nature genetics

SUPPLEMENTARY INFORMATION

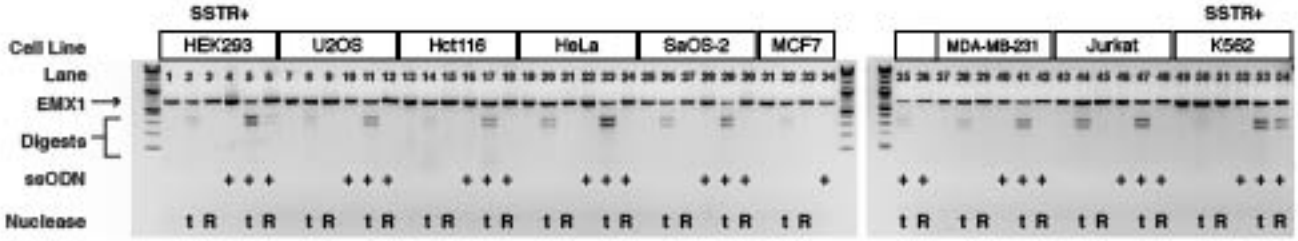
https://doi.org/10.1038/s41588-018-0174-0

In the format provided by the authors and unedited.

CRISPR–Cas9 genome editing in human cells occurs via the Fanconi anemia pathway

Chris D. Richardson[1,2], Katelynn R. Kazane[1,2], Sharon J. Feng[1,2], Elena Zelin[1,2], Nicholas L. Bray[1,2], Axel J. Schäfer[2], Stephen N. Floor[2,3] and Jacob E. Corn[1,2]*

[1]Innovative Genomics Institute, University of California, Berkeley, Berkeley, CA, USA. [2]Department of Molecular and Cell Biology, University of California, Berkeley, Berkeley, CA, USA. [3]Present address: Department of Cell and Tissue Biology, University of California, San Francisco, San Francisco, CA, USA.
*e-mail: jcorn@berkeley.edu

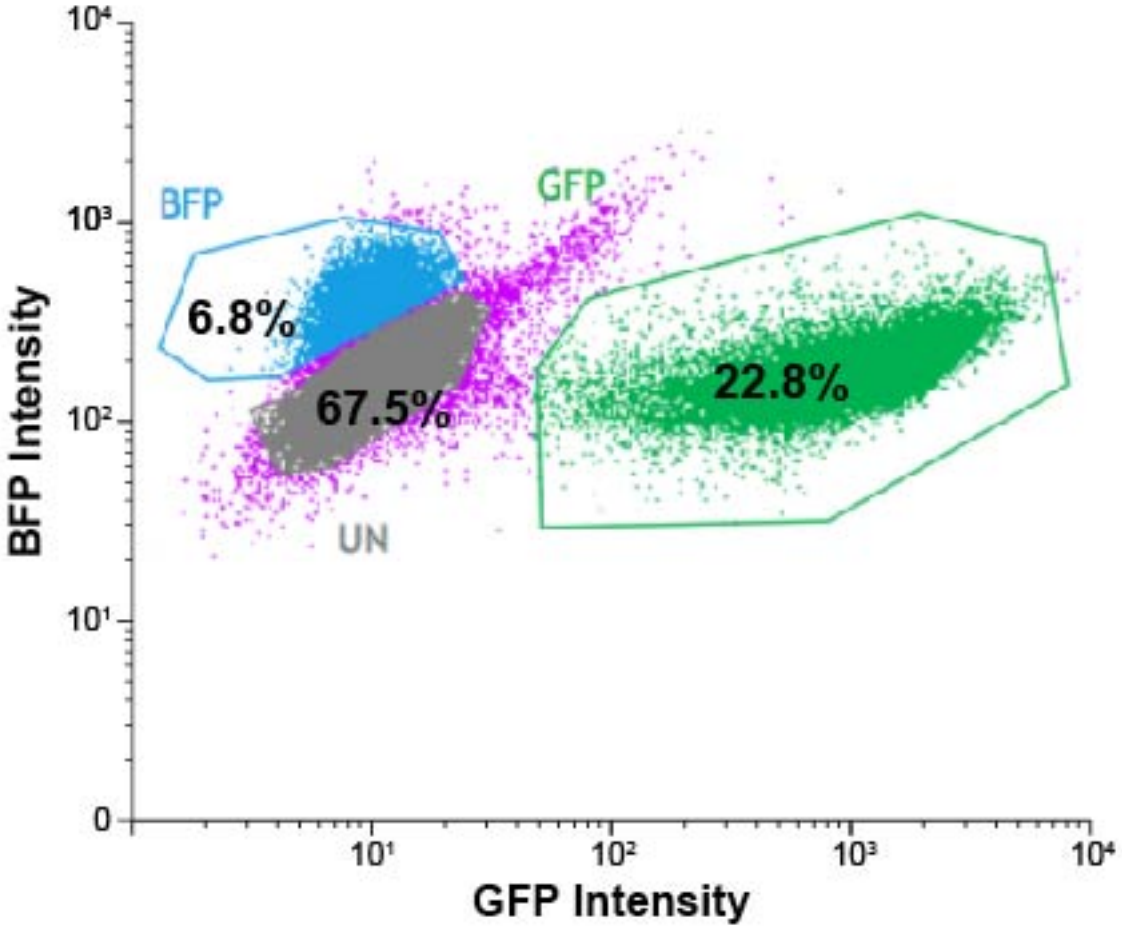


Supplementary Figure 1

SSTR efficiency varies in different human cell lines.

Nine cell lines were edited using RNP targeting the *EMX1* locus either without or with ssODN containing the PciI sequence. The edited locus was amplified, reannealed, and digested using T7 endonuclease I (t), which quantifies gene disruption, or the restriction enzyme PciI (R), which quantifies SSTR. Data presented are representative of independent experiments.

