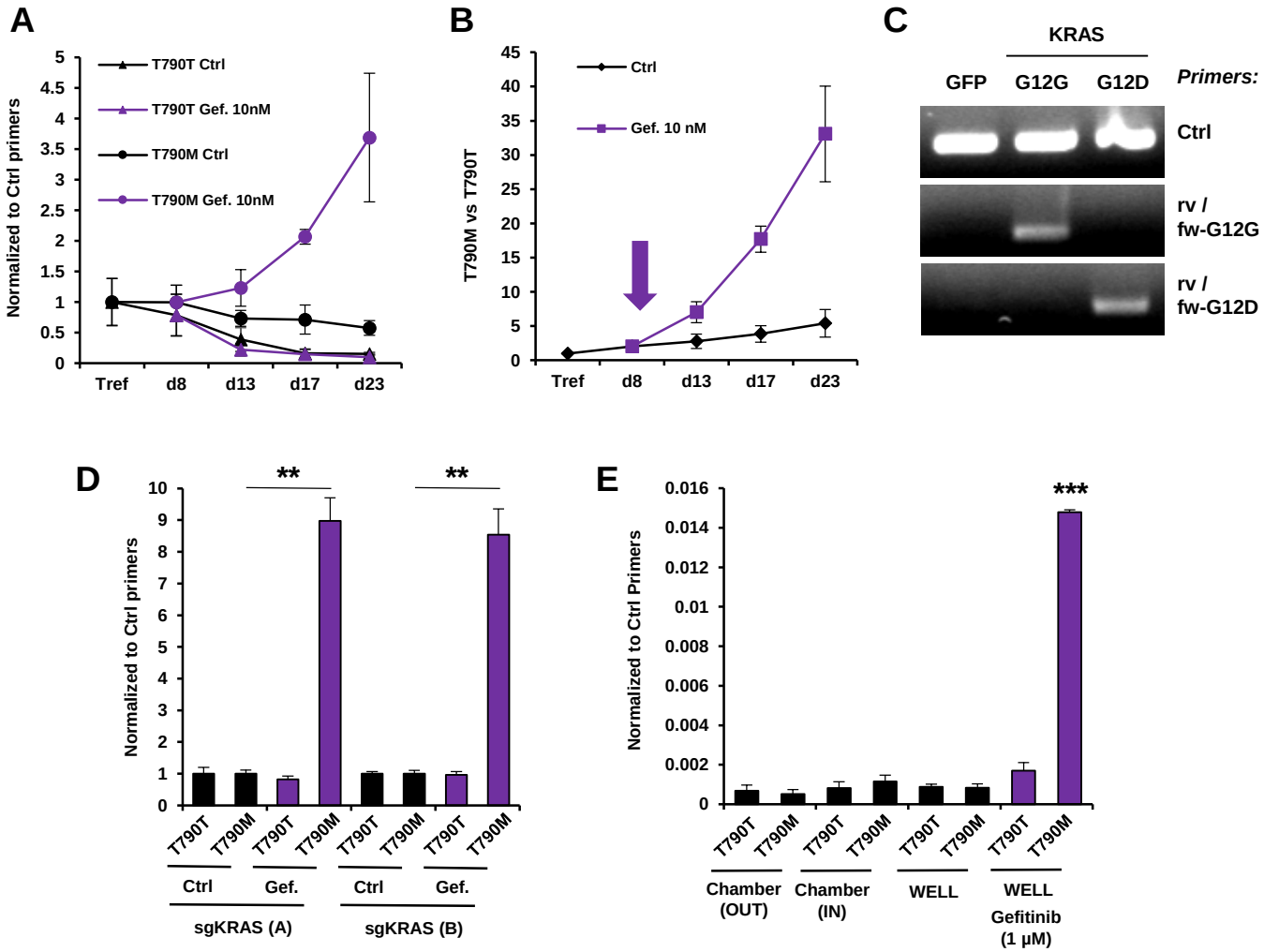


Figure S1, related to Figure 1. CRISPR-Barcoding to Investigate Cancer Resistance to Chemotherapy

(A) The data from Figure 1C are expressed as relative proportion of each barcode normalized to the amount of total DNA measured using qPCR primers amplifying an unmodified region of the EGFR gene.

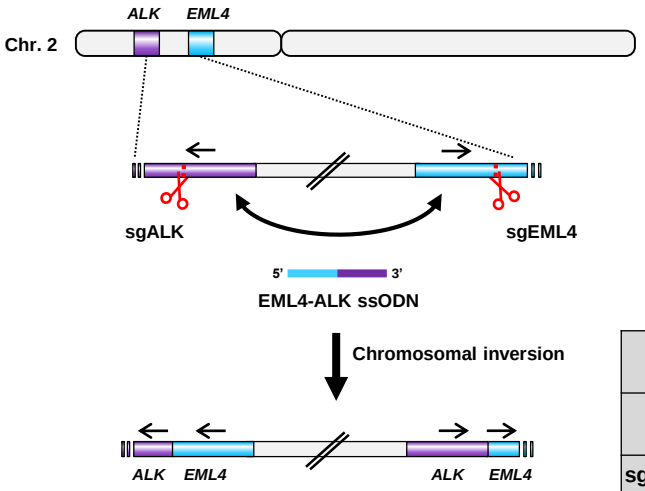
(B) Nine days after transfection for EGFR CRISPR-barcoding using sgEGFR(B), PC9 cells were treated in the presence or the absence of gefitinib (10 nM, arrow). qPCR was performed from gDNA extracted at the indicate time points. The results correspond to the mean values of the T790M- to T790T-EGFR barcode ratio (± SEM; n=3) of one representative of three independent experiments.

(C) PCR using control or barcode-specific primers was performed from gDNA extracted from 293T cells transfected with a control GFP plasmid or the sgKRAS vector with the KRAS-G12G or the KRAS-G12D ssODN.

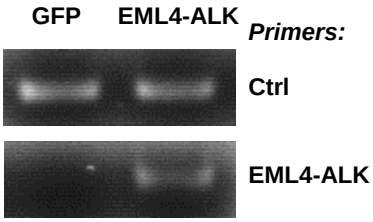
(D)The data from Figure 1F are expressed as relative proportion of each barcode normalized to the amount of total DNA measured using qPCR primers amplifying an unmodified region of the KRAS gene.

(E) Untreated PC9 cells were used in a CRISPR-barcoding invasion assay as described in Figure 4F. qPCR analysis for the EGFR-T790T and -T790M barcodes was performed from cells outside (OUT) or inside (IN) the Boyden chamber, or seeded in parallel in a regular well (WELL). PC9 cells plated in parallel and treated for 3 days with gefitinib (0.5 µM) were used as a control. The mean values (± SEM; n=3) of one representative of two independent experiments were normalized to the total amount of gDNA measured using qPCR primers amplifying an unmodified region of the EGFR gene (Ctrl primers). *** $p<0.001$ compared to untreated cells (Student's t test).

A



B



C

	Ctrl Primers (Ct)	EML4-Alk Primers (Ct)
GFP	19.42	undetermined
sgALK + sgEML4	19.64	26.76
sgALK + sgEML4 + EML4-ALK ssODN	19.61	24.43

