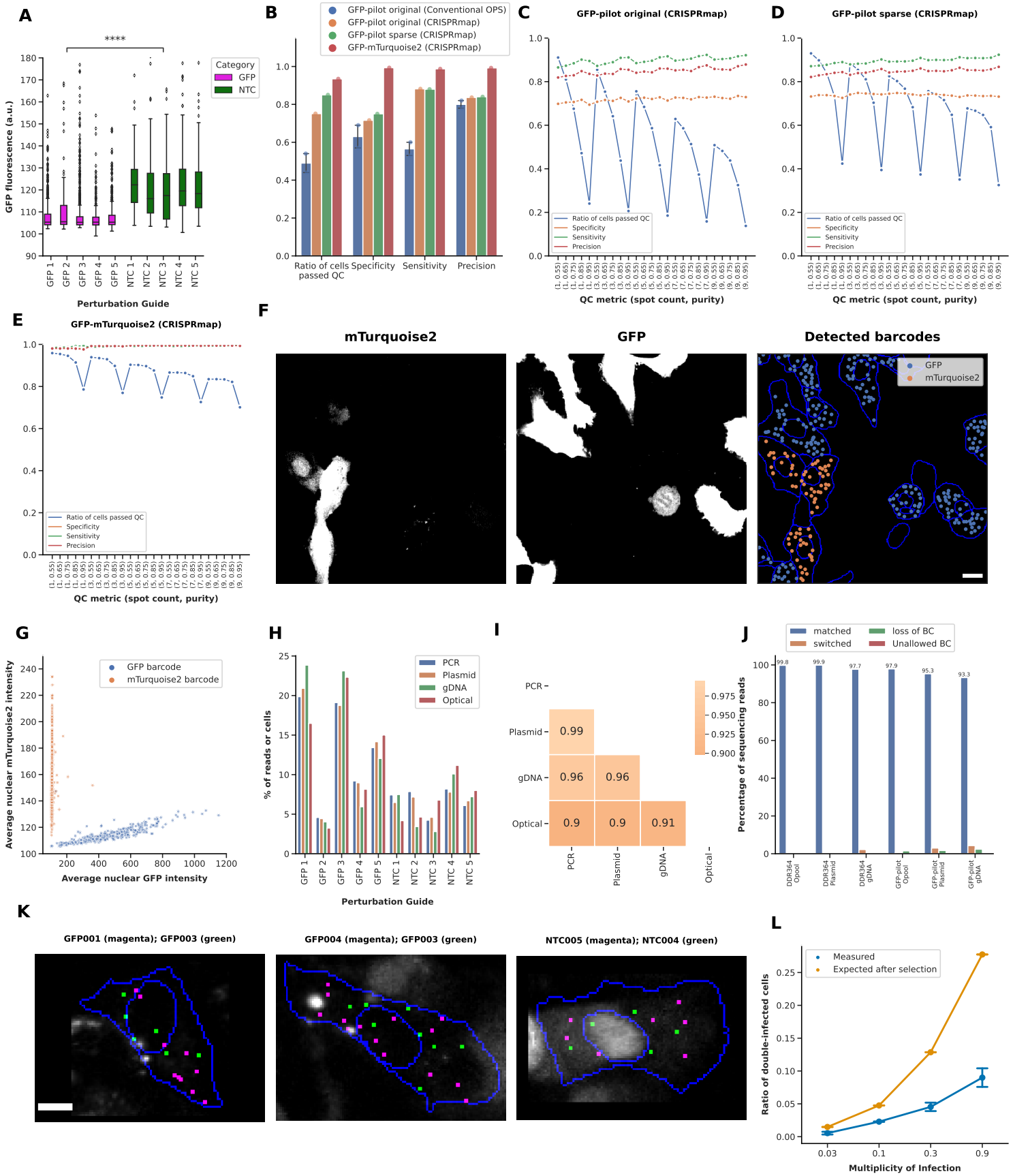
nature biotechnology

Supplementary information

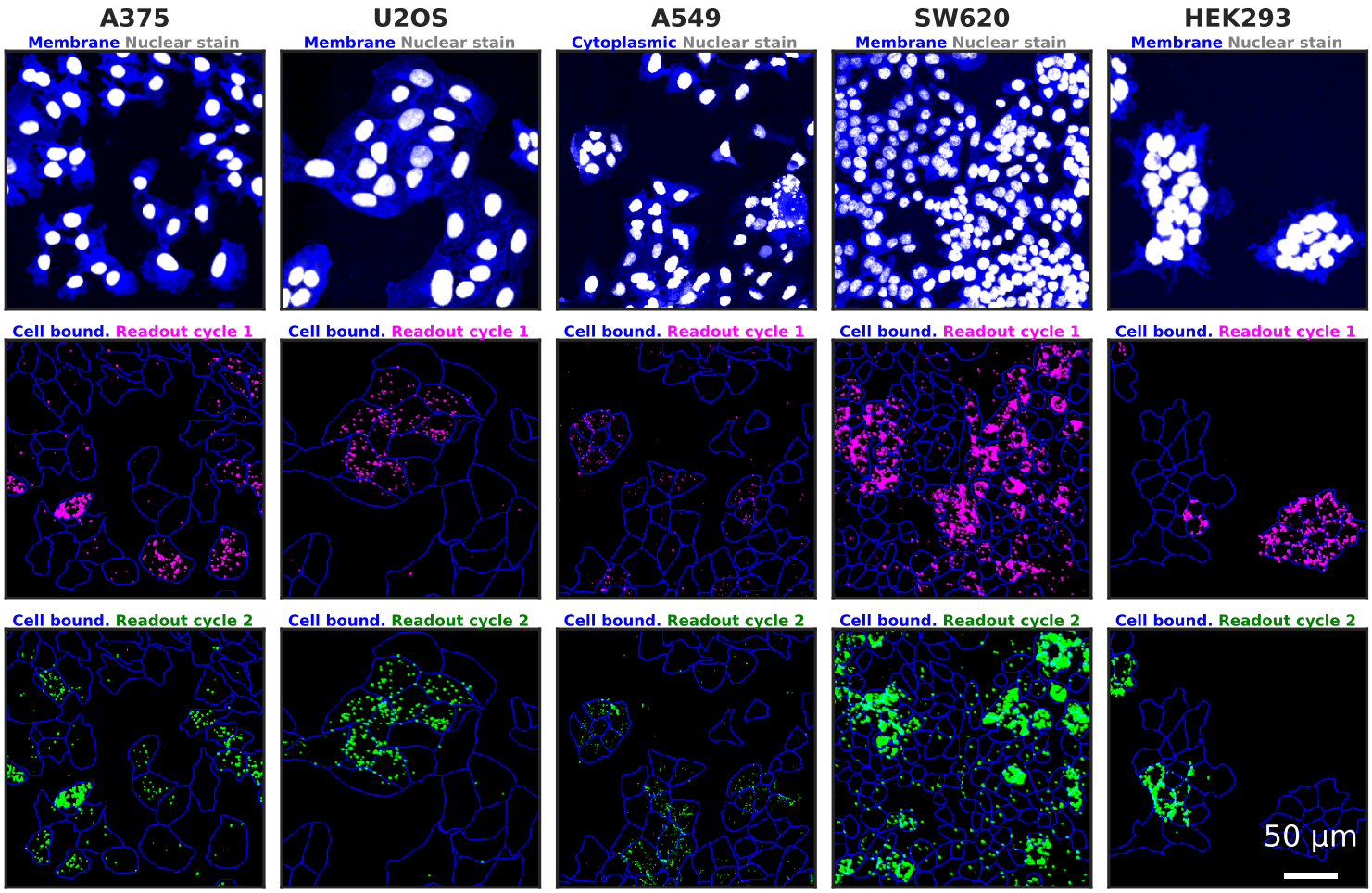
https://doi.org/10.1038/s41587-024-02386-x

Mapping multimodal phenotypes to perturbations in cells and tissue with CRISPRmap

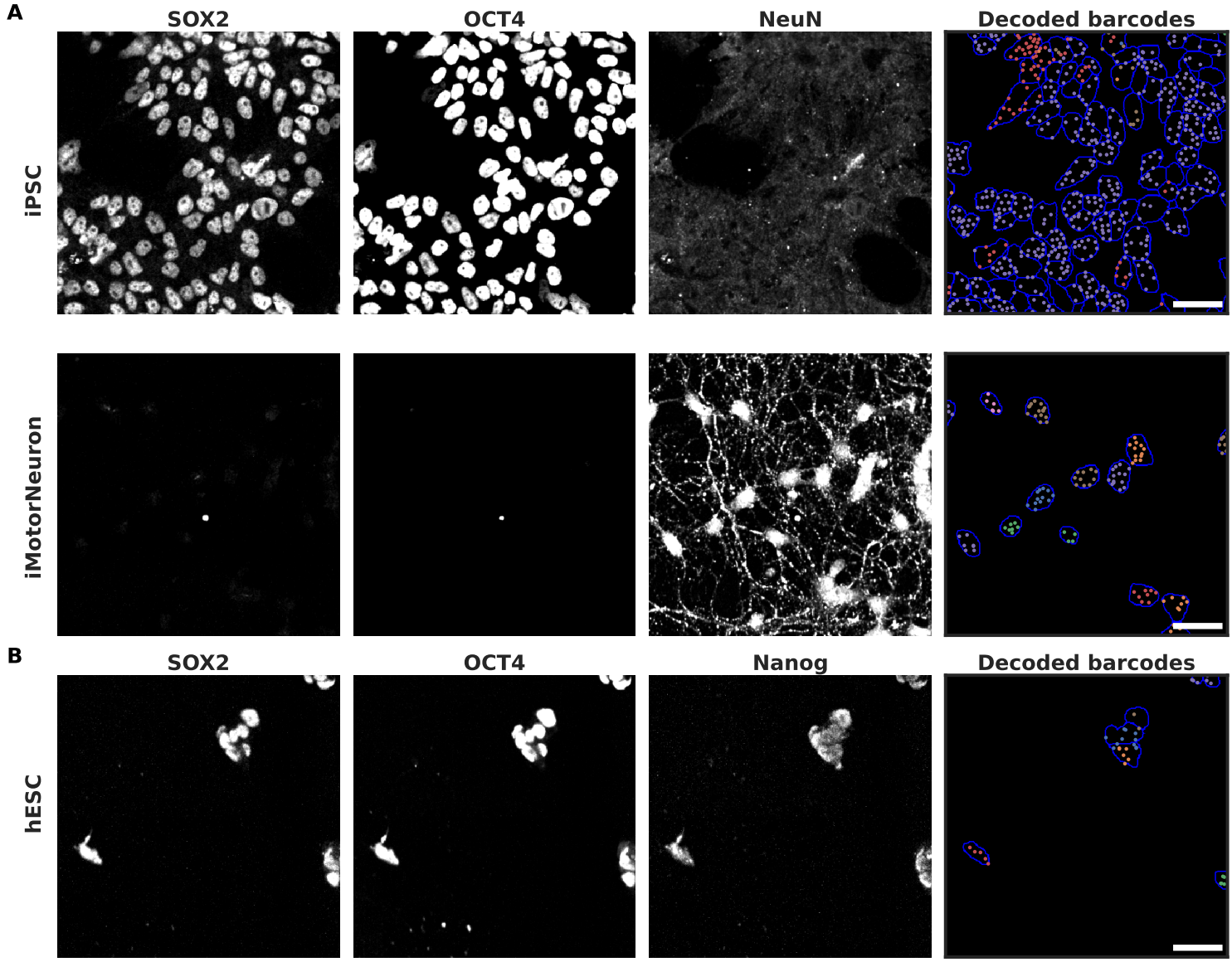
In the format provided by the authors and unedited