A



B

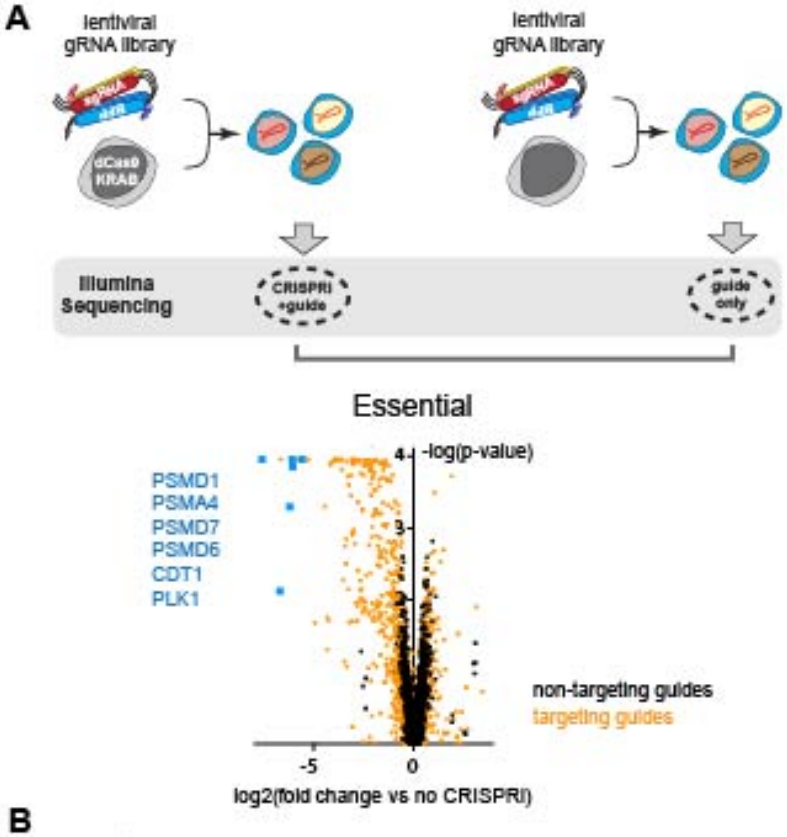
Sample	Sorted Cells
dCas9-KRAB Replicate 1	7,001,595
dCas9-KRAB Replicate 2	7,002,159
K562	7,043,117

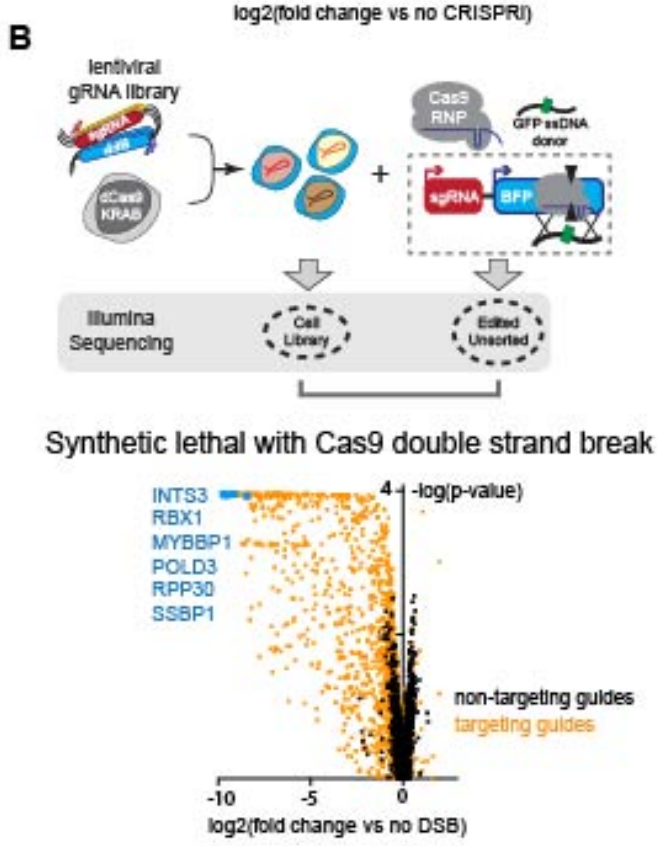
Supplementary Figure 2

BFP, GFP, and non-fluorescent populations are effectively resolved during a pooled CRISPRi screen.

a, Replicate 1 (of *n* = 2 independent experiments) of dCas9-KRAB cells infected with gRNA library and edited at the *BFP* locus. Live single cells were plotted to compare BFP and GFP intensities. Three populations were sorted: $BFP^{+}GFP^{-}$ (BFP), $BFP^{-}GFP^{+}$ (GFP), and

$BFP^{-}GFP^{-}$ (UN). Percentages in each gate are presented for 100,000 cells. **b**, Total cells sorted for each edited population.


A
lentiviral
gRNA library
lentiviral
gRNA library
dCas9
KRAB
Illumina
Sequencing
CRISPRi
+guide
guide
only
Essential
-log(p-value)
PSMD1
PSMA4
PSMD7
PSMD6
CDT1
PLK1
non-targeting guides
targeting guides
-5
0
log2(fold change vs no CRISPRi)

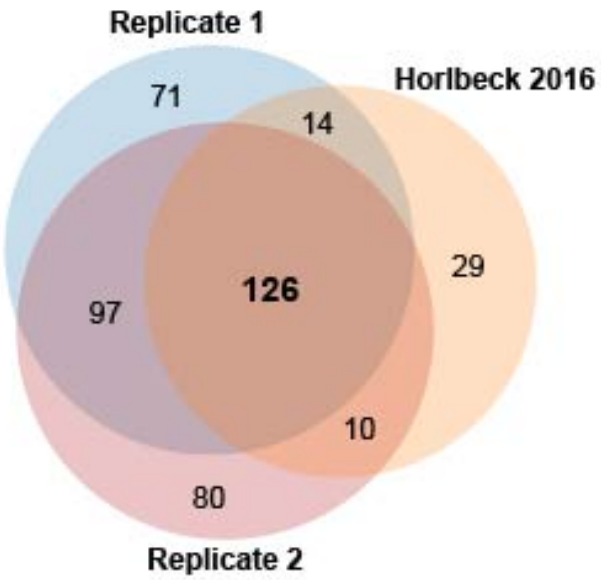

B
lentiviral
gRNA library
dCas9
KRAB
Cas9
RNP
GFP ssDNA
donor
sgRNA
BFP
Illumina
Sequencing
Cell
Library
Edited
Unsorted
Synthetic lethal with Cas9 double strand break
-log(p-value)
INTS3
RBX1
MYBBP1
POLD3
RPP30
SSBP1
non-targeting guides
targeting guides
-10
-5
0
log2(fold change vs no DSB)