D

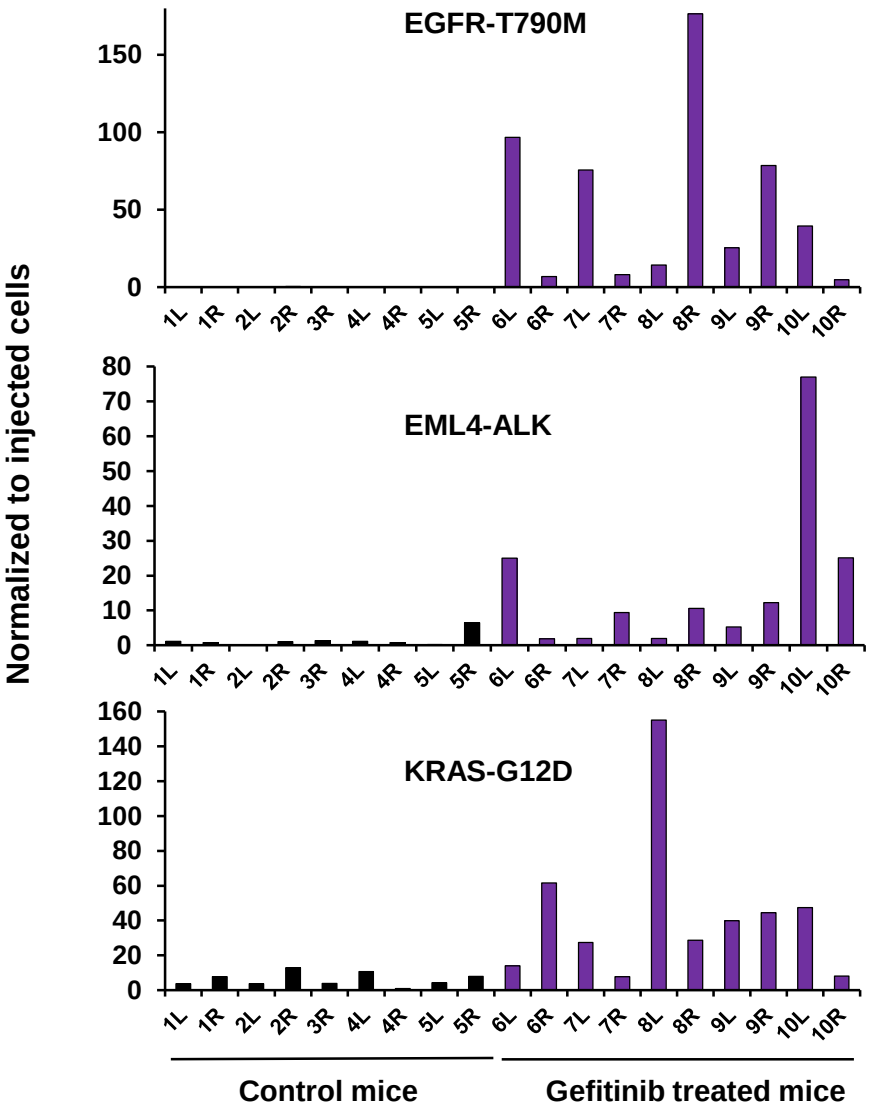


Figure S2, related to Figures 2 and 3. Multiplex Model of NSCLC Resistance to Gefitinib

(A) Schematic representation of the strategy to induce CRISPR-mediated rearrangement of human chromosome 2. The paracentric inversion of ALK and EML4 was induced using two sgRNAs specific for the ALK and EML4 genes and a ssODN corresponding to the EML4-ALK fusion sequence. The direction of ALK and EML4 coding sequences is indicated (arrows).

(B) The specificity of the qPCR primers for the EML4-ALK fusion was assessed using gDNA from 293T cells.

(C) The efficiency of CRISPR-mediated inversion on chromosome 2 in 293T cells was assessed by qPCR in the presence or the absence of EML4-ALK-ssODN. Co-transfection of the donor DNA allowed an increase of efficiency of about 2 Cts.

(D) The data from Figure 3B are represented as three different diagrams corresponding to each mutation.

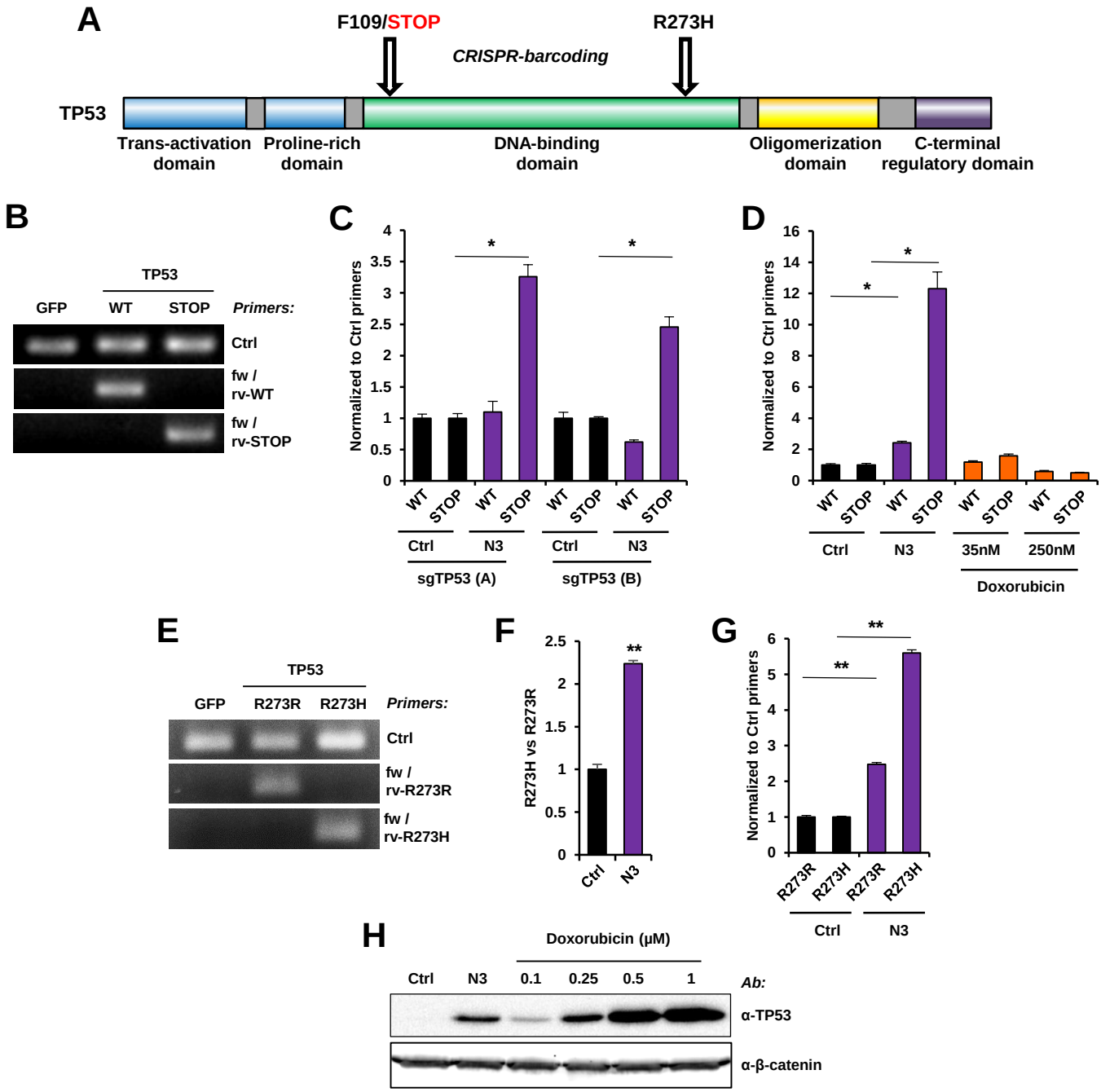


Figure S3, related to Figure 4. Inactivation of the *TP53* Gene

(A) Schematic representation of the TP53 protein. CRISPR-barcoding was used to introduce either a STOP codon at position 109, at the beginning of the DNA-binding domain, or the R273H DNA contact mutation.

(B) The specificity of the qPCR primers for TP53 F109 barcodes was assessed using gDNA from 293T cells transfected with a control GFP plasmid or the CRISPR/Cas9 vector with the TP53-WT or the TP53-STOP ssODN.

(C-D) The data from Figure 4A-B are expressed as relative proportion of each barcode normalized to the amount of total DNA measured using qPCR primers amplifying an unmodified region of the TP53 gene. * $p<0.05$ (Mann-Whitney's test).

(E) The specificity of the TP53-R273R and TP53-R273H qPCR primers was assessed as in B.

(F) HCT-116 cells containing the TP53-R273 barcodes were treated for seven days in the presence or the absence of Nutlin 3 (N3, 10 μM) and the TP53-R273H to TP53-R273R ratio was measured by qPCR. Mean ± SEM; n=5 of one representative of three independent experiments. **$p<0.01$ (Mann-Whitney's test).

(G) The data from Figure S3F are expressed as relative proportion of each barcode normalized to the amount of total DNA measured using qPCR primers amplifying an unmodified region of the TP53 gene. **$p<0.01$ (Mann-Whitney's test).

(H) HCT-116 cells were treated overnight with DMSO (Ctrl), Nutlin-3 (10 μM; N3) or the indicated concentrations of doxorubicin, followed by western blot using anti-TP53 and anti-β-catenin (loading control) antibodies.