A
****
Category
GFP
NTC
GFP fluorescence (a.u.)
GFP 1
GFP 2
GFP 3
GFP 4
GFP 5
NTC 1
NTC 2
NTC 3
NTC 4
NTC 5
Perturbation Guide
B
GFP-pilot original (Conventional OPS)
GFP-pilot original (CRISPRmap)
GFP-pilot sparse (CRISPRmap)
GFP-mTurquoise2 (CRISPRmap)
Ratio of cells passed QC
Specificity
Sensitivity
Precision
C
GFP-pilot original (CRISPRmap)
Ratio of cells passed QC
Specificity
Sensitivity
Precision
QC metric (spot count, purity)
D
GFP-pilot sparse (CRISPRmap)
Ratio of cells passed QC
Specificity
Sensitivity
Precision
QC metric (spot count, purity)
E
GFP-mTurquoise2 (CRISPRmap)
Ratio of cells passed QC
Specificity
Sensitivity
Precision
QC metric (spot count, purity)
F
mTurquoise2
GFP
Detected barcodes
GFP
mTurquoise2
G
GFP barcode
mTurquoise2 barcode
Average nuclear mTurquoise2 intensity
Average nuclear GFP intensity
H
PCR
Plasmid
gDNA
Optical
% of reads or cells
Perturbation Guide
I
PCR
Plasmid
gDNA
Optical
0.99
0.96
0.96
0.9
0.9
0.91
J
matched
switched
loss of BC
Unallowed BC
Percentage of sequencing reads
99.8
99.9
97.7
97.9
95.3
93.3
DDR364 Opool
DDR364 Plasmid
DDR364 gDNA
GFP-pilot Opool
GFP-pilot Plasmid
GFP-pilot gDNA
K
GFP001 (magenta); GFP003 (green)
GFP004 (magenta); GFP003 (green)
NTC005 (magenta); NTC004 (green)
L
Measured
Expected after selection
Ratio of double-infected cells
Multiplicity of Infection

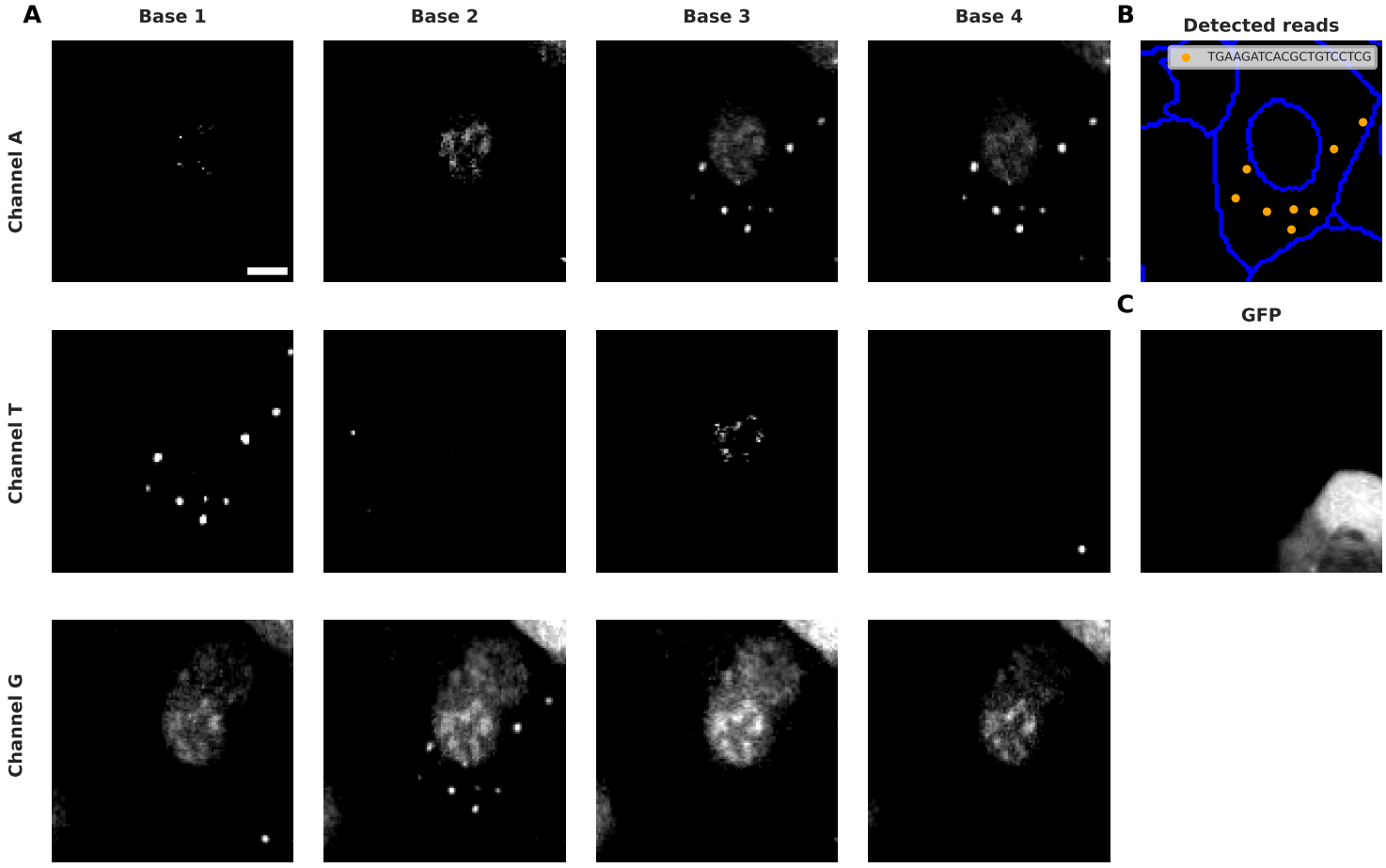
Supplementary Figure 1. Performance of the GFP-targeting pooled optical screen. (A) Genotype-phenotype mapping in cells transduced with the GFP-pilot library. Each GFP-NTC pair was tested, the least significant pair is shown (p = 1.84e-10). Two-sided Mann-Whitney test, *p $<$ 0.05, **p $<$ 0.01, ***p $<$ 0.001, *****p $<$ 0.0001. n = 1 with 4620 cells. Boxes indicate the median and interquartile range (IQR) with whiskers extending 1.5× IQR past the upper and lower quartiles. **(B)** Performance of Conventional OPS on cells with the GFP-pilot library profiled at the original cell density (n = 2, technical replicates, 8104 and 7671 cells), CRISPRmap on cells with GFP-pilot library at the original density (n = 1, 4620 cells) and sparsely seeded (n = 1, 3429 cells), and CRISPRmap on GFP/mTurquoise2-expressing cells with FP-reporting barcodes (GFP/mTurquoise2-reporting cells) sparsely seeded (n = 1, 920 cells). Data are presented as mean values +/- 95% CI. **(C)** Performance of CRISPRmap on cells with GFP-pilot library at the original density under loose to tight QC metrics displayed on the x-axis as (max spot cutoff, purity cutoff). **(D)** As in C) for CRISPRmap on sparsely seeded cells with GFP-pilot library. **(E)** As in C) for CRISPRmap on sparsely seeded GFP/mTurquoise2-reporting cells. **(F)** Genotype-phenotype mapping on GFP/mTurquoise2-reporting cells. Raw mTurquoise2 (left) and GFP (middle) fluorescence are displayed in greyscale. Detected barcodes (right) are shown as spots with cell boundaries outlined in blue. Scale bar, 20 μm. **(G)** Quantification of the average GFP and mTurquoise2 intensity under each nucleus mask colored by barcode identity. **(H)** Quantification of the relative guide abundance across PCR-amplified oligo pool (PCR), plasmid pool (Plasmid), genomic DNA (gDNA) of transduced cells, and optical screens (Optical). **(I)** Pearson correlation of relative guide abundance at PCR, Plasmid, gDNA, and Optical level. **(J)** Quantification of sgRNA-barcode recombination for the DDR364 and GFP-pilot library at PCR, Plasmid and gDNA level. **(K)** Visualization of double-transduced cells. Decoded barcodes of the most (magenta) and second most (green) representing guides in each cell are shown as spots. Raw GFP fluorescence is displayed in greyscale. Cell and nuclear boundaries are outlined in blue. Scale bar, 20 μm. **(L)** Quantification of expected and optically measured ratio of double-transduced cells under different Multiplicity of Infection (MOI). Data are presented as mean values +/- 95% CI. n = 2, technical replicates for each MOI.