Supplementary Figure 3

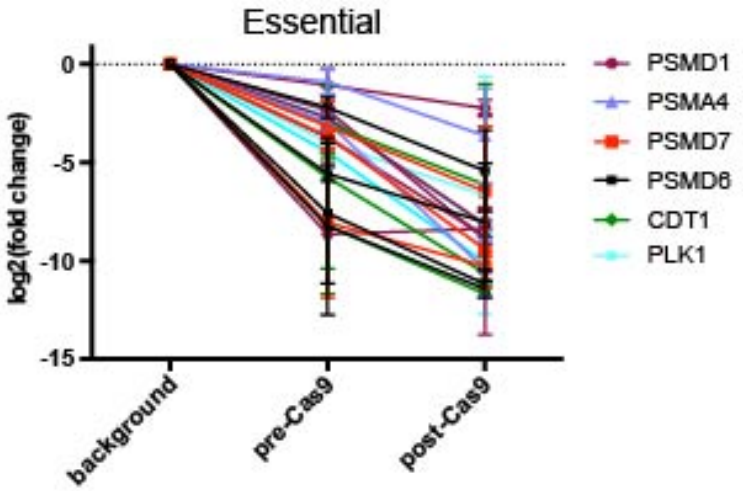
Pooled CRISPRi screens identify DNA repair genes that contribute to multiple phenotypes.

a, Essential genes in the K562 cell background. A volcano plot identifies genes from the DNA repair library enriched or depleted in CRISPRi cells relative to gRNA-only controls. Representative essential genes are highlighted in blue, negative control untargeted gRNAs are shown in black, and targeted gRNAs are shown in orange. Data were generated from $n = 2$ independent experiments. $\log_2$ fold change and Mann–Whitney P values were calculated as described[11]. **b**, Genes synthetic lethal with a DSB. A volcano plot identifies genes from the DNA repair library enriched or depleted in CRISPRi cells treated with Cas9 relative to untreated cell populations. Representative synthetic lethal genes are highlighted in blue, negative control untargeted gRNAs are shown in black, and targeted gRNAs are shown in orange. Data were generated from $n = 2$ independent experiments. $\log_2$ fold change and Mann–Whitney P values were calculated as described[11].

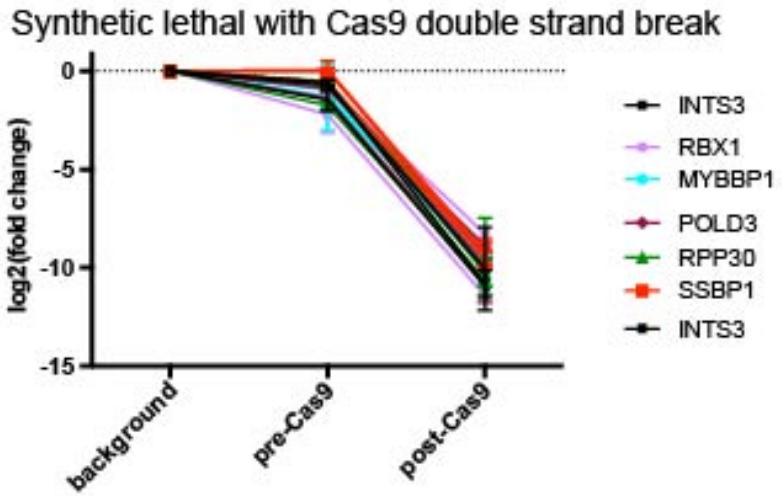
A


Replicate 1
Horlbeck 2016
Replicate 2
71
14
29
97
126
10
80

B


Essential
log2(fold change)
0
-5
-10
-15
background
pre-Cas9
post-Cas9
PSMD1
PSMA4
PSMD7
PSMD6
CDT1
PLK1

C

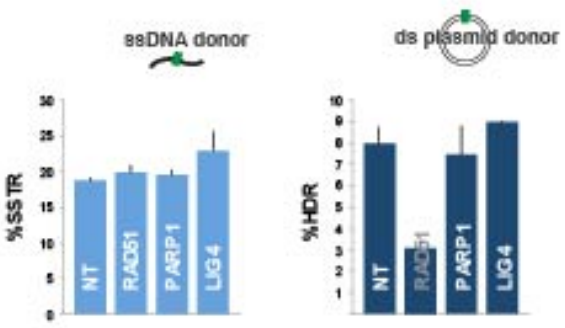

Synthetic lethal with Cas9 double strand break
log2(fold change)
0
-5
-10
-15
background
pre-Cas9
post-Cas9
INTS3
RBX1
MYBBP1
POLD3
RPP30
SSBP1
INTS3

Supplementary Figure 4

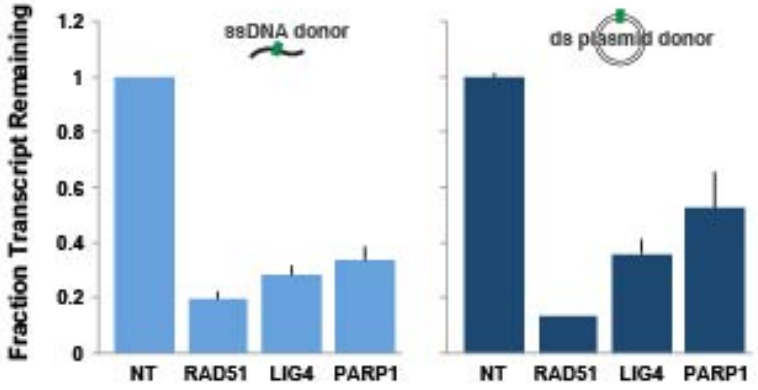
Characterizing non-SSTR phenotypes in pooled screen data.

a, Essential genes identified in this study substantially overlap with previous studies. The distribution of gRNAs in $n = 2$ unedited dCas9-KRAB cell libraries was separately compared to the gRNA distribution in an unedited K562 cell library and gene scores for each target gene were calculated. Essential genes were defined as those showing significant ($P < 0.05$, Mann–Whitney U test) depletion ($\log_2$ (fold change) < –0.5) in the dCas9-KRAB population and compiled into a single list. These essential genes were compared to the Horlbeck K562 dataset[11], filtered for significance ($P < 0.05$) and magnitude of effect (phenotype < –0.2). Overlap between gene lists is presented as a Venn diagram. **b**, Essential genes are progressively lost from the library. Data presented are the mean ± s.d. of $n = 2$ independent experiments. **c**, Genes that are synthetically lethal with a single DSB are abruptly lost from the library following Cas9 treatment. Data presented are the mean ± s.d. of $n = 2$ independent experiments.