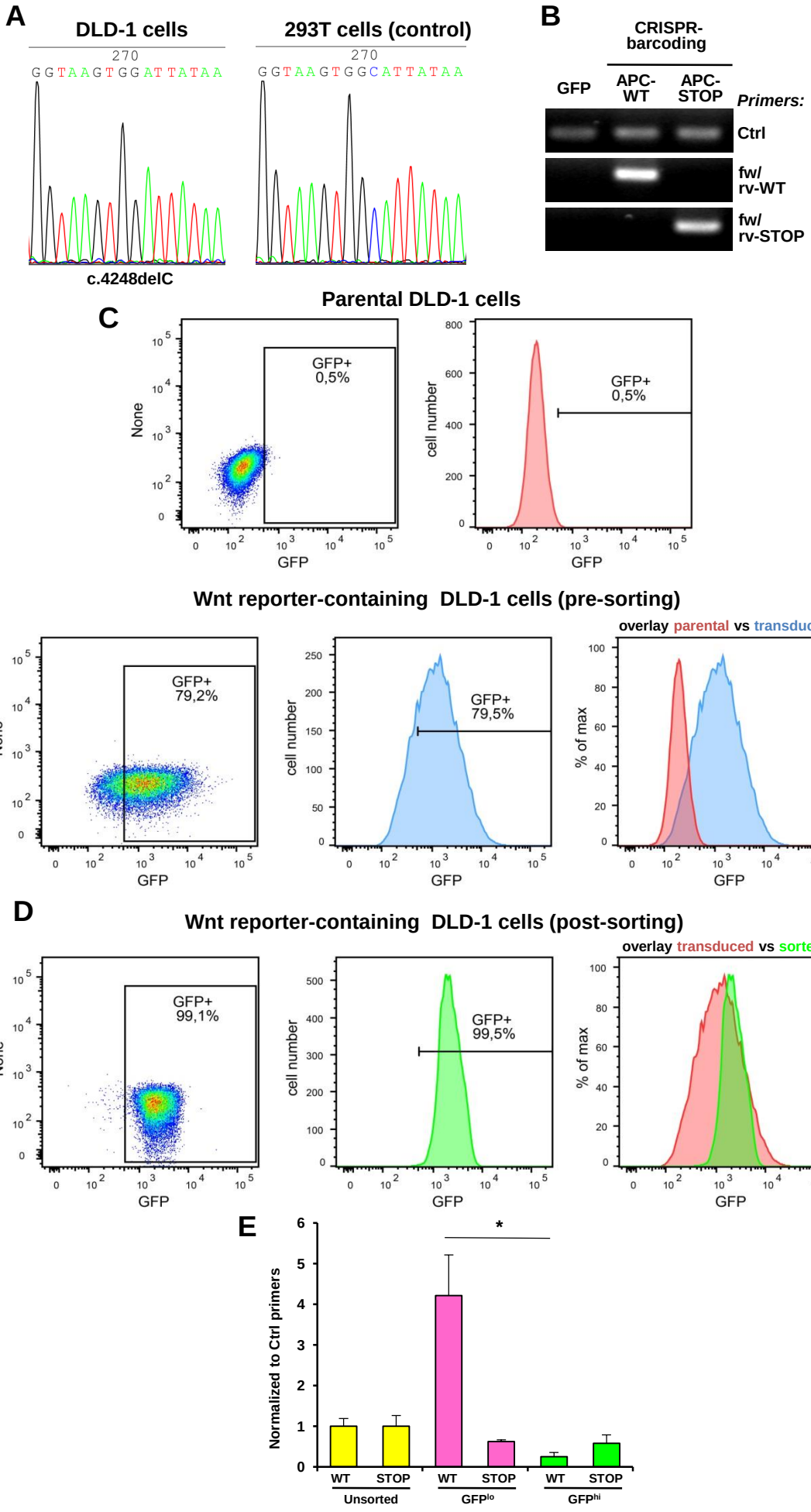

A
DLD-1 cells
293T cells (control)
270
GGTAAGTGGATTATAA
GGTAAGTGGCATTATAA
c.4248delC
B
CRISPR-barcoding
GFP
APC-WT
APC-STOP
Primers:
Ctrl
fw/ rv-WT
fw/ rv-STOP
C
Parental DLD-1 cells
None
GFP+ 0,5%
cell number
GFP
Wnt reporter-containing DLD-1 cells (pre-sorting)
GFP+ 79,2%
GFP+ 79,5%
overlay parental vs transduced
% of max
D
Wnt reporter-containing DLD-1 cells (post-sorting)
GFP+ 99,1%
GFP+ 99,5%
overlay transduced vs sorted
E
Normalized to Ctrl primers
*
WT
STOP
Unsorted
GFPlo
GFPhi

Figure S4, related to Figure 5. FACS Sorting of DLD-1 Cells Containing the Wnt Responsive Reporter

A) Sanger sequencing was performed to confirm the c.4248delC mutation on APC exon 15 in DLD-1 cells. gDNA from 293T cells was used as a control.

(B) The specificity of the qPCR primers for APC barcodes was assessed using gDNA from 293T cells transfected with a control GFP plasmid or the CRISPR/Cas9 vector with the APC-WT or the APC-STOP ssODN.

(C) Parental and Wnt reporter-containing DLD-1 cells were analyzed by flow cytometry before sorting.

(D) The cells in (C) were FACS sorted to isolate GFP^+ cells expressing the Wnt reporter. Flow cytometric analysis of the cells before and after FACS sorting is shown.

(E) The data from Figure 5B are expressed as relative proportion of each barcode normalized to the amount of total DNA measured using qPCR primers amplifying an unmodified region of the TP53 gene. *$p<0.05$ (Student's test).

