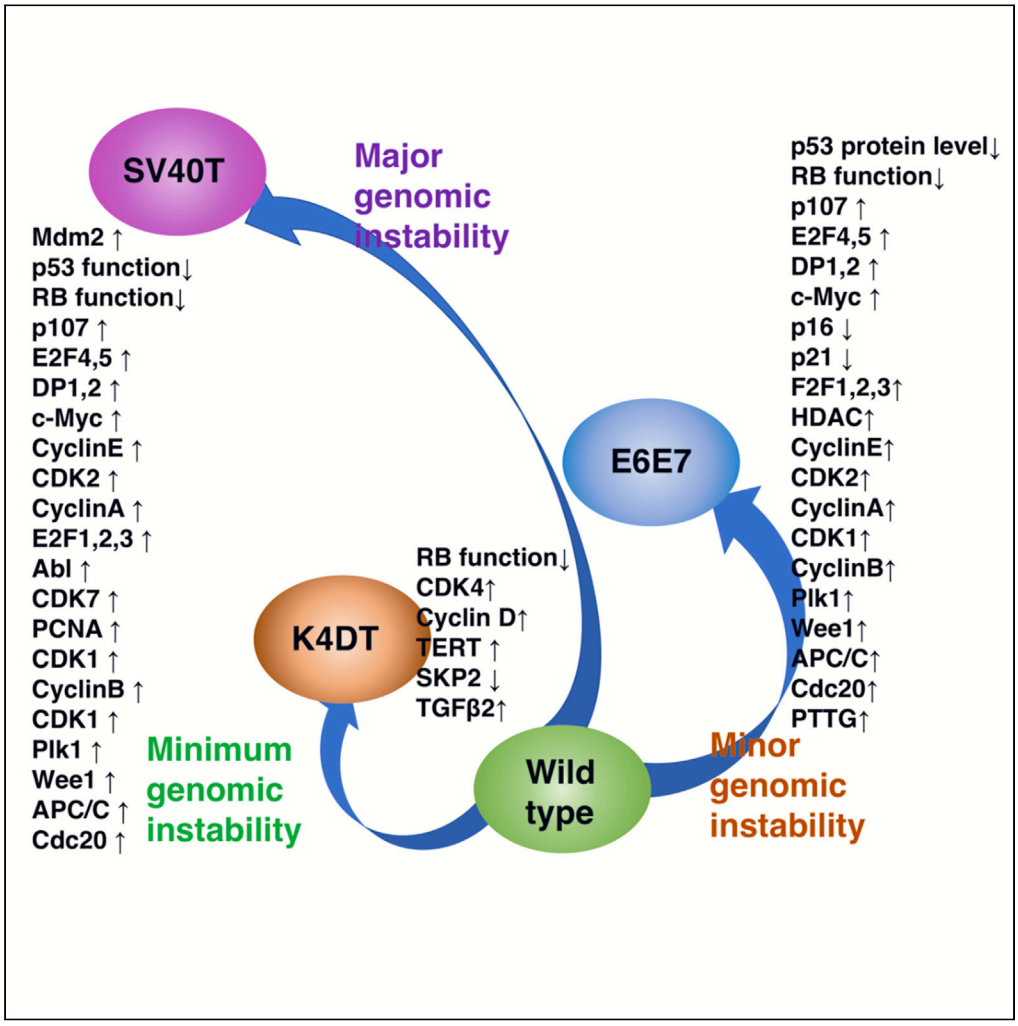
iScience

CellPress
OPEN ACCESS

Article

Combinatorial expression of cell cycle regulators is more suitable for immortalization than oncogenic methods in dermal papilla cells



Tomokazu Fukuda, Kai Furuya, Kouhei Takahashi, ..., Sayo Kashiwagi, Tohru Kiyono, Tsuyoshi Ishii

tomofukuda009@gmail.com (T.F.)
tkiyono@ncc.go.jp (T.K.)
ishii@rohto.co.jp (T.I.)

HIGHLIGHTS

Expression profiles of immortalized dermal papilla cells by the SV40, E6/E7, K4DT method were analyzed

The distance from wild type to K4DT was close, E6E7 was middle, SV40 was most distant

K4DT method showed the most conservative profile when it compared with SV40 and E6E7 method

Fukuda et al., iScience *24*, 101929
January 22, 2021
https://doi.org/10.1016/j.isci.2020.101929

iScience

CellPress
OPEN ACCESS

Article

Combinatorial expression of cell cycle regulators is more suitable for immortalization than oncogenic methods in dermal papilla cells

Tomokazu Fukuda,[1,4,*] Kai Furuya,[1] Kouhei Takahashi,[1] Ai Orimoto,[1] Eriko Sugano,[1] Hiroshi Tomita,[1] Sayo Kashiwagi,[2] Tohru Kiyono,[3,*] and Tsuyoshi Ishii[2,*]

Summary

The immortalized cell is an essential research tool that uses robust growth properties for the functional investigation of gene products. Immortalized mammalian cells have mainly been established using three methods: expression of simian vacuolating virus 40 T antigen (the SV40 method); human papilloma virus-derived oncoprotein E6/E7 (the E6/E7 method); or combinatorial expression of R24C mutant cyclin-dependent kinase 4, cyclin D1, and telomerase reverse transcriptase (the K4DT method). However, it is unclear as to which method is optimal for an *in vitro* model. Here, we compared the biological characteristics and genome-wide expression profiles of immortalized human dermal papilla cells generated by the SV40, E6/E7, or K4DT method. To our knowledge, this is the first study to comprehensively compare expression profiles to determine the optimal immortalization method for maintaining the original nature of the wild-type cells. These data would be valuable to scientists aiming to establish new immortalized cell lines.

Introduction

Cultured cells cannot proliferate infinitely due to shortening of the telomere repeat sequence. This limitation is known as "Heyflick's limit," named for the scientist who discovered this phenomenon (Hayflick, 1965). The process of overcoming this limitation is called ''immortalization,'' and the immortalized condition allows us to continue the cell culture indefinitely. Based on the various types of senescence such as the replicative, premature, stress-induced, oncogene-induced, tumor-suppressor induced senescence, the cell proliferation is limited for the primary cells (Wang et al., 2019). The first immortalized cell line was the HeLa cell line, which was established from human cervical cancer cells (Jones et al., 1971). These immortalized cells became an essential tool for investigating gene function and play a critical role in the screening of candidate chemicals for pharmaceutical drugs. For the immortalization of cells, two methods are the most popular: expression of the oncogenic protein SV40 large T antigen (Tevethia, 1984) and expression of human papilloma virus-derived E6/E7 (Hawley-Nelson et al., 1989). Cellular immortalization is tightly connected to cancer development and cell cycle control since loss of function of negative regulators of cell cycle turnover is required for efficient and infinite cell growth.

SV40 large T antigen is an oncoprotein and was isolated from polyomavirus simian virus 40 (Tevethia and Ozer, 2015). The SV40 oncoprotein induces cellular immortalization by bypassing the negative regulation of retinoblastoma protein (pRB) and the suppression of the p53 tumor suppressor protein (Ahuja et al., 2005). Immortalization using SV40 is also classified as a transformation of mammalian cells from their dominant cell growth capability. Furthermore, SV40 protein is known to bind to various types of transcriptional regulators, such as p300 and CREB-binding protein (Cho et al., 2001).

The E6/E7 oncoprotein was isolated from human papillomavirus (HPV). HPV-derived E6/E7 is also known to induce cervical cancer in humans (Zur Hausen, 2002). The target protein of E6 is the p53 tumor suppressor protein (Münger and Howley, 2002). Expression of the E6 oncoprotein causes the degradation of p53 protein through ubiquitination (Scheffner et al., 1990). Expression of the E7 oncoprotein induces the inactivation of pRB and accelerates the transcriptional activity of the transcription factor E2F1, resulting in the elevation of cell proliferation (Münger et al., 1989). In the SV40 and E6/E7 methods, immortalization is

[1]Graduate School of Science and Engineering, Iwate University, 4-3-5 Ueda, Morioka, Iwate, 020-8551 Japan

[2]Rohto Pharmaceutical Co., Ltd., Basic Research Development Division, 6-5-4 Kunimidai, Kizugawa, Kyoto, 619-0216, Japan

[3]Exploratory Oncology Research & Clinical Trial Center, National Cancer Center, 6-5-1 Kashiwanoha, Kashiwa-city, Chiba, 277-8577, Japan

[4]Lead contact

*Correspondence: tomofukuda009@gmail.com (T.F.), tkiyono@ncc.go.jp (T.K.), ishii@rohto.co.jp (T.I.)

https://doi.org/10.1016/j.isci.2020.101929

achieved through inactivation of two major tumor suppressor pathways, pRB and p53. Because they cause inactivation of the tumor suppressor p53, which is characterized as the "guardian of the genome" (Lane, 1992), the expression of SV40 or E6/E7 oncoproteins sometimes causes genome instability, such as polyploid formation and chromosome gain or loss. Loss of function of the p53 gene causes genomic instability as a result of impairment of the accurate replication ability of the genome, and loss of function of p53 is frequently observed in around half of human malignant cancers (Royds and Iacopetta, 2006), indicating the critical role of p53 for tumor suppression.

In 2011, Shiomi et al. showed that combinatorial expression of R24C mutant CDK4, cyclin D1, and telomerase reverse transcriptase (TERT) results in the efficient immortalization of human-derived myogenic cells (Shiomi et al., 2011). Furthermore, Shiomi et al. showed that these immortalized myogenic cells retained the ability to differentiate into mature cells. Senescent cells halt cell proliferation due to the accumulation of the senescence protein p16 (Brenner et al., 1998). The accumulated cytoplasmic p16 protein binds to endogenous CDK4, which results in decreased kinase activity of the CDK4-cyclin D1 complex. The reduced kinase activity of the CDK4-cyclin D1 complex induces hypo-phosphorylation of the pRB protein, which results in a decrease in the protein level of the transcription factor E2F1, which is free from pRB protein binding. However, expression of R24C mutant CDK4 and cyclin D1 induces constitutively elevated CDK4-cyclinD1 complex activity since p16 protein cannot bind to mutant CDK4 due to this point mutation (Shiomi et al., 2011). The expression of R25C mutant CDK4 and cyclin D1 constitutively enhances cell proliferation in various types of human cells, such as myogenic cells, fibroblasts, corneal epithelial cells (Fukuda et al., 2019a), and dermal papilla cells (DPCs) (Fukuda et al., 2020). An important characteristic of immortalization using R25C mutant CDK4 and cyclin D1 is the intact condition of the p53 tumor suppressor gene, which enables cell expansion without genomic instability. Immortalization using R25C mutant CDK4, cyclin D1, and telomerase reverse transcriptase was named the K4DT method based on the genes used. Furthermore, it was been shown that the K4DT method works in a wide variety of animal cells, such as bovine, swine (Donai et al., 2014), prairie vole, monkey (Kuroda et al., 2015), midget buffalo, rat, and megabat-derived cells (Tani et al., 2019). This method can even be applied to sea turtle-derived cells, based on the highly conserved amino acid sequence and function of cell cycle regulators in animal evolution.

In this study, we comprehensively compare the biological characteristics of immortalized cells generated using three different methods (the SV40, E6/E7, and K4DT methods) on the same cell sources. The fundamental question is which immortalization method is best able to retain the original nature of the wild-type cells, as measured by the whole transcriptome. This study is the first report to describe the comprehensive expression profiles of different immortalization methods, which will contribute to understanding the characteristics of immortalized cells and could accelerate the application of established cells as an *in vitro* model for human diseases.

Results

Biological characteristics of immortalized human dermal papilla cells generated using the K4DT, SV40, and E6/E7 methods

We obtained primary human DPCs from PromoCell (Heidelberg, Germany) through a local distributor (Takara Bio, Shiga, Japan). After the expansion of the primary cells, we introduced either enhanced green fluorescent protein-expressing retrovirus, simian vacuolating virus 40 (SV40)-expressing retrovirus (Fukuda et al., 2012), or human papilloma virus-derived E6/E7-expressing retrovirus. As shown in Figure S1, although the efficiency of gene introduction was only around a few percent, by selecting the successfully introduced cells using antibiotics (G418 for SV40, hygromycin for E6/E7), we successfully established human DPCs that expressed SV40 or E6/E7. We previously reported the establishment of immortalized DPCs with combinatorial expression of R24C mutant cyclin-dependent kinase 4 (CDK4), cyclin D1, and telomerase reverse transcriptase (TERT) (Fukuda et al., 2020). Hereafter, we will refer to immortalized cells with the combinatorial expression of R24C mutant CDK4, cyclin D1, and TERT as K4DT. In Figure 1A, we show the morphology of wild-type DPCs and DPCs established using the SV40, E6/E7, and K4DT methods. Although E6/E7 and K4DT showed uniform cell size, there was a variation in cell size in SV40 cells. We first checked for the successful genomic insertion of the introduced gene with polymerase chain reaction (PCR) analysis. As shown in Figures S2 and S34, wild-type cells did not show any genomic insertion, while SV40, E6/E7, and K4DT cells showed the successful insertion of the expected combinations of expression cassettes. Furthermore, we detected the protein expression of CDK4, cyclin D1, SV40, E7, and tubulin with western blotting. As shown in Figures 1B and S35, while wild-type cells did not show intense signals of

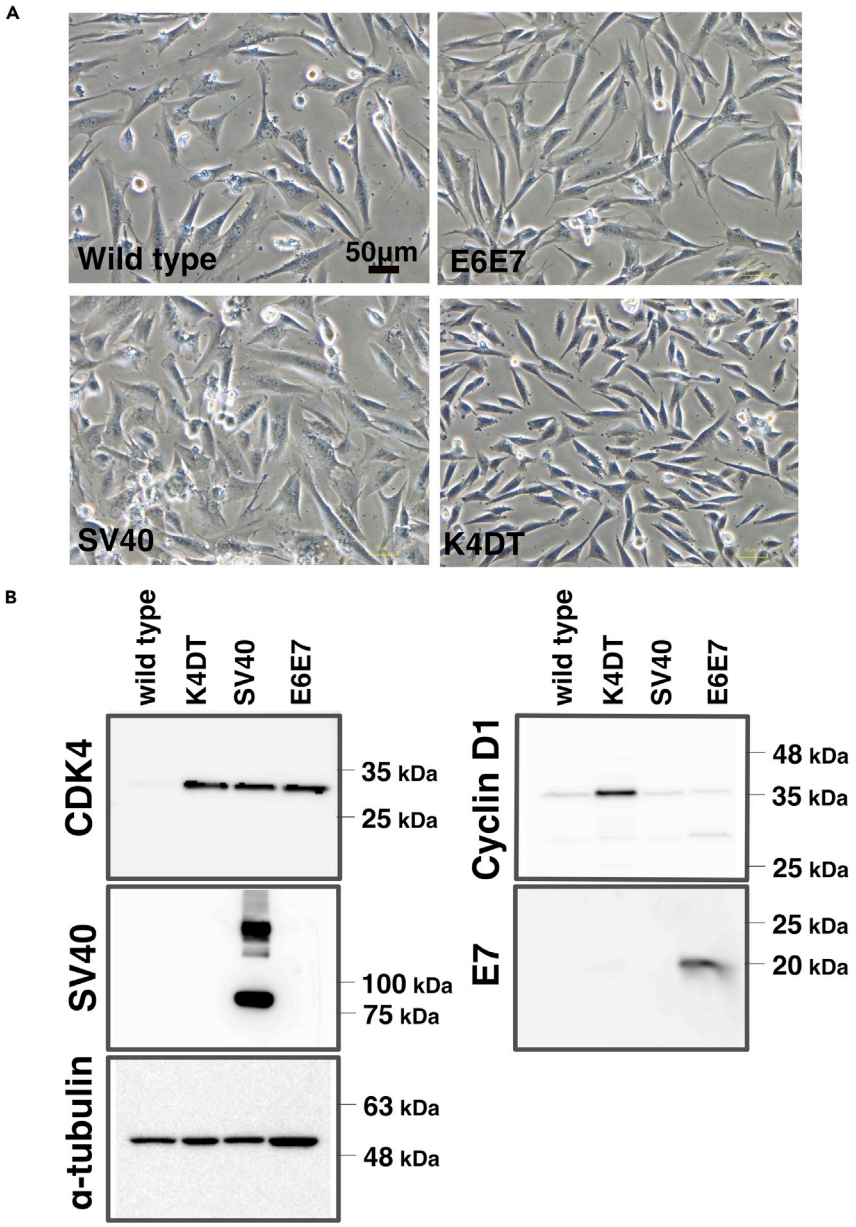


Figure 1. Cell morphology and the detection of introduced genes in wild-type and immortalized dermal papilla cells

(A) Cell morphology of wild-type, E6/E7, SV40, and K4DT immortalized DPCs at passage 3.

(B) Detection of introduced gene products in western blots of wild-type, K4DT, SV40, and E6/E7 cells.

detected proteins, SV40, E6/E7, and K4DT cells showed protein expression of the corresponding introduced proteins. We next analyzed the cell cycles of the wild type, SV40, E6/E7, and K4DT cells. Although wild-type, K4DT, and E6/E7 cells did not show any abnormalities, SV40 cells exhibited a broader G2/M peak compared to wild-type cells, suggesting polyploid abnormalities in SV40 cells (Figure 2).

Whole-gene transcriptome analysis of wild-type and immortalized DPCs

To comprehensively compare the expression pattern of the whole genome, we carried out RNA-Seq analysis using the Illumina Hiseq platform (100 bp paired-end). We show the workflow of the sequencing in Figure 3A. After the removal of the adapter sequences, we carried out mapping of the sequences. The amount

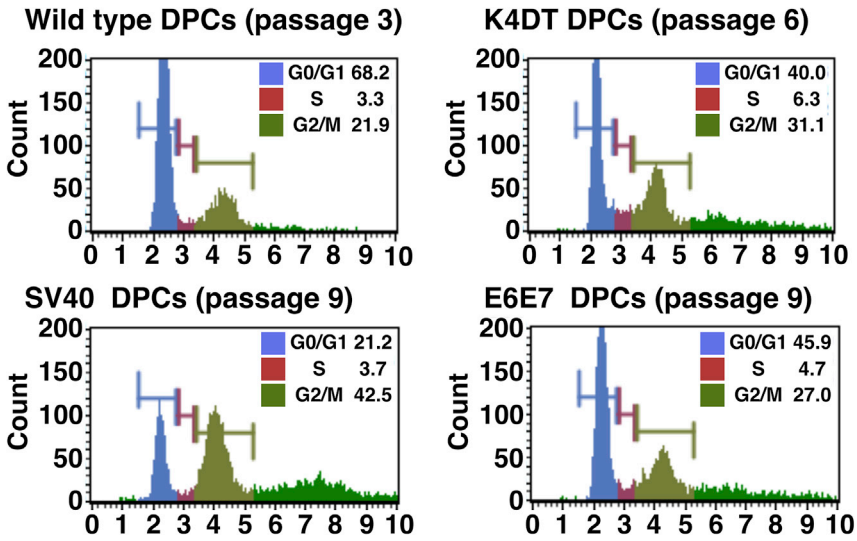


Figure 2. Cell cycle analysis of wild-type, K4DT, SV40, and E6/E7 immortalized DPCs

A representative histogram of cells is shown.

of obtained sequence reads was at least 22 M reads, which is enough to describe the whole transcriptome (Figure 3B) (Fukuda et al., 2019b). To evaluate the reproducibility of the data, we carried out RNA-Seq with triplicate biological replications. The quality of the sequencing reads was determined using the TASTQC program (Figures S3–S6). Almost all the read data existed within the green area, indicating that the quality of the reads was reliable. We mapped the obtained sequence to an up-to-date human reference sequence (GRCh38) with the STAR program. The mapping ratio of reads was more than 95%, suggesting that our mapping strategy was appropriate for detecting the expression profile (Figure 3B). The mapping ratio to human genome sequence (GRh38) was 96.2% (wild type 1), 96.7% (wild type 2), 96.3% (wild type 3), 96.5% (K4DT-1), 96.1% (K4DT-2), 95.8% (K4DT-3), 95.1% (K4DT AR-1), 96.3% (K4DT AR-2), 96.4% (K4DT AR-3), 96.3% (SV40-1), 96.2% (SV40-2), 95.9% (SV40-3), 96.8% (E6E7-1), 96.5% (E6E7-2), 96.6% (E6E7-3). We, furthermore, mapped sequencing reads to total sequencing library of human cDNA (Homo_sapiens.GRCh38.cdna.all.fa) with bowtie2 program. The overall alignment rate of all reads to cDNA was below. 89.79% (wild type 1, 26838188 reads), 90.28% (wild type 2, 25388984 reads), 90.11% (wild type 3, 27158793 reads), 90.74% (K4DT-1, 23884949 reads), 90.74% (K4DT-2, 25895869 reads), 90.14% (K4DT-3, 26615408 reads), 88.95% (K4DT AR-1, 29396085 reads), 90.30% (K4DT AR-2, 26536855 reads) 88.04% (K4DT AR-3, 24322764 reads), 90.40% (SV40-1, 23171547 reads), 90.49% (SV40-2, 31207937 reads), 90.71% (SV40-3, 39095454 reads), 92.24% (E6E7-1, 23119054 reads), 92.57% (E6E7-2, 26509457 reads), 91.38% (E6E7-3, 28433365 reads).

The complete list of expression counts for wild-type, K4DT, SV40, and E6/E7 DPCs is listed on Figshare (https://figshare.com/s/124ed9cd873719d6cb4d). We inputted the expression profile of the whole genome into TCC-GUI for downstream analysis. Due to the limitation of TCC-GUI, we used top 25,000 genes to draw the dendrogram analysis and principal component analysis (PCA) plot to explain the variance of the expression level. First, we carried out dendrogram analysis. As shown in Figures 3C and S7, the results of correlation plot and data matrix among all replicates showed three biological replicates showed reproducible results on the RNA-Seq detection (Figure S7). Furthermore, we processed the data with three-dimensional PCA. From wild-type cells, the distance to K4DT cells was the smallest, the distance to E6/E7 cells was moderate, and the distance to SV40 cells was the largest (Figure 3D). We also showed the detailed locations of three-dimensional PCA analysis with a movie file (Figure 360). Three-dimensional PCA analysis based on 25,000 genes showed cumulative proportion of the variance to be 65.5% (PC1, 33.26%, PC2, 18.964%, PC3, 13.315%), indicating that 3D PCA would be classified as the results of whole-gene analysis. From these data, we concluded that the gene expression profile of K4DT immortalized cells was closest to that of wild-type cells when compared to E6/E7 and SV40 immortalized cells.

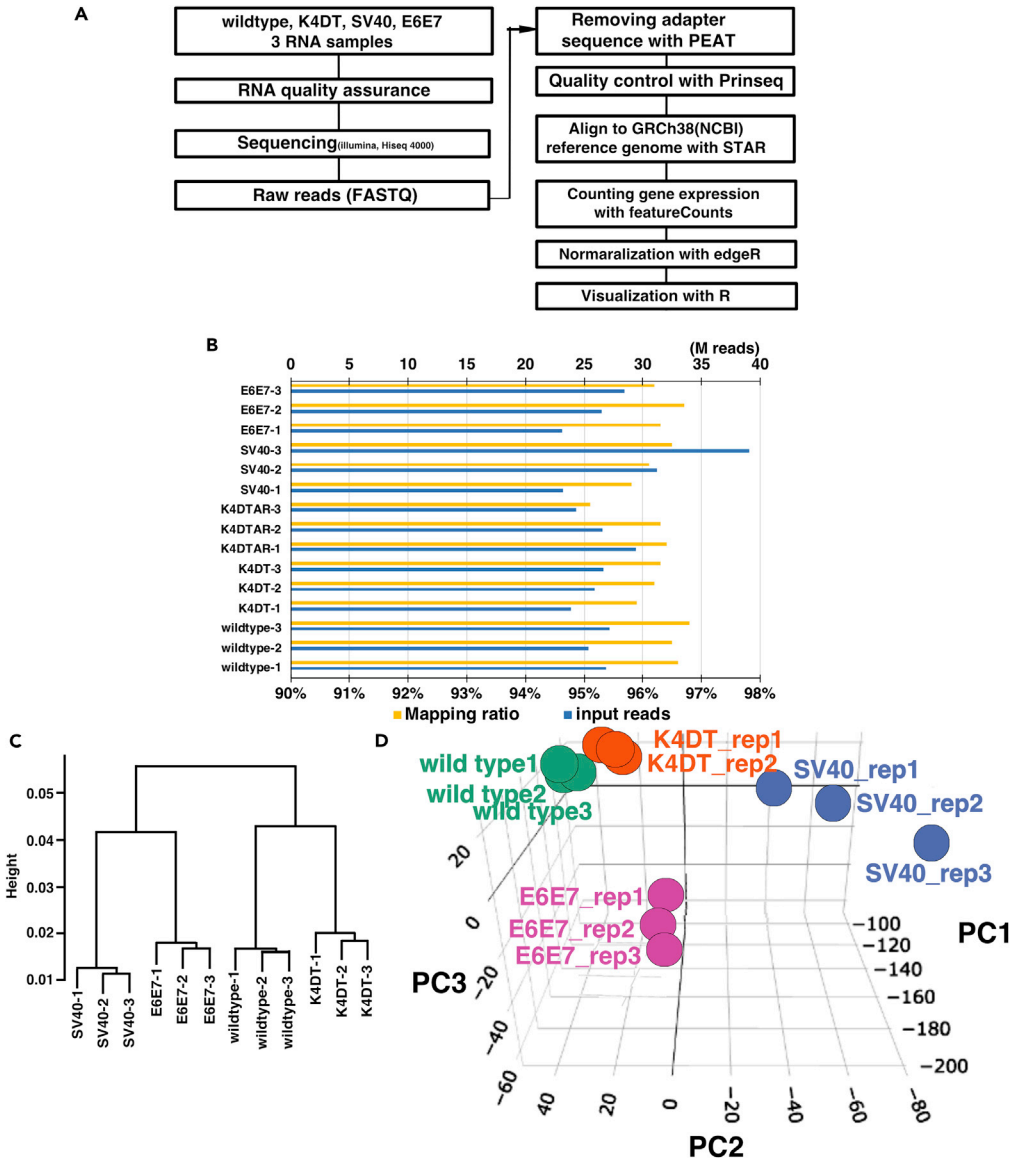


Figure 3. RNA-Seq analysis of wild-type, K4DT, SV40, and E6/E7 immortalized DPCs

Figure360▷ For a Figure360 author presentation of this figure, see https://doi.org/10.1016/j.isci.2020.101929.