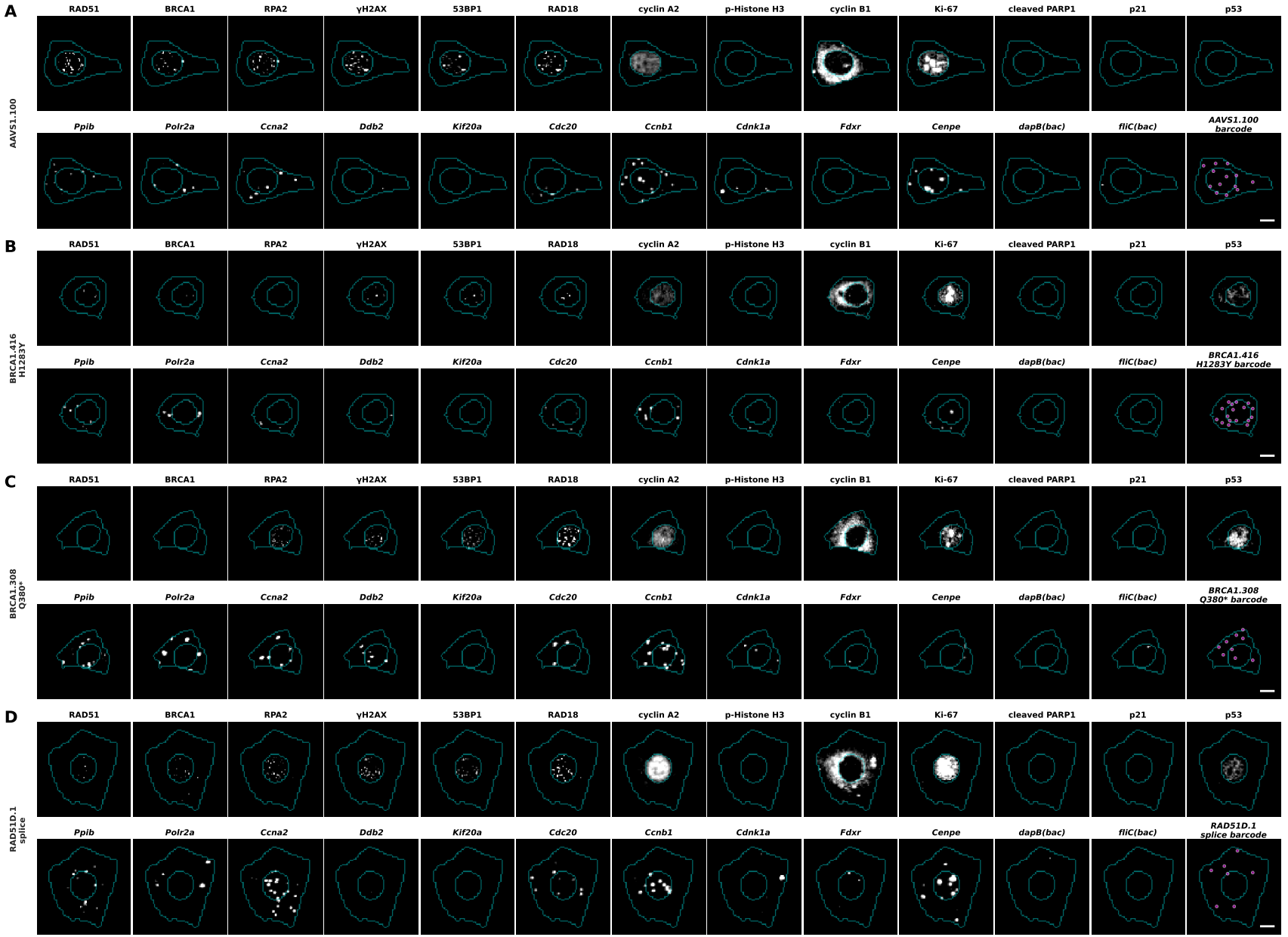
Supplementary Figure 2. CRISPRmap application on multiple cell lines. Visualization of barcode detection on five cell lines: A735, U2OS, A549, SW620, and HEK293 (left to right). The morphology of each cell line is shown (top row) with membrane or cytoplasmic stains (blue) and nuclear stain (white). The fluorescence signal from two CRISPRmap imaging cycles are shown in magenta (middle row) and green (bottom row). Cell boundaries are outlined in blue. Scale bar, 50 μm.



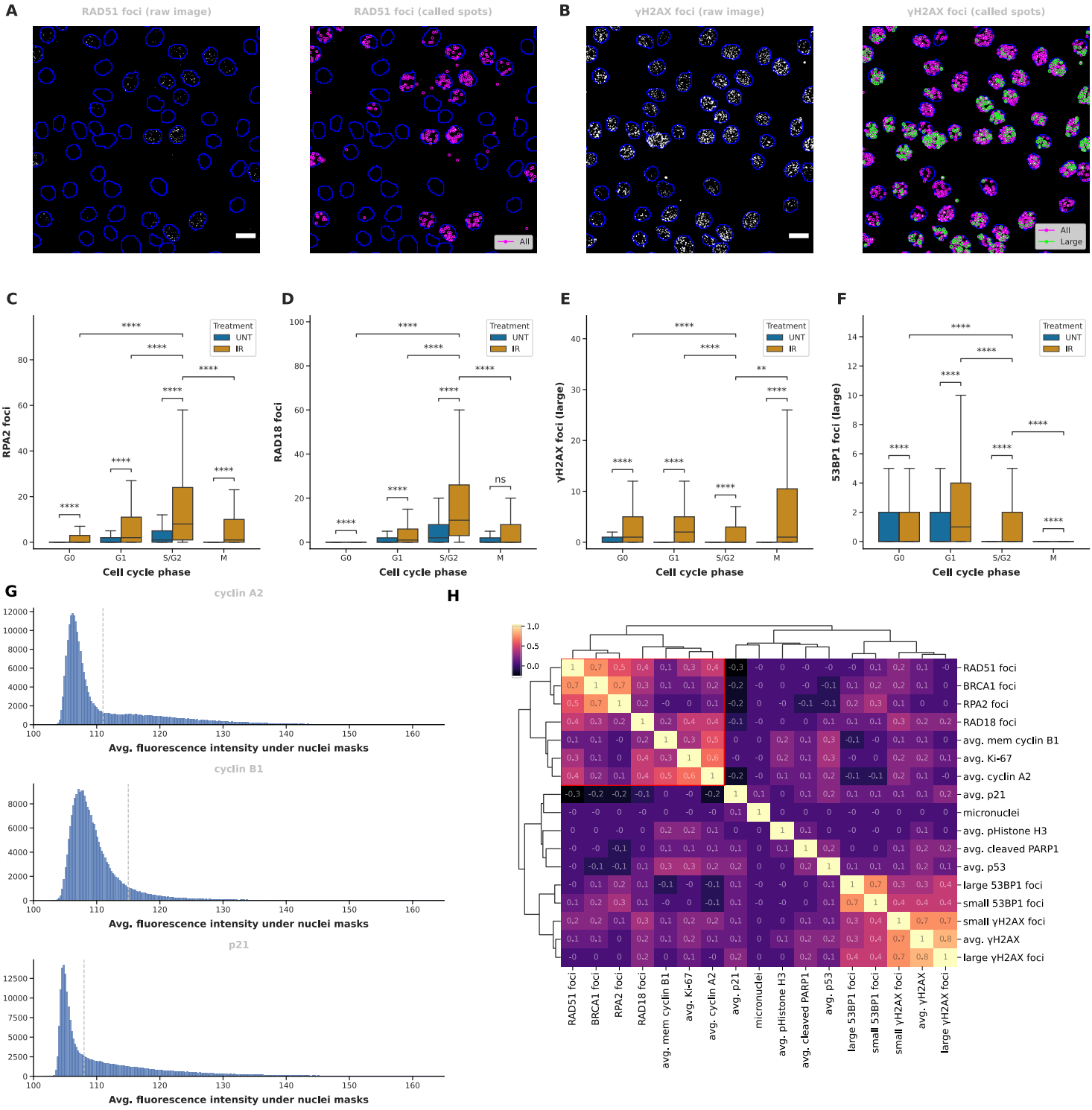
Supplementary Figure 3. CRISPRmap application on multiple cell lines with cell type validation by immunofluorescence. **(A)** Visualization of the cell type marker expression (SOX2: stem cell marker, OCT4: stem cell marker, NeuN: neuronal marker) and detected CRISPRmap barcodes (from left to right) on Induced pluripotent stem cells (iPSCs, top row) and iPSC-derived motor neurons (iMotorNeuron, bottom row). Raw fluorescence signal from the cell type markers were displayed in greyscale. Detected CRISPRmap barcodes were shown as false-colored spots with each color corresponding to a particular barcode, and cell boundaries outlined in blue. Scale bar, 20 μm. **(B)** As in A) for cell type markers (SOX2: stem cell marker, OCT4: stem cell marker, Nanog: stem cell marker) and detected barcodes on human embryonic stem cells (hESCs).



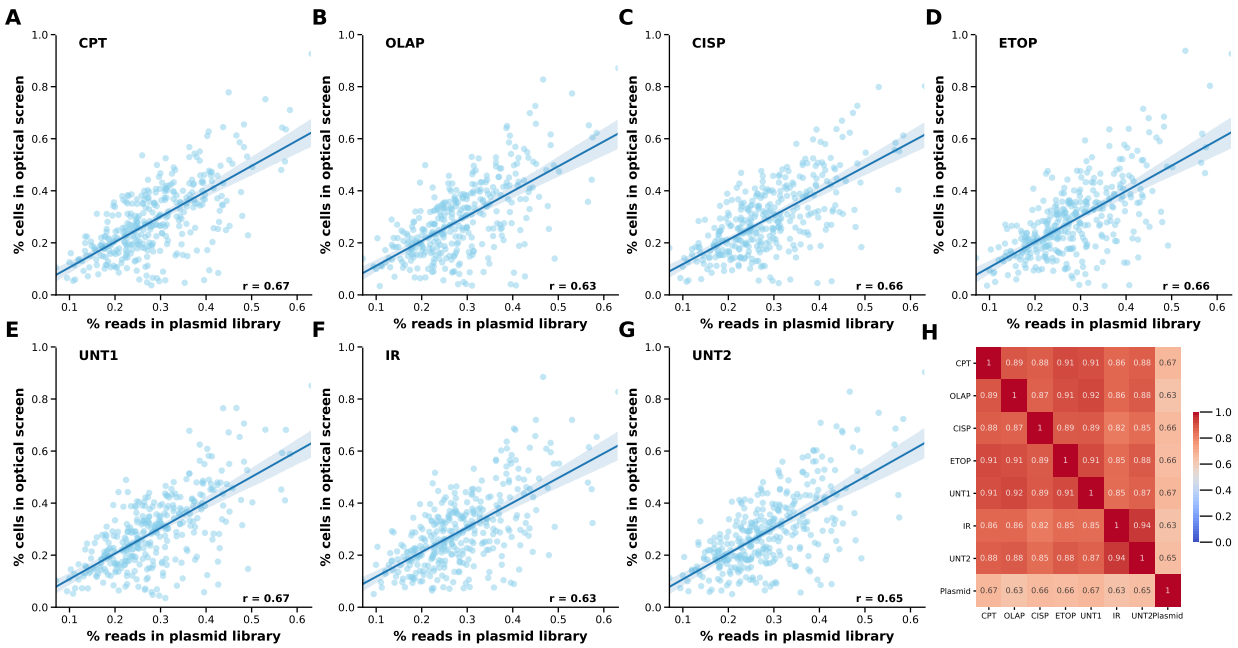
Supplementary Figure 4. Conventional optical pooled screen (OPS) application on HT1080 cells transduced with the GFP-pilot library. (A) Visualization of the guide sequence readout by in situ sequencing, showing a cell with a GFP-targeting guide ("TGAAGATCACGCTGTCCTCG"). Raw fluorescence signal over four readout rounds and three fluorescence channels are displayed in greyscale. Scale bar, 20 μm. **(B)** As in A) for decoded reads, which are displayed as false-colored spots. Cell and nucleus boundaries are outlined in blue. **(C)** As in A) for raw GFP fluorescence shown in gray-scale.



