Molecular Cell, Volume *70*

Supplemental Information

Kinetics and Fidelity of the Repair of Cas9-Induced Double-Strand DNA Breaks

Eva K. Brinkman, Tao Chen, Marcel de Haas, Hanna A. Holland, Waseem Akhtar, and Bas van Steensel

Supplementary Data

Supplementary Table 1, Related to STAR Methods

Name	Sequence	Location (hg38)	Targeting strand
LBR guide #2	5′ GCCGATGGTGAAGTGGTAAG 3′	Chr 1: 225424038-225424057	-
LBR guide #5	5′ GTATTTTAGTGATCAGCCTG 3′	Chr 1: 225424155-225424174	-
LBR guide #6	5′ AGGCTACATTCAATCTCATT 3′	Chr 1: 225424215-225424234	+
LBR guide #7	5′ GAGATTGAATGTAGCCTTTC 3′	Chr 1: 225424212-225424231	-
LBR guide #8	5′ AGAGTGTGTTTACAGTAAGT 3′	Chr 1: 225423869-225423888	-
LBR guide #9	5′ GTGTGAGCTTCTTGGGAACA 3′	Chr 1: 225423715-225423734	+
Intergenic guide	5′ TGGTCTCCTTGTCTGTGTGG 3′	Chr 11: 5561098- 5561117	+
AAVS1 (Mali et al.)	5′ GTCCCCTCCACCCCACAGTG 3′	Chr 19: 55115771- 55115790	-

Supplementary Table 2, Related to STAR Methods

Name	Number	Oligo*
Adaptor primer	EB486	GCGTGGTCGCGGCCGAGGA
Extension primer 3′	EB479	TGGGTGGTTGGCAGAGTTAC
Broken primer 3′	EB487	GAATGTAGCCTTTCTGGCCCTAAAATCCTG
Standard primer 3′	EB488	TCCTACTTGGCATTTTCTATAATTAACCTGA
Extension primer 5′	EB551	CCCTTGGGCATGGAATATAA
Broken primer 5′	EB553	CCTTCCAGCACTTGGCTGACTGTGT
Standard primer 5′	EB555	GATTGAGCTCTTGCTTTGGGTCACATAC
LBR2-fw	EB279	AAATGGCTGTCTTTCCCAGTAA
LBR2-rv	EB361	ACGCAGTGGCTAAATCATCC
AAVS1-fw	EB326	GCTTCTCCTCTTGGGAAGTGTA
AAVS1-rv	EB415	TTTCTGTCTGCAGCTTGTGG
chr11-fw	EB333	AGGAAGACGATGGAGAAGACAG
chr11-rv	EB416	CTTTCCTGCCCATGTTGATT
LBR2-5-6-7-fw-bc	EB386-EB392, EB396-EB398, EB417-EB421	ACACTCTTTCCCTACACGACGCTCTTCCGATCT**NNNNNN**TAAAGCTGGGAGGTGCTGTC
LBR8-fw-bc	EB511, EB518-EB519	ACACTCTTTCCCTACACGACGCTCTTCCGATCT**NNNNNN**AGCTCAATCCTCTGCCTTCA
Lbr9-fw-bc	EB512	ACACTCTTTCCCTACACGACGCTCTTCCGATCT**TTAGGT**AGTAAACCCAGGGGACCAAC
LBR2-5-6-7-8-9-rv1	EB393	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTACGCAGTGGCTAAATCATCC
LBR2-5-6-7-8-9-rv2	EB449	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGTAGCCTTTCTGGCCCTAAAAT
LBR8-rv	EB517	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGCCTGTGGAAAAAGACGAAT
AAVS1-fw-bc	EB451-EB460	ACACTCTTTCCCTACACGACGCTCTTCCGATCT**NNNNNN**AAGGAGGAGGCCTAAGGATG
AAVS1-rv	EB450	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCTTGCTTTCTTTGCCTGGAC
Chr11-fw-bc	EB462-EB471	ACACTCTTTCCCTACACGACGCTCTTCCGATCT**NNNNNN**CAGCATGGAGAGGAAAAGGT
Chr11-rv	EB461	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTAACCTGAGCTCATTGAGGGTT
Illmunia-fw	EB354	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT
Illumina-rv		CAAGCAGAAGACGGCATACGAGAT**NNNNNN**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATC

*N's are the barcode or index sequences

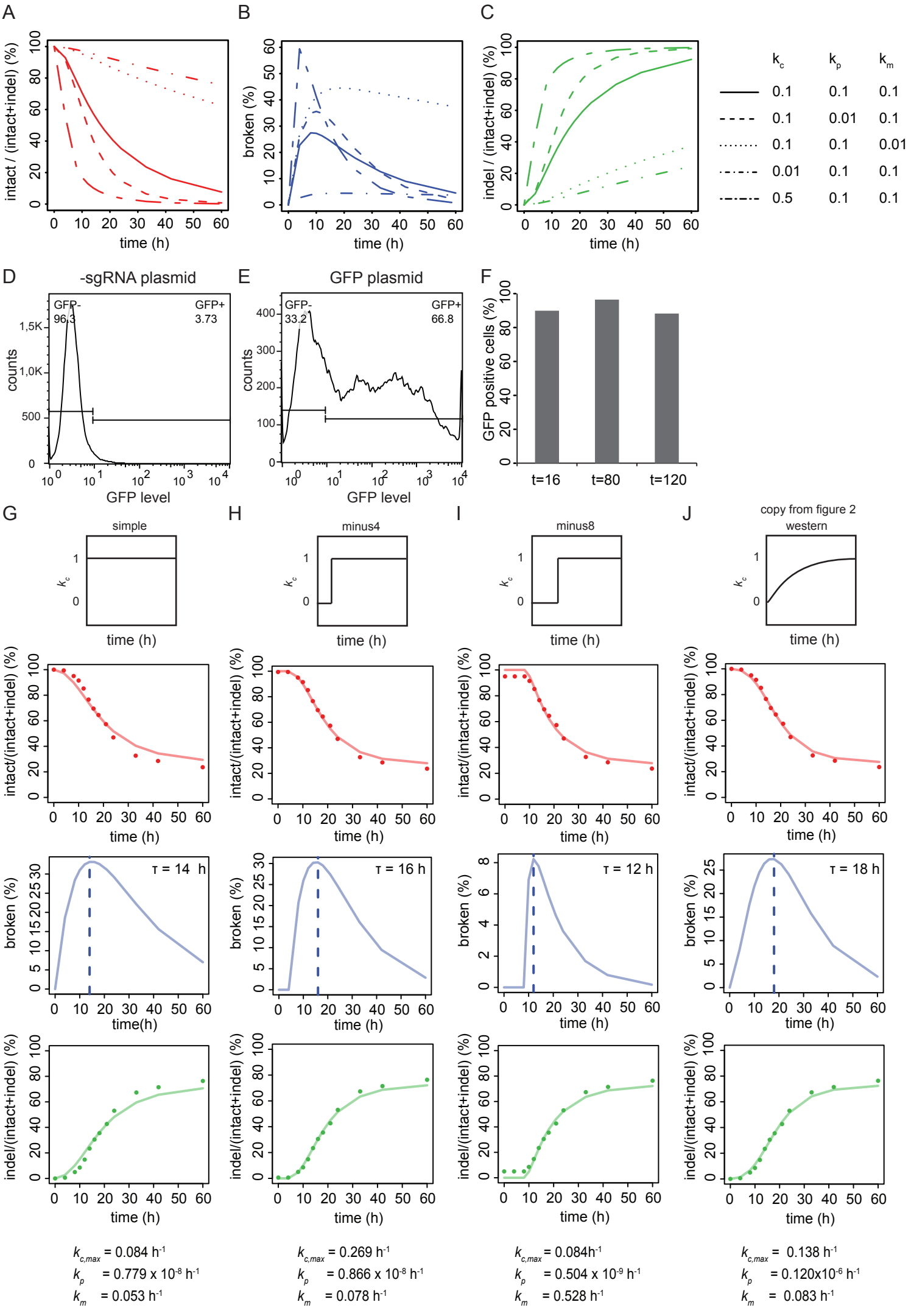
Supplementary Table 3, Related to Figure 5
Fitted parameter values for the LBR2 locus in the absence or presence of NU7441. Parameters were fitted for +1 as well as -7 indels, including different start (t=0h) and end (t=60h) rates for the -7 deletion to model the delayed accumulation. Average values ± standard deviations are shown for n independent time courses, each fitted individually.