(A) Workflow of the RNA-Seq analysis.

(B) The number of sequence reads of each sample obtained after PRINSEQ (blue bars) and the mapping ratio (yellow bars) are shown as a bar graph.

(C) Dendrogram analysis of wild-type, K4DT, SV40, and E6/E7 immortalized DPCs.

(D) Three-dimensional principal component analysis (PCA) based on the expression profiling of wild-type, K4DT, E6/E7, and SV40 immortalized DPCs.

Visualization of differentially expressed genes among wild-type and immortalized DPCs

Using TCC-GUI (Su et al., 2019), we analyzed the expression profiles of the 33,122 human genes (https://doi.org/10.6084/m9.figshare.12197691.v2).

During the analysis of differentially expressed (DE) genes, the edgeR program extracted 15,103 genes as the first candidates (p value of edgeR is less than 0.01). Before applying to the DAVID functional analysis, we carried out the selection of DE genes with different expression threshold (no filtering, more than 300, 500, 700, 900, 1000, 1500, 1700, 2000, 2500, 3000, 10,000 counts) under the p value less than 0.01. The number of

Table 1. Number of differentially expressed (DE) genes on different thresholds on wild-type, K4DT, SV40, E6E7 immortalized human dermal papilla cells

Filtering	Number of DE genes	File name
No filtering, p < 0.01 on EdgeR	15,103	f_list1_1.csv
300 counts on any sample, p < 0.01 on EdgeR	9836	f_list1_2.csv
500 counts on any sample, P < 0.01 on EdgeR	8651	f_list1_3.csv
700 counts on any sample, p < 0.01 on EdgeR	8036	f_list1_4.csv
900 counts on any sample, p < 0.01 on EdgeR	6750	f_list1_5.csv
1000 counts on any sample, p < 0.01 on EdgeR	6350	f_list1_6.csv
1500 counts on any sample, p < 0.01 on EdgeR	4829	f_list1_7.csv
1700 counts on any sample, p < 0.01 on EdgeR	4355	f_list1_8.csv
2000 counts on any sample, p < 0.01 on EdgeR	3786	f_list1_9.csv
2500 counts on any sample, p < 0.01 on EdgeR	3060	f_list1_10.csv
3000 counts on any sample, p < 0.01 on EdgeR	2524	f_list1_11.csv
10,000 counts on any sample, p < 0.01 on EdgeR	547	f_list1_12.csv

genes selected by the above mentioned threshold is summarized in Table 1. The number of genes was changed from 15,103 genes with no filtering to 576 with at least 10,000 counts in any of the sample. The full list of no filtering genes (https://doi.org/10.6084/m9.figshare.13174148.v1), list of at least 300 counts (https://doi.org/10.6084/m9.figshare.13174142.v1), list of at least 500 counts (https://doi.org/10.6084/m9.figshare.13174151.v1), list of at least 700 counts (https://doi.org/10.6084/m9.figshare.13174154.v1), list of at least 900 counts (https://doi.org/10.6084/m9.figshare.13174157.v1), list of at least 1000 counts (https://doi.org/10.6084/m9.figshare.13174166.v2), list of least 1500 counts (https://doi.org/10.6084/m9.figshare.13174172.v1), list of least 1700 counts (https://doi.org/10.6084/m9.figshare.13174175.v1), list of least 2000 counts (https://doi.org/10.6084/m9.figshare.13174178.v1), list of least 2500 counts (https://doi.org/10.6084/m9.figshare.13174202.v1), list of least 3000 counts (https://doi.org/10.6084/m9.figshare.13174205.v1), list of least 10,000 counts (https://doi.org/10.6084/m9.figshare.13174229.v1) are given.

Based on the DE gene list with different expression thresholds (no filtering, more than 300, 500, 700, 900, 1000, 1500, 1700, 2000, 2500, 3000, 10,000 counts), we carried out the pathway analysis on each threshold. The results of top 5 significant pathways in each threshold are summarized in Table 2. The number of times listed within top 5 was highest in cell cycle (10 times), ribosome (9 times), pathway in cancer (5 times). From the viewpoint of association to chromosome instability, we selected cell cycle and pathway in cancer.

We obtained a gene list that corresponded to the cell cycle in DAVID (112 genes). The expression levels of these 112 genes, after normalization, are presented as a heatmap in Figure 5A. As shown in Figure 4A, the normalized heatmap reveals that the difference between wild-type and K4DT cells is small, while SV40 and E6/E7 cells have major differences compared to wild-type cells. The raw expression counts of the 112 cell cycle-related genes in the KEGG of wild-type, K4DT, SV40, and E6/E7 cells are shown using box plots in Figures S8–S14. The threshold for DE genes was a change in expression count of 2X or greater, or 0.5X or less, in immortalized cells, and these DE genes were mapped into the KEGG cell cycle gene pathway map (Figure 5). As shown in Figure 6, although the affected genes in K4DT cells were restricted to cyclin D-CDK4 complex-related networks, SV40 and E6/E7 cells exhibited significant differences in various types of growth signals, such as c-Myc, cyclin E-CDK2, cyclin B-CDK1, and the APC/C-Cdc complex.

We also analyzed the pathway in cancer. The KEGG map of pathway in cancer contains 282 genes. The expression levels of these 282 genes, after normalization, are presented as a heatmap in Figure 4B. As shown in Figure 4B, the difference between wild-type and K4DT cells is relatively close when it is compared with that from SV40 or E6E7 cells. The raw expression counts of the 282 cell cycle-related genes in the KEGG of wild-type, K4DT, SV40, and E6/E7 cells are shown using box plots in Figures S15–S32. DE genes were mapped into the KEGG pathway in cancer map (Figure 6). Mapping on the pathway in cancer revealed

Table 2. List of top five DAVID pathway analyses on differentially expressed genes detected by different thresholds among wild-type, K4DT, SV40, E6E7 immortalized dermal papilla cells.

Filtering	Number of DE genes	DAVID ranking				
		1st	2nd	3rd	4th	5th
No filtering, p < 0.01 on EdgeR	15,103	Pathway in cancer	Small cell lung cancer	Neurotrophin signaling pathway	Chronic myeloid leukemia	Endocytosis
300 counts on any sample, p < 0.01 on EdgeR	9836	Cell cycle	Colorecal cancer	Pancreatic cancer	Chronic myeloid leukemia	Pyrimidine metabolism
500 counts on any sample, p < 0.01 on EdgeR	8651	Cell cycle	small cell lung cancer	Spliceosome	Ribosome	Pancretic cancer
700 counts on any sample, p < 0.01 on EdgeR	8036	Ribosome	Cell cycle	Ubiquitin mediated proteolysis	Chronic myeloid leukemia	Spliceosome
900 counts on any sample, p < 0.01 on EdgeR	6750	Cell cycle	Ribosome	Spliceosome	Ubiquitin mediated proteolysis	Chronic myeloid leukemia
1000 counts on any sample, p < 0.01 on EdgeR	6350	Ribosome	Cell cycle	Spliceosome	Proteasome	Focal adhesion
1500 counts on any sample, p < 0.01 on EdgeR	4829	Ribosome	Cell cycle	Focal adhesion	Proteasome	Pathway in cancer
1700 counts on any sample, p < 0.01 on EdgeR	4355	Ribosome	Focal adhesion	Cell cycle	Proteasome	Pathway in cancer
2000 counts on any sample, p < 0.01 on EdgeR	3786	Ribosome	Focal adhesion	Proteasome	Cell cycle	Pathway in cancer
2500 counts on any sample, p < 0.01 on EdgeR	3060	Ribosome	Focal adhesion	Proteasome	Regulation of actine cytoskeleton	Pathway in cancer
3000 counts on any sample, p < 0.01 on EdgeR	2524	Ribosome	Focal adhesion	ECM-receptor Interaction	Regulation of actin cytoskeleton	Cell cycle
10,000 counts on any sample, p < 0.01 on EdgeR	547	Ribosome	Focal adhesion	ECM-receptor interaction	Pathogenic *Escherichia coli* infection	Arrhythmogenic right ventricular cardiomyopathy (ARVC)

ECM, Extracellular Matrix

that the c-Jun, c-Fos, c-Myc, which locate on the downstream of MAP kinase signaling, has unregulated in SV40 and E6E7 cells. Furthermore, expression levels of FGF and FGF-receptor were upregulated in SV40 and E6E7 cells.

Chromosome condition

Lastly, we carried out chromosome analysis of K4DT, SV40, and E6/E7 cells. We first detected the number of chromosomes in K4DT, SV40, and E6/E7 cells. As shown in Table S1, the K4DT cells had a normal human chromosome count of 46 in 50 mitotic cells. However, SV40 cells showed broad distribution of chromosome numbers, from 71 to 80, indicating that SV40 cells have major chromosomal abnormalities. Furthermore, the E6/E7 cells showed chromosome abnormalities in 10% of mitotic cells. We further analyzed the chromosome pattern of K4DT, SV40, and E6/E7 cells using the G banding method (Table S2). In K4DT cells, one mitotic cell showed translocation of chromosome 14 to 15, suggesting that the normal ratio of chromosome condition is 95%. However, SV40 cells had a normal ratio of 0%, and 20% of E6/E7 cells showed a gain or loss of chromosomes. From these results, we concluded that K4DT immortalization caused the fewest genomic abnormalities in DPCs. In contrast, the expression of SV40 caused major genomic instability, and the expression of E6/E7 induced minor or moderate genomic instability in DPCs (Figure 7B).

Discussion

In this study, we compared the whole genome expression profile among immortalized cells obtained by the expression of SV40, E6/E7 oncoprotein, or combinatorial expression of mutant CDK4, cyclin D1, and TERT (K4DT method). The establishment of immortalized cell lines from the same primary cell enabled us to compare the biological characteristics of the immortalization methods comprehensively. Based on the

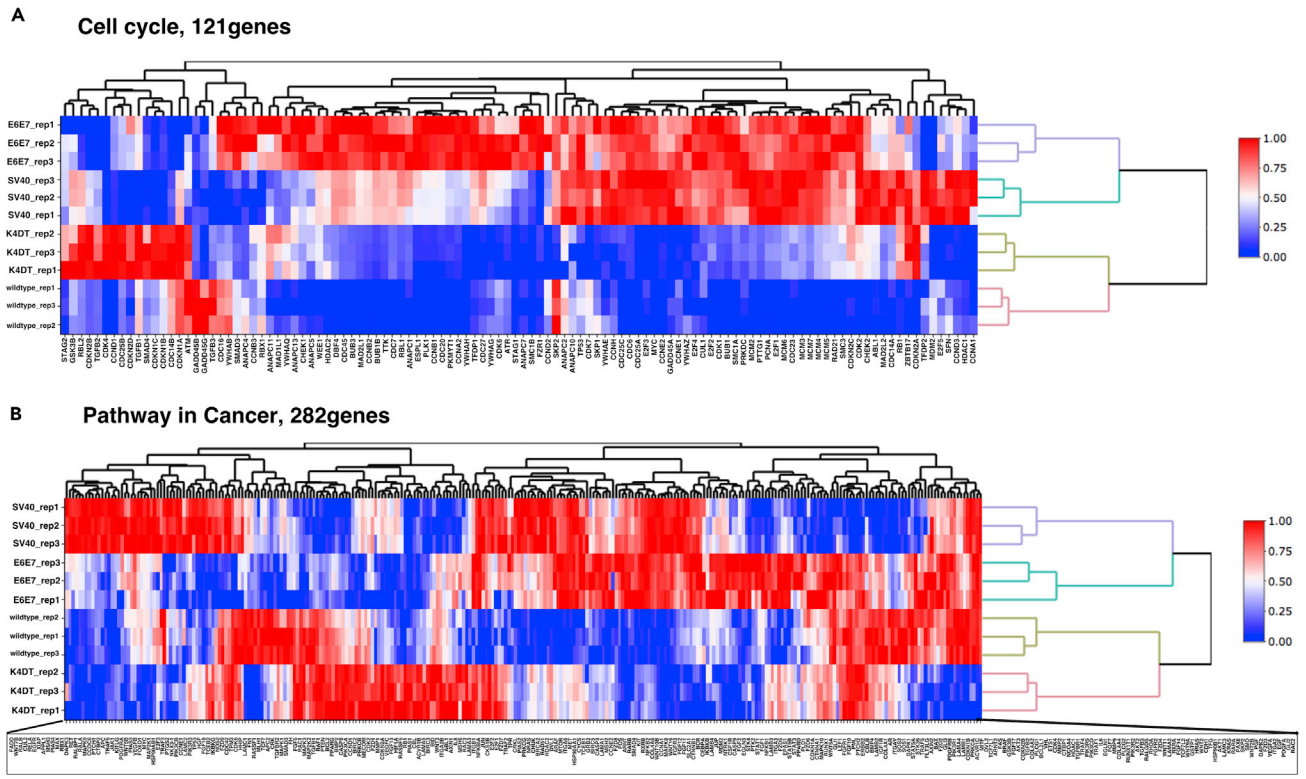


Figure 4. Heapmap analysis

Heatmap of genes listed in cell cycle pathway (A) and pathway in cancer (B) in wild-type, K4DT, SV40, and E6/E7 immortalized DPCs. Red indicates high expression, and blue indicates low expression.

expression profiling of the whole genes, we concluded that K4DT is the most suitable method for maintaining the original nature of the primary cells. Based on previously established immortalized cells, such as HeLa or HEK293T cells, functional assays with these immortalized cells are essential to mimic *in vivo*. However, most scientists would wish to use immortalized cells with relatively intact characteristics that mimic the primary condition. Our manuscript first demonstrated that the combinatorial expression of mutant cyclin-dependent kinase 4 (CDK4), cyclin D1, and telomere reverse transcriptase is the most suitable method for immortalizing human dermal papilla cells while retaining their original nature, compared to the expression of SV40 antigen or human papilloma virus-derived oncoprotein E6/E7. Our data will be attractive to scientists trying to select the best method for *in vitro* cellular modeling of human and animal diseases. The antiproliferative function of p16/tumor suppressor pRB has been listed as the first gate to cause replicative senescence in the cell replication (Wang et al., 2019).

In the pathway analysis with DAVID, cell cycle and pathway in cancer were selected for further analysis. However, the nucleotide excision repair pathway also plays a critical role on the maintenance of the genome. Based on this view point, we also mapped the DE genes on pathway (Figure S33). In good agreement with enhanced cell cycle turnover, upregulation of DNA polymerase complex including PCNA and DNA polymerase epsilon was detected. However, TFIIH complex-, XPE complex-, XPC complex-associated molecules did not show consistent change that suggest loss of function of repair complex. Potential abnormality in function of nucleotide repair will need to be analyzed about accuracy of genomic sequence of SV40 and E6E7 cells.

We also established K4DT immortalized DPCs expressing the androgen receptor (AR) gene using retrovirus gene transfer. Although the exact mechanism is unclear, the endogenous AR expression level became lower after several passages of DPCs (Kwack et al., 2008). In agreement with previous reports of AR suppression, we also detected downregulation of AR in DPCs immortalized using the K4DT method (Fukuda et al., 2020). We are now trying to analyze the whole expression data of K4DT immortalized cells expressing

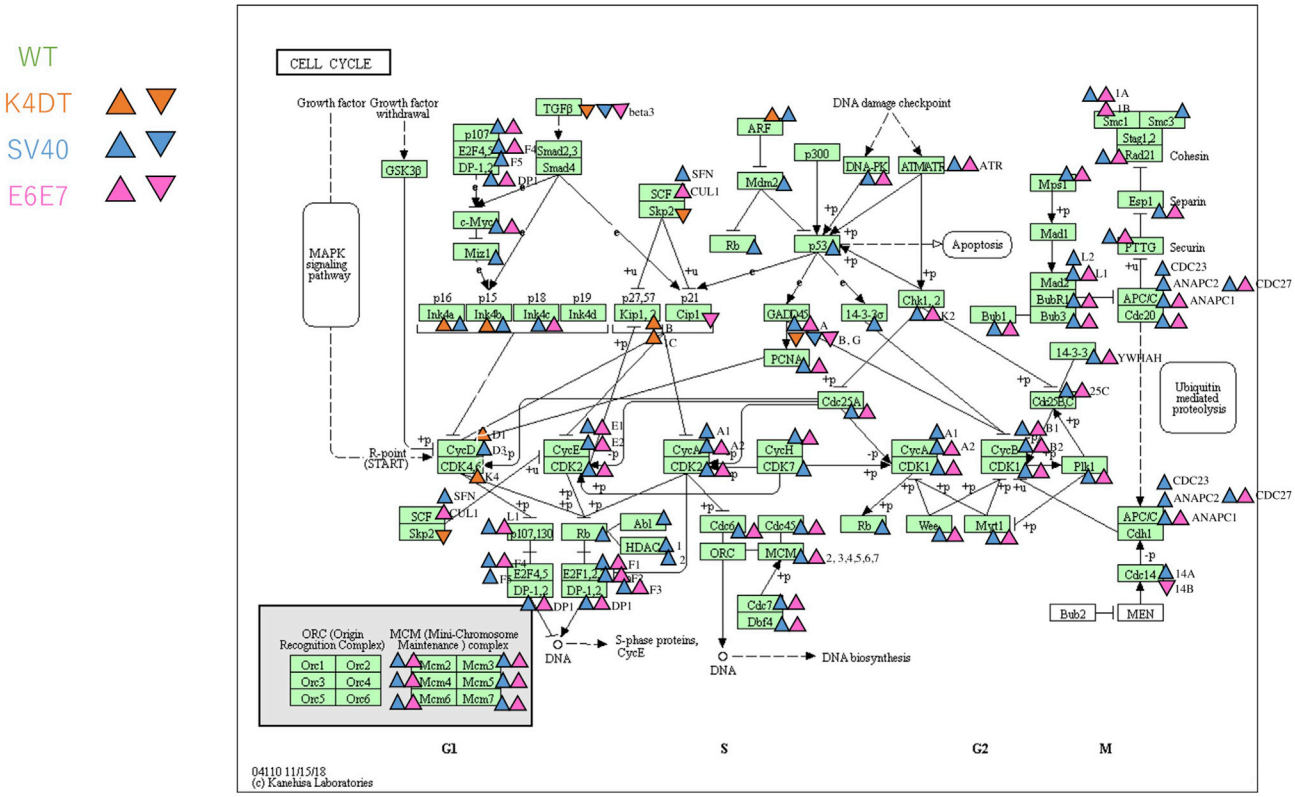


Figure 5. Summary of upregulated and downregulated genes among KEGG cell cycle-related genes

Arrowheads indicate upregulated or downregulated genes in K4DT, E6/E7, or SV40 cells. The genes with more than 200 counts and X2 increase or 1/2 decreased from the wild type were mapped.

AR. The comparison between AR-expressing K4DTs and parent cells would contribute to our understanding of the androgen signaling pathway. A better understanding of the AR signal network would be beneficial for the development of anti-testosterone drugs to treat prostate cancer or androgenic alopecia, which is known to be affected by the activation status of AR.

Limitations of the study

In this study, we carried out the sampling from wild type around passage 3 to 4, while, the sampling from E6E7 cell was finished around passage 7 to 8. In K4DT cells, the sample collection was finished from passage 5 to 6 (see the section of method). Although we could not finish up the exact same number of passages for all cell lines, the accuracy of the passage number is extremely important for accurate evaluation. Furthermore, the accurate data for gene expression and biological triplicates might not be enough for the concise conclusion. Based on our data, more accurate and precise evaluation with larger biological duplication would be necessary to address the node genes that control the biological difference among immortalization method.

Resource availability

Lead contact

Further information and requests for resources and vectors should be directed to and will be fulfilled by the Lead Contact, Tomokazu Fukuda (tomofukuda009@gmail.com).

Data and code availability

The data sets with the next-generation sequencer that reported in this article are uploaded into the DNA Data Bank of Japan (DDBJ, and assigned BioProject accession SSUB015042. The output file for expression counts and lists of differentially expressed genes was listed in Figshare. The complete list of expression counts for wild-type, K4DT, SV40, and E6/E7 DPCs is listed on Figshare (https://figshare.com/s/

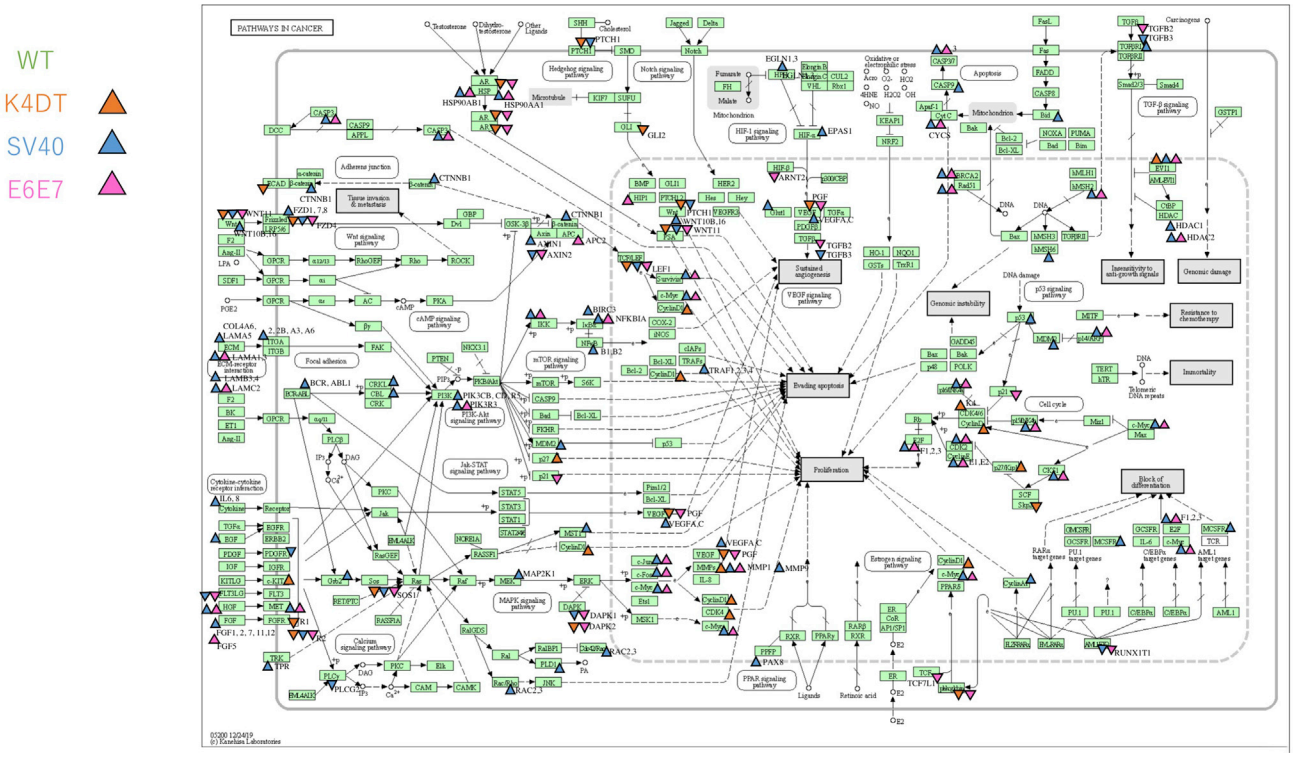


Figure 6. Summary of upregulated and downregulated genes among the KEGG pathway in cancer-related genes

Arrowheads indicate upregulated or downregulated genes in K4DT, E6/E7, or SV40 cells. The genes with more than 200 counts and X2 increase or 1/2 decreased from the wild type were mapped.

124ed9cd873719d6cb4d). The full list of no filtering genes (https://doi.org/10.6084/m9.figshare.13174142.v1), list of at least 300 counts (https://doi.org/10.6084/m9.figshare.13174142.v1), list of at least 500 counts (https://doi.org/10.6084/m9.figshare.13174151.v1), list of at least 700 counts (https://doi.org/10.6084/m9.figshare.13174154.v1), list of at least 900 counts (https://doi.org/10.6084/m9.figshare.13174157.v1), list of at least 1000 counts (https://doi.org/10.6084/m9.figshare.13174166.v2), list of least 1500 counts (https://doi.org/10.6084/m9.figshare.13174172.v1), list of least 1700 counts (https://doi.org/10.6084/m9.figshare.13174175.v1), list of least 2000 counts (https://doi.org/10.6084/m9.figshare.13174178.v1), list of least 2500 counts (https://doi.org/10.6084/m9.figshare.13174202.v1), list of least 3000 counts (https://doi.org/10.6084/m9.figshare.13174205.v1), list of least 10,000 counts (https://doi.org/10.6084/m9.figshare.13174229.v1) are given.

Software and its versions were used for the data analysis.

FastQC, version 0.11.8, was used for quality check of raw FASTQ sequencing file. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

PRINSEQ, version 0.20.4, was used for the removal of low quality reads. http://prinseq.sourceforge.net/

PEAT, version 1.2, was used for the removal of the adapter sequence. https://github.com/jhhung/PEAT.

STAR, version 2.6.1, was used for the mapping. https://github.com/alexdobin/STAR

featureCount, SUBREAD, release 1.6.5, was used for the expression counting. http://subread.sourceforge.net.