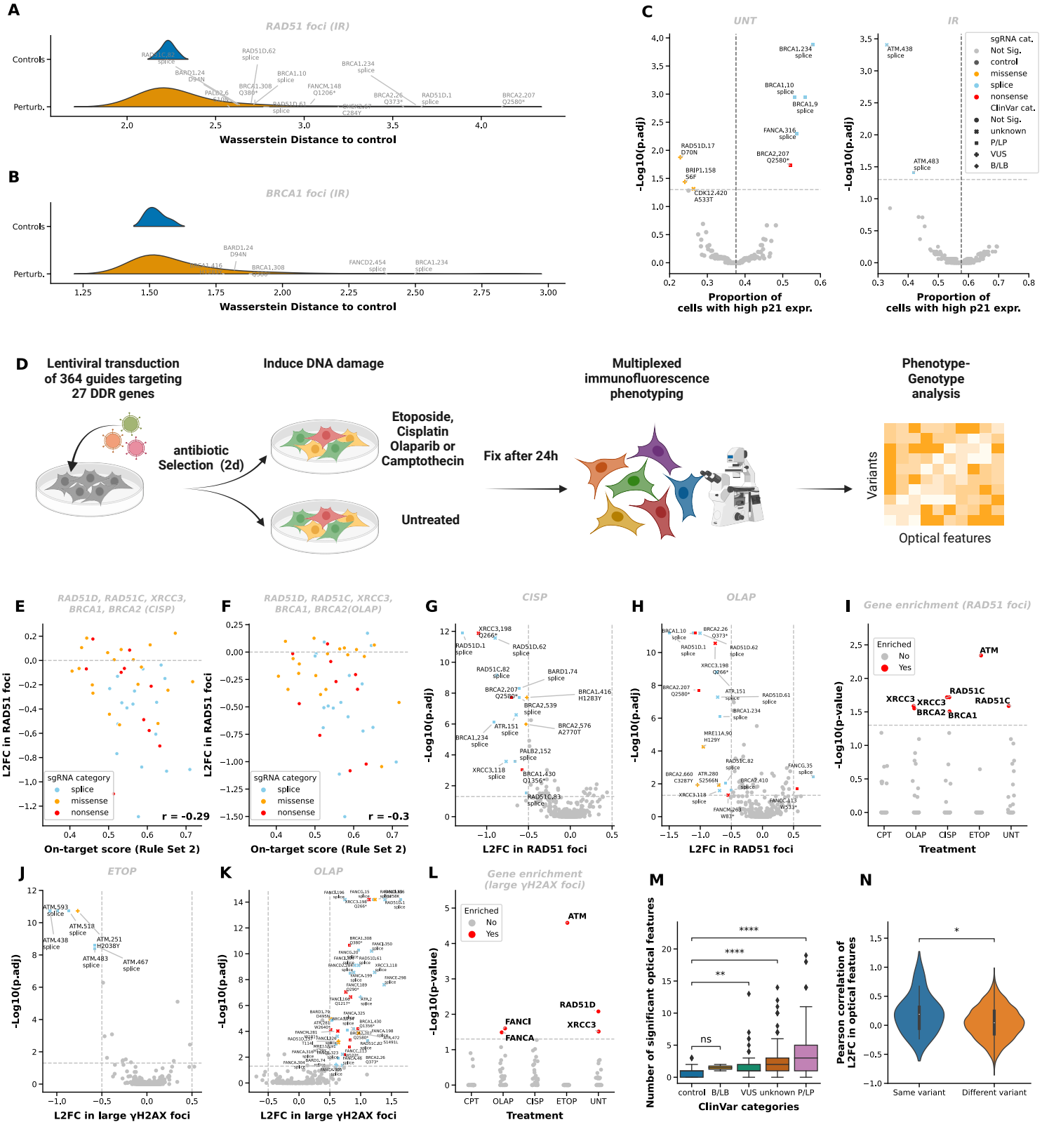
Supplementary Figure 5. Multi-modal profiling of base-edited cells at single-cell resolution. (A) The subcellular distribution of 13 protein stains (top row), 12 transcript detection for 12 genes (bottom row, starting from the most left) and barcode detection (bottom row, most right), are shown for a single cell with the AAVS1.100 guide. Cell and nuclear segmentation are outlined in green, whereas raw antibody signal and transcript detection are in gray-scale. Decoded barcodes are shown as false colored (magenta) spots. Scale bar, 5 µm. **(B)** As in A) for a cell with the BRCA1.416 (H1283Y) guide. **(C)** As in A) for a cell with the BRCA1.308 (Q380*) guide. **(D)** As in A) for a cell with the RAD51D.1 (splice) guide. Only raw data under the cell mask of the cell of interest is plotted.


A
RAD51 foci (raw image)
RAD51 foci (called spots)
All
B
γH2AX foci (raw image)
γH2AX foci (called spots)
All
Large
C
RPA2 foci
Cell cycle phase
G0
G1
S/G2
M
Treatment
UNT
IR
D
RAD18 foci
ns
E
γH2AX foci (large)
F
53BP1 foci (large)
G
cyclin A2
cyclin B1
p21
Cell count
Avg. fluorescence intensity under nuclei masks
H
RAD51 foci
BRCA1 foci
RPA2 foci
RAD18 foci
avg. mem cyclin B1
avg. Ki-67
avg. cyclin A2
avg. p21
micronuclei
avg. pHistone H3
avg. cleaved PARP1
avg. p53
large 53BP1 foci
small 53BP1 foci
small γH2AX foci
avg. γH2AX
large γH2AX foci

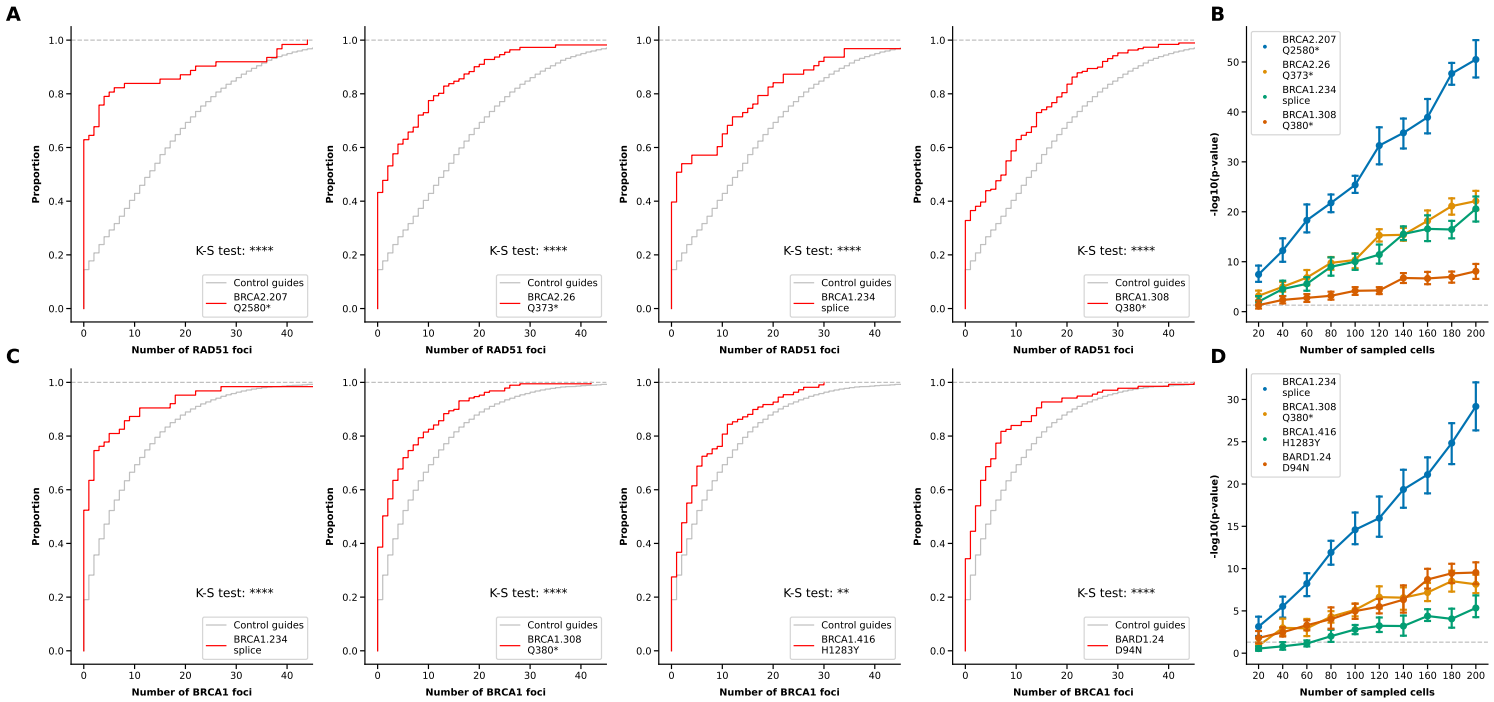
Supplementary Figure 6. Multiplexed immunofluorescence enables cell cycle phase separation and nuclear foci detection. **(A)** Visualization of nuclear foci detection for RAD51. Raw antibody stain for RAD51 is shown in grayscale in both left and right panels. Nuclear boundaries outlined in blue. All computationally detected RAD51 foci are displayed as circles in false color (magenta) on the right panel. Scale bar, 20 μm. **(B)** Same as A) for γH2AX. All computationally detected γH2AX foci are displayed as circles in false color (magenta) on the right panel, whereas computationally detected spots categorized as large γH2AX foci are displayed as circles of false color green. The same region in A) is shown. Scale bar, 20 μm. **(C)** Quantification of the number of RPA2 foci per cell across cell cycle phases in untreated (UNT) (n = 1 with 120,253 cells) and irradiated (IR) cells (n = 1 with 106,116 cells). Boxes indicate the median and interquartile range (IQR) with whiskers extending 1.5× IQR past the upper and lower quartiles (outliers are omitted). Two-sided Mann-Whitney test, *p $<$ 0.05, **p $<$ 0.01, ***p $<$ 0.001, ****p $<$ 0.0001. Exact p-values of G1UNT vs. G1IR, G0UNT vs. G0IR, G2/SUNT vs. G2/SIR, MUNT vs. MIR, G2/SIR vs. MIR, G1IR vs. G2/SIR, G0IR vs. G2/SIR are 0.00e+00, 0.00e+00, 0.00e+00, 2.95e-19, 7.13e-08, 0.00e+00, 0.00e+00, respectively. **(D)** Same as C) for RAD18 foci. p-values are 0.00e+00, 3.17e-77, 0.00e+00, 1.15e-01, 3.49e-15, 0.00e+00, 0.00e+00. **(E)** Same as C) for large γH2AX foci. p-values are 0.00e+00, 0.00e+00, 0.00e+00, 2.53e-36, 2.80e-03, 0.00e+00, 1.32e-124. **(F)** Same as C) for large 53BP1 foci. p-values are 0.00e+00, 5.17e-08, 0.00e+00, 3.04e-28, 8.59e-06, 0.00e+00, 1.01e-21. **(G)** Quantification of protein expression level of cyclin A2, cyclinB1, and p21. A threshold on the average fluorescence intensity under nuclear masks is set to categorize the expression level of the corresponding protein into low (smaller than the threshold) or high (greater than or equal to the threshold) categories. The threshold for each protein is displayed as a dashed line on the corresponding histogram. **(H)** Correlation of optical features measured in this experiment, showing positive correlation among nuclear foci involved in homologous recombination (RAD51, BRCA1, RPA2, RAD18), S/G2 phase cell cycle markers (cyclin A2, cyclin B1) and proliferation marker (Ki-67).