	LBR2 without NU7441 (n=7)	LBR2 with NU7441 (n=4)
$k_{c,max}$ (h^{-1})	0.11 ± 0.02	0.10 ± 0.004
k_p (h^{-1})	$1.3 \pm 2.3 \times 10^{-6}$	$8.6 \pm 5.5 \times 10^{-6}$
k_m (h^{-1})	0.17 ± 0.12	0.06 ± 0.01
k_{+1} (h^{-1})	0.11 ± 0.08	0.02 ± 0.006
$k_{-7,t=0h}$ (h^{-1})	$3.0 \pm 5.8 \times 10^{-3}$	0.02 ± 0.004
$k_{-7,t=60h}$ (h^{-1})	0.09 ± 0.09	0.05 ± 0.005

Supplementary Table 4, Related to Figure 2 and Figure 6
Summary of estimated parameter values. ([a]) Average values ± standard deviations for the LBR2 locus, based on 7 independent time courses, each fitted individually. k_p fraction show the proportion of perfect repair of the overall repair rate ($k_p/(k_p+k_m)$). ([b]) Values estimated by a single model fit using the all data points from n datasets (see Methods). 5-95% confidence margins are indicated in brackets below each value and were obtained by bootstrapping (1,000 cycles).

	LBR2[a] (n=7)	LBR2[b] (n=7)	LBR8[b] (n=6)	AAVS1[b] (n=6)	chr11[b] (n=5)
$k_{c,max}$ (h^{-1})	0.11 ± 0.01	0.084 [0.077 – 0.187]	0.111 [0.017 – 0.246]	0.053 [0.005 – 0.153]	0.066 [0.047 – 0.121]
k_p (h^{-1})	$1.6 \pm 1.6 \times 10^{-5}$	0.002 [1.1×10^{-10} – 0.053]	0.060 [2.0×10^{-10} – 0.091]	2.8×10^{-7} [3.3×10^{-9} – 1.901]	2.4×10^{-7} [2.9×10^{-11} – 0.014]
k_m (h^{-1})	0.15 ± 0.11	0.174 [0.050 – 0.217]	0.019 [0.005 – 0.226]	0.443 [0.150 – 8.850]	0.051 [0.022 – 0.100]
$k_p/(k_p+k_m)$	1.1×10^{-4}	0.011 [9.6×10^{-10} – 0.377]	0.758 [3.8×10^{-9} – 0.805]	1.1×10^{-7} [1.4×10^{-8} – 0.580]	6.2×10^{-10} [5.8×10^{-10} – 0.261]