Bowtie2, version 2.3.4.3, was used for the mapping of reads to cDNA (Homo_sapiens.GRCh38.cdna.all.fa). http://bowtie-bio.sourceforge.net/bowtie2/index.shtml.

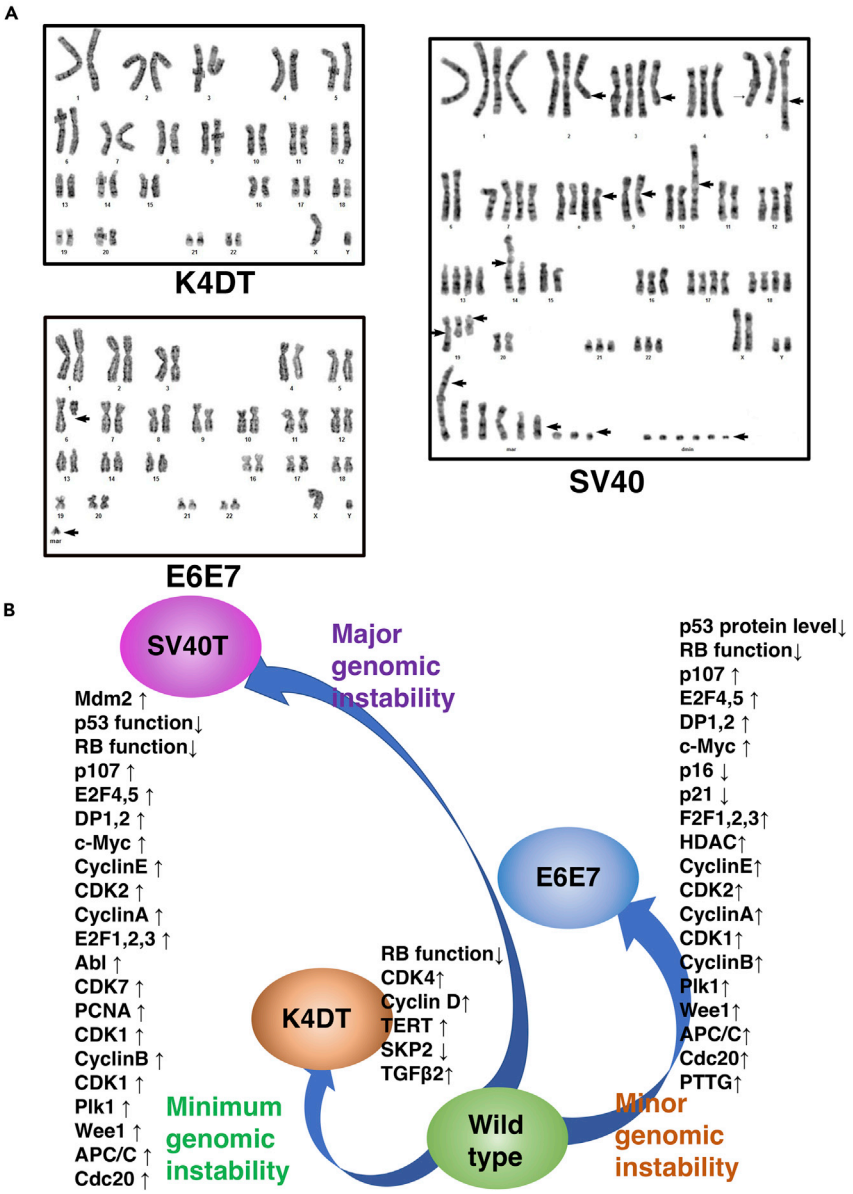


Figure 7. Chromosome analysis and summary of expression profiling of immortalized cells

(A) Chromosome analysis of K4DT, E6/E7, and SV40 cells using the G banding method. Small arrows indicate abnormalities detected by G banding.

(B) Summary of the expression profiling of K4DT, E6/E7, and SV40 immortalized DPCs. The distance between K4DT and wild-type cells is the shortest, that between E6/E7 and wild-type cells is moderate, and that between SV40 and wild-type cells is the greatest. Major genomic instability was observed in SV40 cells.

R package, version 4.0.3, was used for the downstream analysis. https://www.r-project.org.

TCC-GUI tool was used for the downstream analysis. https://github.com/swsoyee/TCC-GUI.

Materials availability

The established cell might be shared with the material transfer agreements of technology office and corresponding authors.

Methods

All methods can be found in the accompanying Transparent methods supplemental file.

Supplemental information

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2020.101929.

Acknowledgments

We thank Dr. Hiroyuki Miyoshi (RIKEN, BioResource Center) for providing the packaging plasmids for the recombinant virus. We thank Dr. Taku Ozaki (Graduate School of Science and Technology, Iwate University) for mentoring to our students. We also thank Dr. Koji Kadota (Tokyo University) for the kind help to overcome the problem on the downstream analysis. We thank to Mr. Ryoji Fukuda and Mss. Hisako Fukuda for their encouragement and supports throughout of this study.

Author contribution

T.F., E.S., H.T., S.K., Y.T., and T.I., contributed to the experimental design. T.F., K.F., K.T., and A.O., performed experiments. T.K. contributed to the construction of expression vectors and supported the recombinant viral introduction and establishment of immortalized cells. T.F. wrote the manuscript.

Declaration of interests

The authors declare no conflict of interests.

Received: May 29, 2020
Revised: November 5, 2020
Accepted: December 8, 2020
Published: January 22, 2021

References

Ahuja, D., Sáenz-Robles, M.T., and Pipas, J.M. (2005). SV40 large T antigen targets multiple cellular pathways to elicit cellular transformation. Oncogene *24*, 7729–7745.

Brenner, A.J., Stampfer, M.R., and Aldaz, C.M. (1998). Increased p16 expression with first senescence arrest in human mammary epithelial cells and extended growth capacity with p16 inactivation. Oncogene *17*, 199–205.

Cho, S., Tian, Y., and Benjamin, T.L. (2001). Binding of p300/CBP co-activators by polyoma large T Antigen. J. Biol. Chem. *276*, 33533–33539.

Donai, K., Kiyono, T., Eitsuka, T., Guo, Y., Kuroda, K., Sone, H., Isogai, E., and Fukuda, T. (2014). Bovine and porcine fibroblasts can be immortalized with intact karyotype by the expression of mutant cyclin dependent kinase 4, cyclin D, and telomerase. J. Biotechnol. *176*, 50–57.

Fukuda, T., Katayama, M., Yoshizawa, T., Eitsuka, T., Mizukami, H., Nakagawa, K., Ito, H., Komagata, H., Song, S., Roh, S., et al. (2012). Efficient establishment of pig embryonic fibroblast cell lines with conditional expression of the simian vacuolating virus 40 large T fragment. Biosci. Biotechnol. Biochem. *76*, 1372–1377.

Fukuda, T., Gouko, R., Eitsuka, T., Suzuki, R., Takahashi, K., Nakagawa, K., Sugano, E., Tomita, H., and Kiyono, T. (2019a). Human-derived corneal epithelial cells expressing cell cycle regulators as a new resource for in vitro ocular toxicity testing. Front. Genet. *10*, 587.

Fukuda, T., Doi, K., Donai, K., Takahashi, K., Kobayashi, H., Hirano, T., Nishimori, K., and Yasue, H. (2019b). Data descriptor: Global transcriptome analysis of pig induced pluripotent stem cells derived from six and four reprogramming factors. Sci. Data *6*, 190034.

Fukuda, T., Takahashi, K., Takase, S., Orimoto, A., Eitsuka, T., Nakagawa, K., and Kiyono, T. (2020). Human derived immortalized dermal papilla cells with a constant expression of testosterone receptor. Front. Cell Dev. Biol. *8*.

Hawley-Nelson, P., Vousden, K.H., Hubbert, N.L., Lowy, D.R., and Schiller, J.T. (1989). HPV16 E6 and E7 proteins cooperate to immortalize human foreskin keratinocytes. EMBO J. *8*, 3905–3910.

Hayflick, L. (1965). The limited in vitro lifespan of human diploid strains. Exp. Cell Res. *37*, 614–636.

Jones, H.W., McKusick, V.A., Harper, P.S., and Wuu, K.D. (1971). The HeLa cell and a reappraisal of its origin. Obstet. Gynecol. *38*, 945–949.

Kuroda, K., Kiyono, T., Eitsuka, T., Isogai, H., Takahashi, K., Donai, K., Isogai, E., and Fukuda, T. (2015). Establishment of cell lines derived from the genus macaca through controlled expression of cell cycle regulators. J. Cell. Biochem. *116*, 205–211.

Kwack, M.H., Sung, Y.K., Chung, E.J., Im, S.U., Ahn, J.S., Kim, M.K., and Kim, J.C. (2008). Dihydrotestosterone-inducible dickkopf 1 from balding dermal papilla cells causes apoptosis in follicular keratinocytes. J. Invest. Dermatol. *128*, 262–269.

Lane, D.P. (1992). p53, guardian of the genome. Nature *358*, 15–16.

Münger, K., and Howley, P.M. (2002). Human papillomavirus immortalization and transformation functions. In Virus Res. 213–228.

Münger, K., Werness, B.A., Dyson, N., Phelps, W.C., Harlow, E., and Howley, P.M. (1989). Complex formation of human papillomavirus E7 proteins with the retinoblastoma tumor suppressor gene product. EMBO J. *8*, 4099–4105.

Royds, J.A., and Iacopetta, B. (2006). p53 and disease: when the guardian angel fails. Cell Death Differ *13*, 1017–1026.

Scheffner, M., Werness, B.A., Huibregtse, J.M., Levine, A.J., and Howley, P.M. (1990). The E6 oncoprotein encoded by human papillomavirus types 16 and 18 promotes the degradation of p53. Cell *63*, 1129–1136.

Shiomi, K., Kiyono, T., Okamura, K., Uezumi, M., Goto, Y., Yasumoto, S., Shimizu, S., and Hashimoto, N. (2011). CDK4 and cyclin D1 allow human myogenic cells to recapture growth property without compromising differentiation potential. Gene Ther. *18*, 857–866.

Su, W., Sun, J., Shimizu, K., and Kadota, K. (2019). TCC-GUI: a Shiny-based application for differential expression analysis of RNA-Seq count data. BMC Res. Notes *12*, 133.

Tani, T., Eitsuka, T., Katayama, M., Nagamine, T., Nakaya, Y., Suzuki, H., Kiyono, T., Nakagawa, K., Inoue-Murayama, M., Onuma, M., et al. (2019). Establishment of immortalized primary cell from the critically endangered Bonin flying fox (Pteropus pselaphon). PLoS One *14*, e0221364.

Tevethia, M.J. (1984). Immortalization of primary mouse embryo fibroblasts with SV40 virions, viral DNA, and a subgenomic DNA fragment in a quantitative assay. Virology *137*, 414–421.

Tevethia, M.J., and Ozer, H.L. (2015). SV40-mediated immortalization. Methods Mol. Biol. *165*, 185–199.

Wang, Y., Chen, S., Yan, Z., and Pei, M. (2019). A prospect of cell immortalization combined with matrix microenvironmental optimization strategy for tissue engineering and regeneration 06 Biological Sciences 0601 Biochemistry and Cell Biology. Cell Biosci *9*, 1–21.

Zur Hausen, H. (2002). Papillomaviruses and cancer: from basic studies to clinical application. Nat. Rev. Cancer *2*, 342–350.

iScience, Volume *24*

Supplemental Information

Combinatorial expression of cell cycle regulators is more suitable for immortalization than oncogenic methods in dermal papilla cells

Tomokazu Fukuda, Kai Furuya, Kouhei Takahashi, Ai Orimoto, Eriko Sugano, Hiroshi Tomita, Sayo Kashiwagi, Tohru Kiyono, and Tsuyoshi Ishii

TRANSPARENT METHODS SUPPLEMENTAL FILE

Cell culture and preparation of recombinant retrovirus

Primary male-derived human dermal papilla cells (DPCs) were obtained from Promocell through the local distributor. The primary cells were maintained with the recommended medium, Follicle Dermal Papilla Cell Growth Medium (C-26501). The primary cells were maintained in 35-mm cell culture dishes (Thermo Fisher Scientific, Waltham, Massachusetts, USA). Recombinant retroviruses expressing SV40 and E6/E7 protein were prepared with transient expression in 293T cells. QCXIN-based SV40-expressing retrovirus was used. A pCLMSCV-based retrovirus expressing E6/E7 oncoprotein was used. For transient packaging, the packaging plasmids PCL-gag_pol and pCMV-VSV-G (provided from Dr. Hiroyuki Miyoshi, RIKEN Bioresource Center, Tsukuba, Japan) and the retrovirus plasmid were introduced into 293T cells with lipofection. The supernatant containing retrovirus was recovered and filtered with a 0.45-μm disk filter (Sartorius, Göttingen, Germany), and then concentrated with Polyethylene Glycol 6000-based concentration solution. The recombinant viruses were

infected into DPCs with 4 μg/ml of polybrene. The surrogate marker (Enhanced green fluorescent protein, EGFP)-expressing retrovirus was prepared with the same packaging method. After the retroviral infection, the cells were selected with G418 or Hygromycin. The detailed method for establishing immortalized DPCs using the K4DT method was described in our previous publication (Fukuda et al., 2020).

The passage number of sample collection

For wild type cell, DNA extraction (passage 4), RNA extraction (passage 4), cell cycle analysis (passage 3), protein extraction for western blot (passage 4), chromosome analysis (passage 4) was carried out. For K4DT cell, DNA extraction (passage 6), RNA extraction (passage 6), cell cycle analysis (passage 6), protein extraction for western blot (passage 5), chromosome analysis (passage 5) was done. For SV40 cell, DNA extraction (passage 5), RNA extraction (passage 7), cell cycle analysis (passage 9), protein extraction for western blotting (passage 9), chromosome analysis (passage 9) was finished. For E6E7 cell, DNA extraction (passage 7), RNA extraction (passage 7), cell cycle analysis (passage 8), protein extraction for western blotting (passage 7), chromosome analysis (passage 7) was analyzed.

Genomic PCR

For the extraction of genomic DNA, we used the NucleoSpin Tissue Kit (cat. no. 740952, TaKaRa Bio, Shiga, Japan). Amplification reaction was carried out with 100 ng of template DNA, 1X PCR buffer of KOD-FX neo (KFX- 201; Toyobo, Osaka, Japan), 0.4 mM dNTPs (KFX-201, Toyobo), 0.5 U KOD-FX neo (KFX-201, Toyobo), and 0.3 mM of each primer), based on protocol provided from the manufacturer of PCR enzyme. The reaction conditions were as follows; Cyclin D1, CDK4 and TSC2, pre-denaturation at 94°C for 2 min, denaturation at 94°C for 10 sec, and extension at 68°C for 1 min. for 40 cycles (denaturation and extension), SV40, pre-denaturation at 94°C for 2 min, denaturation at 94°C for 10 sec, and extension at 55°C for 30 sec. for 40 cycles, E6E7, pre-denaturation at 94°C for 2 min, denaturation at 94°C for 10 sec, and extension at 65°C for 30 sec. for 40 cycles. TSC2 was used as the internal control. We visualized PCR products by electrophoresis in 1% agarose / Tris–acetate–EDTA (ethylenediaminetetraacetic acid) gels and stained with ethidium bromide. For the detection of CyclinD1, TF806(5'-GGCACCAAAATCAACGGGACTTT-3') and TF807(5'-TTCCTCGCAGACCTCCAGCA-3'), for the detection of CDK4, TF806(5'-

GGCACCAAAATCAACGGGACTTT-3') and TF808(5'-ACGAACTGTGCTGATGGGAAGGC-3'), for the detection of TERT, TF806(5'-GGCACCAAAATCAACGGGACTTT-3') and TF809(5'-AGCTCCTTCAGGCAGGACACCT-3'), for the detection of SV40, TF1068(5'-AGCCCAGCCACTATAAGTACCAT-3') and TF1069(5'-CTAGCTCAAAGTTCAGCCTGTCCAAG-3'), for the detection of E6E7, TF1060(5'-AAGCAAAGATTCCATAATATAAGGGGTCGGT-3') and TF1061(5'-TGAGAACAGATGGGGCACACAATTCCT-3'), for the detection of TSC2 gene as the internal control, TF963(5'-AAACCGAGCCCCATTTGACC-3') and TF964(5'-TGGTCGTAGCGGAATCGAGGAT-3') were used.

Cell cycle analysis, western blotting, and RNA extraction

Cell cycle analysis was carried out using the Muse cell cycle assay kit (cat. No. MCH100106, Merck Millipore, Billerica, MA, USA) and a Cell Analyzer (Merck Millipore) on wild type, SV40 cell, E6/E7 cell, and K4DT dermal papilla cells. Fixation and analysis was performed according to the manufacturer's instructions. Total RNA from wild type, SV40, E6/E7, and K4DT cells was obtained with the NucleoSpin RNA

kit (Takara Bio). 5 μg of total RNA was used for the reverse transcriptase reaction. The cDNAs were applied into an Illumina next generation sequencer (Hiseq2500). Western blotting was performed for the detection of introduced proteins, such as CDK4, cyclin D1, SV40, and E7 oncoprotein. Alpha-tubulin was detected as a control for the western blotting. The cells were lysed in a solution containing 50 mM Tris-HCl (pH7.4), 0.15 M NaCl, 1% Triton X100, 2.5 mM/ml deoxycholate, and protease inhibitor cocktail (Nakarai Tesque, Kyoto, Japan). The detailed protocol of western blotting is described in our previous publication (Fukuda et al., 2008). Rabbit anti-human cyclin D1 antibody (1:5000 dilution, code no. 553, Medical & Biological Laboratories Co., LTD (MBL), Nagoya, Japan), mouse anti-human CDK4 antibody (1:200 dilution, cat. No. sc-56277, Santa Cruz Biotechnology, Dallas, TX, USA), mouse anti-SV40 antibody (1:200 dilution, cat. No. sc-147, Santa Cruz Biotechnology), mouse anti-E7 antibody (1:100 dilution, cat No. sc-65711, Santa Cruz Biotechnology), and mouse anti-α-tubulin antibody (1:1000, cat. no. sc-32293, Santa Cruz Biotechnology) were used as primary antibodies. As the secondary antibodies, goat anti-mouse IgG labeled with horseradish peroxidase (HRP) (1:2000, code no. 330, MBL) or goat anti-rabbit IgG with HRP (1:2000, code no. 458,

MBL) was used. Signals were detected using Signal Enhancer HIKARI for western blotting and ELISA (code no. 02270-81, Nacalai Tesque), and the images were detected using an ImageQuant LAS-4000 Mini system (Fujifilm, Tokyo, Japan).

RNA-Seq analysis

We evaluated the quality of raw sequencing reads with FastQC. The paired-end sequencing data was obtained from three biological replicates. We analyzed data from 12 samples of wild type, SV40, E6/E7, and K4DT immortalized DPCs. The raw reads data were processed with PEAT for the removal of the adaptor sequence. The low quality reads were removed with PRINSEQ. The remaining data after PRINSEQ was mapped to a human reference genome (GRCh38, NCBI) with STAR. Gene expression was counted with the feature Count program. The gene expression counts data was normalized with the edgeR program based on the TMM normalization method. The differentially expressed genes were identified with the TCC-GUI program, which was provided by the website of Dr. Koji Kadota (Tokyo University, Tokyo, Japan) (https://infinityloop.shinyapps.io/TCC-GUI/). The parameters of TCC-GUI are described below. The normalization method was the TMM method, differentially

expressed genes were counted with EdgeR, the filtering threshold for low count genes was set at 30, the FDR cut off was set at 0.1, and elimination of potential differentially expressed genes was set at 0.05. For PCA analysis, we carried out the analysis with log(X+1) transform minus, Center minus, Scale ON. For heatmap analysis, the distance measure was the Euclidean method, and the agglomeration method was completed under the condition of log(X+1) transform minus.

Downstream pathway analysis

After the mapping of sequencing reads, we obtained expression counts for 33,122 genes. From the results of differentially expressed gene analysis using TCC-GUI, we further extracted 15103 DE genes that have P value less than 0.01. The list of DE genes was submitted to the pathway analysis online tool DAVID. The expression levels of genes listed in cell cycle and pathway in cancer were extracted from heatmap analysis and visualization of the expression level with TCC-GUI.

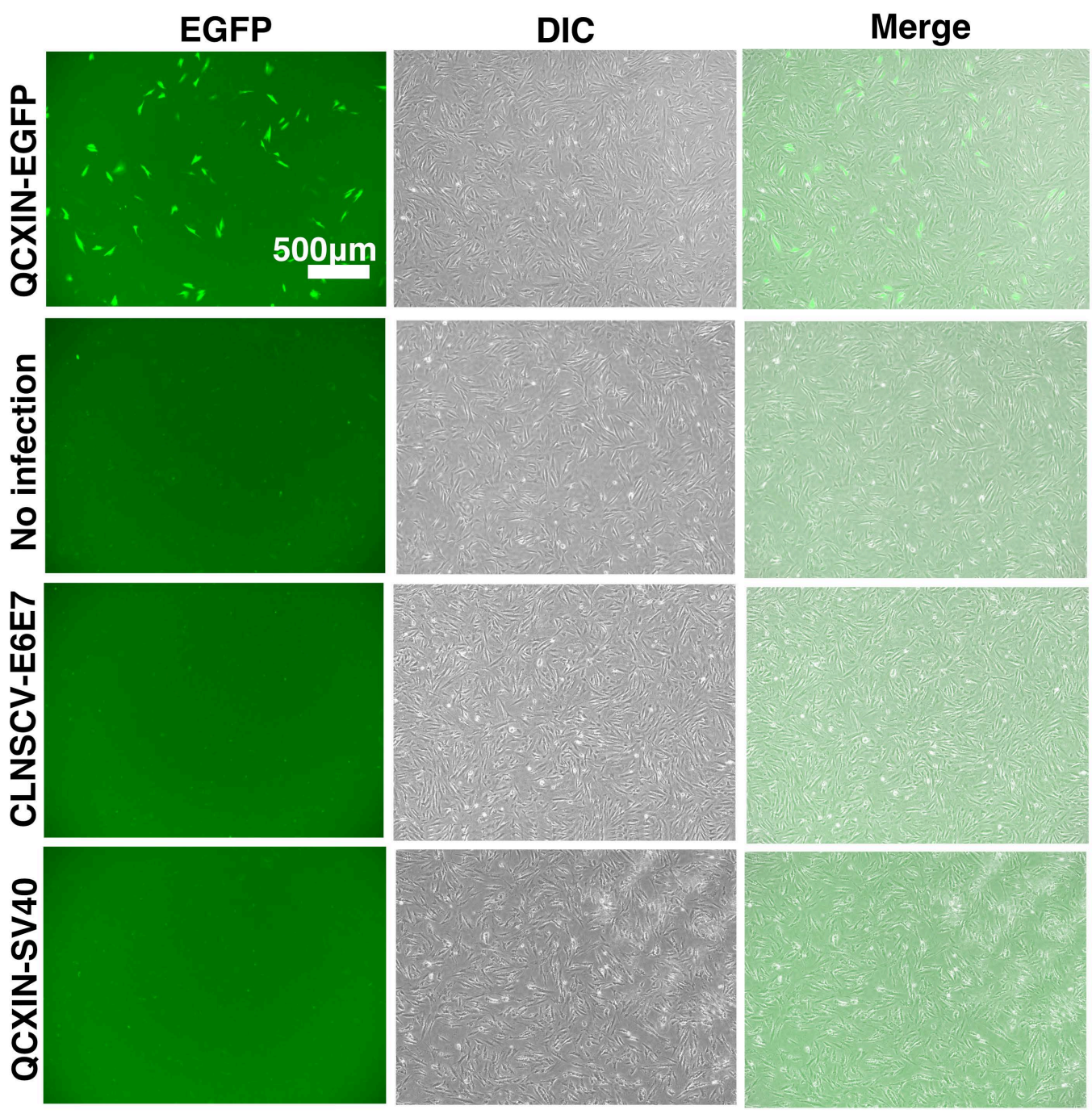


Fig. S1 Gene introduction efficiency of enhanced green fluorescence protein expressing retrovirus to primary dermal papilla cells (DPCs) and cell morphology of intact cell, E6E7 expressing DPCs, and SV40 expressing DPCs. Bar indicate 500μm.