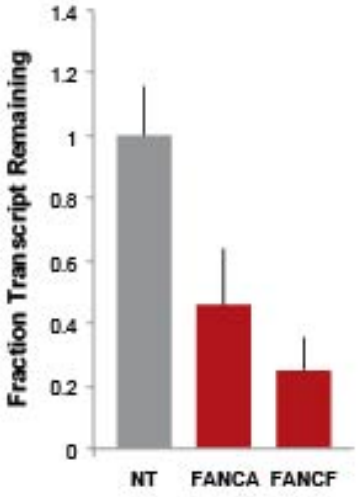
A


ssDNA donor
ds plasmid donor
%SSTR
%HDR
NT
RAD51
PARP1
LIG4

B


Fraction Transcript Remaining
ssDNA donor
ds plasmid donor
1.2
1
0.8
0.6
0.4
0.2
0
NT
RAD51
LIG4
PARP1

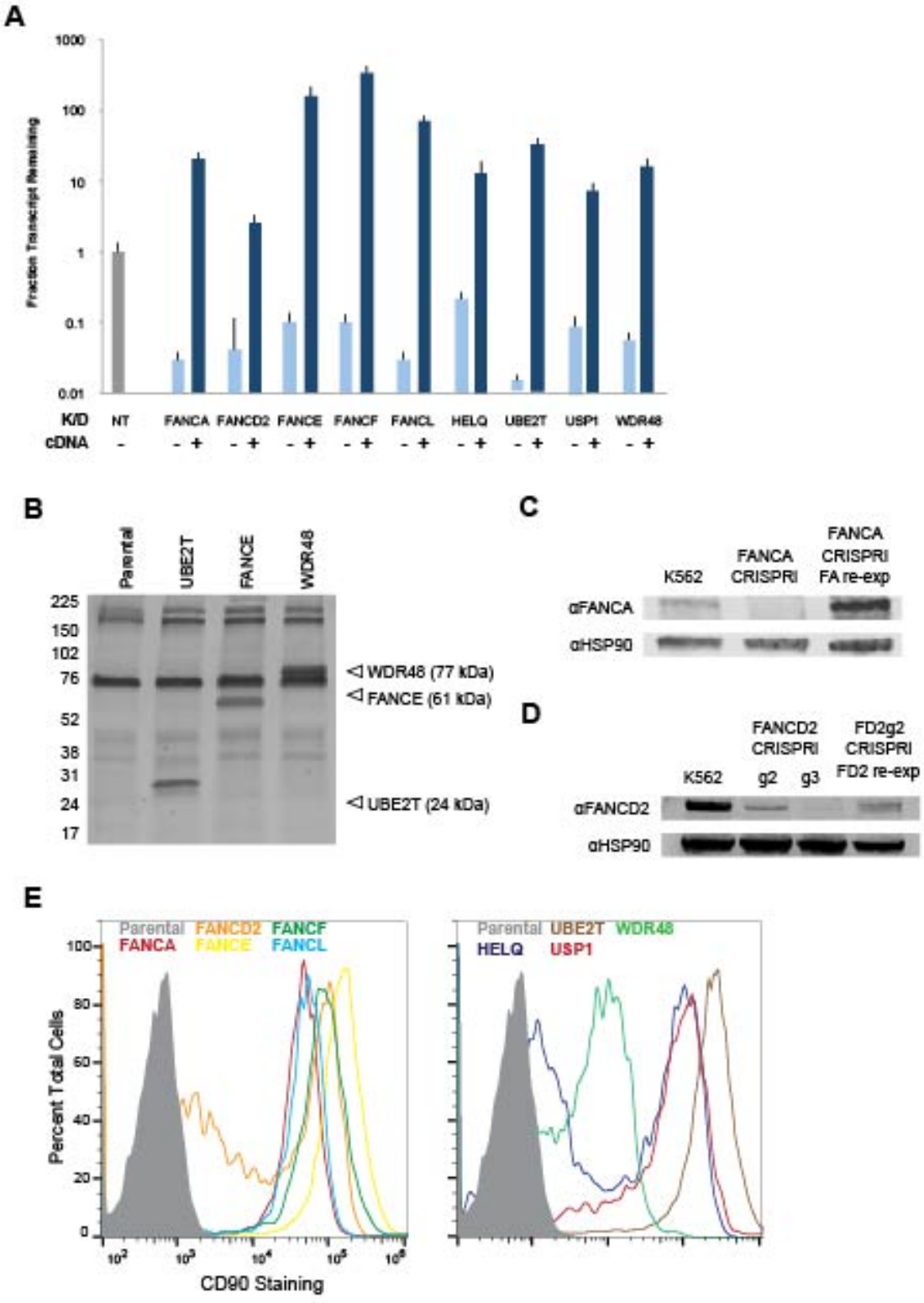
C


Fraction Transcript Remaining
1.4
1.2
1
0.8
0.6
0.4
0.2
0
NT
FANCA
FANCF

Supplementary Figure 5

Effective knockdown of DNA repair factors using siRNA.

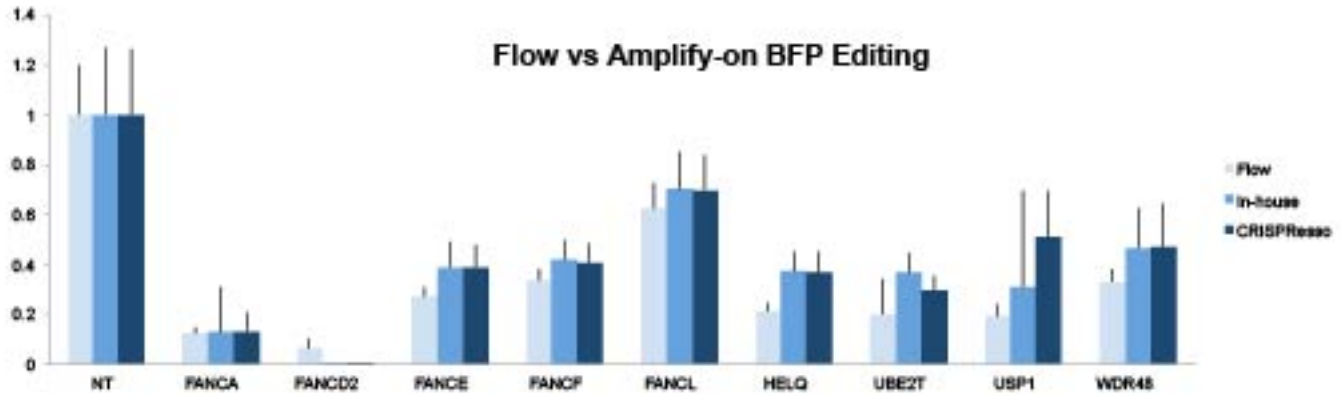
a, Knockdown and editing of DSB repair factors reveals genetic differences between HDR and SSTR. K562 cells expressing a *BFP* reporter were treated with the indicated siRNAs and edited by co-administration of RNP targeting the *BFP* reporter with ssODN or dsDonor containing a *BFP*-to-*GFP* mutation. Data presented are the mean ± s.d. of $n = 2$ independent experiments. **b**, siRNA of *RAD51*, *PARP1*, and *LIG4* in the K562 cells from **a**. Cells were siRNA treated for 48 h prior to editing. Fold depletion of the siRNA-target transcript over controls (*ACTB*, *GAPDH*) was measured by qPCR. Data presented represent the transcriptional state of the cells at the time of editing. The mean ± s.d. was calculated from $n = 2$ independent experiments. **c**, siRNA knockdown of FANCA and FANCF in human dermal fibroblasts. Cells were siRNA treated prior to editing (Fig. 2c) as described in **b**.