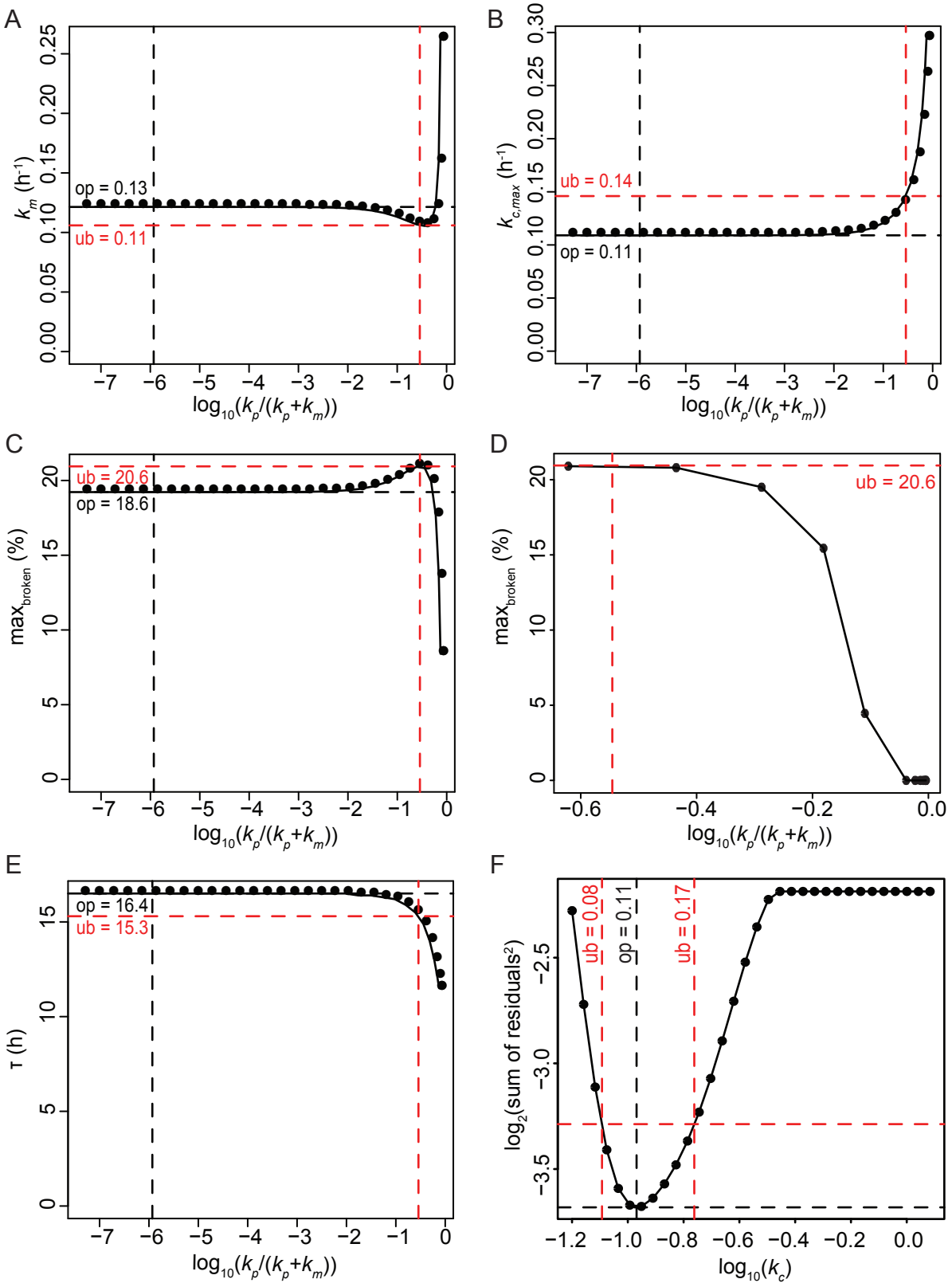
Supplementary Figure S1



Supplementary Figure 1, Related to Figure 1

(**A-C**) Simulations of the ODE model showing the relative abundance of the three states (intact, broken, indel) over time for various sets of rate constants. (**D-E**) Representative flow cytometry plots of a transfection of a control plasmid (**D**) or GFP expression plasmid (**E**) in same experiment as sgRNA-LBR2 transfection for a time series. Cells were collected 48 h after transfection with indicated plasmid. Histograms of GFP fluorescence intensity is plotted. Proportions of GFP-negative and GFP-positive cells are indicated. (**F**) Separate flow cytometry experiment of a transfection of GFP plasmid, imaged 16, 80 or 120 hours after transfection (n=1). Note that the proportion of GFP-positive cells is virtually stable. **(G-J)** Various tested ODE models that differ in the assumed onset and accumulation curve of Cas9 activity over time: (**G**) instant onset at t = 0 h; (**H**) instant onset at t = 4 h; (**I** instant onset at t = 8 h; (**J**) gradual onset quantified by Western blots. The latter is the model used in all analyses. Relative abundances of the intact and indel fractions are plotted (dots) together with the model fit (solid lines). Broken fraction is estimated from the model.