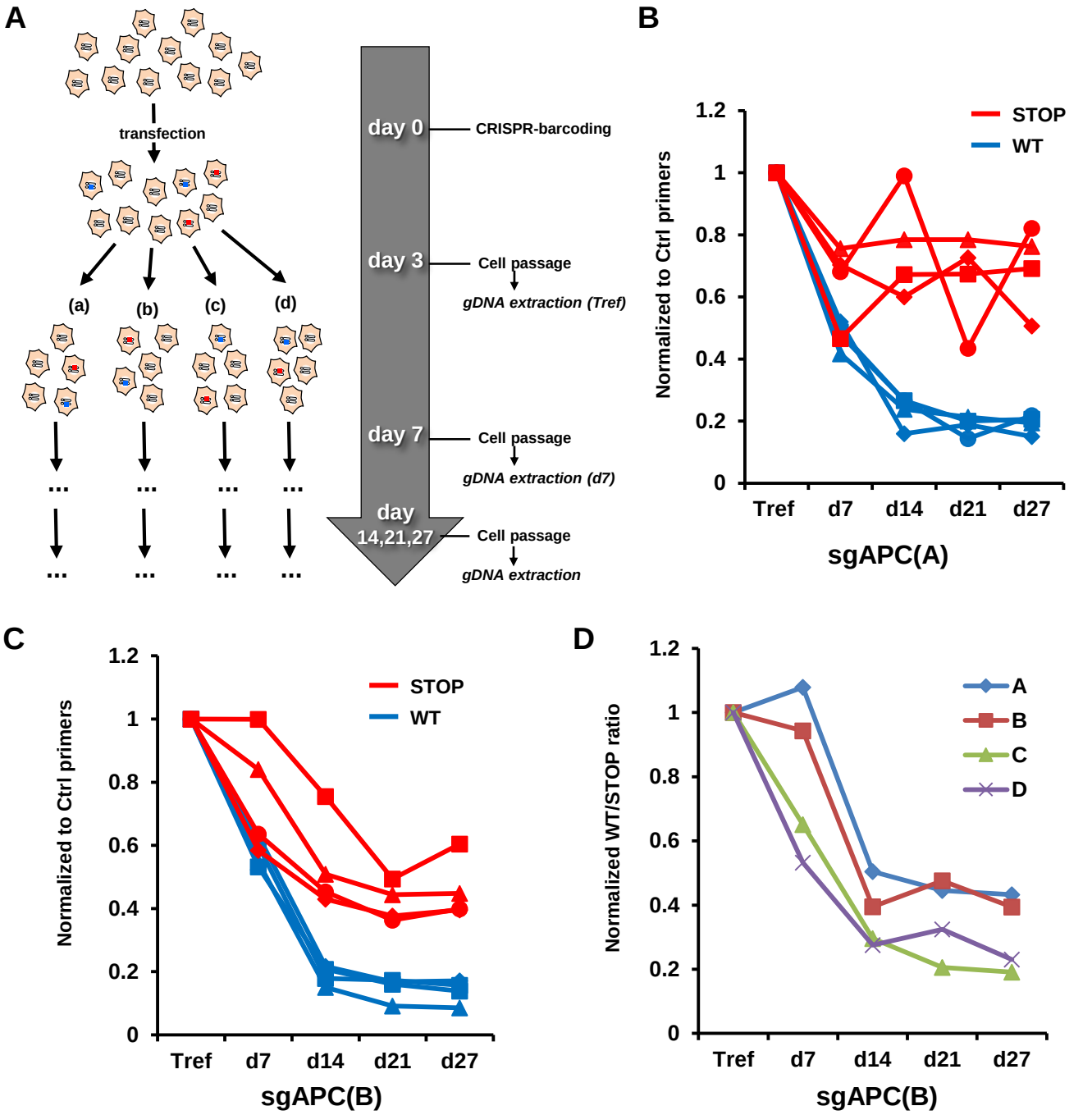
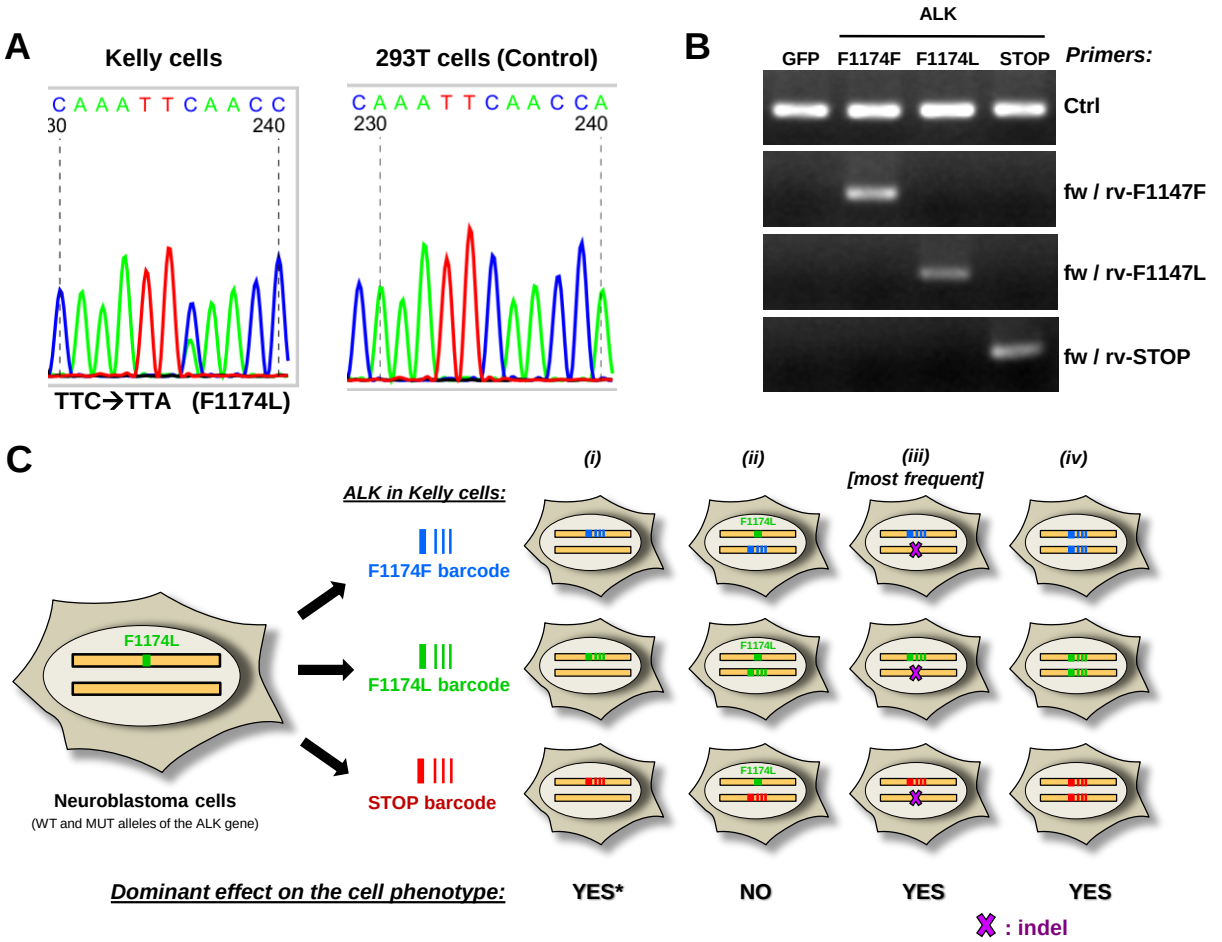


Figure S5, related to Figure 5. APC CRISPR-Barcoding in DLD-1 Cells

(A) Schematic representation of the CRISPR-barcoding approach to investigate the effects of *APC* gene repair on DLD-1 cell growth.

(B-C) APC CRISPR-barcoding was performed using two disctinct sgRNA (sgAPC(A) and sgAPC(B)), and the APC-WT and APC-STOP qPCR values were normalized to the total amount of gDNA, measured using qPCR primers amplifying an unmodified region of the APC gene (Ctrl primers). The results correspond to the APC-WT to APC-STOP ratio of one representative of three independent experiments.

(D) The data in B are represented as APC-WT to APC-STOP barcode ratio.



F

	APC-WT	APC-MUT
mean	0.6198	1.365
sd	*0.1855*	*0.4576*

	p53-WT	p53-STOP
mean	50.919	13.637
sd	*10.901*	*4.009*

	ALK-F1174F	ALK-F1174L	ALK-STOP
mean	62.584	18.276	8.8747
sd	*34.939*	*11.78*	*1.286*

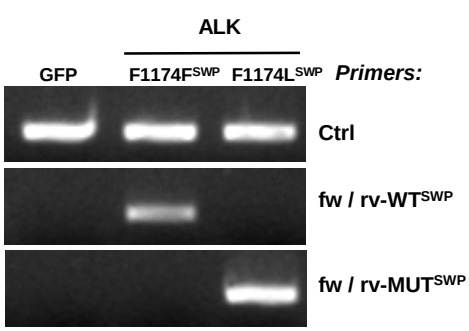


Figure S6, related to Figure 6. Repair of *ALK* Oncogenic Mutation in Kelly cells

(A) Sanger sequencing was performed to confirm the ALK-F1174L mutation in Kelly cells. gDNA from 293T cells was used as a control.

(B) PCR using control or barcode-specific primers was performed from gDNA extracted from 293T cells transfected with a control GFP plasmid or the sgALK vector, together with either the ALK-F1174F, ALK-F1174L or ALK-STOP ssODN.

(C) Diagram illustrating the different possible genotypes corresponding to the ALK-F1174 codon following CRISPR-barcoding transfection of Kelly cells. The predicted effects on the cell phenotype are indicated. (*): dominant effect referred to the F1174L mutation

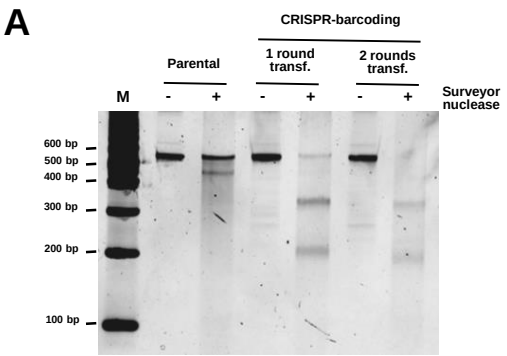
(D) ALK CRISPR-barcoding in Kelly cells using sgALK(B). The relative proportion of the ALK-F1174F, ALK-F1174L and ALK-STOP barcodes was assessed by qPCR at the indicated time points. The results correspond to the mean values (±SEM, n=4), normalized to the amount of DNA measured using qPCR targeting an unmodified region of the ALK gene (Ctrl primers), of one representative of three independent experiments.