Fig.S1

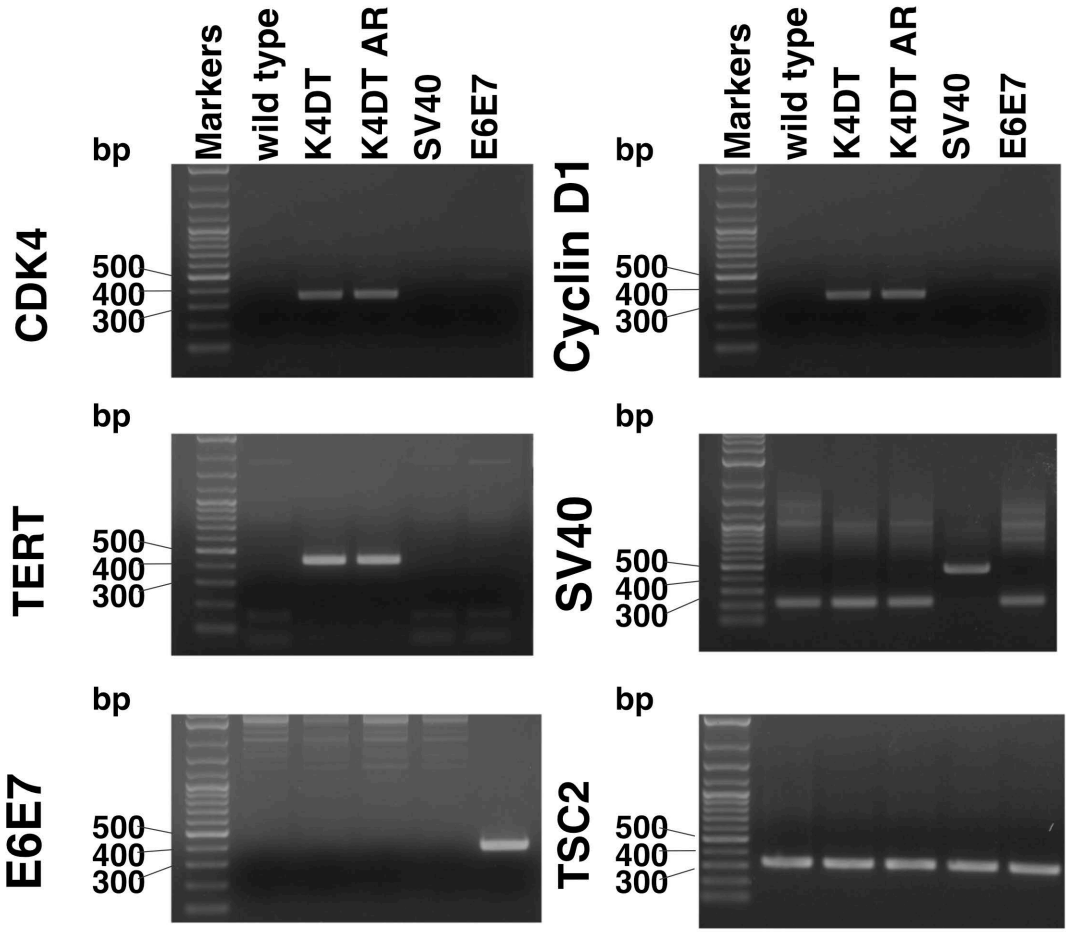


Fig. S2 Detection of inserted expression cassettes with PCR analysis. The primers used for the detection was listed in Materials and Method section.

Fig.S2

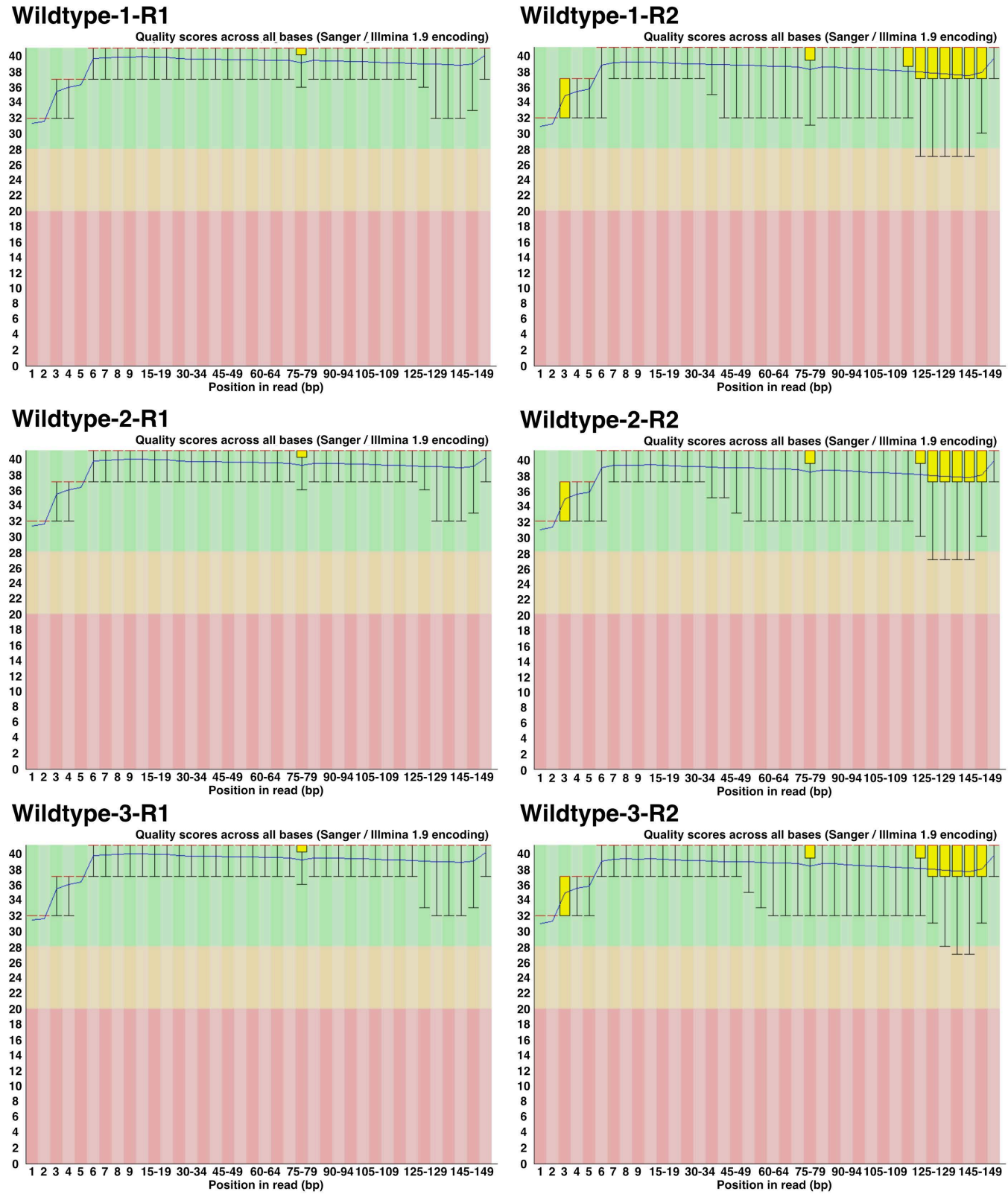


Fig. S3 Results of FastQC evaluation of raw sequencing reads from wild type cells after the PEAT adaptor removal and PRINSEQ. R1 and R2 indicate that reads are paired end.

Fig.S3

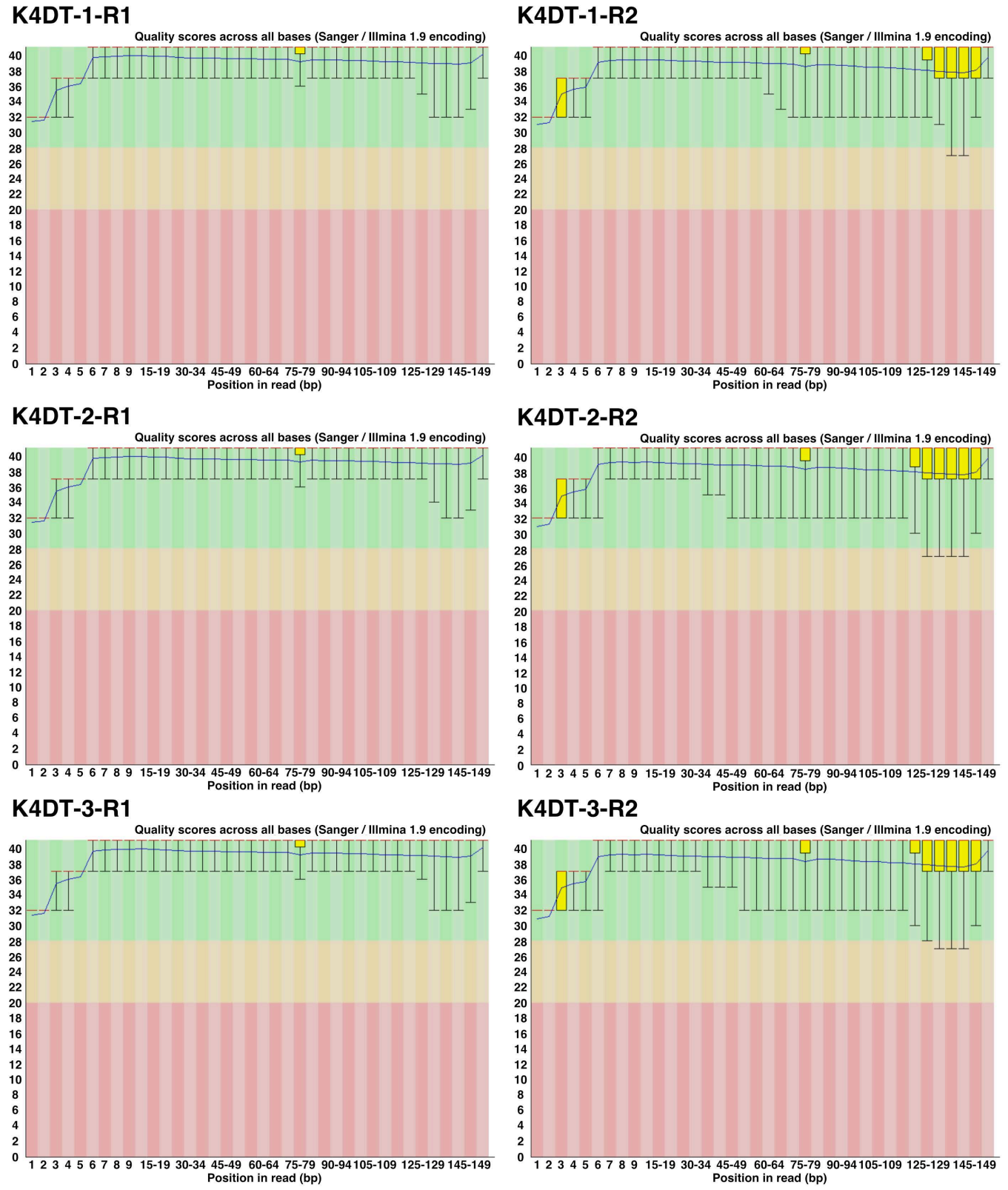


Fig. S4 Results of FastQC evaluation of raw sequencing reads from K4DT immortalized cells after the PEAT adaptor removal and PRINSEQ. R1 and R2 indicate that reads are paired end.

Fig.S4

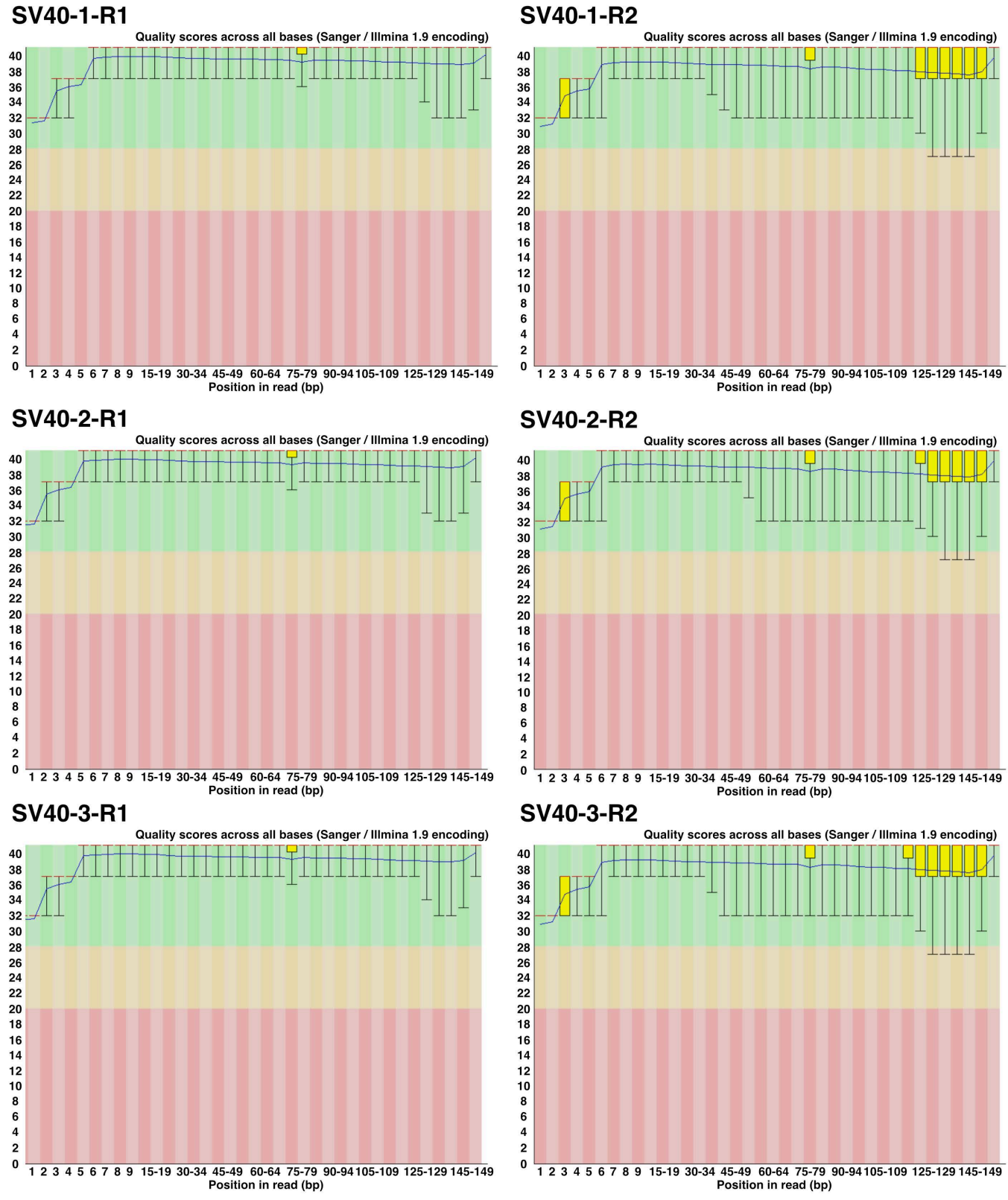


Fig. S5 Results of FastQC evaluation of raw sequencing reads from SV40 immortalized cells after the PEAT adaptor removal and PRINSEQ. R1 and R2 indicate that reads are paired end.

Fig.S5

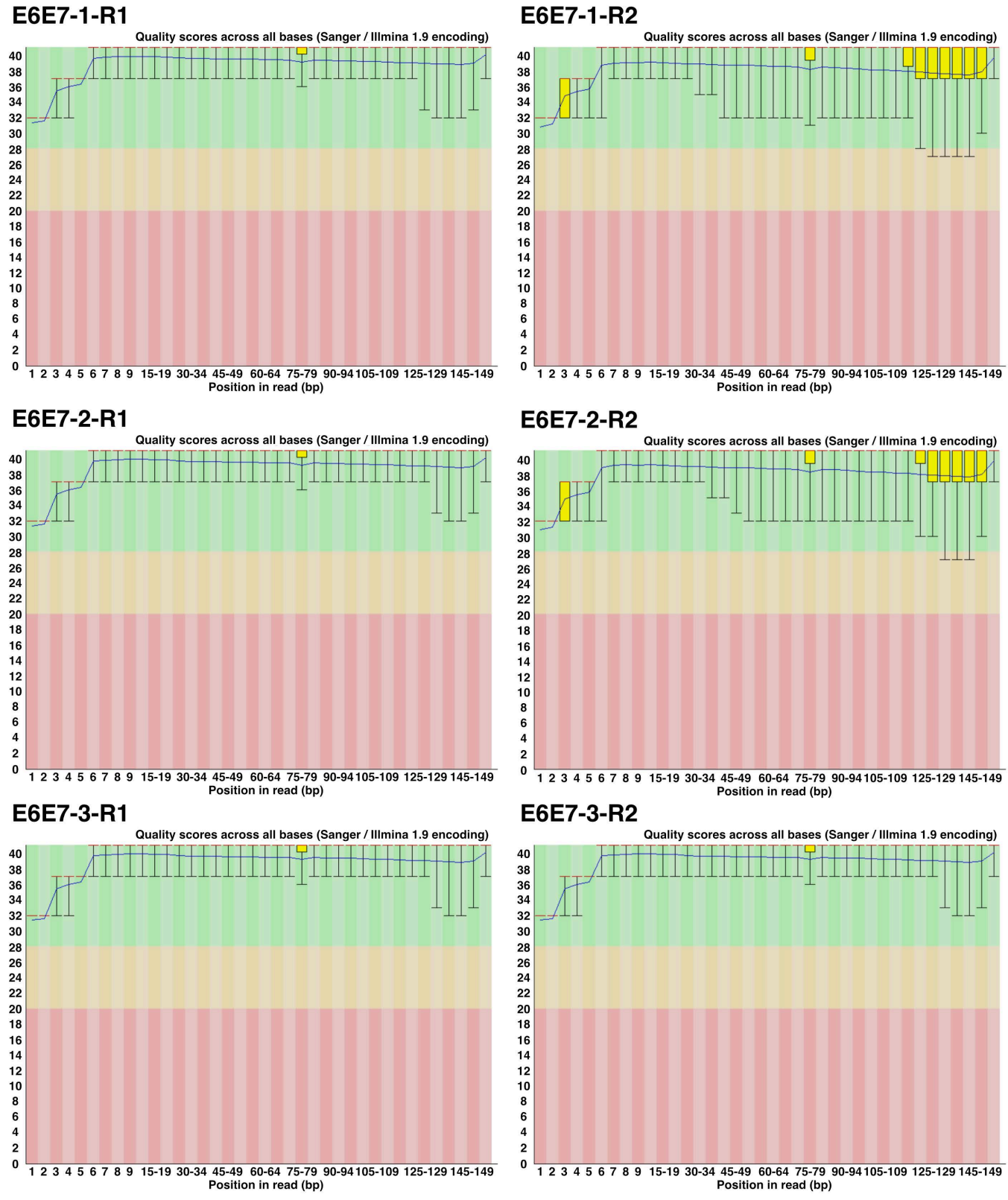


Fig. S6 Results of FastQC evaluation of raw sequencing reads from E6/E7 immortalized cells after the PEAT adaptor removal and PRINSEQ. R1 and R2 indicate that reads are paired end.

Fig.S6

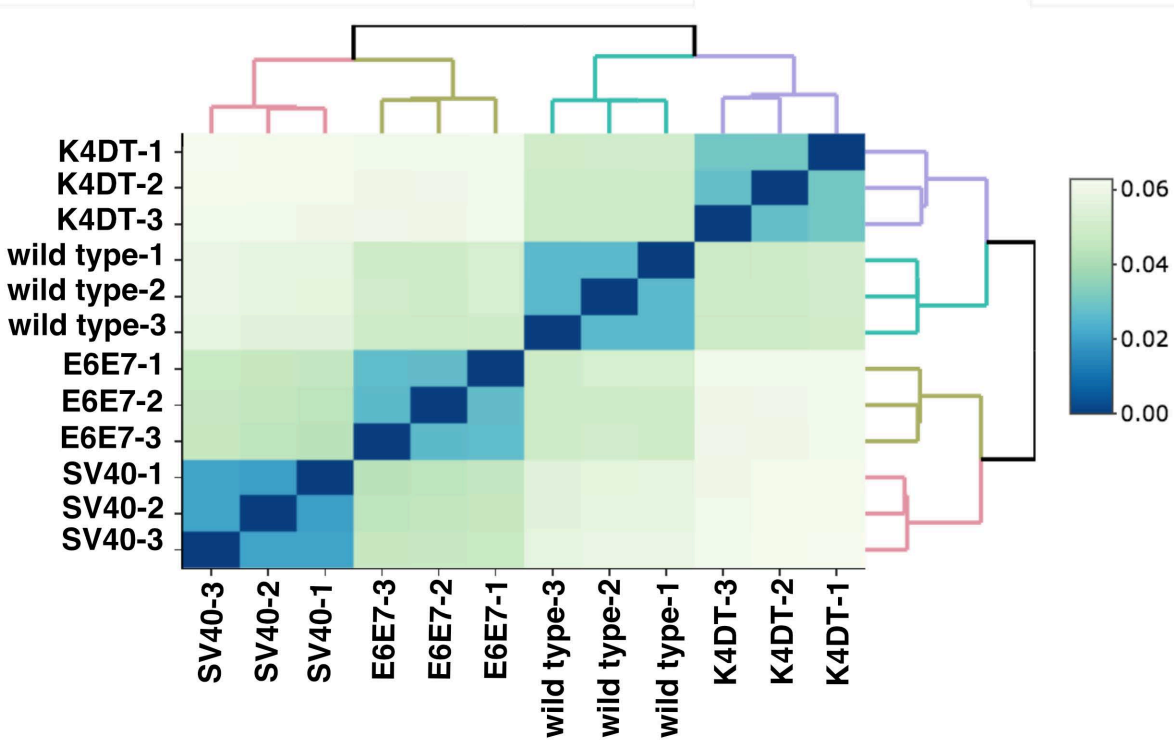


Fig. S7 Results of correlation matrix of whole expression counts of wild type, K4DT, E6/E7, SV40 DPCs. 1, 2, 3 means triplicated samples. The triplicated read data formed unique clusters.

Fig.S7

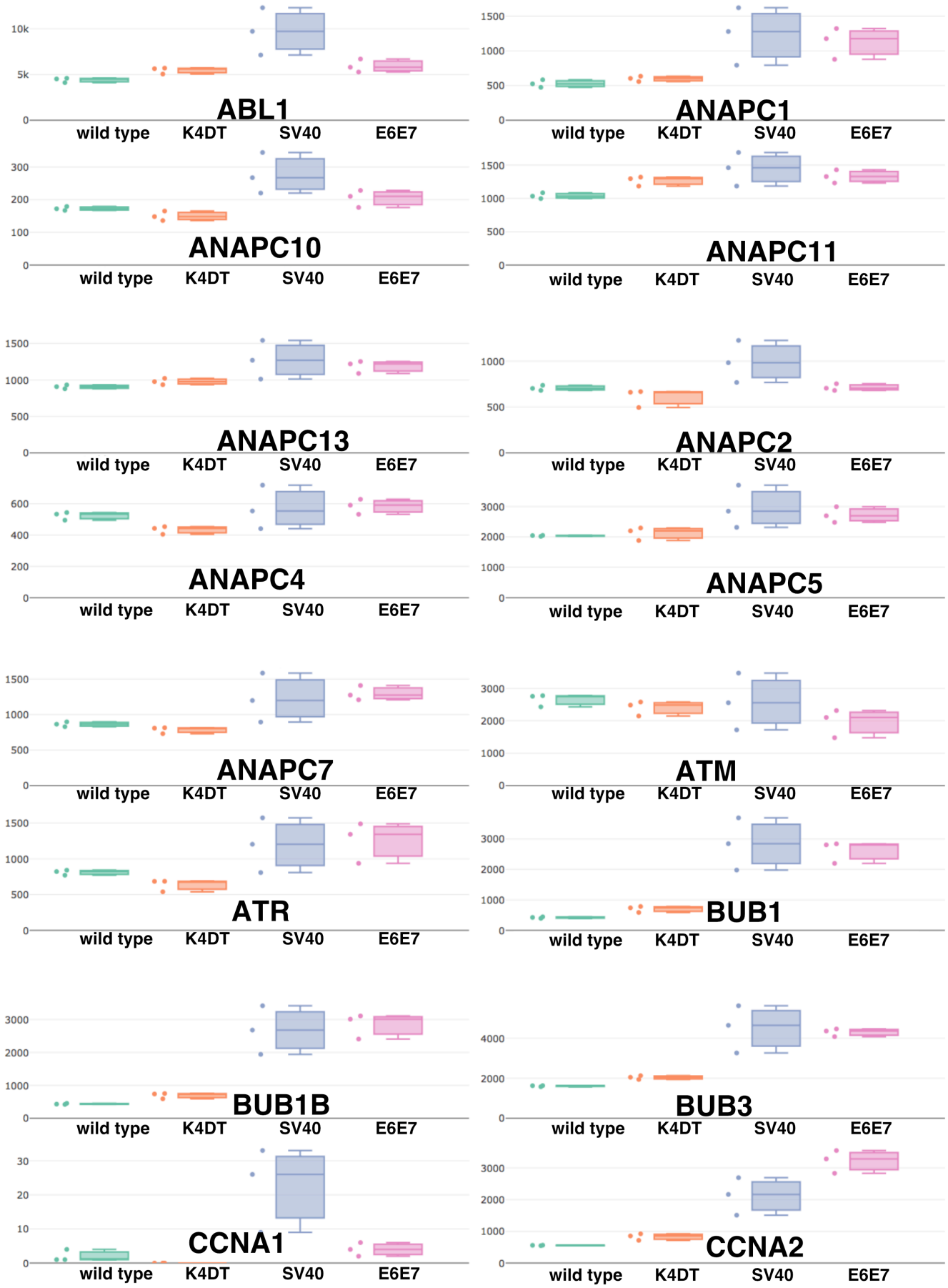

ABL1
ANAPC1
ANAPC10
ANAPC11
ANAPC13
ANAPC2
ANAPC4
ANAPC5
ANAPC7
ATM
ATR
BUB1
BUB1B
BUB3
CCNA1
CCNA2
wild type
K4DT
SV40
E6E7

Fig.S8

Fig. S8 Raw expression counts of cell cycle related genes listed in cell cycle of KEGG. The expression counts are listed with box bar plots. The expression counts of ABL1, ANAPC1, ANAPC10, ANAPC11, ANAPC13, ANAPC2, ANAPC4, ANAPC5, ANAPC7, ATM, ATR, BUB1, BUB1B, BUB3, CCNA1, CCNA2 were shown.