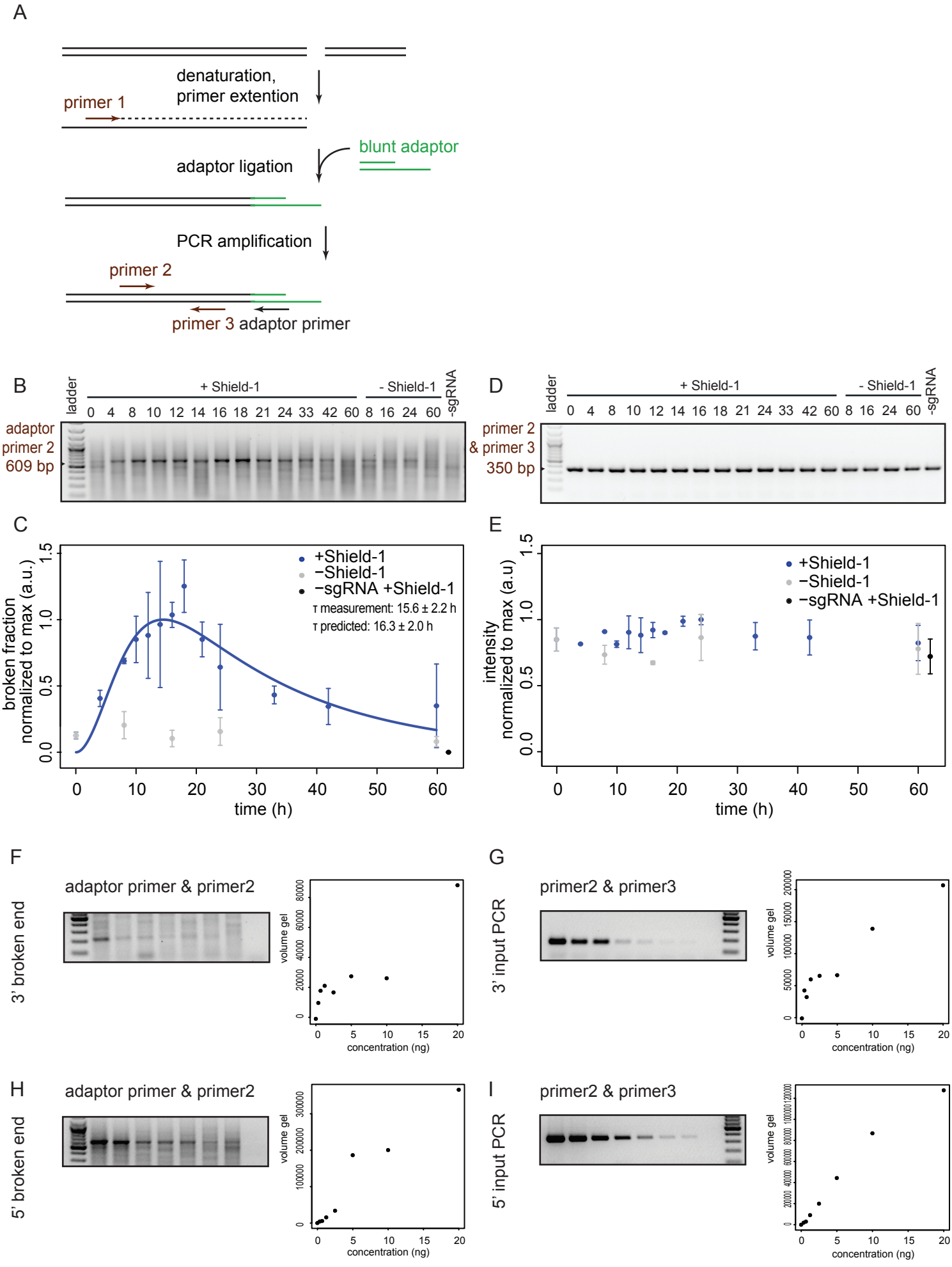
Supplementary Figure S2



Supplementary Figure 2, Related to Figure 2

Changes of the parameters $k_{c,max}$ (**A**), k_m (**B**), the amount predicted broken fraction (**C, D**) and τ (**E**) in a parameter sweep survey in which we imposed different fixed perfect/mutagenic repair ratios. Optimal fit is shown in black and significant upper bound in red ($P < 0.1$, F-test). Within this confidence interval the parameters show only modest changes. (**F**) Residuals of the fit to the data points are plotted for a parameter sweep survey of a fixed $k_{c,max}$ and changing k_p and k_m rates.