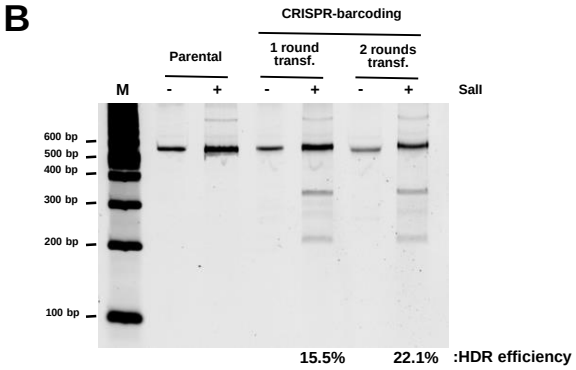
(E) The data shown in (D) were normalized to the ALK-F1174L barcode.

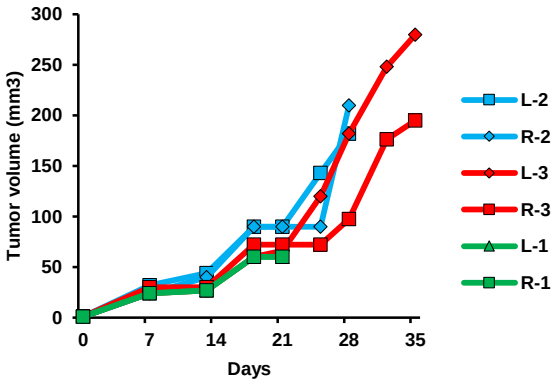
(F) To test for a potential effect of CRISPR/Cas9 induced mutations on mRNA stability, the relative proportions of APC, p53 and ALK barcodes were measured by qPCR in the gDNA and the cDNA of 293T cells transfected for CRISPR-barcoding. The mean values (arbitrary units; ± s.d. ; n=4) of the different barcodes normalized to the total amount of gDNA or cDNA are represented.

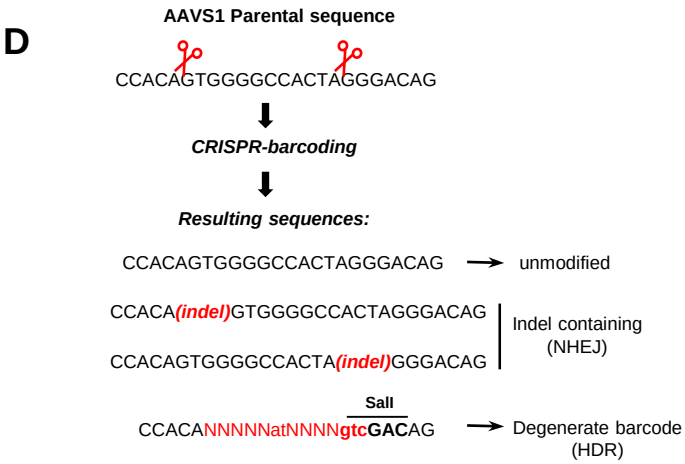
(G) The specificity of the primers used to detect ALK “swapped” barcodes (SWP) was assessed as in B.

*$p<0.05$ compared to the F1174L control barcode. (Mann-Whitney’s test)


A
CRISPR-barcoding
Parental
1 round transf.
2 rounds transf.
M
- + - + - +
Surveyor nuclease
600 bp
500 bp
400 bp
300 bp
200 bp
100 bp


B
CRISPR-barcoding
Parental
1 round transf.
2 rounds transf.
M
- + - + - +
SalI
600 bp
500 bp
400 bp
300 bp
200 bp
100 bp
15.5%
22.1%
:HDR efficiency


C
Tumor volume (mm3)
300
250
200
150
100
50
0
0
7
14
21
28
35
Days
L-2
R-2
L-3
R-3
L-1
R-1


D
AAVS1 Parental sequence
CCACAGTGGGGCCACTAGGGACAG
CRISPR-barcoding
Resulting sequences:
CCACAGTGGGGCCACTAGGGACAG
unmodified
CCACA(indel)GTGGGGCCACTAGGGACAG
CCACAGTGGGGCCACTA(indel)GGGACAG
Indel containing (NHEJ)
SalI
CCACANNNNNatNNNNgtcGACAG
Degenerate barcode (HDR)

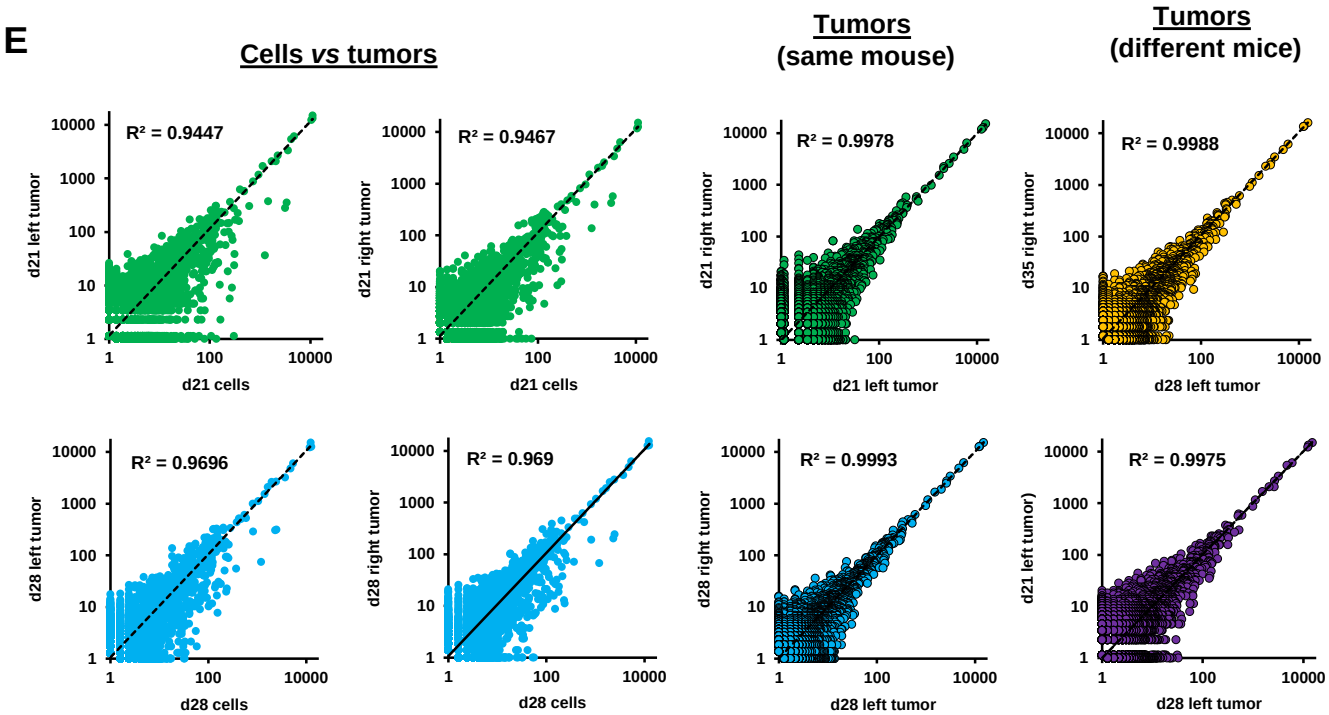

E
Cells vs tumors
Tumors (same mouse)
Tumors (different mice)
R² = 0.9447
d21 left tumor
d21 cells
R² = 0.9467
d21 right tumor
d21 cells
R² = 0.9978
d21 right tumor
d21 left tumor
R² = 0.9988
d35 right tumor
d28 left tumor
R² = 0.9696
d28 left tumor
d28 cells
R² = 0.969
d28 right tumor
d28 cells
R² = 0.9993
d28 right tumor
d28 left tumor
R² = 0.9975
d21 left tumor)
d28 left tumor

Figure S7, related to Figure 7. Analysis of Intratumor Heterogeneity through Highly Complex CRISPR-Barcodes

(A) AAVS1 surveyor assay from gDNA derived from BT474 cells before (parental) or after one or two rounds of CRISPR-barcoding transfection using two different sgRNAs.

(B) AAVS1 RFLP assay using SalI restriction enzyme and gDNA derived from BT474 cells before (parental) or after one or two rounds of CRISPR-barcoding transfection. HDR efficiency was assessed by comparing the intensity of the bands corresponding to the undigested and SalI-digested amplicons.

(C) Measurement of BT474 tumor volumes in SCID mice.

(D) Schematic representation of the different types of sequences resulting from AAVS1 CRISPR-barcoding.