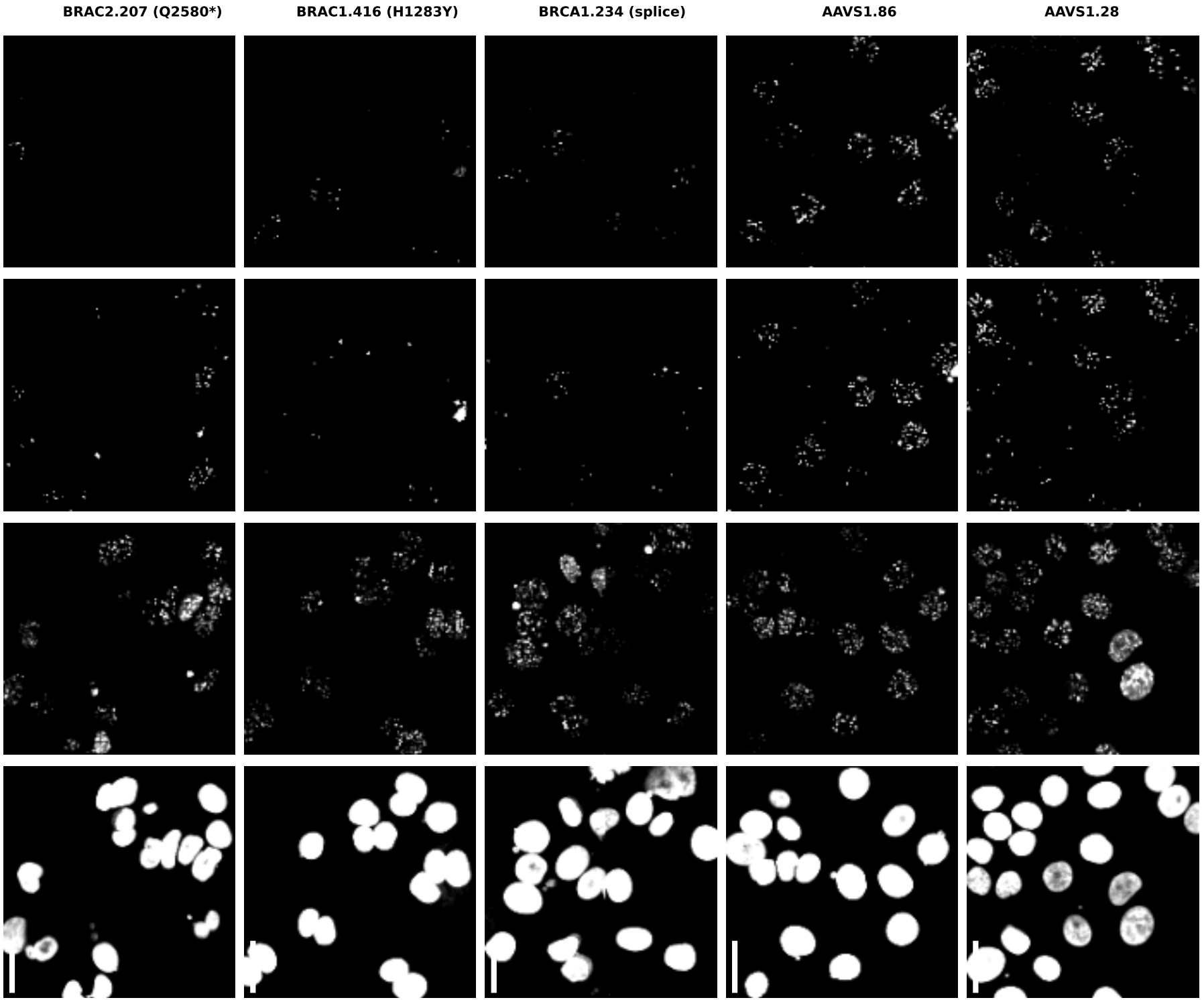
Supplementary Figure 7. Library representation in the CRISPRmap base-editing screening on MCF7 cells treated with ionizing radiation or DNA damaging agents. (A) Correlation of the relative abundance of guides in the DDR364 library based on the percentage of sequencing reads in plasmid library and the percentage of barcode-assigned cells in the optical screen under the camptothecin (CPT) treatment. Pearson correlation (r) is 0.67. Line, linear regression fit; shaded area, 95% confidence interval of the fit. **(B)** As in A) for cells under olaparib (OLAP) treatment. Pearson correlation (r) is 0.63. **(C)** As in A) for cells under cisplatin (CISP) treatment. Pearson correlation (r) is 0.66. **(D)** As in A) for cells under etoposide (ETOP) treatment. Pearson correlation (r) is 0.66. **(E)** As in A) for untreated cells (UNT1) as the control for the four DNA damaging agents in A) - D). Pearson correlation (r) is 0.67. **(F)** As in A) for cells under ionizing radiation (IR). Pearson correlation (r) is 0.63. **(G)** As in A) for untreated cells (UNT2) as the control for the IR-treated cells in F). Pearson correlation (r) is 0.65. **(H)** Visualization of the pearson correlation among the relative guide abundance in cells treated under the seven conditions in A) - G) and the plasmid library.


A
RAD51 foci (IR)
Controls
Perturb.
Wasserstein Distance to control
B
BRCA1 foci (IR)
Controls
Perturb.
Wasserstein Distance to control
C
UNT
IR
-Log10(p.adj)
Proportion of cells with high p21 expr.
sgRNA cat.
Not Sig.
control
missense
splice
nonsense
ClinVar cat.
Not Sig.
unknown
P/LP
VUS
B/LB
D
Lentiviral transduction of 364 guides targeting 27 DDR genes
antibiotic Selection (2d)
Induce DNA damage
Etoposide, Cisplatin Olaparib or Camptothecin
Untreated
Fix after 24h
Multiplexed immunofluorescence phenotyping
Phenotype-Genotype analysis
Variants
Optical features
E
RAD51D, RAD51C, XRCC3, BRCA1, BRCA2 (CISP)
L2FC in RAD51 foci
On-target score (Rule Set 2)
sgRNA category
splice
missense
nonsense
r = -0.29
F
RAD51D, RAD51C, XRCC3, BRCA1, BRCA2(OLAP)
L2FC in RAD51 foci
On-target score (Rule Set 2)
r = -0.3
G
CISP
-Log10(p.adj)
L2FC in RAD51 foci
H
OLAP
-Log10(p.adj)
L2FC in RAD51 foci
I
Gene enrichment (RAD51 foci)
-Log10(p-value)
Enriched
No
Yes
CPT
OLAP
CISP
ETOP
UNT
Treatment
J
ETOP
-Log10(p.adj)
L2FC in large γH2AX foci
K
OLAP
-Log10(p.adj)
L2FC in large γH2AX foci
L
Gene enrichment (large γH2AX foci)
-Log10(p-value)
Enriched
No
Yes
Treatment
M
Number of significant optical features
ns
control
B/LB
VUS
unknown
P/LP
ClinVar categories
N
Pearson correlation of L2FC in optical features
Same variant
Different variant

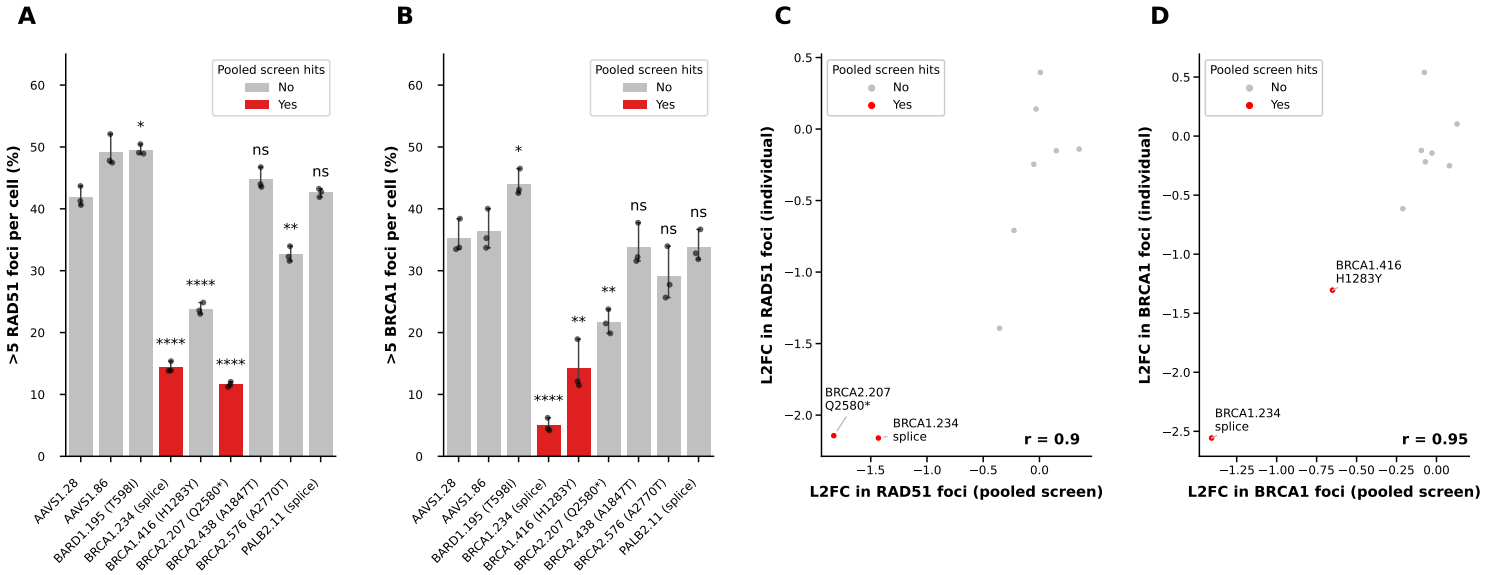
Supplementary Figure 8. Performance of CRISPRmap base-editing screening on MCF7 cells treated with ionizing radiation or DNA damaging agents. (A) Wasserstein distance of cells with DDR gene-targeting guides (Perturb) or control guides (Controls) to control cells for RAD51 foci in irradiated cells. Hits of RAD51 foci identified in the pooled screening are marked. **(B)** same as A) for BRCA1 foci. **(C)** Volcano plots showing variants yielding significant changes in the proportion of untreated (left) or irradiated (right) cells with high p21 expression. Significance, p.adj < 0.05. Two-sided Beta-Binomial test. The proportion of control cells with high p21 is marked with the vertical line. **(D)** Experimental workflow (Methods). Image made in Biorender. **(E)** Correlation between L2FC in RAD51 foci in CISP-treated cells and the Rule Set 2 (RS2) score. All guides targeting RAD51D, RAD51C, XRCC3, BRCA1, and BRCA2 are shown. The Pearson correlation (r) equals 0.29. **(F)** same as E) for OLAP-treated cells. The Pearson correlation (r) equals 0.30. **(G)** Volcano plot showing variants yielding significant changes in RAD51 foci in CISP-treated cells. Guides targeting DDR genes with RS2 score $>= 0.5$, all AAVS1-targeting and non-targeting control (NTC) guides are shown. Two-sided KS test, *p.adj < 0.05, **p.ajd < 0.01, ***p.adj < 0.001, ****p.adj < 0.0001. Figure legend as in C). **(H)** same as G) for OLAP-treated cells. **(I)** Gene enrichment analysis in variants that result in significant changes in RAD51 foci. One-sided (greater) Fisher exact test. Significance, p.adj < 0.05. (J) Same as G) for large γH2AX foci in CISP-treated cells. **(K)** same as J) for OLAP-treated cells. **(L)** same as I) for large γH2AX foci. **(M)** Quantification of the number of optical features with statistical significance scored by variants from different ClinVar categories. Boxes indicate the median and IQR with whiskers extending 1.5× IQR past the upper and lower quartiles. Two-sided Mann-Whitney test, *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001. The p-values (from left to right) are 2.28e-01, 3.69e-03, 3.83e-09, 2.07e-11. **(N)** Comparison of the Pearson correlation of L2FC in optical features between guides leading to the same amino acid (AA) change and different AA changes (p = 4.48e-02). Boxes as in M). Two-sided independent t-test, *p.adj < 0.05, **p.adj < 0.01, ***p.adj < 0.001, ****p.adj < 0.0001.