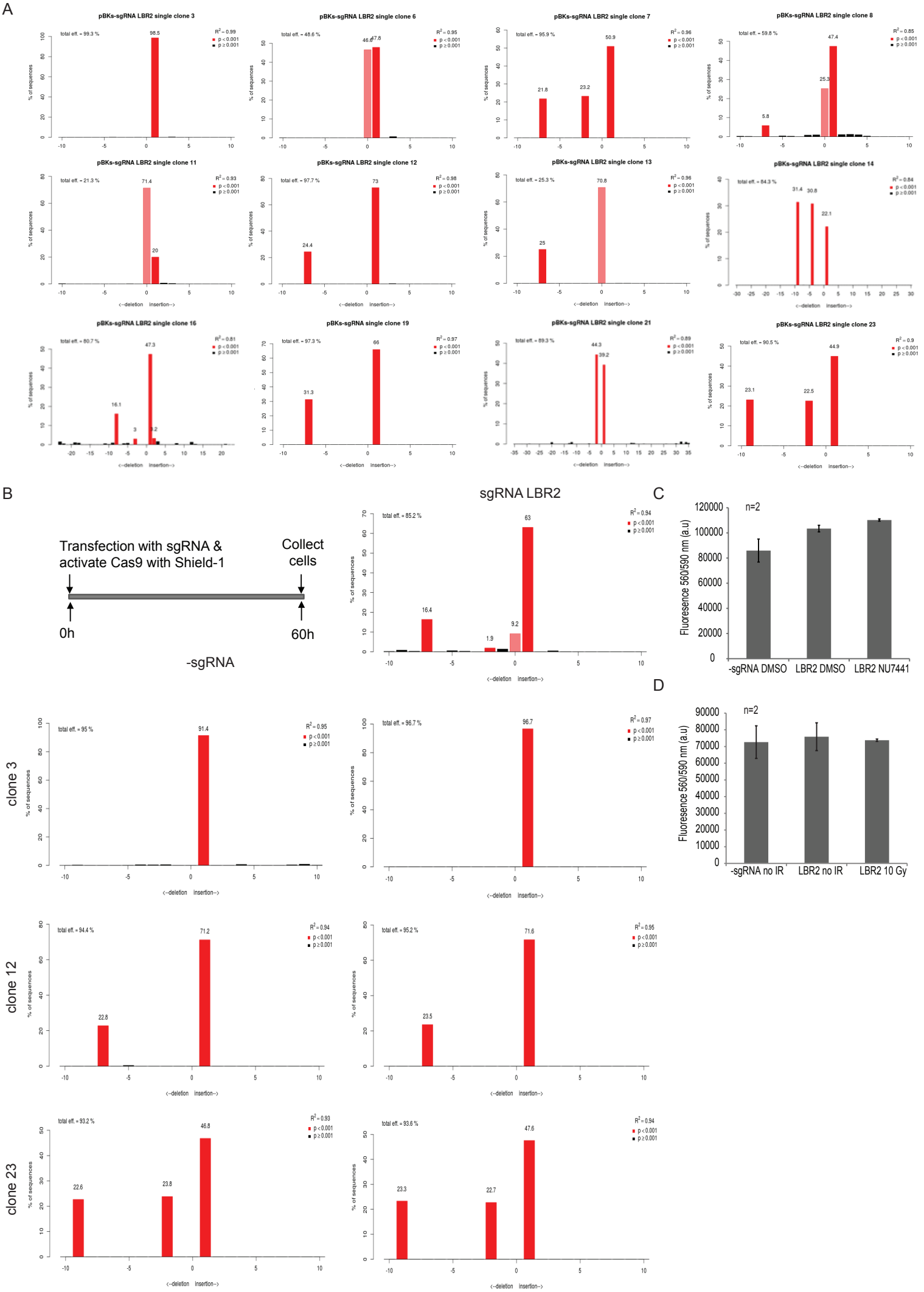
Supplementary Figure S3



Supplementary Figure 3, , Related to Figure 3

(**A**) Schematic view of the LM-PCR assay to detect the 5 prime broken ends after DSB induction. (**B**) Representative agarose gel of the LM-PCR products of a time series. The expected product is 609 bp in size. (**C**) The broken fraction measured as band intensities (data from 3 LM-PCR experiments spanning 3 different time series; values are mean ± SD). Solid blue line shows an ODE curve fit to the LM-PCR data to determine τ (blue shading, mean ± SD). (**D**) Representative agarose gel of the input PCR products of a time series. The expected product is 350 bp in size (**E**) The input PCR fraction measured as band intensities (data from 2 PCR experiments spanning 2 different time series; values are mean ± SD). (**F-I**) Test of the linearity of the used primer pairs, indicated in the figure. T=16 hour sample was diluted to range of concentrations. Agarose gel and quantification is plotted.