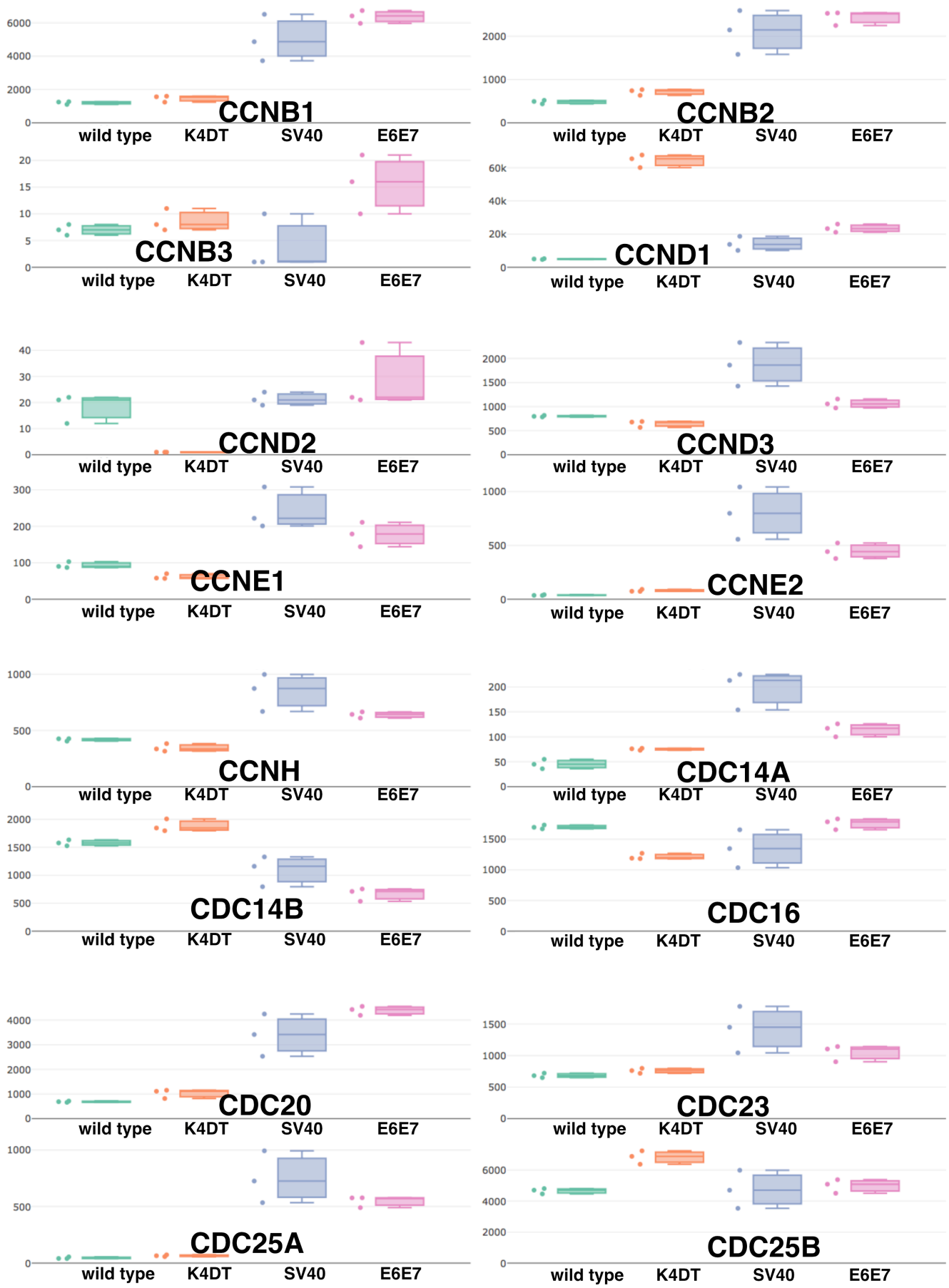

CCNB1
CCNB2
CCNB3
CCND1
CCND2
CCND3
CCNE1
CCNE2
CCNH
CDC14A
CDC14B
CDC16
CDC20
CDC23
CDC25A
CDC25B
wild type
K4DT
SV40
E6E7

Fig.S9

Fig. S9 Raw expression counts of cell cycle related genes listed in cell cycle of KEGG. The expression counts are listed with box bar plots. The expression counts of CCNB1, CCNB2, CCNB3, CCND1, CCND2, CCND3, CCNE1, CCNE2, CCNH, CDC14A, CDC14B, CDC16, CDC20, CDC23, CDC25A, CDC25B were shown

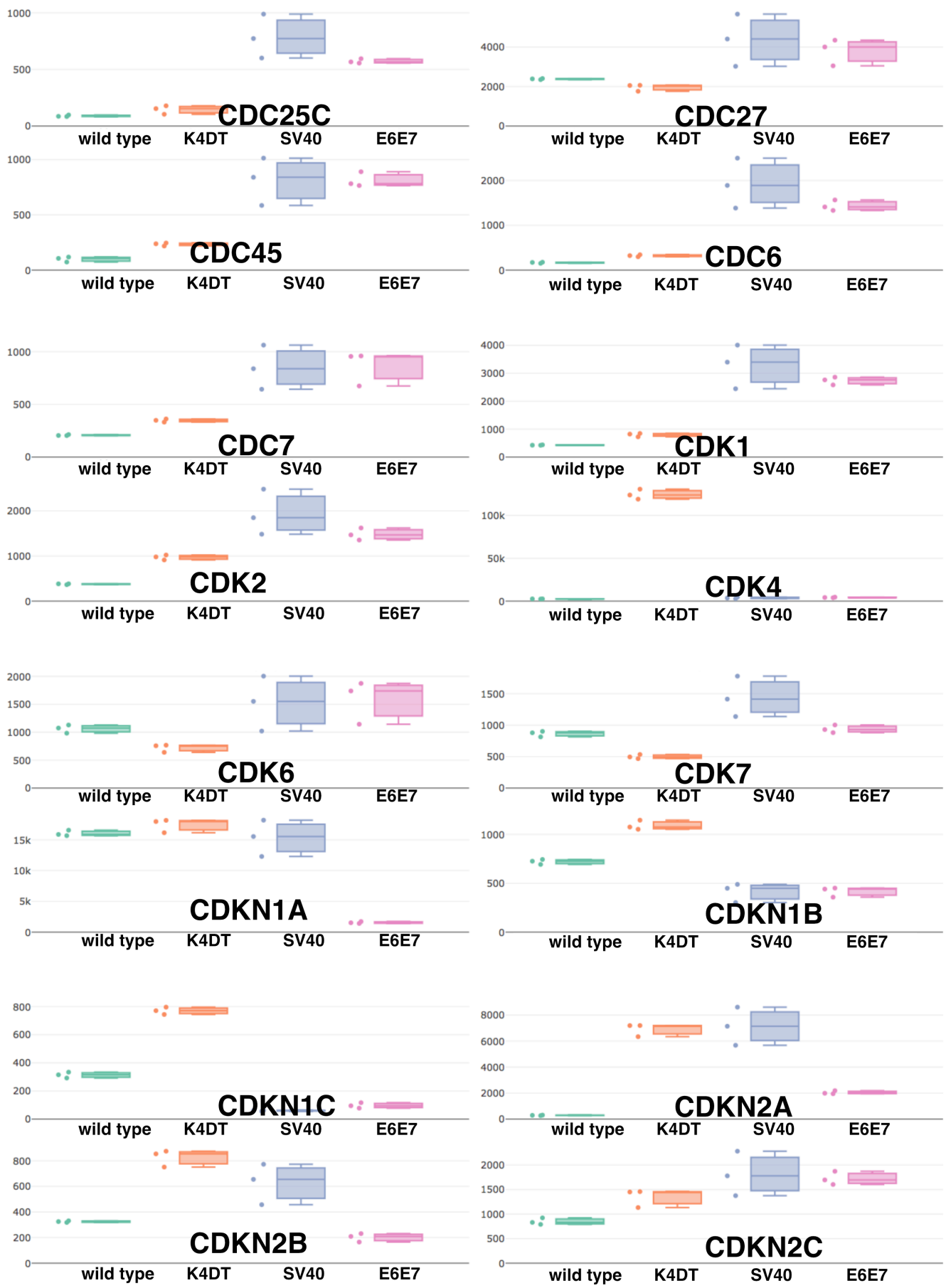

CDC25C
CDC27
CDC45
CDC6
CDC7
CDK1
CDK2
CDK4
CDK6
CDK7
CDKN1A
CDKN1B
CDKN1C
CDKN2A
CDKN2B
CDKN2C
wild type
K4DT
SV40
E6E7

Fig.S10

Fig. S10 Raw expression counts of cell cycle related genes listed in cell cycle of KEGG. The expression counts are listed with box bar plots. The expression counts of CDC25C, CDC27, CDC45, CDC6, CDC7, CDK1, CDK2, CDK4, CDK6, CDK7, CDKN1A, CDKN1B, CDKN1C, CDKN2A, CDKN2B, CDKN2C were shown.

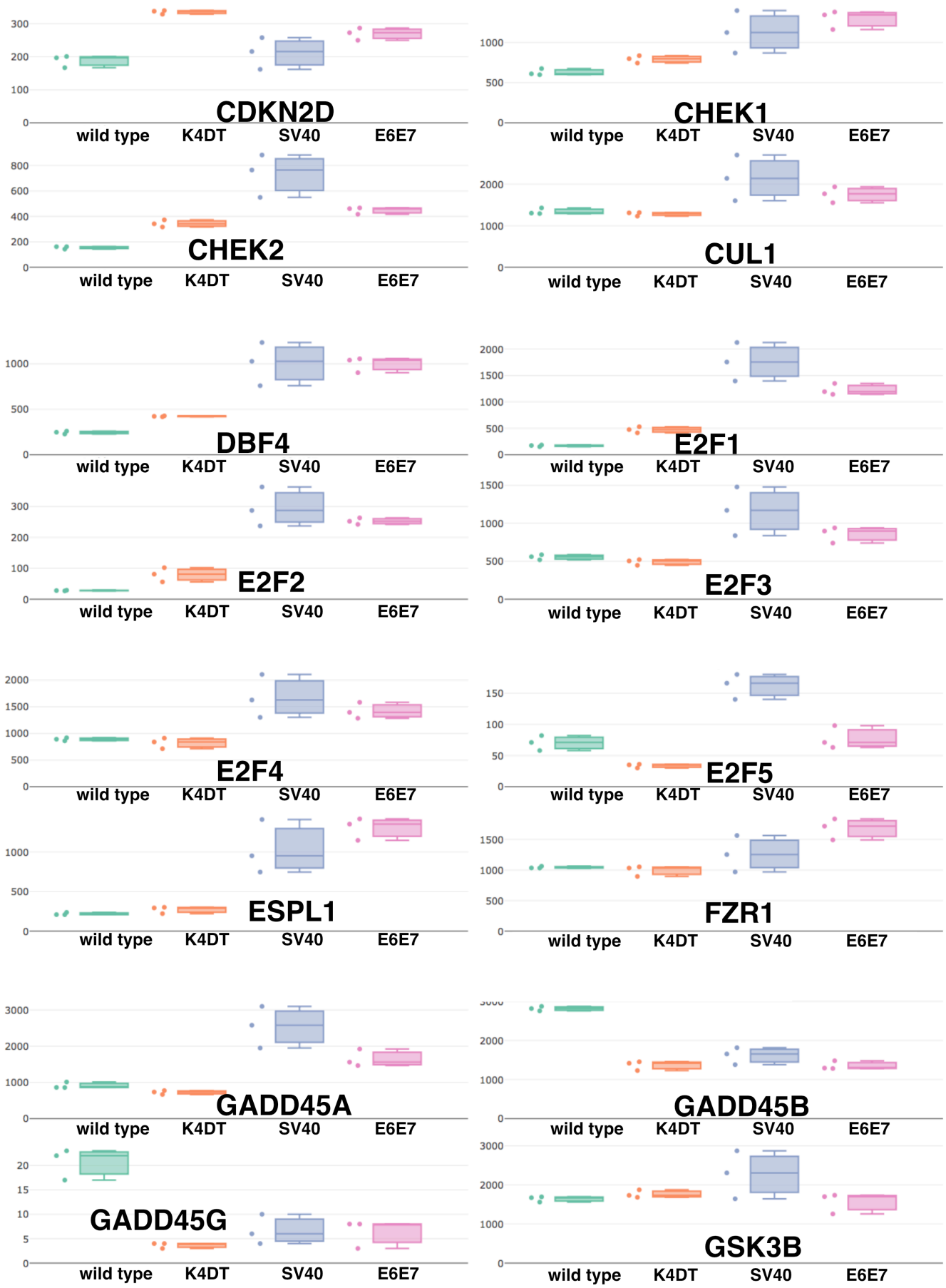

CDKN2D
CHEK1
CHEK2
CUL1
DBF4
E2F1
E2F2
E2F3
E2F4
E2F5
ESPL1
FZR1
GADD45A
GADD45B
GADD45G
GSK3B
wild type
K4DT
SV40
E6E7

Fig.S11

Fig. S11 Raw expression counts of cell cycle related genes listed in cell cycle of KEGG. The expression counts are listed with box bar plots. The expression counts of CDKN2D, CHEK1, CHEK2, CUL1, CDF4, E2F1, E2F2, E2F3, E2F4, E2F5, ESPL1, FZR1, GADD45A, GADD45B, GADD45G, GSK3B were shown.

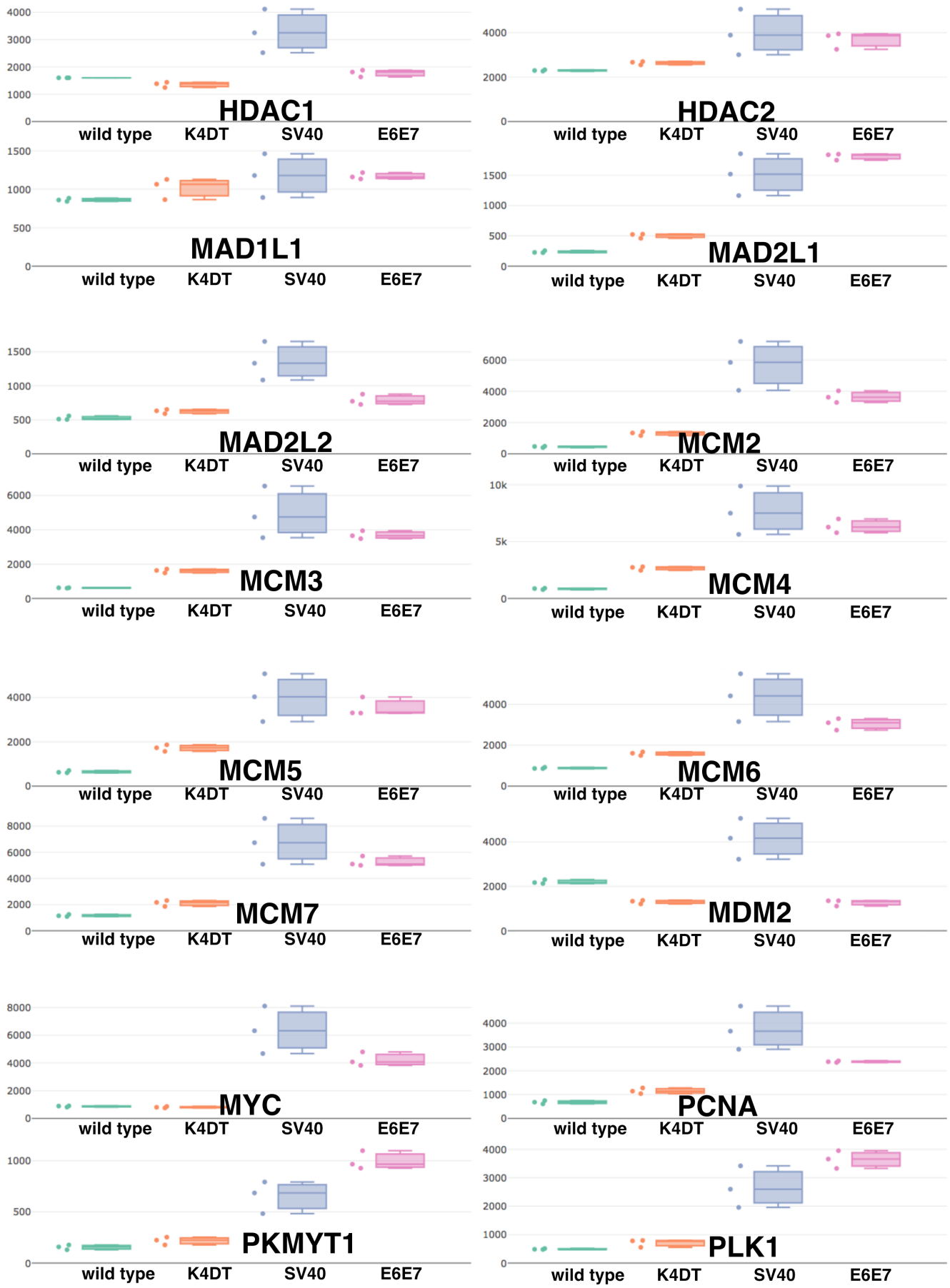

HDAC1
wild type
K4DT
SV40
E6E7
HDAC2
MAD1L1
MAD2L1
MAD2L2
MCM2
MCM3
MCM4
MCM5
MCM6
MCM7
MDM2
MYC
PCNA
PKMYT1
PLK1

Fig.S12

Fig. S12 Raw expression counts of cell cycle related genes listed in cell cycle of KEGG. The expression counts are listed with box bar plots. The expression counts of HDAC1, HDAC2, MAD1L1, MAD2L1, MAD2L2, MCM2, MCM3, MCM4, MCM5, MCM6, MCM7, MDM2, MYC, PCNA, PKMYT1, PLK1 were shown.

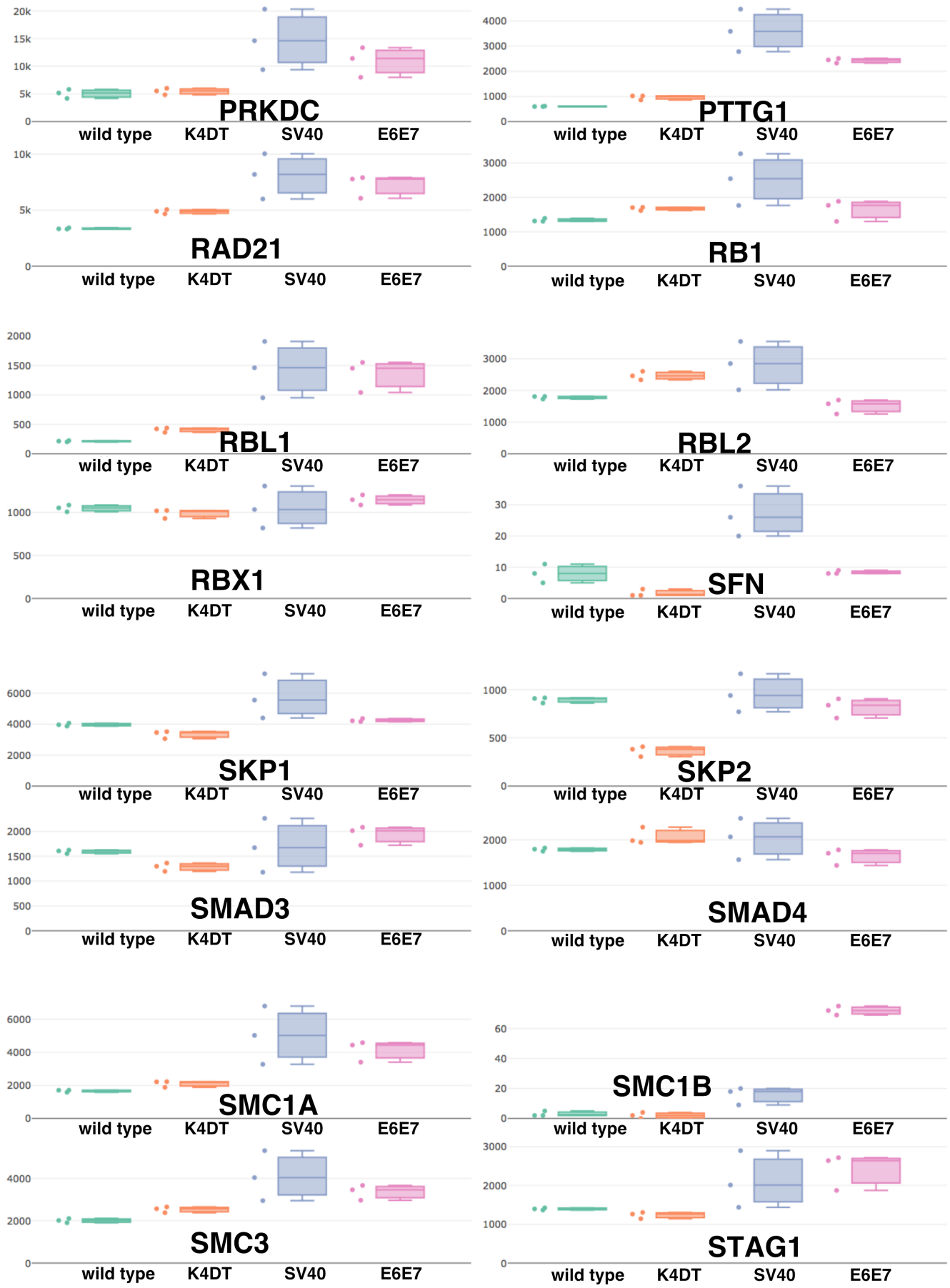

PRKDC
wild type
K4DT
SV40
E6E7
PTTG1
RAD21
RB1
RBL1
RBL2
RBX1
SFN
SKP1
SKP2
SMAD3
SMAD4
SMC1A
SMC1B
SMC3
STAG1

Fig.S13

Fig. S13 Raw expression counts of cell cycle related genes listed in cell cycle of KEGG. The expression counts are listed with box bar plots. The expression counts of PRKDC, PTTG1, RAD21, RB1, RBL1, RBL2, RBX1, SFN, SKP1, SKP2, SMAD3, SMAD4, SMC1A, SMC1B, SMC3, STAG1 were shown.

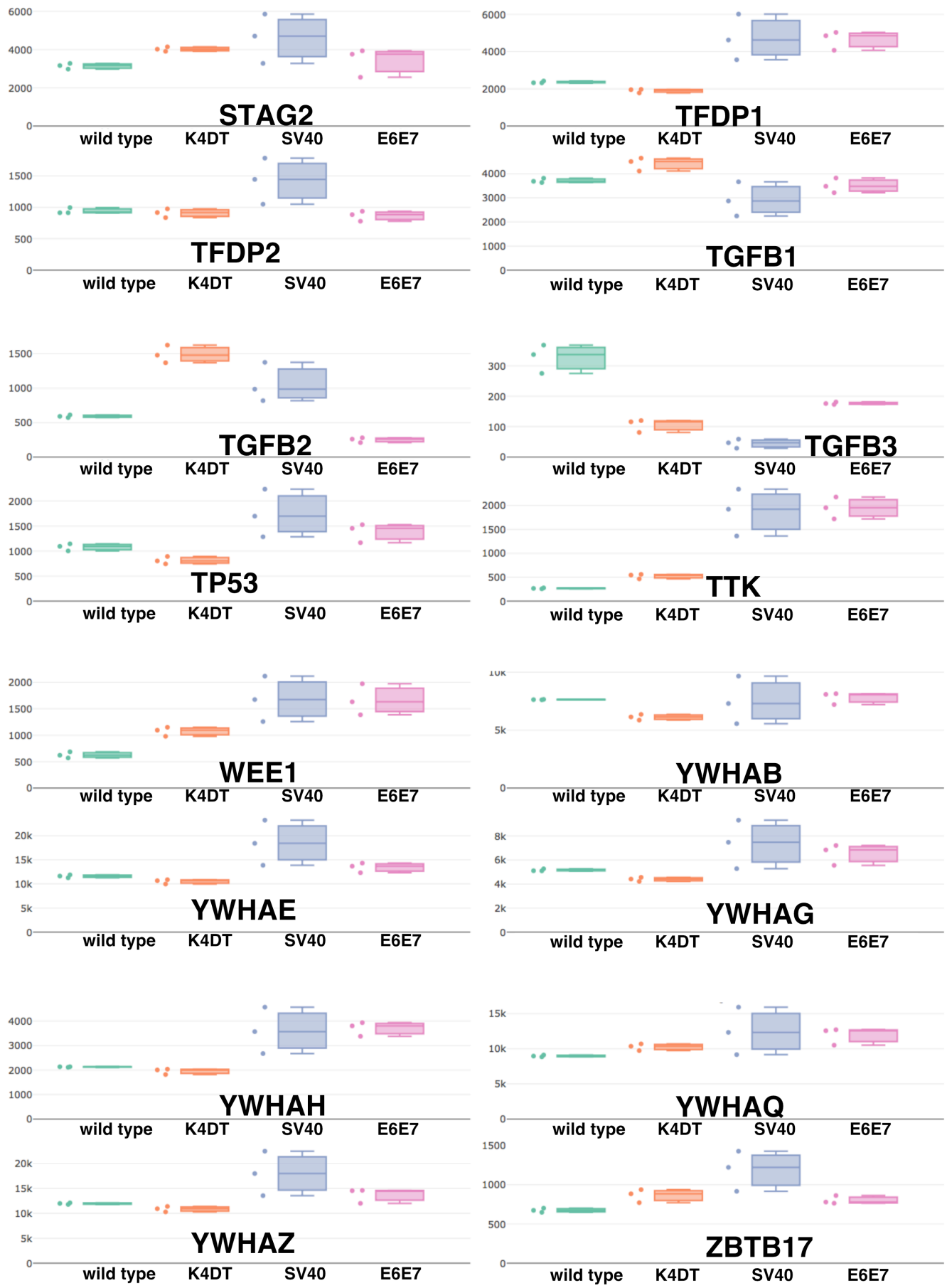

STAG2
TFDP1
TFDP2
TGFB1
TGFB2
TGFB3
TP53
TTK
WEE1
YWHAB
YWHAE
YWHAG
YWHAH
YWHAQ
YWHAZ
ZBTB17
wild type
K4DT
SV40
E6E7

Fig.S14

Fig. S14 Raw expression counts of cell cycle related genes listed in cell cycle of KEGG. The expression counts are listed with box bar plots. The expression counts of STAG2, TFDP1, TFDP2, TGFB1, TGFB2, TGFB3, TP53, TTK, WEE1, YWHAB, YWHAE, YWHAG, YWHAH, YWHAQ, YWHAZ, ZBTB17 were shown.