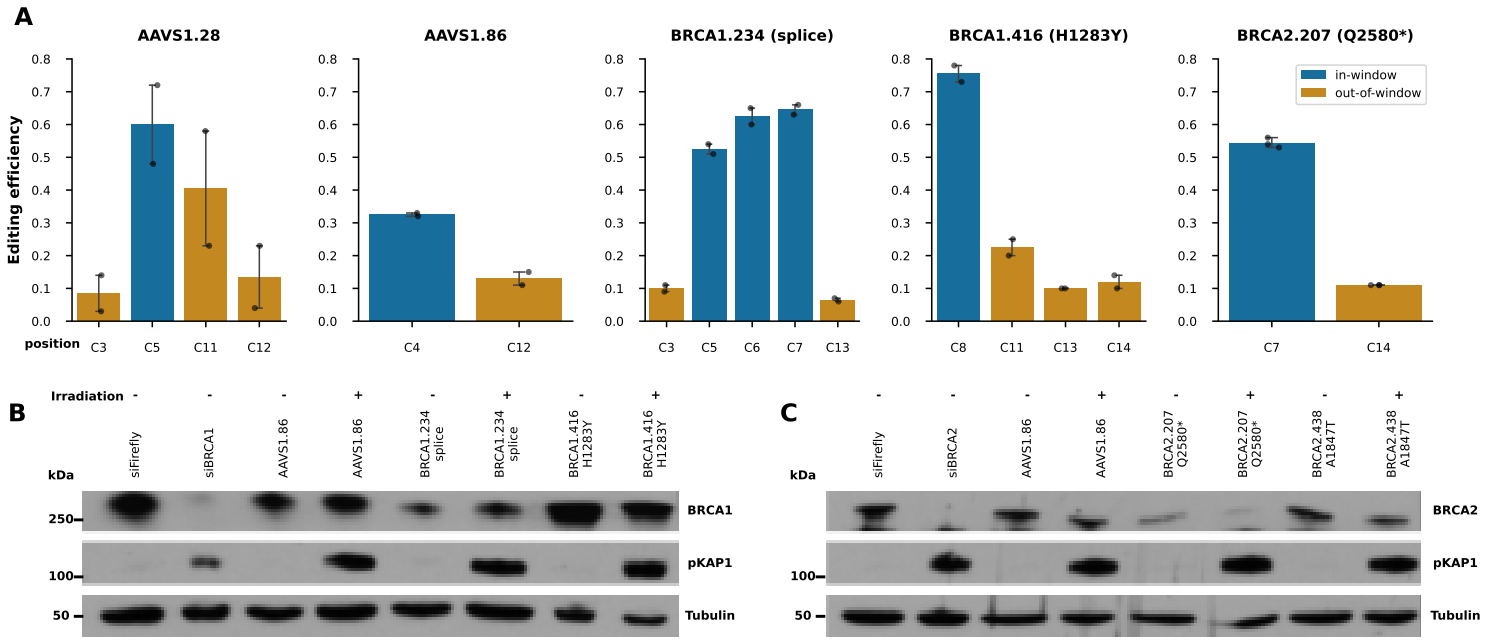
Supplementary Figure 9. Foci number distribution and random sampling test of variants resulting in significant loss of RAD51 or BRCA1 foci in the MCF7 cells treated with ionizing radiation. (A) Quantification of the RAD51 foci count in cells with the BRCA2.207 (Q2580*), BRCA2.26 (Q373*), BRCA1.234 (splice), and BRCA1.308 (Q380*) variant (left to right), showing significantly reduced RAD51 foci compared to the cells with AAVS1-targeting and non-targeting (control) variants. Cells in S/G2-phase treated with ionizing radiation are shown. Two-sided KS test, *p.adj $<$ 0.05, **p.adj $<$ 0.01, ***p.adj $<$ 0.001, ****p.adj $<$ 0.0001. The p-values (from left to right) are 2.55e-15, 2.70e-12, 1.19e-06, 5.71e-07. **(B)** Quantification of the sample size effect for the four variants in A) on the statistical significance in RAD51 foci changes. 10 random samples were taken from the screen data for each variant at each sample size and the p-value was calculated by the two-sided K-S test, showing higher statistical significance in all four variants at larger sizes of random samples. P-value cutoff of 0.05 is shown in grey (dashed). Data are presented as mean values +/- 95% CI. n=10 technical replicates. **(C)** As in A) for the BRCA1 foci count in cells with the BRCA1.234 (splice), BRCA1.308 (Q380*), BRCA1.416 (H1283Y), and BARD1.24 (D94N) variant. The p-values (from left to right) are 5.12e-09, 2.39e-08, 4.88e-03, 3.71e-06. **(D)** As in B) for the four variants in C) on the statistical significance in BRCA1 foci changes.


BRAC2.207 (Q2580*)
BRAC1.416 (H1283Y)
BRCA1.234 (splice)
AAVS1.86
AAVS1.28
RAD51
BRCA1
γH2AX
DAPI

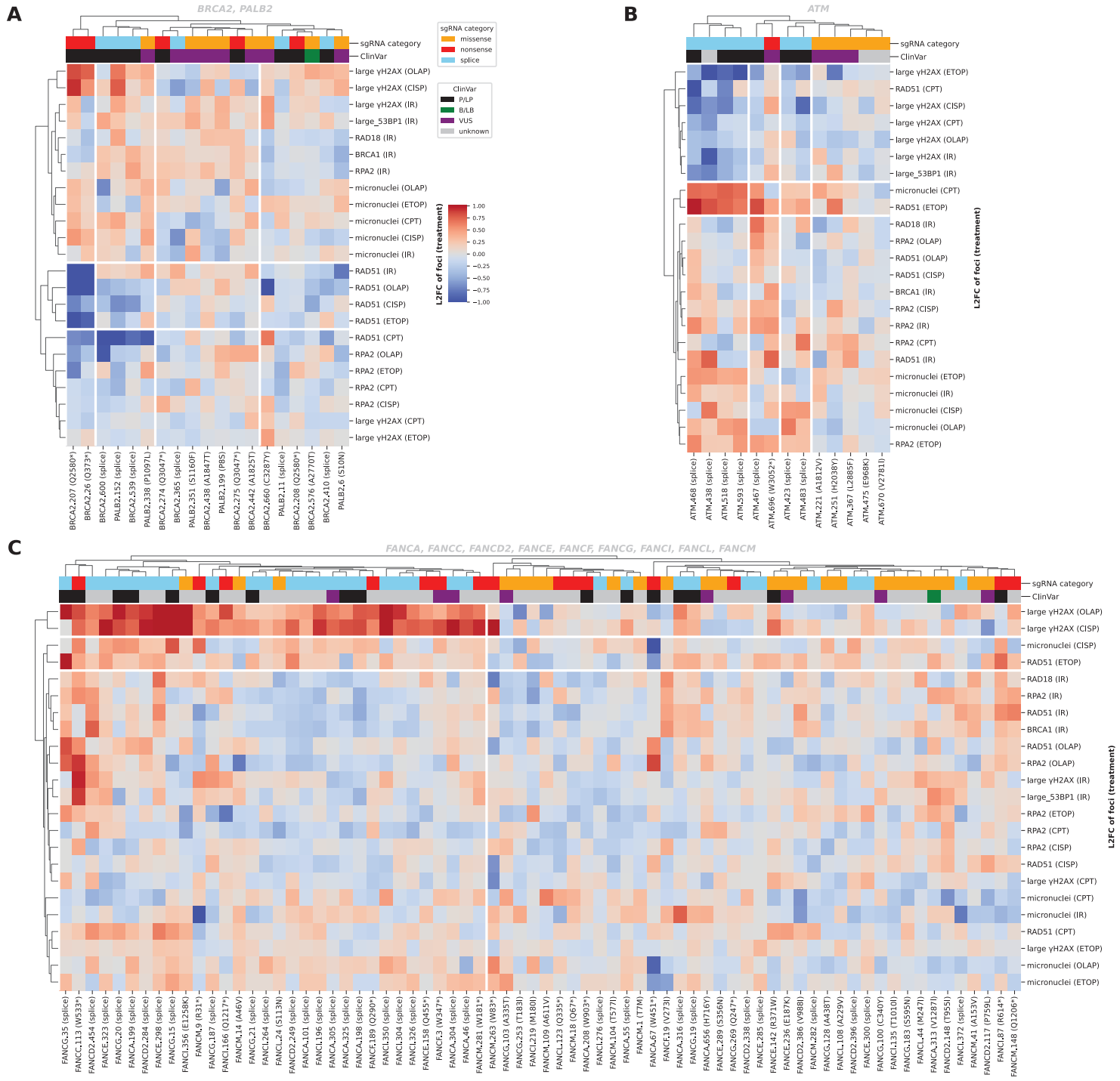
Supplementary Figure 10. Phenotypic validation of hits identified in the pooled screen on MCF7 cells with individually transduced guides. Visualization of the RAD51, BRCA1 and γH2AX antibody staining (left to right) in MCF7 cells individually transduced with the AAVS1.28, AAVS1.86, BRCA1.234 (splice), BRCA1.416 (H1283Y) and BRCA2.207 (Q2580*) guide (top to bottom) after the treatment of ionizing radiation. DAPI staining is shown on the right most column. Raw fluorescence signal of RAD51, BRCA1, γH2AX and DAPI are shown in greyscale. Scale bar, 20 μm.