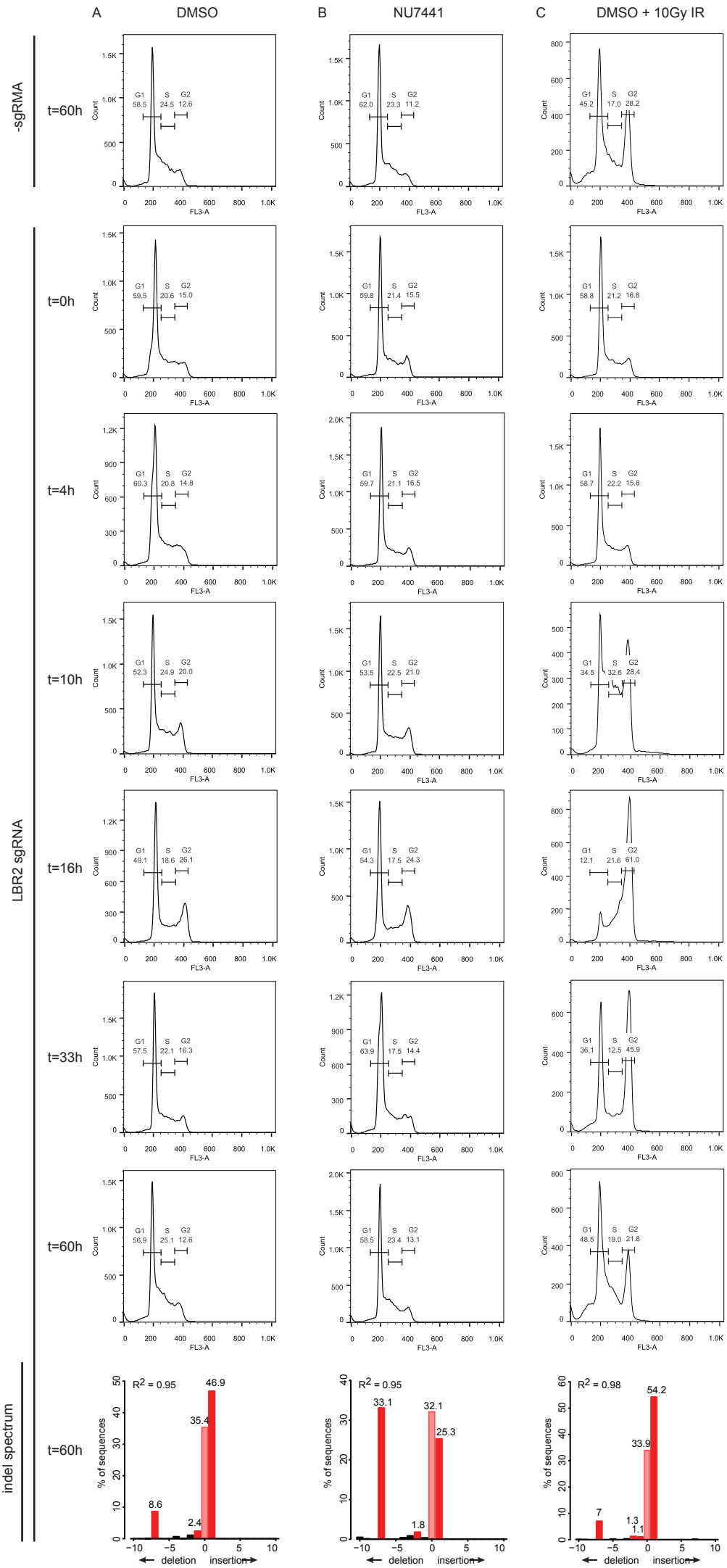
Supplementary Figure S4



Supplementary Figure 4, , Related to Figure 4

(**A**) Indel spectra determined by TIDE in twelve out of twenty cell clones derived from sgRNA-LBR2 treated cells that gained a mutation. Note that K562 cells are tri- to tetraploid, hence the individual cell clones can have multiple peaks for allele specific mutations. (**B**) Three cell clones with already acquired indels by sgRNA-LBR2 were re-transfected with or without a sgRNA-LBR2 expressing plasmid. The panels show the distribution and frequencies of indels as determined by TIDE 60 hours after Cas9/sgRNA induction. (**C**) Cell viability assay for the cells 48 hours after addition of Shield-1 and either DMSO or NU7441. (**D**) Cell viability assay for the cells without a sgRNA, with sgRNA-LBR2 and with additional damage by 10 Gy. Data in (C-D) are mean ± standard deviation from two technical replicates.