A
Fraction Transcript Remaining
1000
100
10
1
0.1
0.01
K/D NT FANCA FANCD2 FANCE FANCF FANCL HELQ UBE2T USP1 WDR48
cDNA - - + - + - + - + - + - + - +
B
Parental
UBE2T
FANCE
WDR48
225
150
102
76
52
38
31
24
17
WDR48 (77 kDa)
FANCE (61 kDa)
UBE2T (24 kDa)
C
K562
FANCA CRISPRI
FANCA CRISPRI FA re-exp
αFANCA
αHSP90
D
K562
FANCD2 CRISPRI g2
g3
FD2g2 CRISPRI FD2 re-exp
αFANCD2
αHSP90
E
Parental
FANCD2
FANCF
FANCA
FANCE
FANCL
Parental
UBE2T
WDR48
HELQ
USP1
Percent Total Cells
CD90 Staining

Supplementary Figure 6

Transcriptional manipulation of DNA repair factors using CRISPRi.

a, Stable CRISPRi cells targeting the indicated gene with or without re-expression of a cDNA of the indicated gene were harvested (Fig. 2a,b) and fold depletion of the indicated transcripts over control transcripts (*ACTB*, *GAPDH*) was measured by qPCR. The mean ± s.d. was calculated from *n* = 2 independent experiments. **b**, Epitope tags reveal robust re-expression of FA proteins. HA blots of parental and cDNA re-expression cell lines are presented along with predicted sizes for re-expressed genes. Data are representative of *n* = 2 independent experiments. **c**, Blotting for FANCA confirms knockdown/re-expression data. Top row, FANCA blot of parental (K562) cells, FANCA knockdown CRISPRi cells, and FANCA knockdown CRISPRi cells with re-expressed FANCA. Bottom row, loading controls for each sample. Data are representative of *n* = 2 independent experiments. **d**, Blotting for FANCD2 confirms knockdown/re-expression data. Top row, FANCD2 blot of parental (K562) cells, two FANCD2 knockdown cell lines, and FANCD2 knockdown g2 with re-expressed FANCD2. Bottom row, loading controls for each sample. Data are representative of *n* = 2 independent experiments. **e**, CD90 abundance validates transcription of cDNA constructs. Re-expression constructs express Thy1.1 co-transcriptionally with the cDNA, which was monitored by live-cell flow cytometry. Data are representative of *n* = 2 independent experiments.

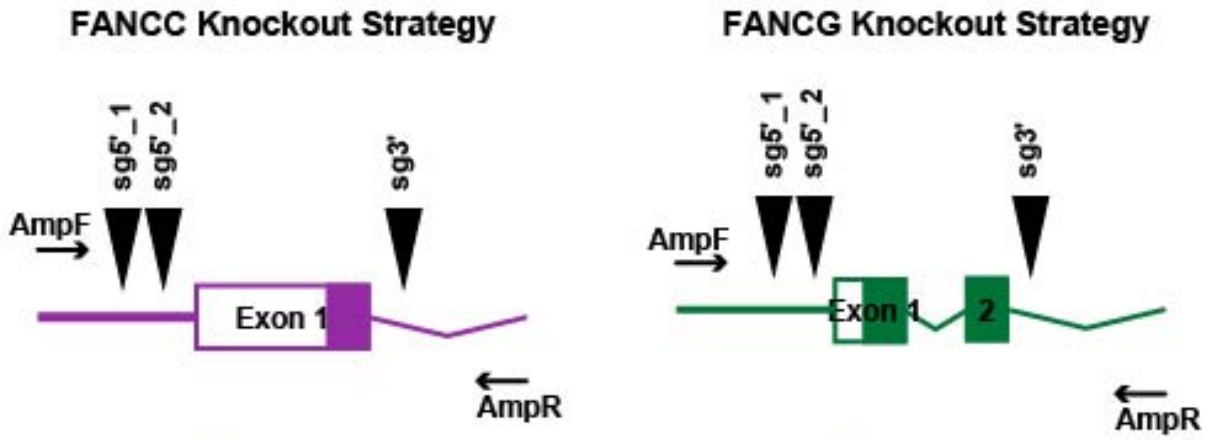


Supplementary Figure 7

Flow cytometry and amplicon sequencing produce similar editing rates.

Editing at the *BFP* locus was quantified by flow cytometry (light blue), amplicon sequencing followed by in-house bioinformatic analysis (medium blue), or amplicon sequencing followed by CRISPRessoPOOL (Pinello, L. et al. Analyzing CRISPR genome-editing experiments with CRISPResso. *Nat. Biotechnol.* **34**, 695–697 (2016)) analysis (dark blue). The mean ± s.d. was generated from *n* = 2 independent experiments.

A



Guide Pair 1 (●)

Name	Sequence
FANCC-sg5'_1	CACCCATCAATGGGTCGTGA
FANCC-sg3'	GACTGAAGTGCCGATGGATT

Guide Pair 2 (●)