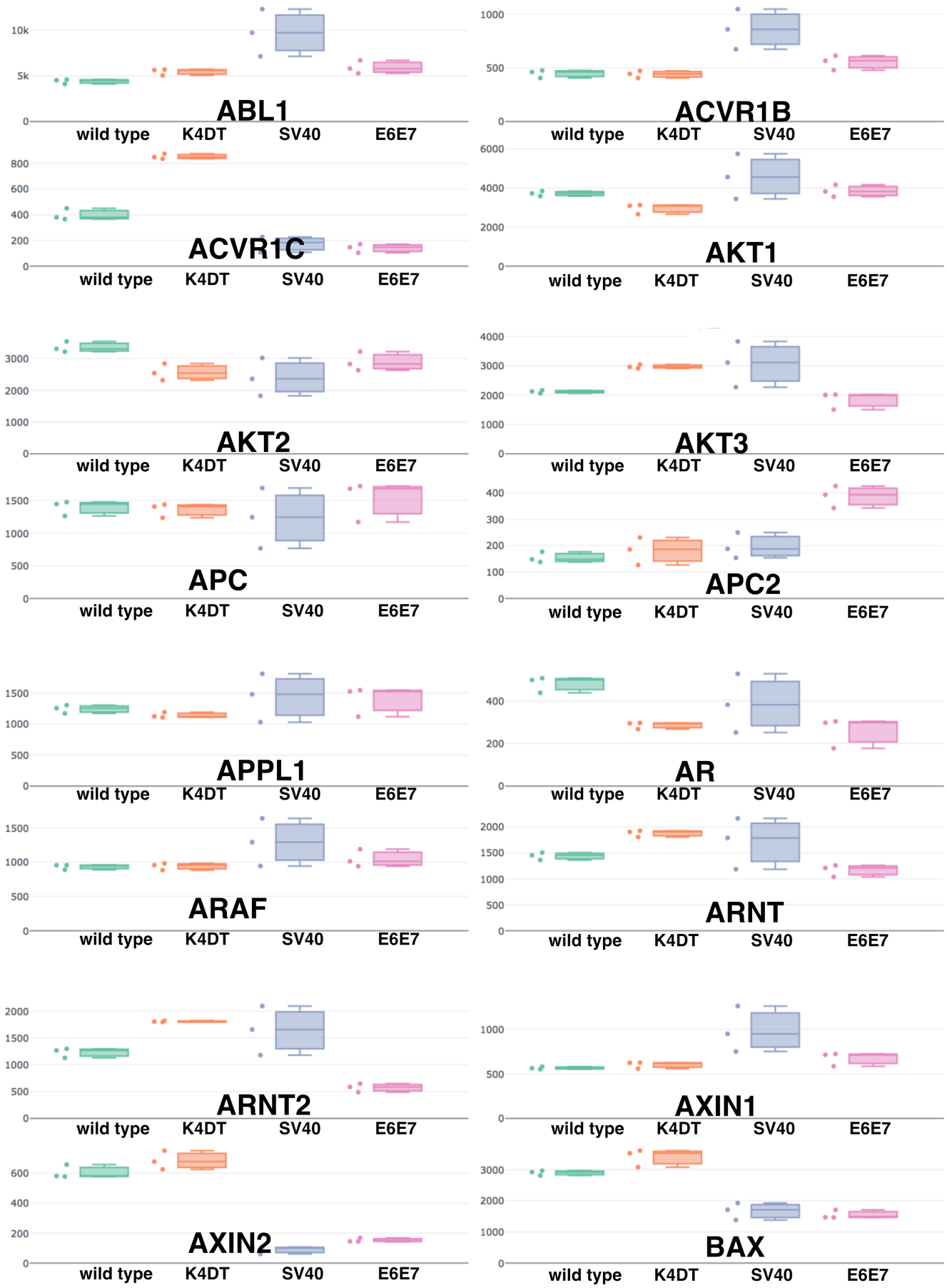

ABL1
ACVR1B
ACVR1C
AKT1
AKT2
AKT3
APC
APC2
APPL1
AR
ARAF
ARNT
ARNT2
AXIN1
AXIN2
BAX
wild type
K4DT
SV40
E6E7

Fig.S15

Fig. S15 Raw expression counts of pathway in cancer related genes listed in pathway in cancer of KEGG. The expression counts are listed with box bar plots. The expression counts of ABL1, ACVR1B, ACVR1C, AKT1, AKT2, AKT3, APC, APC2, APPL1, AR, ARAF, ARNT, ARNT2, AXIN1, AMIN2, BAX were shown.

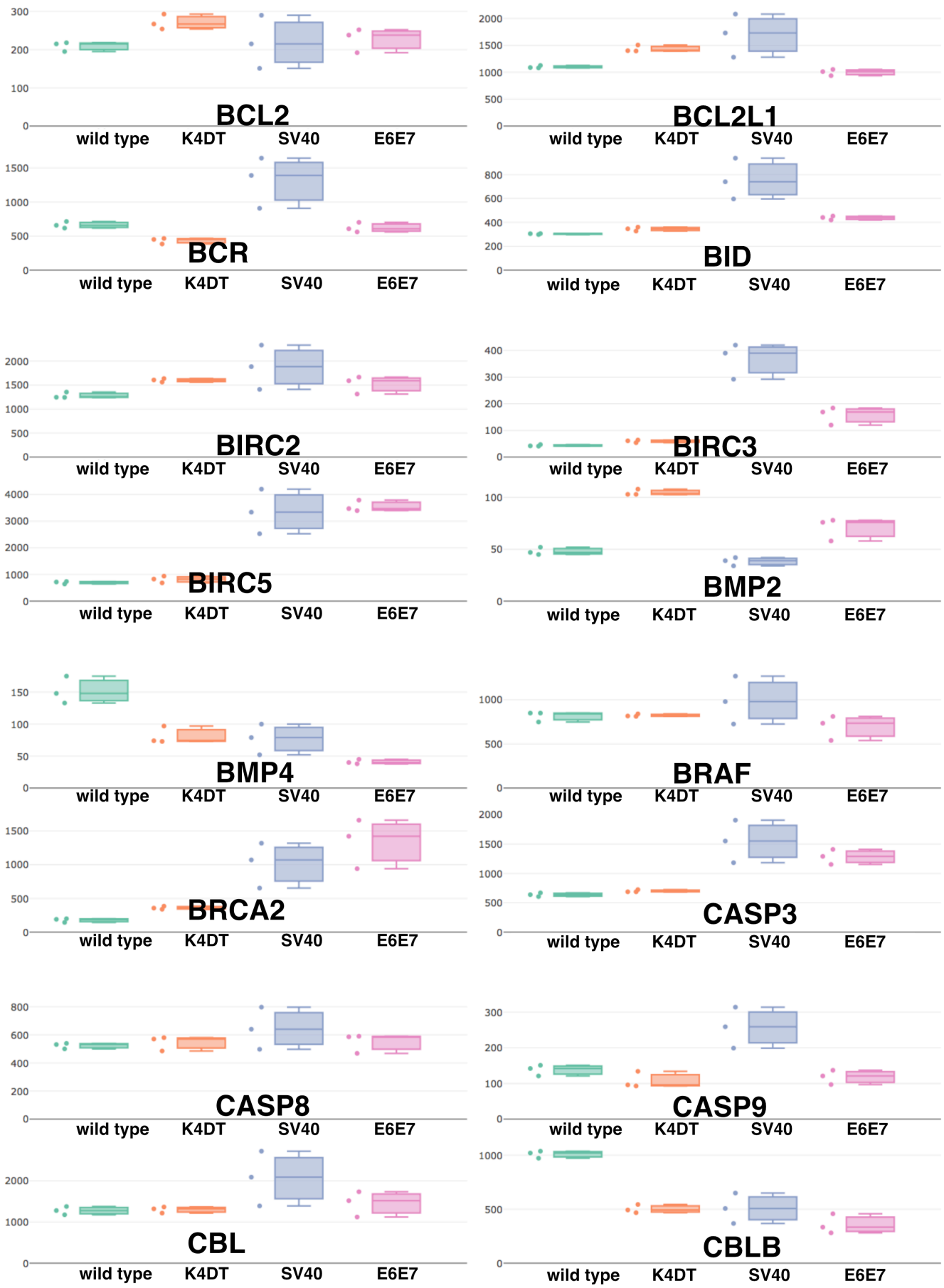

BCL2
BCL2L1
BCR
BID
BIRC2
BIRC3
BIRC5
BMP2
BMP4
BRAF
BRCA2
CASP3
CASP8
CASP9
CBL
CBLB
wild type
K4DT
SV40
E6E7

Fig.S16

Fig. S16 Raw expression counts of pathway in cancer related genes listed in pathway in cancer of KEGG. The expression counts are listed with box bar plots. The expression counts of BCL2, BCL2L1, BCR, BID, BIRC2, BIRC3, BIRC5, BMP2, BMP4, BRAF, BRCA2, CASP3, CASP8, CASP9, CBL, CBLB, were shown.

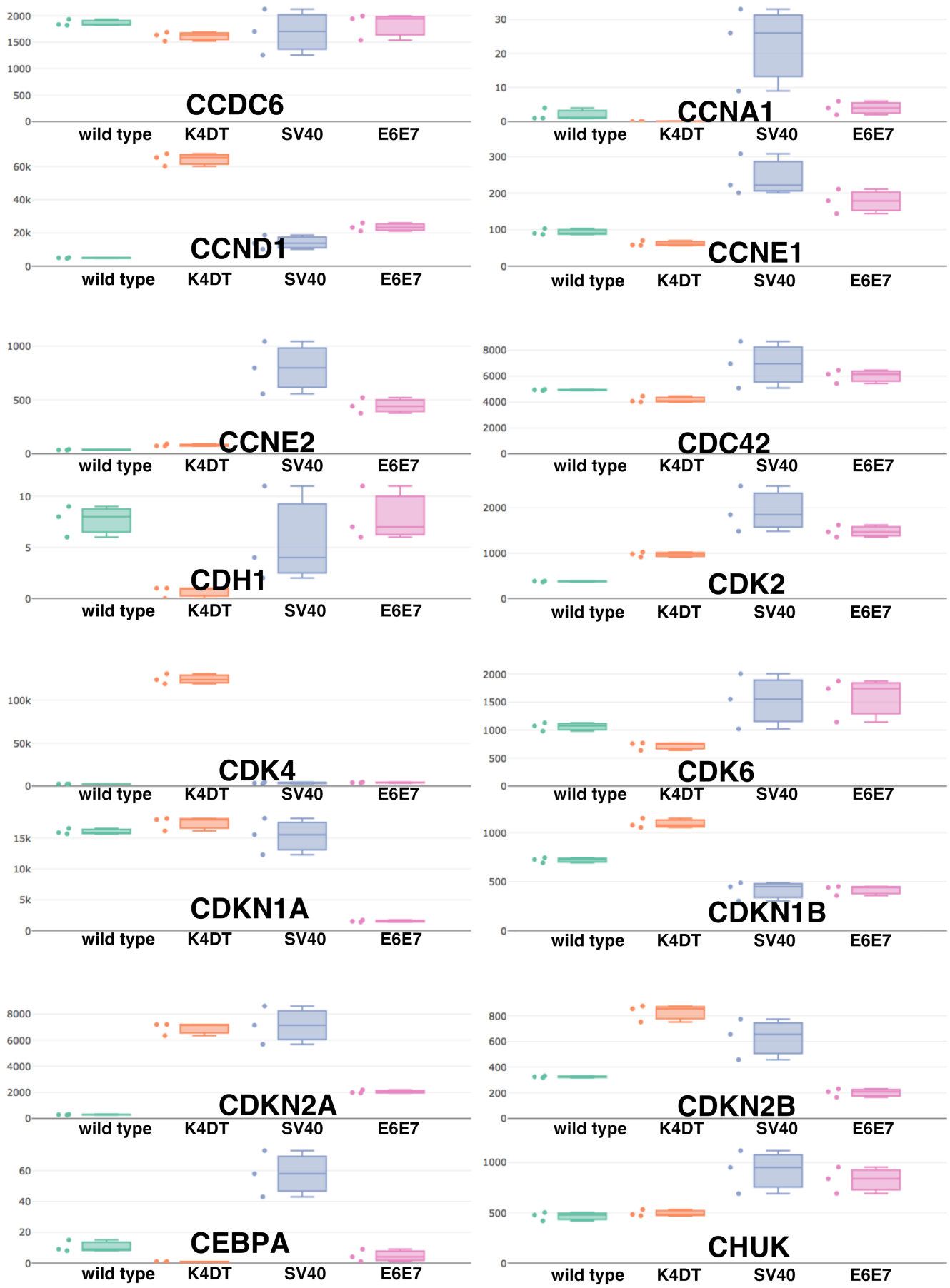

CCDC6
CCNA1
CCND1
CCNE1
CCNE2
CDC42
CDH1
CDK2
CDK4
CDK6
CDKN1A
CDKN1B
CDKN2A
CDKN2B
CEBPA
CHUK
wild type
K4DT
SV40
E6E7

Fig.S17

Fig. S17 Raw expression counts of pathway in cancer related genes listed in pathway in cancer of KEGG. The expression counts are listed with box bar plots. The expression counts of CCDC6, CCNA1, CCND1, CCNE1, CCNE2, CDC42, CDH1, CDK2, CDK4, CDK6, CDKN1A, CDKN1B, CDKN2A, CDKN2B, CEBPA, CHUK were shown.

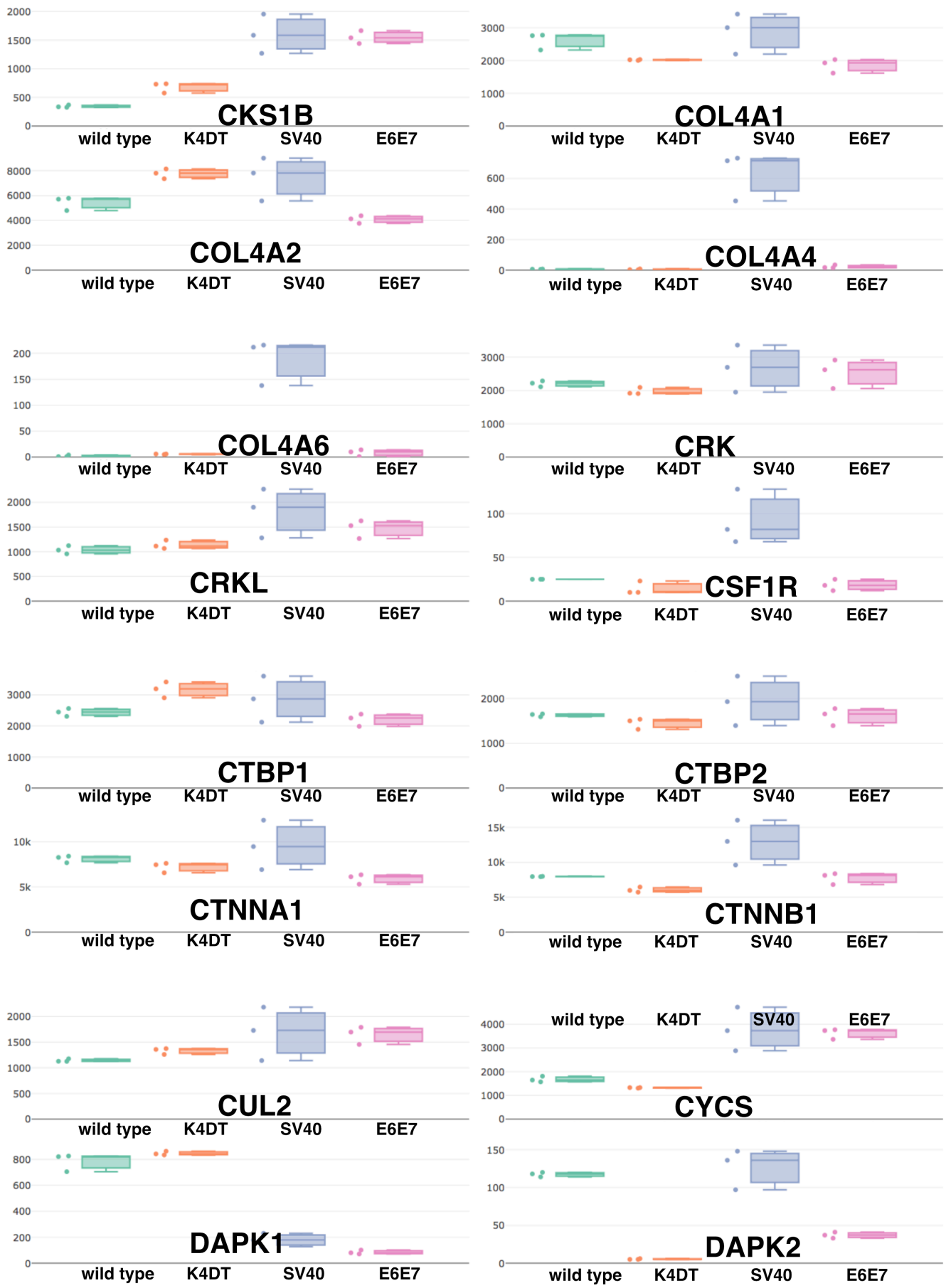

CKS1B
COL4A1
COL4A2
COL4A4
COL4A6
CRK
CRKL
CSF1R
CTBP1
CTBP2
CTNNA1
CTNNB1
CUL2
CYCS
DAPK1
DAPK2
wild type
K4DT
SV40
E6E7

Fig.S18

Fig. S18 Raw expression counts of pathway in cancer related genes listed in pathway in cancer of KEGG. The expression counts are listed with box bar plots. The expression counts of CKS1B, COL4A1, COL4A2, COL4A4, COL4A6, CRK, CRKL, CSF1R, CTBP1, CTBP2, CTNNA1, CTNNB1, CUL2, CUL2, DAPK1, DAPK2 were shown.

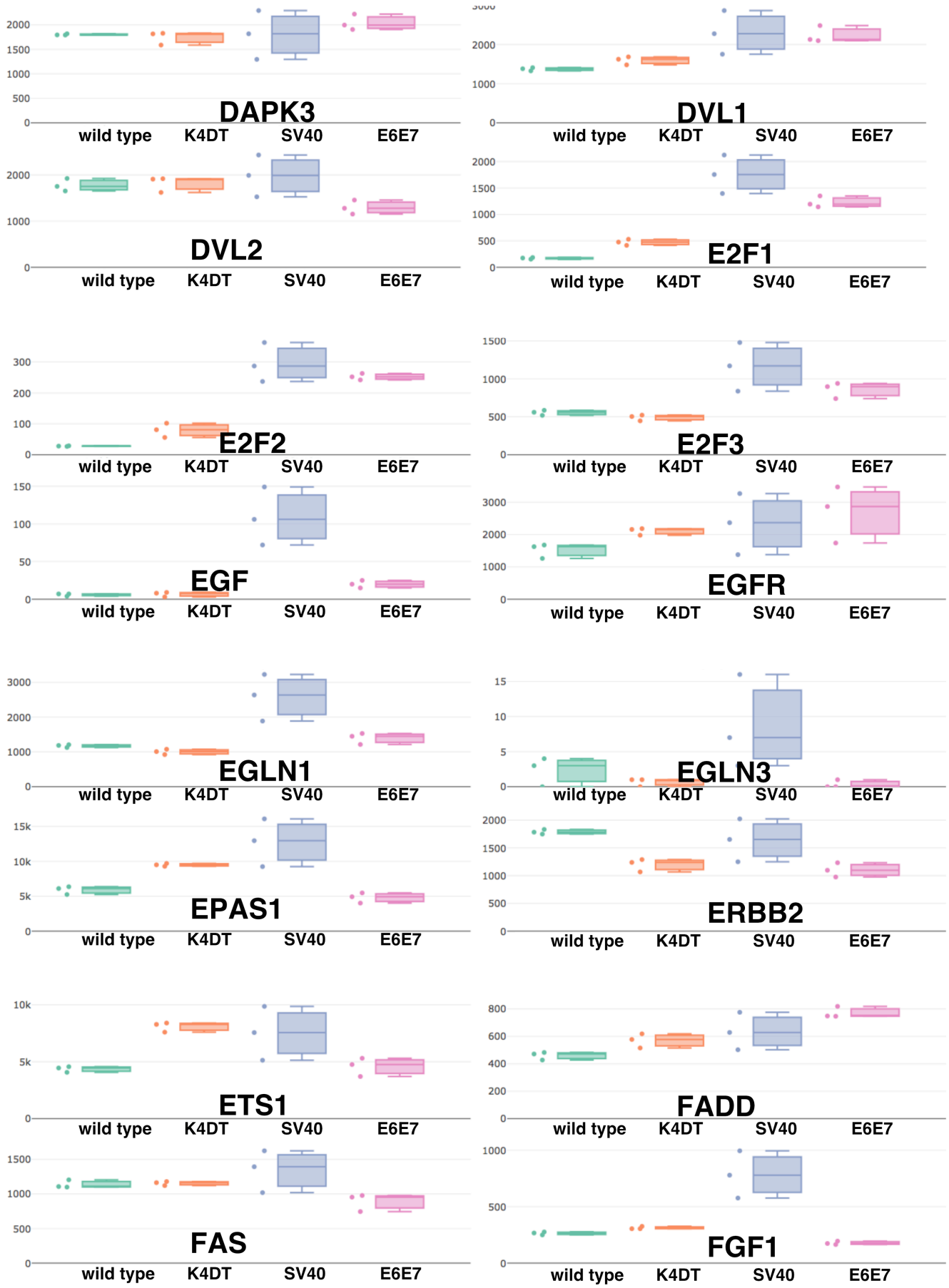

DAPK3
DVL1
DVL2
E2F1
E2F2
E2F3
EGF
EGFR
EGLN1
EGLN3
EPAS1
ERBB2
ETS1
FADD
FAS
FGF1
wild type
K4DT
SV40
E6E7

Fig.S19

Fig. S19 Raw expression counts of pathway in cancer related genes listed in pathway in cancer of KEGG. The expression counts are listed with box bar plots. The expression counts of DAPK2, DVL1, DVL1, DVL2, E2F1, E2F2, E2F3, EGF, EGFR, EGLN1, EGLN3, EPAS1, ERBB2, ETS1, FADD, FAS, FGF1 were shown.

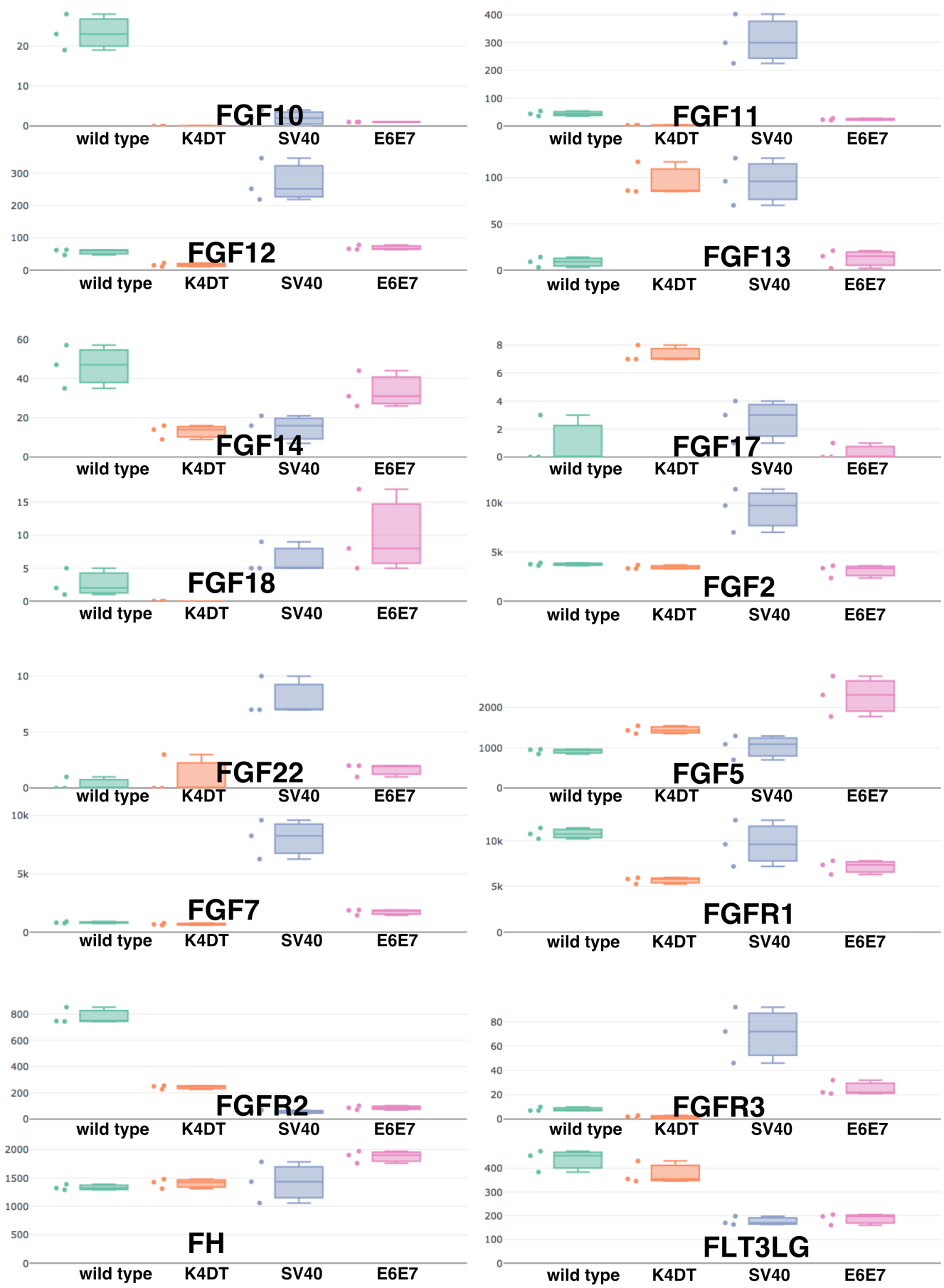

FGF10
FGF11
FGF12
FGF13
FGF14
FGF17
FGF18
FGF2
FGF22
FGF5
FGF7
FGFR1
FGFR2
FGFR3
FH
FLT3LG
wild type
K4DT
SV40
E6E7

Fig.S20

Fig. S20 Raw expression counts of pathway in cancer related genes listed in pathway in cancer of KEGG. The expression counts are listed with box bar plots. The expression counts of FGF10, FGF11, FGF12, FGF13, FGF14, FGF17, FGF18, FGF2, FGF22, FGF5, FGF7, FGFR1, FGFR2, FGFR3, FH, FLT3LG were shown.