(E) gDNA was derived from BT474 tumors or BT474 cells maintained in culture, and the AAVS1 barcodes were analyzed by deep sequencing. The plots show comparisons of the normalized frequency of each barcode from the indicated samples (BT474 tumors versus corresponding BT474 cells in culture; BT474 tumors originated in the same or different animals). Linear regressions were calculated, and the corresponding coefficients of determination (r-squared) are indicated.

Table S1, related to Figures 1, 4, 5, 6 and 7. HDR efficiency in the different cell models used in the study.

Cell line	Gene	HDR Efficiency
PC9	EGFR	~0.6 % (a)
PC9	KRAS	~1 % (a)
HCT-116	TP53	~3.5 % (a)
DLD-1	APC	~9 % (a)
Kelly	ALK	~0.4 % (a)
BT474	AAVS1	~22 % (b)
PC9	AAVS1	~8 % (b)

(a) As assessed by deep sequencing (percentage of reads containing the barcodes vs the endogenous sequence).

(b) As assessed by RFLP (to increase efficiency, the cells were subjected to 2-3 rounds of transfection using two different sgRNAs).

Table S2, related to the Experimental Procedures. List of sgRNA target sequences and ssODNs used for CRISPR-barcoding.

sgRNA target	
Name	**Sequence 5'-3'**
sgAPC(A)	AGGGCTATCTGGAAGATCAC
sgAPC(B)	CCATGCAGTGGAATGGTAAG
sgTP53(A)	GGGCAGCTACGGTTTCCGTC
sgTP53(B)	CAGAATGCAAGAAGCCCAGA
sgTP53(R273)	ACTGGGACGGAACAGCTTTG
sgALK(A)	TGCAGCGAACAATGTTCTGG
sgALK(B)	AGAACATTGTTCGCTGCATT
sgEGFR(A)	CTGCGTGATGAGCTGCACGG
sgEGFR(B)	GAGCTGCGTGATGAGCTGCA
sgKRAS(A)	GAATATAAACTTGTGGTAGT
sgKRAS(B)	CTTGTGGTAGTTGGAGCTGG
sgEML4	GACCTGAACAGCAAGTTTGT (a)
sgALK(Chr. Inv.)	GGCCTTGCTGAAACTTCCTT (a)
sgAAVS1(A)	GTCCCCTCCACCCCACAGTG
sgAAVS1(B)	GTCACCAATCCTGTCCCTAG
ssODN	
Name	**Sequence 5'-3' (b)**
APC-WT	CACTTGATAGTTTTGAGAGTCGTTCGATTGCCAGCTCCGTTCAGAGTGA ACCATGCAGTGGAATGGTAAGTGcgATcATtAGtCCgAGcGAcCTTCCAGAT AGCCCTGGACAAACCATGCCACCAAGCAGAAGTAAAACACCTCCACCAC CTCCTCAAACAG
APC-MUT	CACTTGATAGTTTTGAGAGTCGTTCGATTGCCAGCTCCGTTCAGAGTGA ACCATGCAGTGGAATGGTAAGTtGacTcgagAattcAGTGATCTTCCAGATAG CCCTGGACAAACCATGCCACCAAGCAGAAGTAAAACACCTCCACCACCT CCTCAAACAG
TP53-WT	CTCCTACACCGGCGGCCCCTGCACCAGCCCCCTCCTGGCCCCTGTCAT CTTCTGTCCCTTCCCAGAAAACCTACCAGGGCAGCTAtGGcTTtCGcCTcG GaTTCTTGCATTCTGGGACAGCCAAGTCTGTGACTTGCACGGTCAGTTG CCCTGAGGGGCTGGCTTCCATGAGACTTCAATGCCTGG
TP53-STOP	CTCCTACACCGGCGGCCCCTGCACCAGCCCCCTCCTGGCCCCTGTCAT CTTCTGTCCCTTCCCAGAAAACCTACCAGGGCAGCTACGGTTgagaattcT CTTCTTGCATTCTGGGACAGCCAAGTCTGTGACTTGCACGGTCAGTTGC CCTGAGGGGCTGGCTTCCATGAGACTTCAATGCCTGG
TP53-R273H	TTTTAAATGGGACAGGTAGGACCTGATTTCCTTACTGCCTCTTGCTTCTC TTTTCCTATCCTGAGTAGTGGTAATCTACTGGGcCGaAAtAGtTTcGAaGTa CaTGTTTGTGCCTGTCCTGGGAGAGACCGGCGCACAGAGGAAGAGAAT CTCCGCAAGAAAGGGGAGCCTCACCACGAGCTGCCCC
TP53-R273R	TTTTAAATGGGACAGGTAGGACCTGATTTCCTTACTGCCTCTTGCTTCTC TTTTCCTATCCTGAGTAGTGGTAATCTACTGGGACGcAAtAGCTTTGAaGT cCGaGTaTGcGCaTGcCCTGGGAGAGACCGGCGCACAGAGGAAGAGAAT CTCCGCAAGAAAGGGGAGCCTCACCACGAGCTGCCCC

ALK-F1174F	GTAACTTTGTATCCTGTTCCTCCCAGTTTAAGATTTGCCCAGACTCAGCT CAGTTAATTTTGGTTACATCCCTCTCTGCTCTGCAGCAAgTTtAAtCAtCAa AACATaGTTCGCTGCATTGGaGTGAGCCTGCAATCCCTGCCCCGGTTCA TCCTGCTGGAGCTCATGGCGGG
ALK-F1174L	GTAACTTTGTATCCTGTTCCTCCCAGTTTAAGATTTGCCCAGACTCAGCT CAGTTAATTTTGGTTACATCCCTCTCTGCTCTGCAGCAAATTaAAtCAtCA GAAtATcGTcCGcTGCATTGGaGTGAGCCTGCAATCCCTGCCCCGGTTCA TCCTGCTGGAGCTCATGGCGGG
ALK-STOP	GTAACTTTGTATCCTGTTCCTCCCAGTTTAAGATTTGCCCAGACTCAGCT CAGTTAATTTTGGTTACATCCCTCTCTGCTCTGCAGCAAATagtgaattatcGA ACATTGTTCGCTGCATTGaGGTGAGCCTGCAATCCCTGCCCCGGTTCAT CCTGCTGGAGCTCATGGCGGG
ALK-F1174L-SWP	GTAACTTTGTATCCTGTTCCTCCCAGTTTAAGATTTGCCCAGACTCAGCT CAGTTAATTTTGGTTACATCCCTCTCTGCTCTGCAGCAAgTTaAAtCAtCAa AACATaGTTCGCTGCATTGGaGTGAGCCTGCAATCCCTGCCCCGGTTCA TCCTGCTGGAGCTCATGGCGGG
ALK-F1174F-SWP	GTAACTTTGTATCCTGTTCCTCCCAGTTTAAGATTTGCCCAGACTCAGCT CAGTTAATTTTGGTTACATCcCTCTCTGCTCTGCAGCAAATTtAAtCAtCAG AAtATcGTcCGtTGCATTGGaGTGAGCCTGCAATCCCTGCCCCGGTTCATC CTGCTGGAGCTCATGGCGGG
EGFR-T790T	CTCCCTCCCTCCAGGAAGCCTACGTGATGGCCAGCGTGGACAACCCCCC ACGTGTGCCGCCTGCTGGGCATCTGCCTCACCTCtACaGTcCAaCTgATtA CcCAGCTCATGCCCTTCGGCTGCCTCCTGGACTATGTCCGGGAACACAA AGACAATATTGGCTCCCAGTACCTGCTCAACTGGTGTGTG
EGFR-T790M	CTCCCTCCCTCCAGGAAGCCTACGTGATGGCCAGCGTGGACAACCCCC ACGTGTGCCGCCTGCTGGGCATCTGCCTCACCTCtACtGTaCAGCTtATaA tGCAaCTgATGCCCTTCGGCTGCCTCCTGGACTATGTCCGGGAACACAAA GACAATATTGGCTCCCAGTACCTGCTCAACTGGTGTGTG
KRAS-G12G	TAACCTTATGTGTGACATGTTCTAATATAGTCACATTTTCATTATTTTTATT ATAAGGCCTGCTGAAAATGACTGAATATAAgtTgGTaGTgGTTGGcGCaGG aGGCGTAGGCAAGAGTGCCTTGACGATACAGCTAATTCAGAATCATTTT GTGGACGAATATGATCCAACAATAGAGGTAAA
KRAS-G12D	TAACCTTATGTGTGACATGTTCTAATATAGTCACATTTTCATTATTTTTATT ATAAGGCCTGCTGAAAATGACTGAATATAAgtTaGTtGTcGTTGGtGCcGac GGCGTAGGCAAGAGTGCCTTGACGATACAGCTAATTCAGAATCATTTTG TGGACGAATATGATCCAACAATAGAGGTAAA
EML4-ALK	AGTGAACACAGTTGTGTTGTTCAATTTTTAAGGTATTTTTAGATGATAAAT ATTGATGTAAGTGGAGACAGTTGACCTGAACAGCAAGTTCTTAGGCTCC ATGGCACCCAGGGTGCTTCCACCCAACCTTCCCTCCCTCCCTCGTTCAC GTGGGGTTATACTTGCAACACAGTCTGCTGG
AAVS1-deg	GTGTCCCCGAGCTGGGACCACCTTATATTCCCAGGGCCGGTTAATGTG GCTCTGGTTCTGGGTACTTTTATCTGTCCCCTCCACCCCACANNNNNatN NNNgtcGACAGGATTGGTGACAGAAAAGCCCCATCCTTAGGCCTCCTCCT TCCTAGTCTCCTGATATTGGGTCTAACCCCCACCTCCTGTTAGGCA