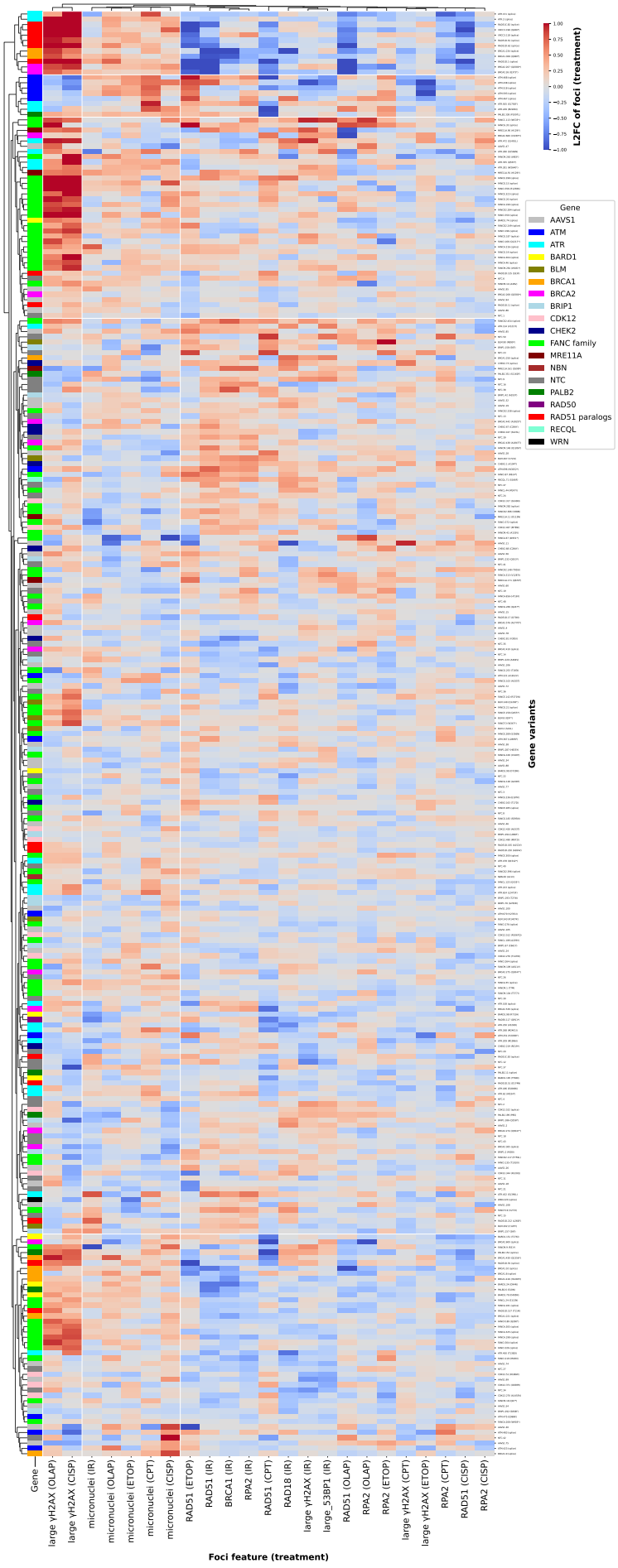
Supplementary Figure 11. Validation of optical phenotypes captured in the pooled screen on MCF7 cells with individually transduced guides. **(A)** Quantification of the percentage of cells with >5 RAD51 foci in cells that were individually transduced with each of the nine guides and treated by ionizing irradiation. Hits identified in the screen are marked in red, showing significantly lower ratios of cells with >5 RAD51 foci compared to the ratios in cells with the AAVS1 control guides. Data are presented as mean values +/- 95% CI. n = 3 biological replicates. Two-sided independent t-test, *p.adj $<$ 0.05, **p.adj $<$ 0.01, ***p.adj $<$ 0.001, ****p.adj $<$ 0.0001. The exact p-values (from left to right) are 1.95e-02, 1.34e-05, 7.57e-05, 6.21e-06, 1.01e-01, 1.40e-03, 4.88e-01, respectively. **(B)** As in A) for BRCA1 foci. The exact p-values (from left to right) are 1.201e-02, 6.13e-05, 1.85e-03, 2.32e-03, 6.18e-01, 1.10e-01 , 5.47e-01, respectively. **(C)** Correlation of RAD51 foci log2-fold change (L2FC) resulted from guides delivered in the pooled screen and guides transduced individually (individual). Hits identified in the pooled screen are marked in red. The Pearson correlation (r) is 0.90. **(D)** same as C) for BRCA1 foci. The Pearson correlation (r) is 0.95.



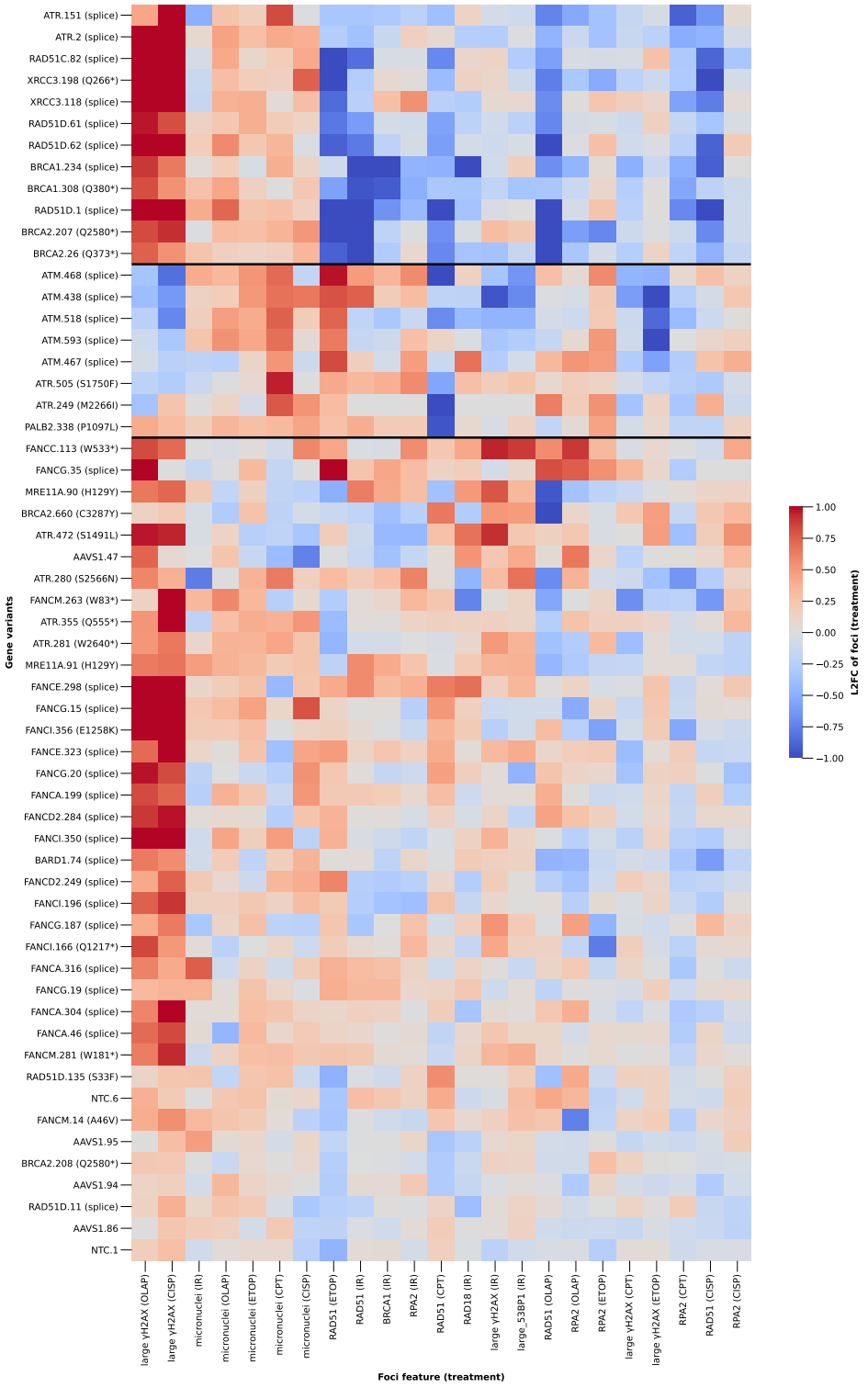
Supplementary Figure 12. Genomic and protein-level validation of hits identified in the pooled screen on MCF7 cells with individually transduced guides. (A) Quantification of the base-editing efficiency in MCF7 cells individually transduced with the AAVS1.28 (n = 2, biological replicates), AAVS1.86 (n = 2, biological replicates), BRCA1.234 (splice) (n = 2, biological replicates), BRCA1.416 (H1283Y) (n = 2, biological replicates) and BRCA2.207 (Q2580*) (n = 3, biological replicates) guide (left to right) based on sanger sequencing of the PCR-amplified genomic locus targeted by each guide. Editing efficiency on C bases both inside (in-window) and outside (out-of-window) of the expected editing window were quantified, showing good on-target (in-window) base editing. The scale of editing efficiency is 0 (no editing) to 1 (complete editing). The C base position is listed as the n-th base from the first base of the gRNA targeted sequence. Data are presented as mean values +/- 95% CI. **(B)** Immunoblots on guides targeting BRCA1 showing the reduction of full-length BRCA1 protein in the BRCA1.234 (splice) variant, compared to the missense variants BRCA1.416 (H1283Y), and the AAVS1.86 variant. Cells transduced with Firefly siRNA (siFirefly) and BRCA1 siRNA (siBRCA1) were included to show the specificity of BRCA1 protein detection. Cells were treated with or without irradiation and the induction of DNA damage is shown with the phospho-KAP1 (pKAP1) staining. Tubulin is used as the loading control. **(C)** As in B) for BRCA2 variants showing the reduction of full-length BRCA2 protein in the nonsense variant BRCA2.207 (Q2580*), compared to the missense variant BRCA2.438 (A1847T) and the AAVS1.86 variant.