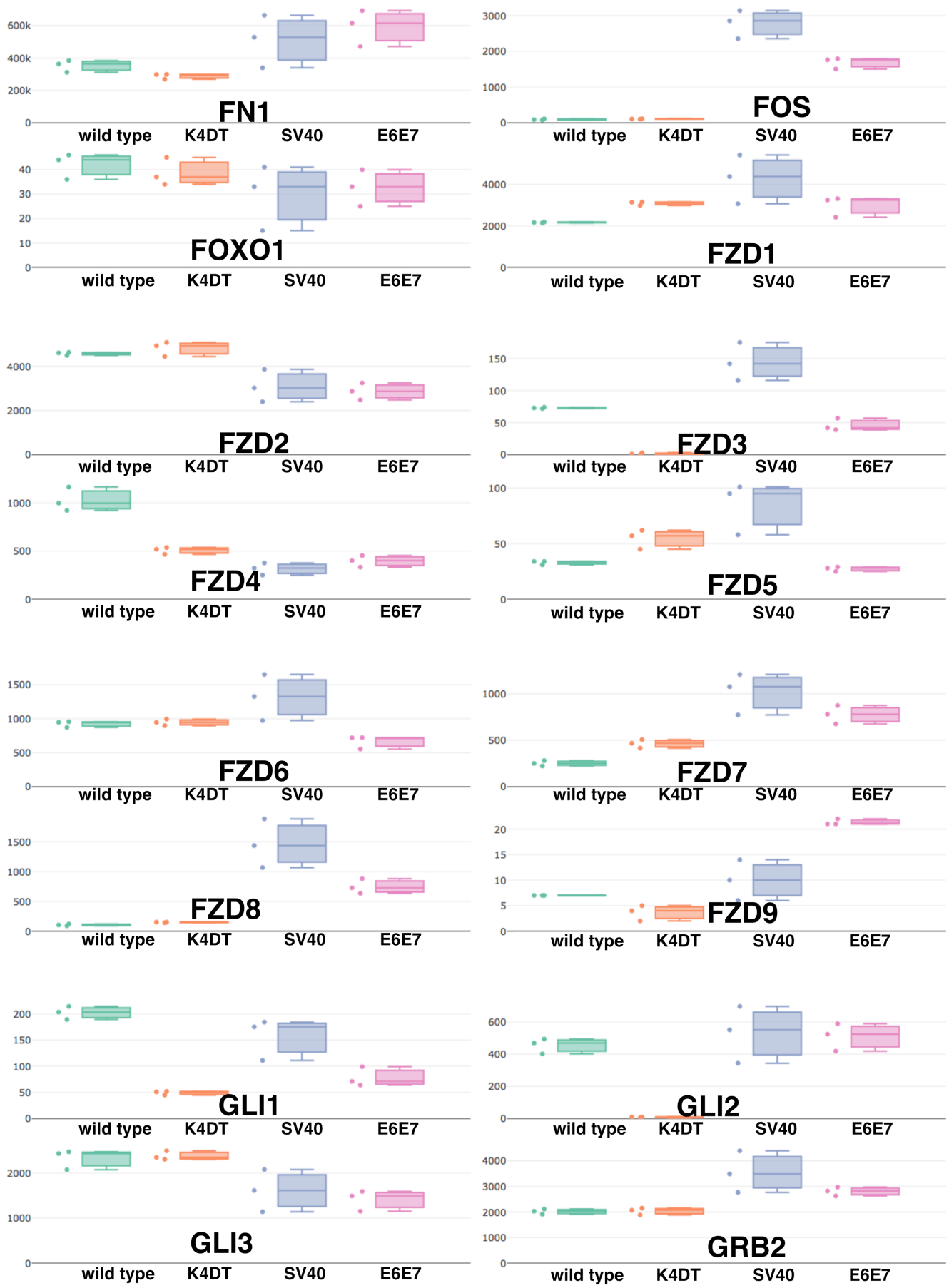

FN1
FOXO1
FOS
FZD1
FZD2
FZD3
FZD4
FZD5
FZD6
FZD7
FZD8
FZD9
GLI1
GLI2
GLI3
GRB2
wild type
K4DT
SV40
E6E7

Fig.S21

Fig. S21 Raw expression counts of pathway in cancer related genes listed in pathway in cancer of KEGG. The expression counts are listed with box bar plots. The expression counts of FN1, FOS, FOXO1, FZD1, FZD2, FZD3, FZD4, FZD5, FZD6, FZD7, FZD8, FZD9, GLI1, GLI2, GLI3, GRB2 were shown

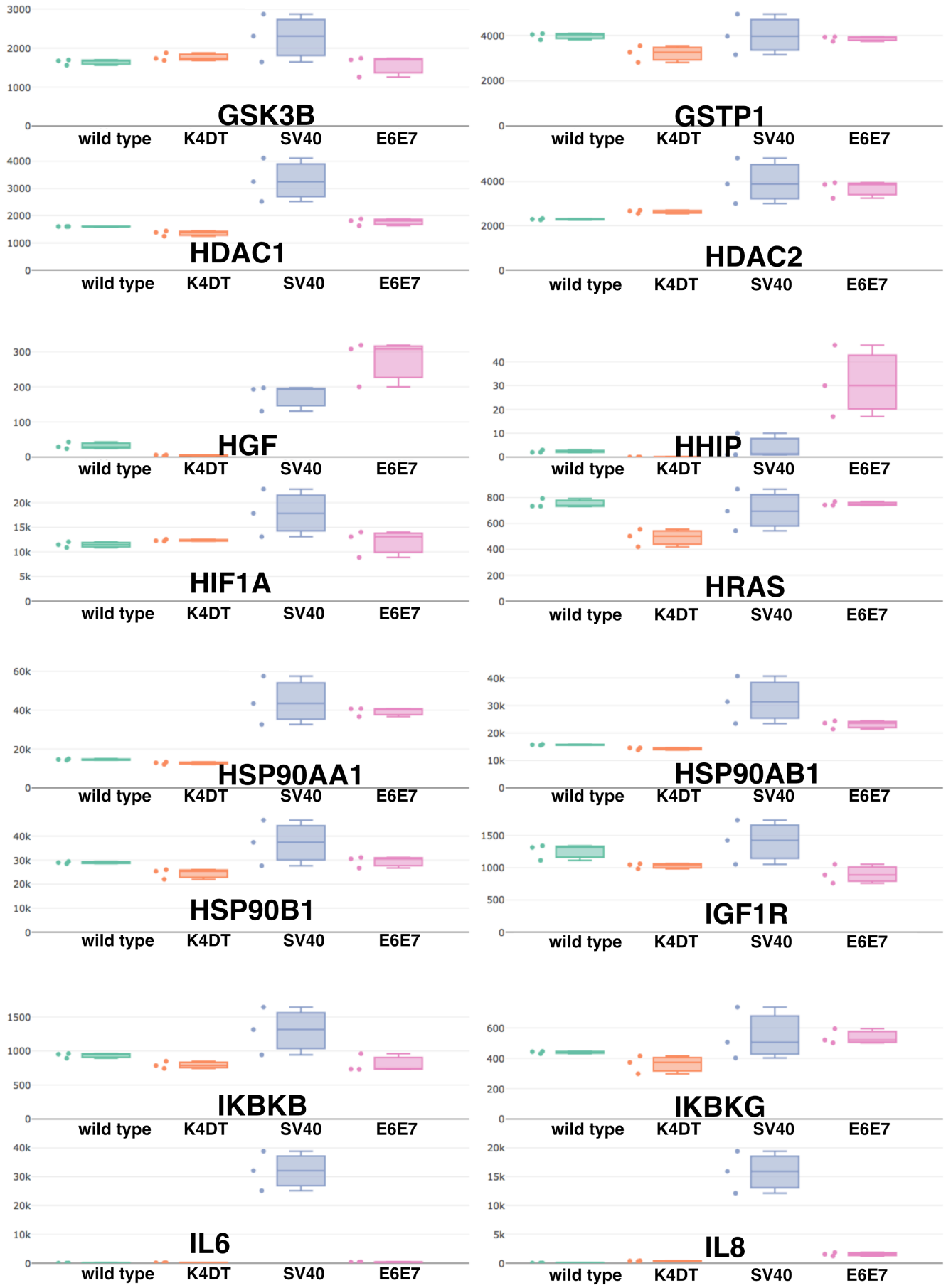

GSK3B
GSTP1
HDAC1
HDAC2
HGF
HHIP
HIF1A
HRAS
HSP90AA1
HSP90AB1
HSP90B1
IGF1R
IKBKB
IKBKG
IL6
IL8
wild type
K4DT
SV40
E6E7

Fig.S22

Fig. S22 Raw expression counts of pathway in cancer related genes listed in pathway in cancer of KEGG. The expression counts are listed with box bar plots. The expression counts of GSK3B, GSTP1, HDAC1, HDAC2, HGF, HHIP, HIF1A, HRAS, HSP90AA1, HSP90AB1, HSP90B1, IGF1R, IKBKB, IKBKG, IL6, IL8 were shown.

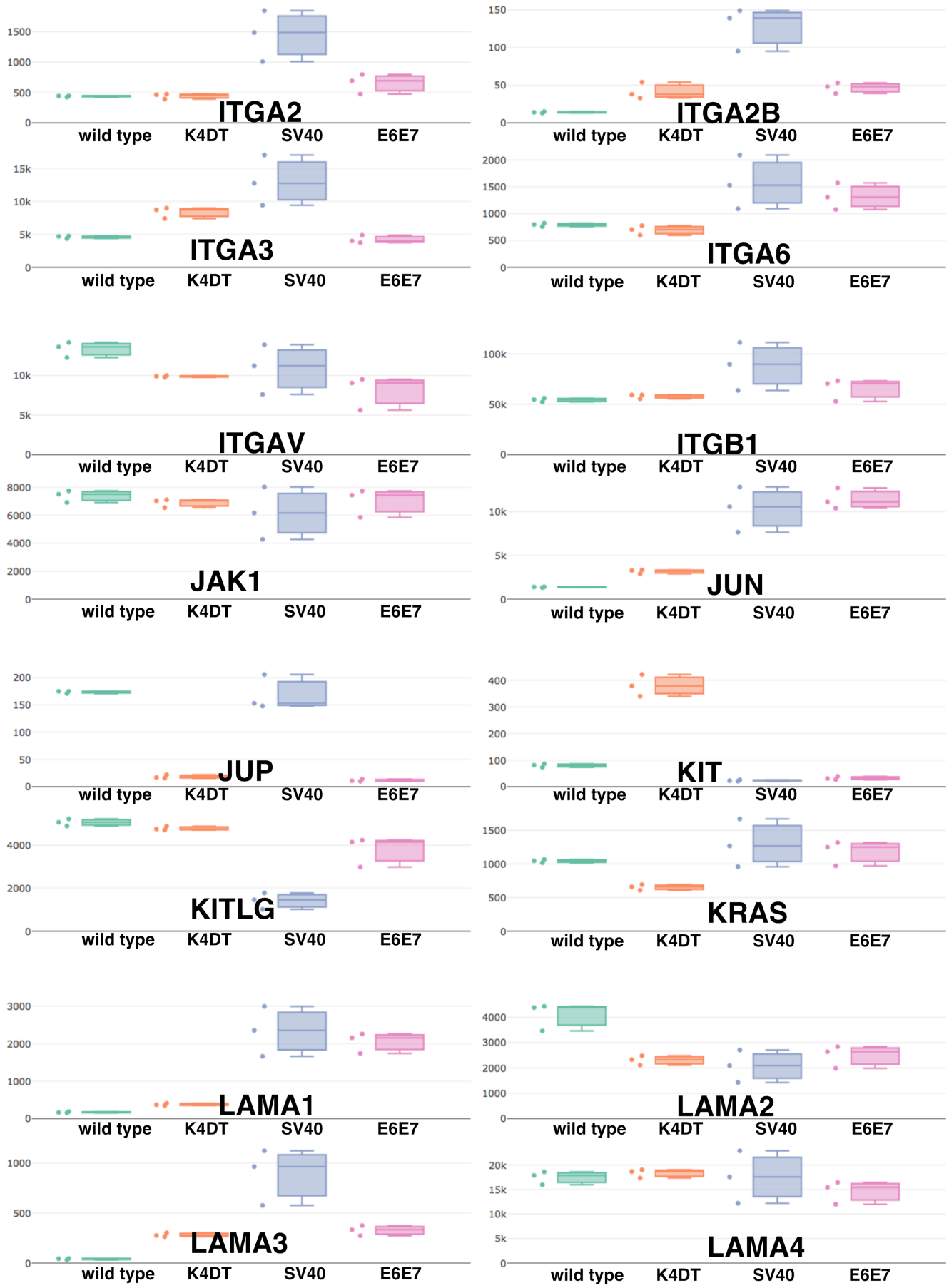

ITGA2
ITGA2B
ITGA3
ITGA6
ITGAV
ITGB1
JAK1
JUN
JUP
KIT
KITLG
KRAS
LAMA1
LAMA2
LAMA3
LAMA4
wild type
K4DT
SV40
E6E7

Fig.S23

Fig. S23 Raw expression counts of pathway in cancer related genes listed in pathway in cancer of KEGG. The expression counts are listed with box bar plots. The expression counts of ITGA2, ITGA2B, ITGA3, ITGA6, ITGAV, ITGB1, JAK1, JUN, JUP, KIT, KITLG, KRAS, LAMA1, LAMA2, LAMA3, LAMA4 were shown.

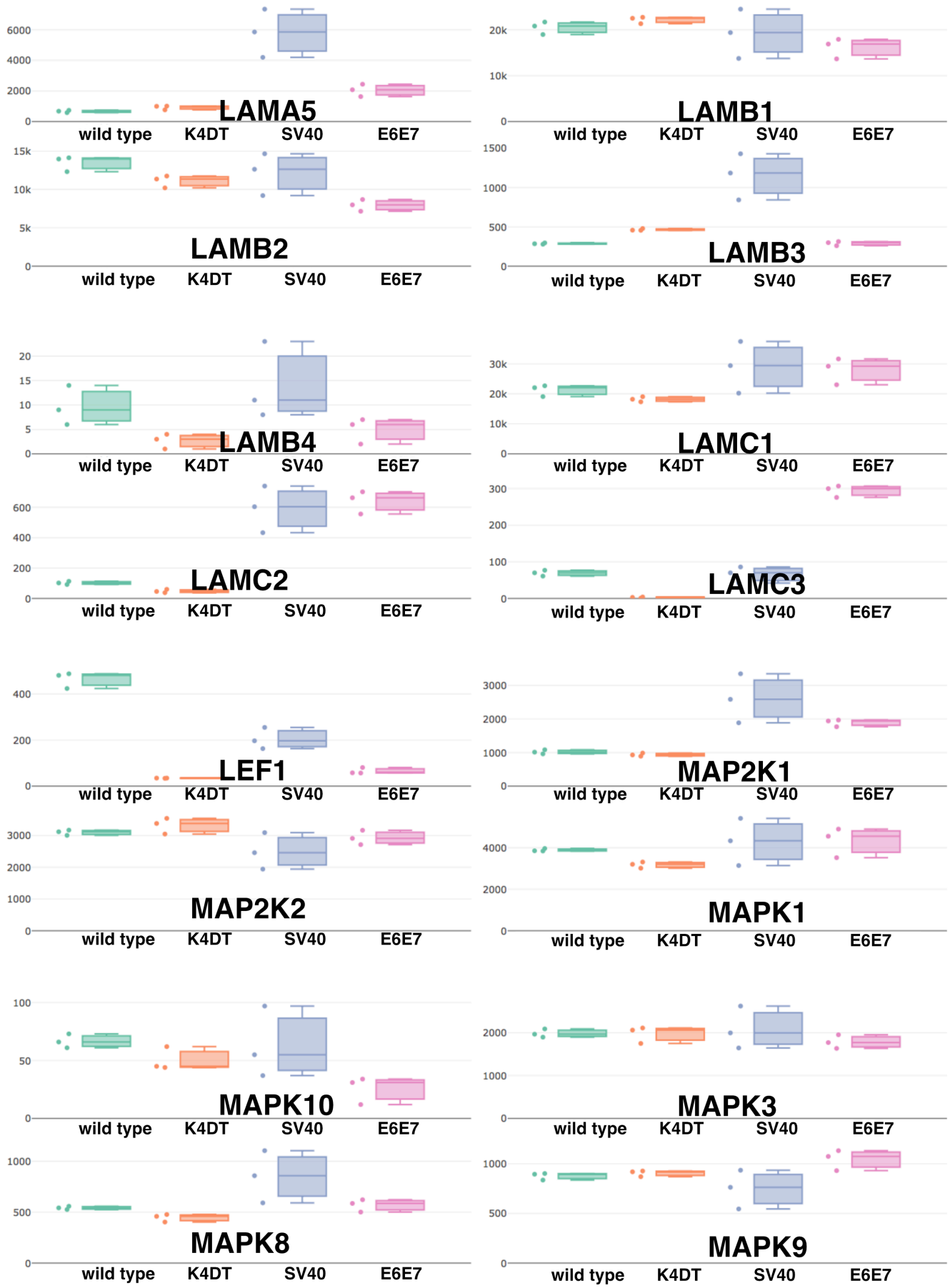

LAMA5
LAMB1
LAMB2
LAMB3
LAMB4
LAMC1
LAMC2
LAMC3
LEF1
MAP2K1
MAP2K2
MAPK1
MAPK10
MAPK3
MAPK8
MAPK9
wild type
K4DT
SV40
E6E7

Fig.S24

Fig. S24 Raw expression counts of pathway in cancer related genes listed in pathway in cancer of KEGG. The expression counts are listed with box bar plots. The expression counts of LAMA5, LAMB1, LAMB2, LAMB3, LAMB4, LAMC1, LAMC2, LAMC3, LEF1, MAP2K1, MAP2K2, MAPK1, MAPK10, MAPK3, MAPK8, MAPK9 were shown.

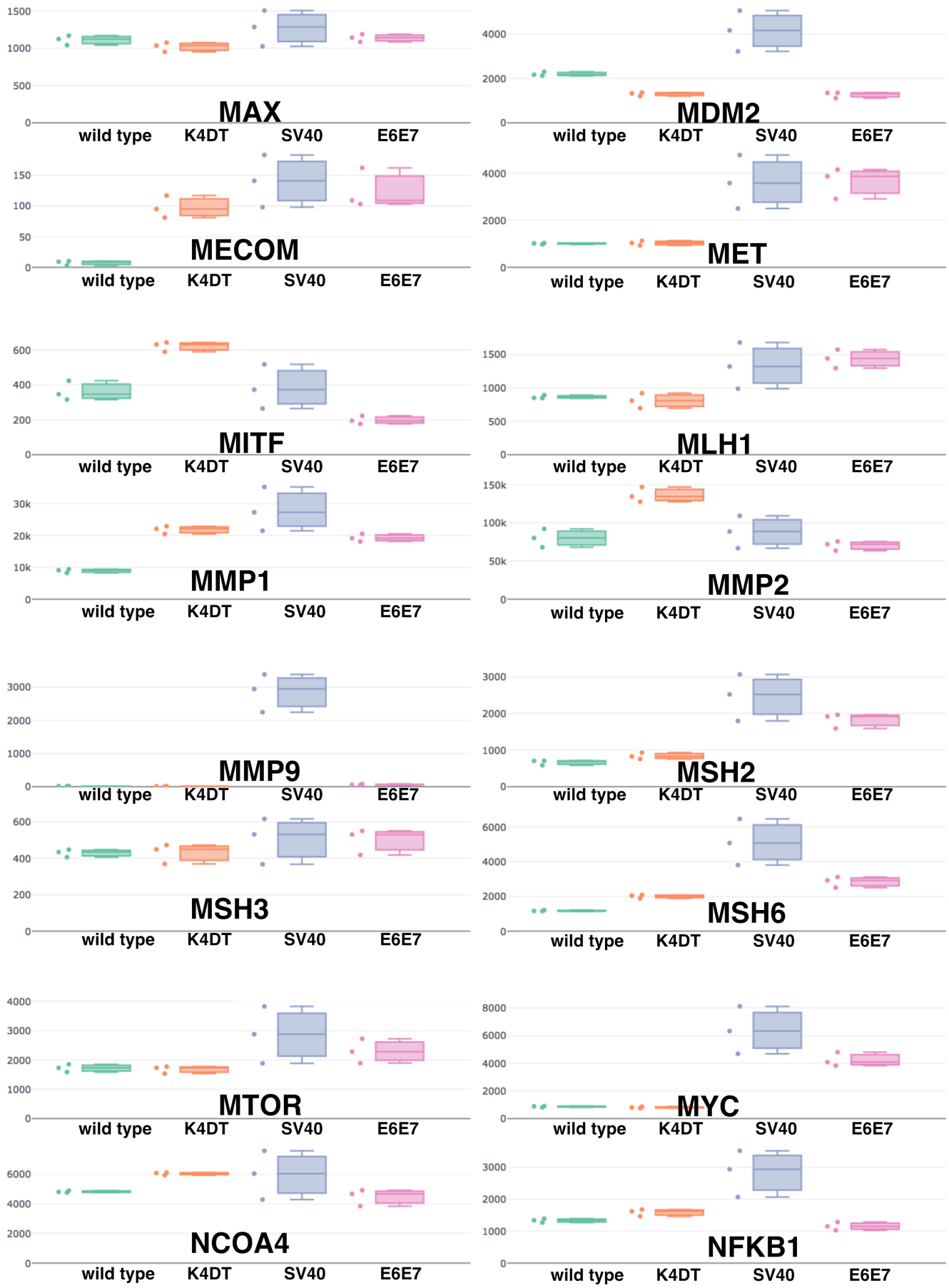

MAX
MDM2
MECOM
MET
MITF
MLH1
MMP1
MMP2
MMP9
MSH2
MSH3
MSH6
MTOR
MYC
NCOA4
NFKB1
wild type
K4DT
SV40
E6E7

Fig.S25

Fig. S25 Raw expression counts of pathway in cancer related genes listed in pathway in cancer of KEGG. The expression counts are listed with box bar plots. The expression counts of MAX, MDM2, MECOM, MET, MITF, MLH1, MMP1, MMP2, MMP9, MSH2, MSH3, MSH6, MTOR, MYC, NCOA4, NFKB1 were shown.

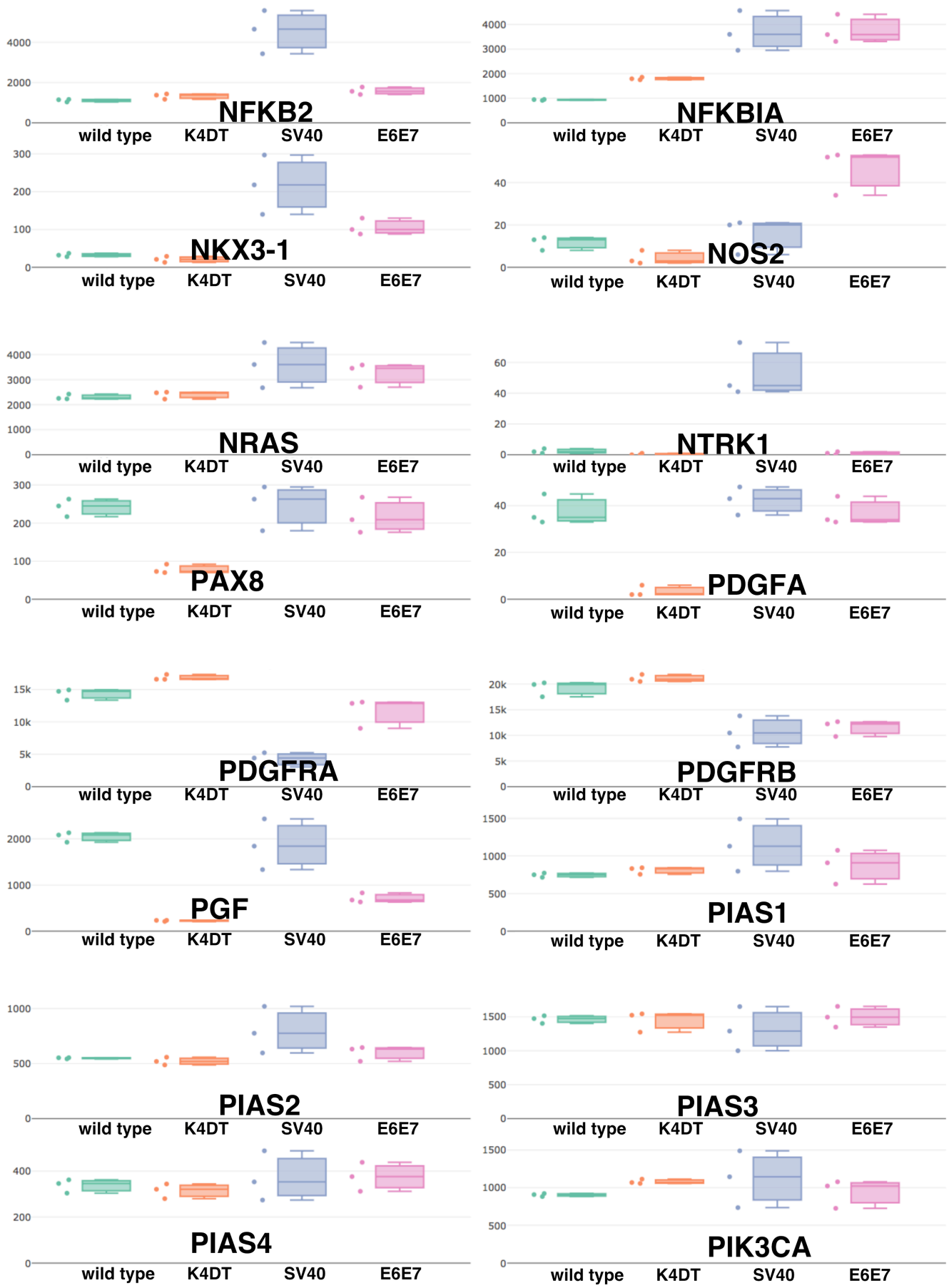

NFKB2
NFKBIA
NKX3-1
NOS2
NRAS
NTRK1
PAX8
PDGFA
PDGFRA
PDGFRB
PGF
PIAS1
PIAS2
PIAS3
PIAS4
PIK3CA
wild type
K4DT
SV40
E6E7

Fig.S26

Fig. S26 Raw expression counts of pathway in cancer related genes listed in pathway in cancer of KEGG. The expression counts are listed with box bar plots. The expression counts of NFKB2, NFKBIA, NKX3-1, NOS2, NRAS, NTRK1, PAX8, PDGFA, PDGFRA, PDGFRB, PGF, PIAS1, PIAS2, PIAS3, PIAS4, PIK3CA were shown.