(a) Sequence from Choi and Meyerson (2014), Nat Commun 5, 3728
(b) Mutations compared to the endogenous sequence are indicated in lowercase letters

Table S3, related to the Experimental Procedures. List of primers used for qPCR, surveyor assay and deep sequencing.

qPCR primers	
Name	**Sequence 5'-3'**
APC_Ctrl_FW	GCTGAAGATCCTGTGAGCGA
APC_Ctrl_RV	TTGTGCCTGGCTGATTCTGA
APC_cbc_FW	AGACCCCACTCATGTTTAGCAG
APC_cbc_WT_RV	GGTCGCTCGGACTAATGATCG
APC_cbc_STOP_RV	GGAAGATCACTGAATTCTCGAGTCA
TP53_Ctrl_FW	GACCCAGGGTTGGAAGTGTC
TP53_Ctrl_RV	CTAGGATCTGACTGCGGCTC
TP53_cbc_FW	CAGGTCCAGATGAAGCTCCCA
TP53_cbc_WT_RV	GAATCCGAGGCGAAAGCCA
TP53_cbc_STOP_RV	TCCCAGAATGCAAGAAGAGAATTCTC
TP53_cbc_R273_FW	CAAGGGTGGTTGGGAGTAGAT
TP53_cbc_R273R_RV	GGGCATGCGCATACTCGG
TP53_cbc_R273H_RV	CAAACATGTACTTCGAAACTATTTCGG
ALK_Ctrl_FW/ ALK_cbc_FW	GCCCATGTTTACAGAATGCCTTT
ALK_Ctrl_RV	CCCCAATGCAGCGAACAATG
ALK_cbc_WT_RV	GCAGCGAACTATGTTTTGATGATTAAAC
ALK_cbc_MUT_RV/ ALK_cbc_WT_SWP_RV	CAATGCAACGGACGATATTCTGA
ALK_cbc_STOP_RV	CAGCGAACAATGTTCGATAATTCACT
ALK_cbc_MUT_SWP_RV	GCAGCGAACTATGTTTTGATGATTtAAC
EGFR_Ctrl_FW	TGCTTCCCCCATTCAGGACT
EGFR_Ctrl_RV	CTCCTTGCACCTCCTCACTG
EGFR_cbc_RV	CCTTCCCTGATTACCTTTGCGA
EGFR_cbc_T790T_FW	CACCTCTACAGTCCAACTGATTACC
EGFR_cbc_T790M_FW	CTCTACTGTACAGCTTATAATGCAACTG
KRAS_Ctrl_FW	ATCCTTTGAGAGCCTTTAGCCG
KRAS_Ctrl_RV	CCAGTTGACTGCAGACGTGTA
KRAS_cbc_G12G_RV	CTGCGCCAACCACTACCAAC
KRAS_cbc_G12D_RV	GTCGGCACCAACGACAACT
KRAS_cbc_FW	TATTAAAAGGTACTGGTGGAGTATTTGAT
EML4_ALK_FW	CTCGTGGTAACATCAGAACAGAGA
EML4_ALK_RV	TGCCATGGAGCCTAAGAACTTG
PCR primers for surveyor assay and RFLP	
Name	**Sequence 5'-3'**
APC fw	AGCTGAAGATCCTGTGAGCG
APC rv	AATGGCTCATCGAGGCTCAG
AAVS1 fw	CATCTTCCAGGGGTCCGAGA
AAVS1 rv	TCAGTGAAACGCACCAGACG
PCR primers for APC deep sequencing	
Name	**Sequence 5'-3'**
APC 1_FW	GATTGCCAGCTCCGTTCAG
APC 1_RV	ATTTACTGCAGCTTGCTTAGGT
APC 2_FW	TGCTCAGACACCCAAAAGT
APC 2_RV	GTTTTACTTCTGCTTGGTGGC

Other PCR primers used for deep sequencing	
Name	**Sequence 5'-3'**
AAVS1ds_FW	*AATGATACGGCGACCACCGAGATCT*ACACTCTTTCCCTACACGACGCTCTTCCGATCTGTTAATGTGGCTCTGGTTCTGG*
AAVS1ds_RV	*CAAGCAGAAGACGGCATACGAGAT***CGTGAT**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGGTGGGGGTTAGACCCAATATC*
EGFR_FW	*AATGATACGGCGACCACCGAGATCT*ACACTCTTTCCCTACACGACGCTCTTCCGATCTCACGTGTGCCGCCTGCTGG
EGFR_RV	*CAAGCAGAAGACGGCATACGAGAT***CGTACG**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCCTTCCCTGATTACCTTTGCGA
KRAS_FW	*AATGATACGGCGACCACCGAGATCT*ACACTCTTTCCCTACACGACGCTCTTCCGATCTTTTTTATTATAAGGCCTGCTGAAAATG
KRAS_RV	*CAAGCAGAAGACGGCATACGAGAT***CCACTC**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTAGAATGGTCCTGCACCAGTAA
TP53_FW	*AATGATACGGCGACCACCGAGATCT*ACACTCTTTCCCTACACGACGCTCTTCCGATCTCCTGTCATCTTCTGTCCCTTCC
TP53_RV	*CAAGCAGAAGACGGCATACGAGAT***GCTACC**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTAGAAATGCAGGGGGATACGG
ALK_FW	*AATGATACGGCGACCACCGAGATCT*ACACTCTTTCCCTACACGACGCTCTTCCGATCTAATTTTGGTTACATCCCTCTCTGC
ALK_RV	*CAAGCAGAAGACGGCATACGAGAT***ATCAGT**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTAAGGACTTGAGGTCTCCCCC

The flowcell sequences are in italics, the Illumina sequencing primers are underlined and the sequence of the indexes used is in bold. For AAVS1, each sample was amplified using a RV primer containing a distinct Illumina index.

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Invasion Assay

PC9 cells containing the EGFR-T790M and -T790T barcodes pre-treated or not with gefitinib (0.5 µM) for 48 hr were seeded in matrigel 6-well invasion chambers (8 µM pores; BD Biosciences) at a density of $1.5X10^6$ cells per chamber, in the presence or the absence of gefitinib (0.5 µM) and incubated for 24 hr. As a control, a fraction of the cells was seeded in parallel in regular wells. The cells inside and outside the chambers were then trypsinized and gDNA was extracted for qPCR.

Mouse Xenografts

PC9 cells containing the EGFR-T790, KRAS-G12 and EML4-ALK barcodes were resuspended in PBS and subcutaneously inoculated in the left and right flank ($2 X 10^6$ cells per site) of male SCID mice (NCI). The size of the tumors was measured by caliper every 3-4 days. Sixteen days after injection, when the tumors were palpable, half of the mice were treated 5 days a week with gefitinib (25 mg/kg/day) by gavage. When the size of at least one of the tumors reached the arbitrary volume of 550 mm^3, the mice were sacrificed and the tumors were dissected. Each tumor was cut in two halves, one for gDNA extraction and one for storing. BT474 cells ($5 X 10^6$ cells per site) were mixed with matrigel and inoculated bilaterally in the mammary fat pad of female SCID mice previously implanted with 17β-estradiol pellets (Innovative Research of America). Mice were sacrificed 3, 4 and 5 weeks after injection and gDNA was extracted from the tumors for deep sequencing analysis. At each time point, gDNA was extracted from the same batch of BT474 cells kept in culture. All animal experiments were approved and performed according to the relevant regulatory standards set by Mount Sinai's Animal Care and Use Committee.