Name	Sequence
FANCC-sg5'_2	TCCCTCCTTCACCCATCAAT
FANCC-sg3'	GACTGAAGTGCCGATGGATT

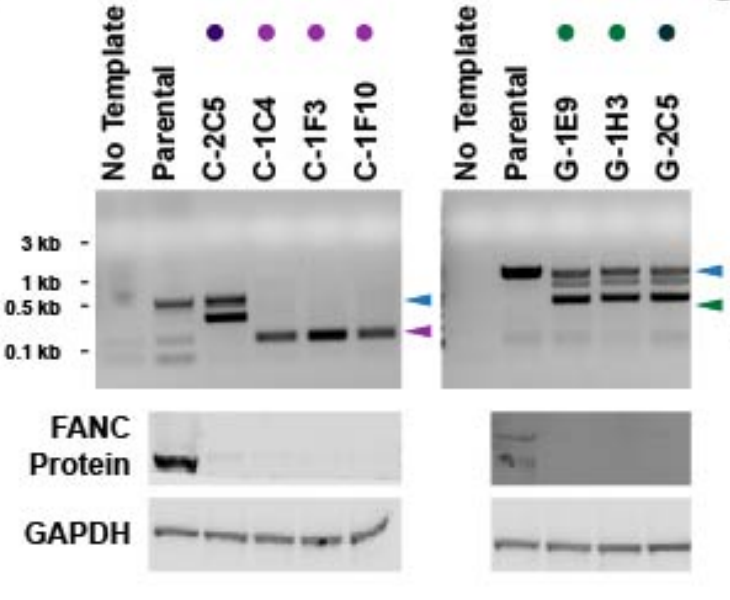
Guide Pair 1 (●)

Name	Sequence
FANCG-sg5'_1	ACCCTTCCTAAGTCCGCTTC
FANCG-sg3'	GGGATCGATATAATTACTCG

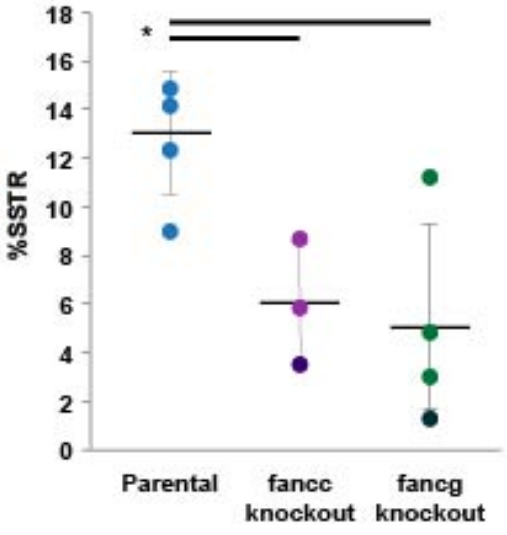
Guide Pair 2 (●)

Name	Sequence
FANCG-sg5'_2	CTAAGTCCGCTTCTGGTCTC
FANCG-sg3'	GGGATCGATATAATTACTCG

B



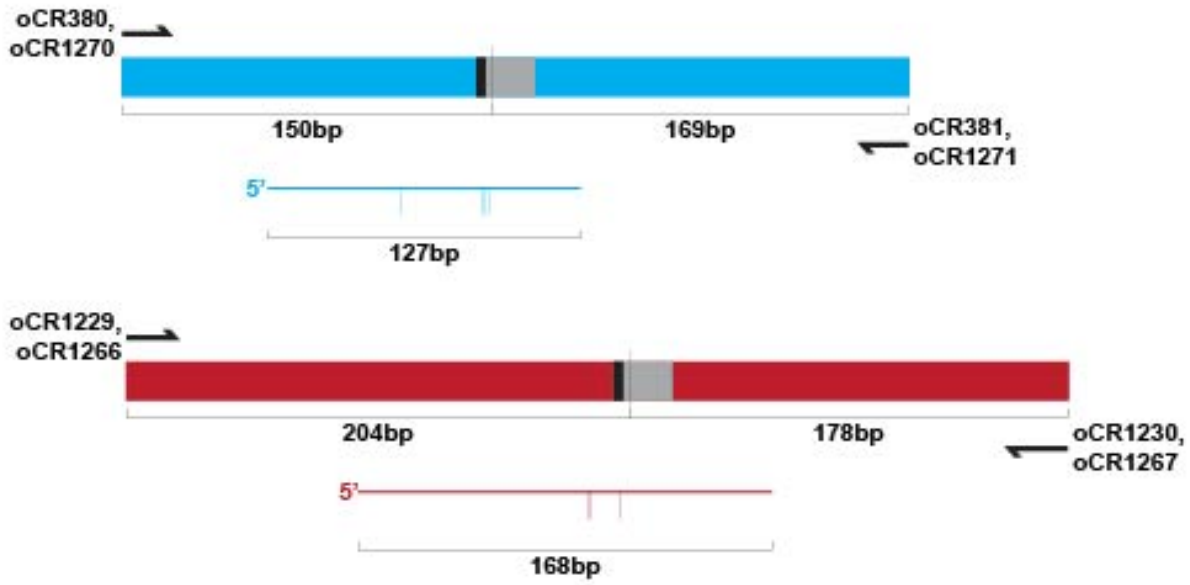
C



Supplementary Figure 8

Knockout of *FANCC* and *FANCG* disrupts SSTR.