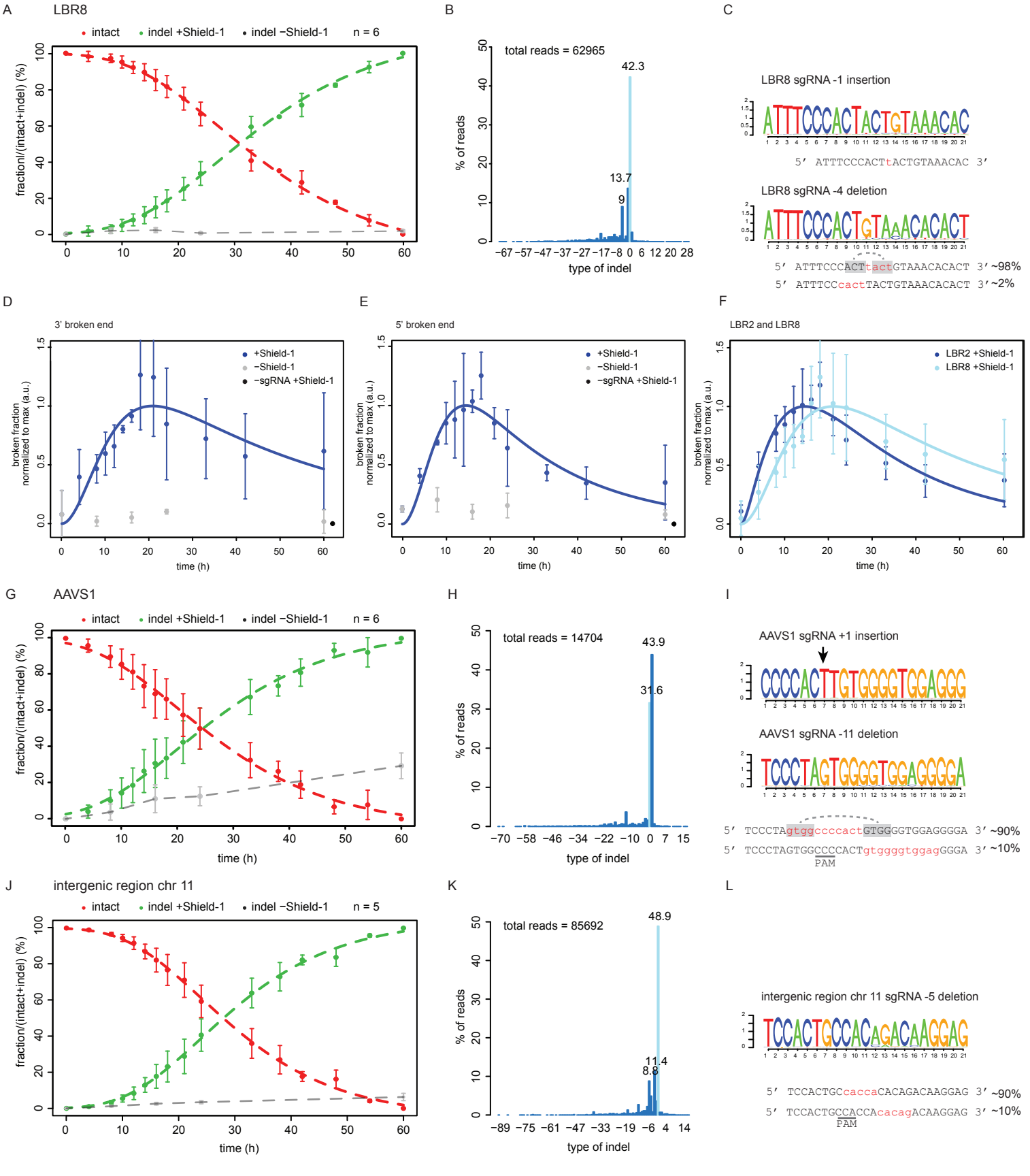
Supplementary Figure S5



Supplementary Figure 5, Related to Figure 5

Flow cytometry DNA content profiles of K562#17 transfected with sgRNA-LBR2 in the absence and presence of NU7441 to investigate whether a cell cycle arrest is triggered by Cas9-induced DSB. As a positive control, cells were treated with 10 Gy of ionizing radiation (right-hand column). Time points after Cas9 induction are indicated. Horizontal axis ("FL3-A") indicates DNA content. Bottom row shows indel spectra of the same cell pools at t = 60 h, confirming that DSBs were formed and repaired in these cells.

Supplementary Figure S6



Supplementary Figure 6, Related to Figure 6

Time series experiments of 3 additional loci; an additional sequence in the *LBR* gene (LBR8, A-E), *AAVS1* gene (G-I) and an intergenic region on chromosome 11 (J-L). (A,G,J) Relative fractions of intact (red) and indel (green) as a function of time. The dashed lines show sigmoid fits of the data points. Indel fraction in absence of Shield-1 is shown in grey. *n* indicates the number of time series per locus and error bars represent the SD. (B,H,K) Distribution of the type of indels as determined by high-throughput sequencing at t = 60h. (C, I, L) Nucleotide compositions of particularly abundant indels. (D-E) Broken fraction detection in a time series experiment in presence of sgRNA-LBR8 of the 3' (D) and 5' end (E) of the break. (F) Comparison of measured broken fractions for sgRNA-LBR2 (n=7) and sgRNA-LBR8 (n=5) (average of all 3' and 5' measurements combined).