Lentiviral Production and Inhibitors

The lentiviral Wnt reporter was generated by cloning 14 Wnt-responsive elements upstream of a minimal promoter (Grumolato et al., 2013), followed by a destabilized GFP (Matsuda and Cepko, 2007). The construct also contains the puromycin resistance gene driven by a different promoter. Lentivirus was produced as previously described (Grumolato et al., 2010). Briefly, 293T cells were cotransfected with the Wnt-responsive lentiviral vector, pCMV Δ8.91 and pMD VSV-G plasmids. Two days after transfection, the conditioned medium was collected, purified by centrifugation, supplemented with 8 µg/mL polybrene, and added for overnight incubation to freshly plated DLD-1 cells. Two days after infection, the cells were selected in 2 mg/mL puromycin.

Cell Sorting by FACS

DLD-1 cells transduced with the Wnt-responsive reporter were sorted based on the level of GFP expression. Briefly, harvested cells were resuspended in sterile FACS-sorting buffer (D-PBS medium, 2 mM EDTA, 0.5% BSA), filtered on 30 µm cell-strainer (BD Biosciences) to remove cell aggregates and isolated by FACS using a FACSAria-III Cytometer (BD Biosciences). DLD-1 cells containing the Wnt reporter were first FACS sorted to remove the fraction of GFP-negative cells, and the resulting cell population was transfected for APC CRISPR-barcoding. Five days after transfection, GFP^{hi} (defined as the 10% of the cells displaying the highest fluorescence) and GFP^{lo} cells (defined as the 10% of the cells displaying the lowest fluorescence) were isolated by FACS, and gDNA was extracted from sorted and un-sorted cells for qPCR analysis. Cells were maintained at 4°C during the entire sorting procedure. Reanalysis of sorted fractions consistently showed that more than 95% of the sorted cells were viable and corresponded to the defined GFP^{hi} or GFP^{lo} gates. Analysis and graphical representation of the data were performed using FlowJo software (Tree stars).

Deep Sequencing

PCR was performed from DLD-1 cell gDNA samples using Herculase II Fusion DNA Polymerase and two couples of primers (Table S3), designed to amplify two overlapping regions of the APC gene spanning 245 and 220 bp, respectively, both including the CRISPR/Cas9 targeted sequence. The samples were prepared from 1 µg of purified amplicons using SPRIworks System I automaton (Beckman). Each couple of amplicons was indexed with the same adapter sequence (Craig et al., 2008), and the enriched libraries were sequenced on a Genome Analyzer IIx (Illumina) with 76 bp paired-end reads. Image analysis and base calling were performed by Real-Time Analysis (RTA 1.9) and CASAVA software (v1.8.2, Illumina). The mean read count passing filters for each sample was at least 160,000, and 94% of the sequenced bases have a QScore above 30. Motifs of interest were counted in the FastQ files.

For AAVS1 barcode sequencing, gDNA was isolated from BT474 cells and tumors tissues using the Epicenter protocol (MC85200) (Jayaprakash et al., 2015). DNA was eluted in 100 µl of TE buffer and checked for quality and quantity using a 1% agarose and qubit fluorometer respectively (Qubit dsDNA BR Assay Kit). The eluted DNA was further heated at 70° C for 30 min to inactivate any leftover proteinase K that was

introduced in the DNA isolation protocol. One microgram of DNA was used for PCR amplifications. 25 cycle-PCR was performed using primers containing the flow cell sequences and an index for in silico segregation of the reads. The total volume of PCR reaction was 100 ul using OneTaq® 2X Master Mix (NEB M0482S). For PC9 cells, DNA was extracted using the NucleoSpin Tissue kit and 25 cycle PCR amplification was performed from 0.8 µg of DNA using Herculase II Fusion DNA Polymerase (Agilent Technologies). The PCR products were run on a 2% low melt agarose and a 300 bp product was size selected and eluted from the gel. The PCR product was then quantified using an Agilent bioanalzyer and qubit fluorometer. Deep sequencing was then performed using the Illumina platform (75 bp SE sequencing), and each sample receiving >250,000 reads.

Surveyor and RFLP Assays

The sequence of the primers used to amplify a ~800 bp region of the APC and AAVS1 genes encompassing the CRISPR/Cas9 targeted sites is provided in Table S3. PCR was performed using Herculase II Fusion DNA Polymerase and the amplicon was purified using PCR clean-up kit (Macherey-Nagel) and eluted with 30 µL of TE buffer. Surveyor assay was performed as previously described(Ran et al., 2013). Briefly, 15 µL of purified-PCR products were then mixed with Taq polymerase PCR buffer and subjected to a re-annealing process to enable DNA hetereoduplex formation. After re-annealing, products were treated with SURVEYOR nuclease and SURVEYOR enhancer S (Transgenomic). The remaining 15 µL of the purified amplicon was used as a control. For RFLP assay, the AAVS1 amplicons were digested in the presence or in the absence of SalI (Promega). For both surveyor and RFLP assays, the reaction products were analyzed on a 10% Acrylamide-TBE gel stained with SYBR-Gold (Life Technologies) and imaged with a Chemidoc gel imaging system (Bio-Rad).

RNA Extraction and Reverse Transcription

Total RNA was isolated using the Tri-Reagent (Sigma-Aldrich) and the RNeasy kit (Qiagen), DNase digested and reverse-transcribed using the Improm-II Reverse Transcription System (Promega).

Immunoblot

Cells were washed once with phosphate-buffered saline (PBS) and lysed on ice in lysis buffer containing 50 mM Hepes pH 7.6, 150 mM NaCl, 5 mM EDTA, 1% Nonidet P-40, 20 mM NAF, 2 mM sodium orthovanadate, supplemented with protease inhibitor mini tablets (Thermo Scientific). Lysates were cleared by centrifugation at 14,000 g during 15 min at 4°C and protein concentrations were determined by using the Bio-Rad protein assay (Bio-Rad). Sodium dodecyl sulfate (SDS) loading buffer was added to equal amounts of lysate, followed by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) and transfer to polyvinylidene fluoride (PVDF) membrane (Thermo Scientific). The membranes were probed using Mouse anti-TP53 (Cell Signaling) and mouse anti-β-catenin (BD Biosciences) antibodies and analyzed by chemiluminescence using ECL Western Blotting Substrate (Thermo Scientific).

Supplemental references

Craig, D.W., Pearson, J.V., Szelinger, S., Sekar, A., Redman, M., Corneveaux, J.J., Pawlowski, T.L., Laub, T., Nunn, G., Stephan, D.A., *et al.* (2008). Identification of genetic variants using bar-coded multiplexed sequencing. Nat. Methods *5*, 887-893.

Grumolato, L., Liu, G., Haremaki, T., Mungamuri, S.K., Mong, P., Akiri, G., Lopez-Bergami, P., Arita, A., Anouar, Y., Mlodzik, M., *et al.* (2013). beta-Catenin-independent activation of TCF1/LEF1 in human hematopoietic tumor cells through interaction with ATF2 transcription factors. PLoS Genet. *9*, e1003603.

Grumolato, L., Liu, G., Mong, P., Mudbhary, R., Biswas, R., Arroyave, R., Vijayakumar, S., Economides, A.N., and Aaronson, S.A. (2010). Canonical and noncanonical Wnts use a common mechanism to activate completely unrelated coreceptors. Genes Dev. *24*, 2517-2530.

Jayaprakash, A.D., Benson, E.K., Gone, S., Liang, R., Shim, J., Lambertini, L., Toloue, M.M., Wigler, M., Aaronson, S.A., and Sachidanandam, R. (2015). Stable heteroplasmy at the single-cell level is facilitated by intercellular exchange of mtDNA. Nucleic Acids Res. *43*, 2177-2187.

Matsuda, T., and Cepko, C.L. (2007). Controlled expression of transgenes introduced by in vivo electroporation. Proc. Natl. Acad. Sci. USA *104*, 1027-1032.

Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. Nat. Protoc. *8*, 2281-2308.