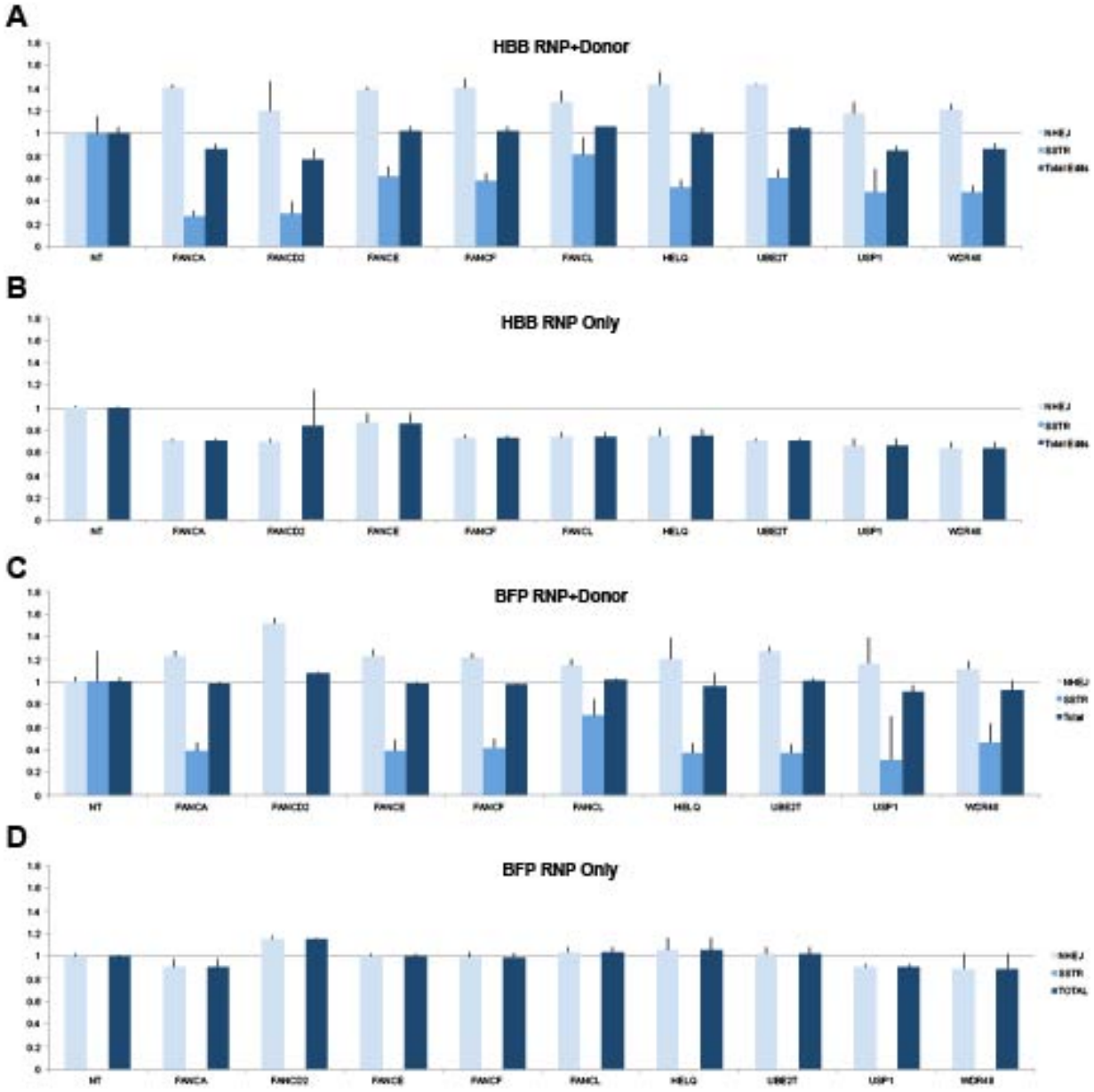
a, Editing strategy used to make clonal knockouts of *FANCC* and *FANCG*. Parental cells were edited with paired nucleases targeting the indicated sequences. Knockout was verified using amplification primers outside the edited locus. **b**, Paired cutting with CRISPR–Cas9 effectively removes the first exon of *FANCC* and *FANCG*. Diagnostic PCR (top) and western blotting (bottom) was performed to verify exon removal and elimination of gene product for *FANCC* or *FANCG* in clonal cell lines. Predicted sizes for genomic PCR amplifying the first exon of *FANCC* or *FANCG* (blue triangle; *FANCC*, 600 bp; *FANCG*, 950 bp) or for predicted excision outcomes (*FANCC*, 190 bp, maroon arrow; *FANCG*, 470 bp, green arrow) are presented alongside gels. Data are representative of $n = 2$ independent experiments. **c**, Cells lacking *FANCC* or *FANCG* show diminished levels of SSTR. SSTR levels were measured in multiple, individually cloned *FANCC* ($n = 4$) or *FANCG* ($n = 3$) knockout cell lines and plotted alongside values for clonally derived wild-type cell lines ($n = 5$). The editing rate for each sample (circle) is presented along with the center value (mean, black bar) and s.d. (error bars) for each group. $^*P < 0.05$ (Welch's *t* test).



Supplementary Figure 9

Schematic of amplicon sequencing used at the *BFP* or *HBB* loci.

Protospacer (gray) and PAM (black) are presented within amplified sequence. ssODN size and orientation are also displayed. Vertical hash marks indicate locations where ssODN and genomic sequences differ.

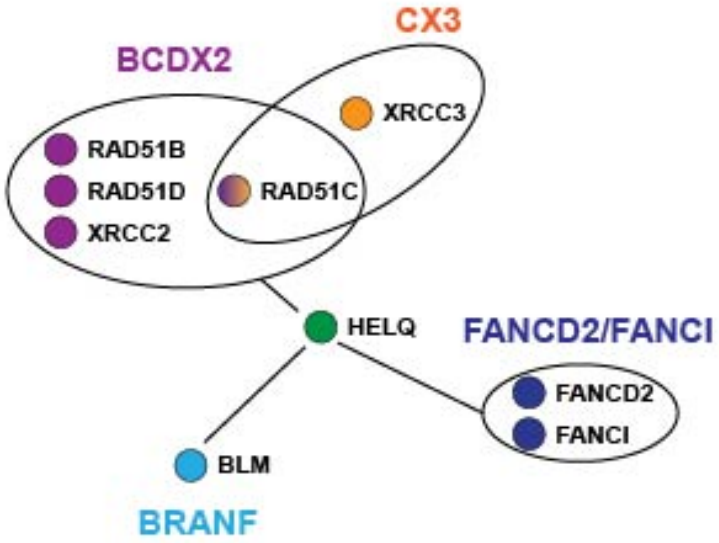


Supplementary Figure 10

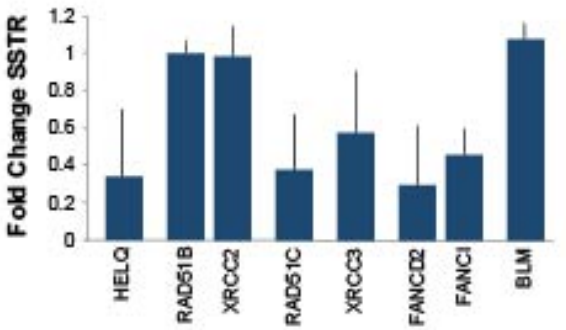
Total editing remains consistent when SSTR is disrupted.

Indel (NHEJ, light blue), SSTR (medium blue), and total (dark blue) editing events were quantified for the indicated CRISPRi cell lines and normalized to untargeted controls. **a**, RNP targeting the *HBB* locus with donor ssODN. **b**, RNP targeting the *HBB* locus without ssODN. **c**, RNP targeting the *BFP* locus with donor ssODN. **d**, RNP targeting the *BFP* locus without ssODN. Data presented are the mean ± s.d. of $n = 2$ independent experiments.