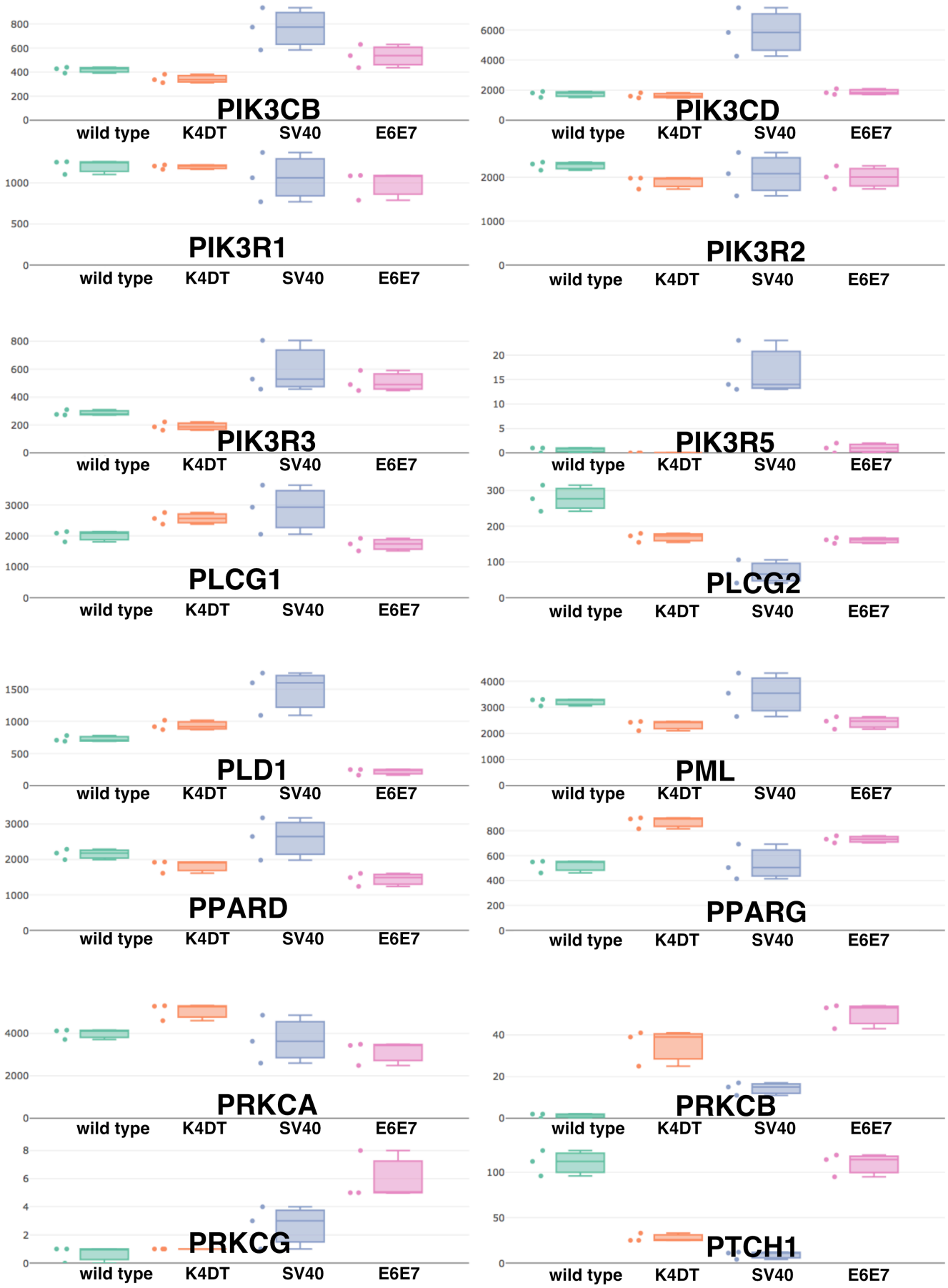

PIK3CB
PIK3CD
PIK3R1
PIK3R2
PIK3R3
PIK3R5
PLCG1
PLCG2
PLD1
PML
PPARD
PPARG
PRKCA
PRKCB
PRKCG
PTCH1
wild type
K4DT
SV40
E6E7

Fig.S27

Fig. S27 Raw expression counts of pathway in cancer related genes listed in pathway in cancer of KEGG. The expression counts are listed with box bar plots. The expression counts of PIK3CB, PIK3CD, PIK3R1, PIK3R2, PIK3R3, PIK3R5, PLCG1, PLCG2, PLD1, PML, PPARD, PPARG, PRKCA, PRKCB, PRKCG, PTHCH1 were shown.

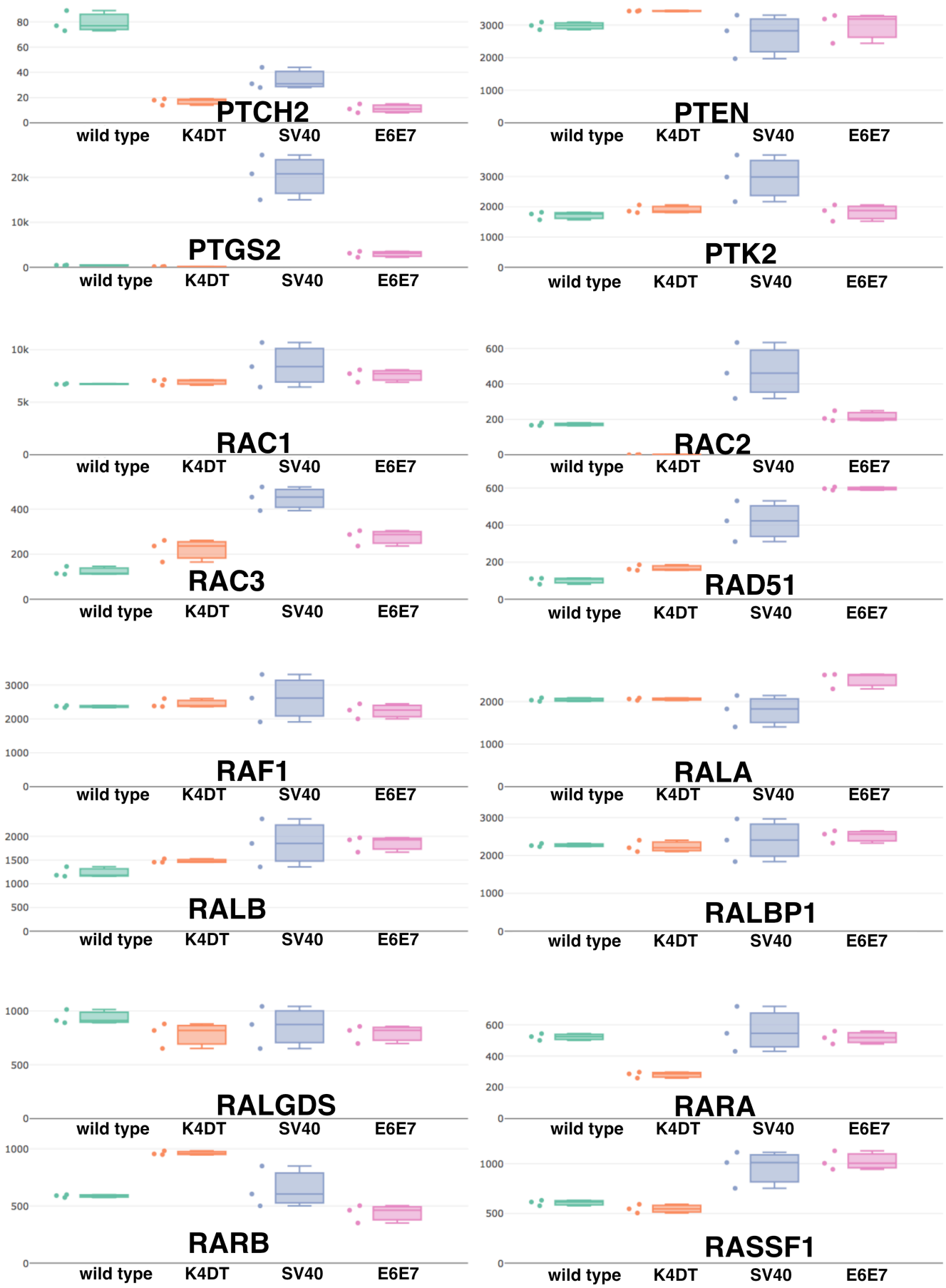

PTCH2
PTEN
PTGS2
PTK2
RAC1
RAC2
RAC3
RAD51
RAF1
RALA
RALB
RALBP1
RALGDS
RARA
RARB
RASSF1
wild type
K4DT
SV40
E6E7

Fig.S28

Fig. S28 Raw expression counts of pathway in cancer related genes listed in pathway in cancer of KEGG. The expression counts are listed with box bar plots. The expression counts of PTCH2, PTEN, PTGS2, PTK2, RAC1, RAC2, RAC3, RAD51, RAF1, RALA, RALB, RALBP1, RALGDS, RARA, RARB, RASSF1 were shown.

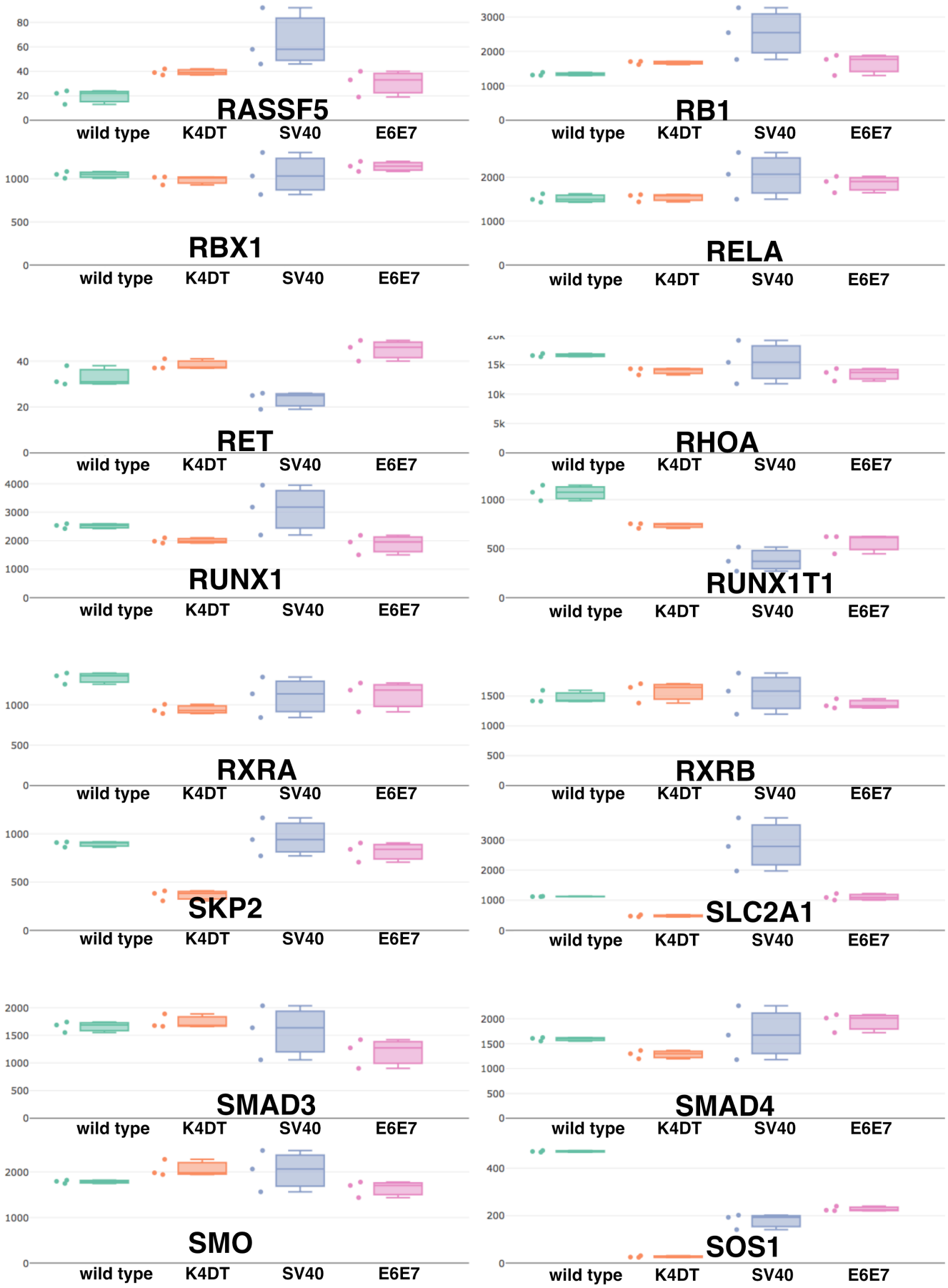

RASSF5
RB1
RBX1
RELA
RET
RHOA
RUNX1
RUNX1T1
RXRA
RXRB
SKP2
SLC2A1
SMAD3
SMAD4
SMO
SOS1
wild type
K4DT
SV40
E6E7

Fig.S29

Fig. S29 Raw expression counts of pathway in cancer related genes listed in pathway in cancer of KEGG. The expression counts are listed with box bar plots. The expression counts of RASSF5, RB1, RBX1, RELA, RET, RHOA, RUNX1, RUNX1T1, RXRA, RXRB, SKP2, SLC2A1, SMAD3, SMAD4, SMO, SOS1 were shown.


SOS2
STAT1
STAT3
STAT5A
STAT5B
STK36
STK4
TCEB1
TCEB2
TCF7
TCF7L1
TCF7L2
TFG
TGFB1
TGFB2
TGFB3
wild type
K4DT
SV40
E6E7

Fig.S30

Fig. S30 Raw expression counts of pathway in cancer related genes listed in pathway in cancer of KEGG. The expression counts are listed with box bar plots. The expression counts of SOS2, STAT1, STAT3, STAT5A, STAT5B, STK36, STK4, TCEB1, TCEB2, TCF7, TCF7L1, TCFL2, TFG, TGFB1, TGFB2, TGFB3 were shown.

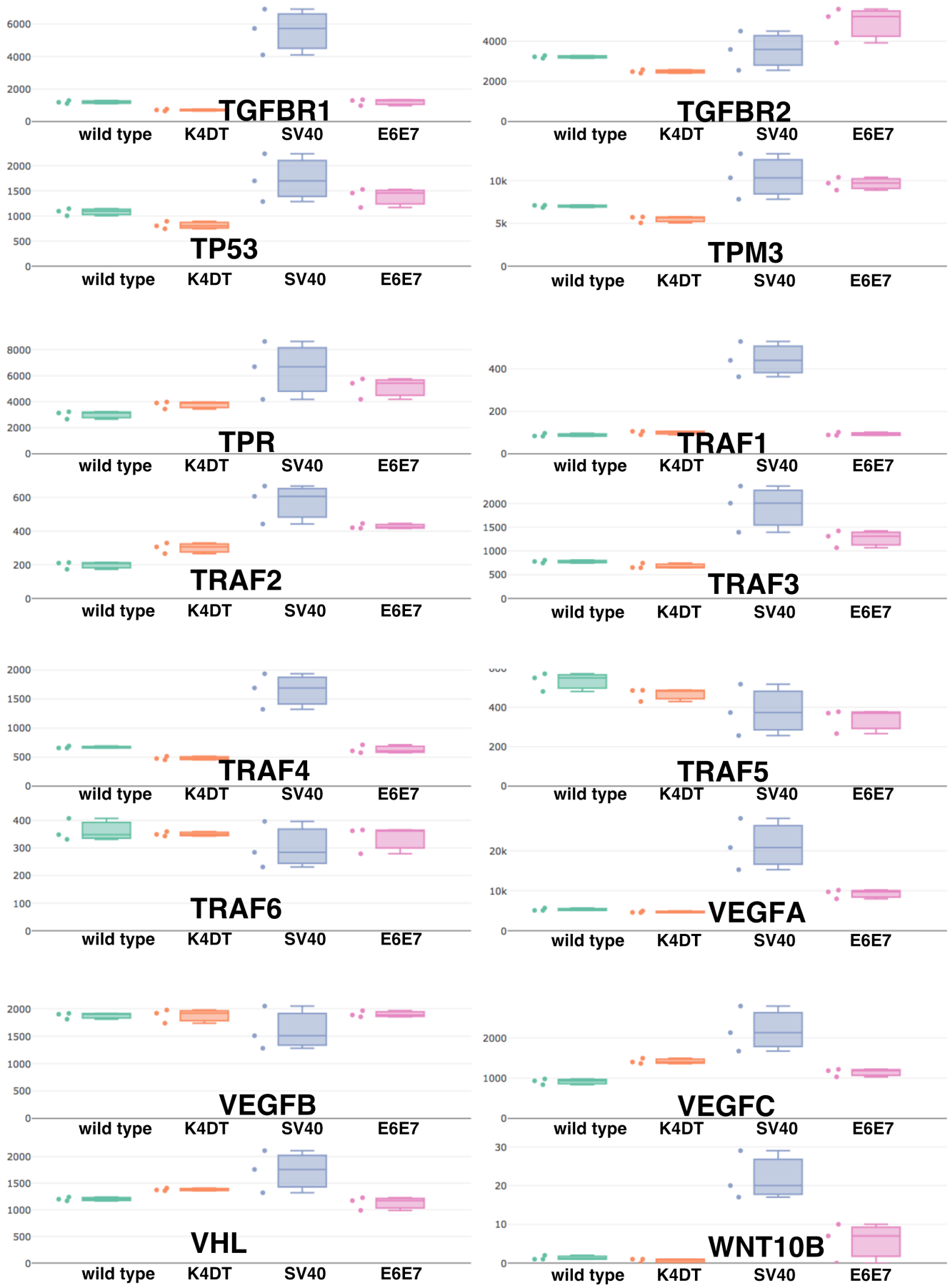

TGFBR1
TGFBR2
TP53
TPM3
TPR
TRAF1
TRAF2
TRAF3
TRAF4
TRAF5
TRAF6
VEGFA
VEGFB
VEGFC
VHL
WNT10B
wild type
K4DT
SV40
E6E7

Fig.S31

Fig. S31 Raw expression counts of pathway in cancer related genes listed in pathway in cancer of KEGG. The expression counts are listed with box bar plots. The expression counts of TGFBR1, TGFBR2, TP53, TPM3, TPR, TRAF1, TRAF2, TRAF3, TARAF4, TRAF5, TRAF6, VEGFA, VEGFB, VEGFC, VHL, WNT10B were shown.


WNT11
WNT16
WNT2
WNT2B
WNT3
WNT5A
WNT5B
WNT7B
WNT9A
XIAP
wild type
K4DT
SV40
E6E7

Fig.S32

Fig. S32 Raw expression counts of pathway in cancer related genes listed in pathway in cancer of KEGG. The expression counts are listed with box bar plots. The expression counts of WNT11, WNT16, WNT2, WNT2B, WNT3, WNT5A, WNT5B, WNT7B, WINT9A, XIAP were shown.

WT

K4DT

SV40

E6E7

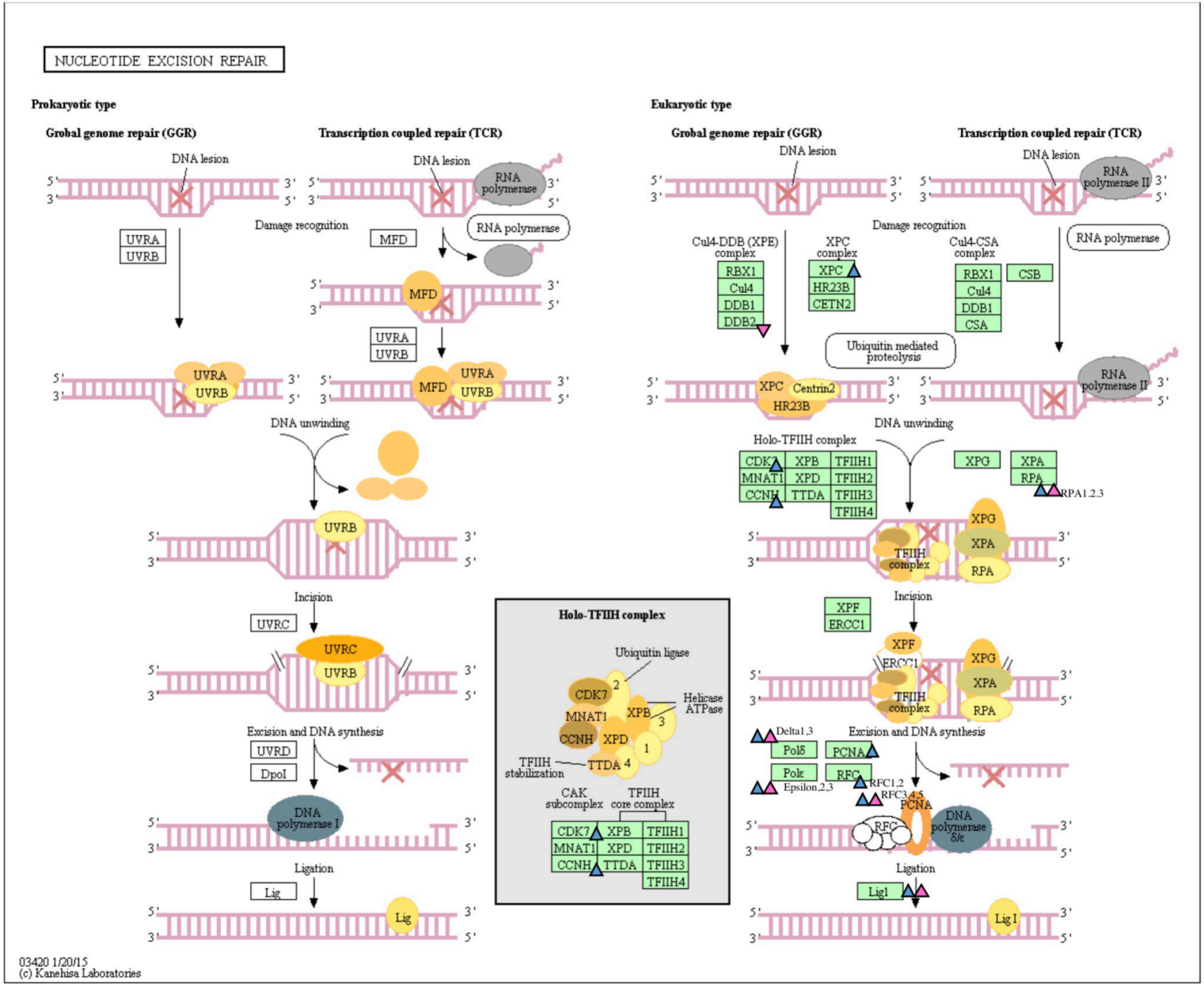


Fig.S33

Fig. S33 Summary of upregulated and downregulated genes among KEGG pathway of nucleotide excision repair. Arrowheads indicate upregulated or downregulated genes in K4DT, E6/E7, or SV40 cells. The genes with more than 200 counts and X2 increase or 1/2 decreased from wild type were mapped.

Table S1. Analysis of chromosome number of immortalized dermal papilla cells with K4DT, SV40T, E6E7 method.

Cell	Chromosome number												
	44	45	46	47	71	72	74	75	76	77	78	79	80
K4DT	0	0	50	0	0	0	0	0	0	0	0	0	0
SV40	0	0	0	0	2	1	2	3	9	17	10	2	1
E6E7	1	4	45	0	0	0	0	0	0	0	0	0	0

Number of samples is 50.

Table S2. G-banding analysis of immortalized dermal papilla cells with K4DT, SV40T, E6E7 method.

Cell	Normal G-band	Abnormalities	Normal ratio
K4DT	19	1, Translocation of 14 to 15	95%
SV40	0	20, Tetraploid, addition or loss of multiple chromosomes	0%
E6E7	16	4, deletion or addition of chromosome	80%

Number of samples is 20.