Reimer D, Fiegl H, Schmid G, Wiedemair A, Rössler J, Marth C, Zeimet AG. Evaluation of folate receptor 1 (FOLR1) mRNA expression, its specific promoter methylation and global DNA hypomethylation in type I and type II ovarian cancers. BMC Cancer. 2016; 16:589.

38. Ayyildiz T, Dolar E, Adim SB, Eminler AT, Yerci O. Lack of prognostic significance of SOCS-1 expression in colorectal adenocarcinomas. Asian Pac J Cancer Prev. 2014; 15:8469–74.

39. Ilboudo A, Nault JC, Dubois-Pot-Schneider H, Corlu A, Zucman-Rossi J, Samson M, Le Seyec J. Overexpression of phosphatidylinositol 4-kinase type IIIα is associated with undifferentiated status and poor prognosis of human hepatocellular carcinoma. BMC Cancer. 2014; 14:7–7.

40. Freeman DJ, Li AG, Wei G, Li HH, Kertesz N, Lesche R, Whale AD, Martinez-Diaz H, Rozengurt N, Cardiff RD, Liu X, Wu H. PTEN tumor suppressor regulates p53 protein levels and activity through phosphatase-dependent and

-independent mechanisms. Cancer Cell. 2003; 3:117–30.

41. Deng C, Zhang P, Harper JW, Elledge SJ, Leder P. Mice lacking p21CIP1/WAF1 undergo normal development, but are defective in G1 checkpoint control. Cell. 1995; 82:675–84.

42. Botcheva K, McCorkle SR, McCombie WR, Dunn JJ, Anderson CW. Distinct p53 genomic binding patterns in normal and cancer-derived human cells. Cell Cycle. 2011; 10:4237–49.

43. Berns EM, Klijn JG, van Putten WL, van Staveren IL, Portengen H, Foekens JA. c-myc amplification is a better prognostic factor than HER2/neu amplification in primary breast cancer. Cancer Res. 1992; 52:1107–13.