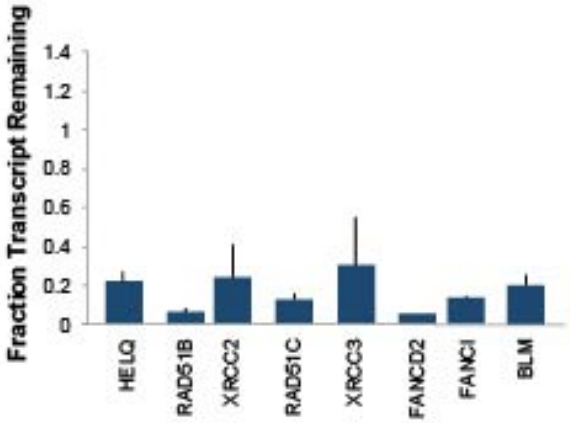
A


CX3
BCDX2
XRCC3
RAD51B
RAD51D
XRCC2
RAD51C
HELQ
FANCD2/FANCI
FANCD2
FANCI
BLM
BRANF

B


Fold Change SSTR
1.2
1
0.8
0.6
0.4
0.2
0
HELQ
RAD51B
XRCC2
RAD51C
XRCC3
FANCD2
FANCI
BLM

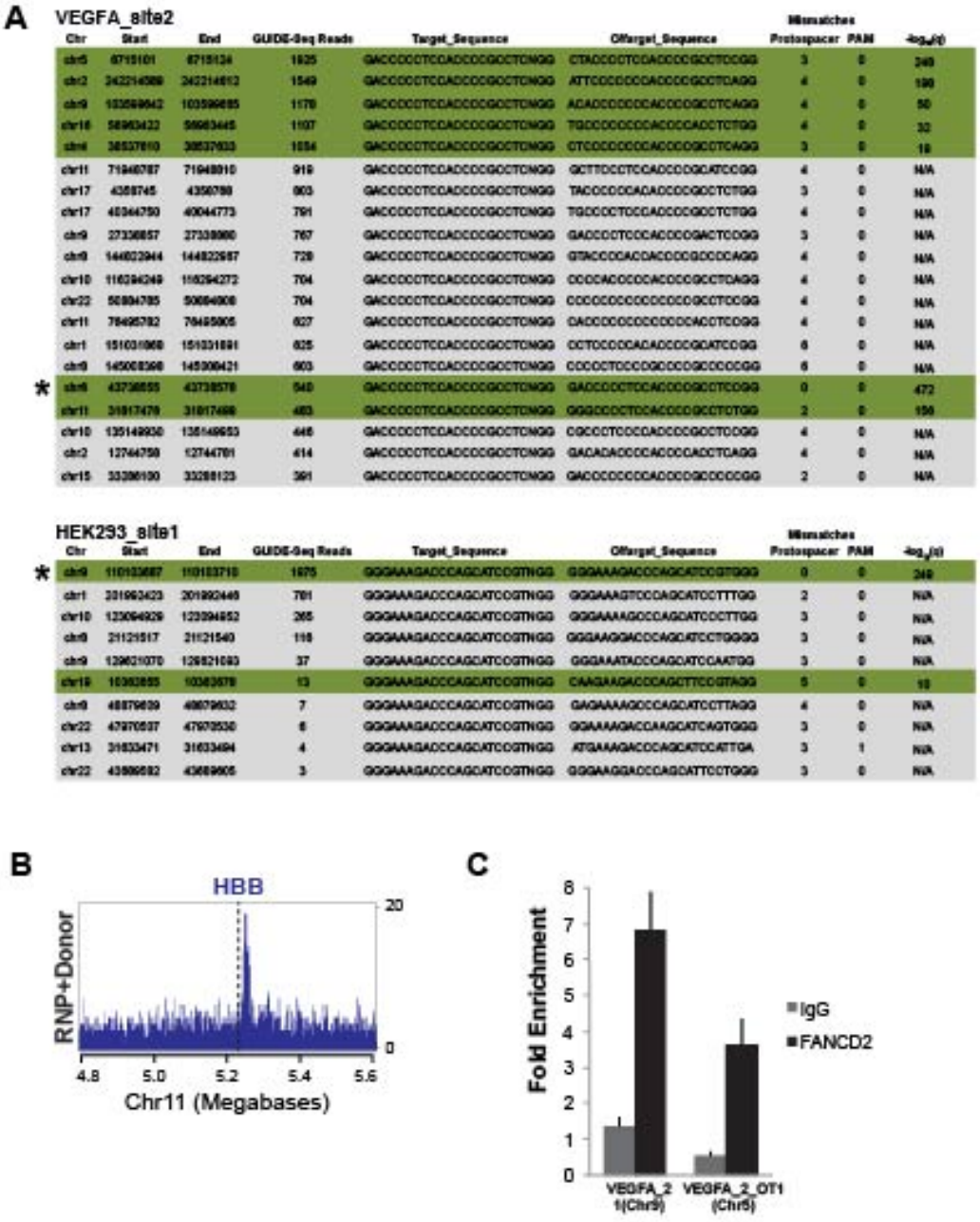
C


Fraction Transcript Remaining
1.4
1.2
1
0.8
0.6
0.4
0.2
0
HELQ
RAD51B
XRCC2
RAD51C
XRCC3
FANCD2
FANCI
BLM

Supplementary Figure 11

HELQ interaction partners have a role in SSTR.

a, An interaction map of HELQ reproduced from an earlier study (Adelman, 2013) illustrates direct physical interactions between HELQ and other complexes and reported interactions from the BIOGRID, STRING and MINT databases. **b**, HELQ interaction partners have a role in SSTR. K562 cells were siRNA or CRISPRi treated against the indicated target genes prior to editing at the *BFP* locus. Data presented are the mean ± s.d. of $n = 2$ independent experiments. **c**, HELQ interaction partners can be effectively depleted by siRNA or CRISPRi. Fold depletion of the target transcript over controls (*ACTB*, *GAPDH*) was measured by qPCR. The mean ± s.d. was calculated from $n = 2$ independent experiments.



Supplementary Figure 12

FANCD2 ChIP-seq identifies on- and off-target Cas9 cut sites.

a, FANCD2 ChIP-seq overlaps with Guide-seq[36] sites. The table describes the top 20 cut sites for the VEGFA_site2 sgRNA or the 10 annotated cut sites for the HEK293_site1 sgRNA. Peaks identified by MACS2 for *n* = 2 FANCD2 ChIP-seq datasets (green highlighting) are presented along with calculated *q* values. **b**, FANCD2 is enriched at the *HBB* cut site in HEK293T cells. Presented data are ChIP-seq pileups generated from comparing RNP + donor (treatment; *n* = 2 independent experiments) and Apo (control; *n* = 2 independent experiments) samples. **c**, FANCD2 is enriched at VEGFA_site2 on- and off-target loci in HEK293T cells. Presented data are ChIP-qPCR fold enrichments generated from non-specific IgG ChIP and FANCD2 ChIP at the indicated loci. The mean ± s.d. was calculated

from n = 2 independent experiments.