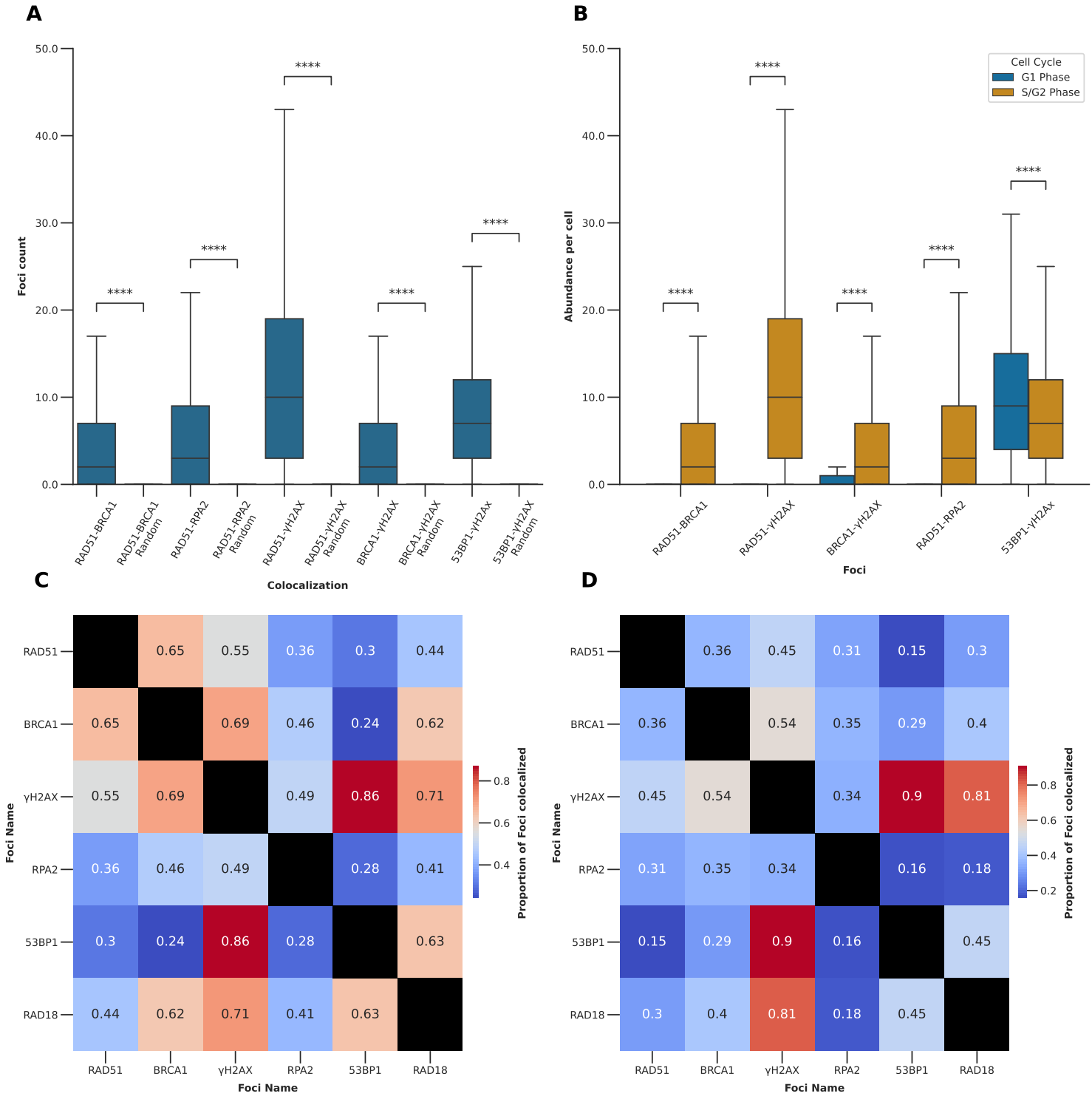
Supplementary Figure 13. Variant analysis of functionally relevant genes and variant clusters with treatment-specific optical signatures. (A) Clustering of guides targeting BRCA2 and PALB2, showing a cluster of pathogenic nonsense variants showing a string reduction of RAD51 across all treatments and another cluster showing reduction of RAD51 in CPT-, OLAP-, and CISP-treated cells mainly composed of pathogenic splice variants. Other clusters show mild phenotypes. Log2-fold change (L2FC) in each optical phenotype in corresponding treatment conditions are shown as rows in the heatmap. Cells in all cell cycle phases are included, untreated cells are not included. All guides with on-target score ≥ 0.5 were included in the clustering and shown in the heatmap. Columns were cut at a depth of 2 and rows were cut a depth of 3 based on the dendrogram. Color scale is -1 to 1. **(B)** same as A) for guides targeting ATM, showing a cluster with reduced large γH2AX foci and micronuclei in all treatments, increased micronuclei in CPT-treated cells and increased RAD51 foci in ETOP-treated cells composed of splice variants, and other clusters with milder phenotypes. **(C)** same as A) for guides targeting all genes in Fanconi anemia (FA) pathway, showing a cluster with increased large γH2AX foci in OLAP- and CISP-treated cells, increased micronuclei in ETOP-treated cells and increased RAD51 foci in ETOP-treated cells, composed of most splice and nonsense variants targeting these genes. The other cluster composed of mainly missense variants shows mild phenotypes.


L2FC of foci (treatment)
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
Gene
AAVS1
ATM
ATR
BARD1
BLM
BRCA1
BRCA2
BRIP1
CDK12
CHEK2
FANC family
MRE11A
NBN
NTC
PALB2
RAD50
RAD51 paralogs
RECQL
WRN
Gene variants
Gene
large γH2AX (OLAP)
large γH2AX (CISP)
micronuclei (IR)
micronuclei (OLAP)
micronuclei (ETOP)
micronuclei (CPT)
micronuclei (CISP)
RAD51 (ETOP)
RAD51 (IR)
BRCA1 (IR)
RPA2 (IR)
RAD51 (CPT)
RAD18 (IR)
large γH2AX (IR)
large_53BP1 (IR)
RAD51 (OLAP)
RPA2 (OLAP)
RPA2 (ETOP)
large γH2AX (CPT)
large γH2AX (ETOP)
RPA2 (CPT)
RAD51 (CISP)
RPA2 (CISP)
Foci feature (treatment)

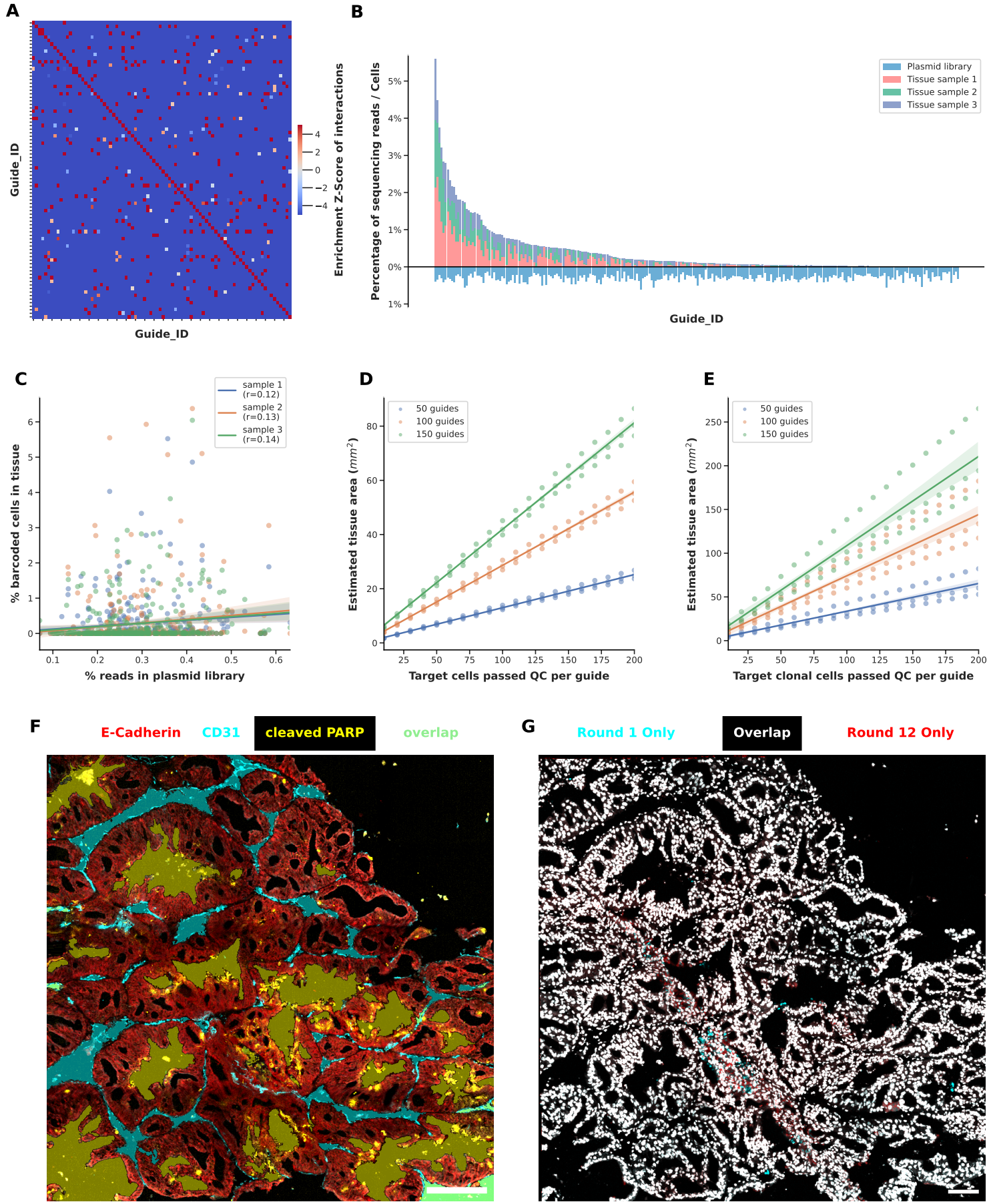
Supplementary Figure 14. Variant analysis of clustering at the library level shows local structure of enrichment in gene targets. Clustering of all targeting guides with a Rule Set2 on-target score (RS2) > 0.5 and the non-targeting control guides (273 guides in total) on foci features across five treatments, showing local structures of enrichment in guides targeting the same (groups of) genes. Guides (rows) are colored by their gene targets. RAD51 paralogs (RAD51D, RAD51C and XRCC3) are considered as a group ("RAD51 paralogs") and FANC family genes (FANCA, FANCC, FANCD2, FANCE, FANCF, FANCG, FANCI, FANCL, FANCM) are grouped as "FANC family".



Supplementary Figure 15. Variant analysis of clustering at the library level identifies clusters with local structure of enrichment in gene targets. Enlarged view of the top 3 clusters of Supplementary Figure 14. The top most cluster shows a significant enrichment of guides targeting RAD51 paralogs (p = 1.1e-05, hypergeometric test), featured by downregulation of RAD51 foci across all treatments and upregulation of γH2AX foci in CISP- and OLAP-treated cells. Strong enrichment of ATM variants is observed by the second top cluster (p = 5.4e-06), where γH2AX foci are significantly reduced in ETOP-treated cells and mostly down-regulated in CISP- and OLAP-treated cells. FANC family gene variants are observed to be enriched in the third top cluster (p = 3.2e-05) which features an upregulation of γH2AX foci in OLAP-and CISP-treated cells.



Supplementary Figure 16. Colocalization of DNA Damage Repair proteins across cell cycle. (A) Quantification of the number of colocalized DNA Damage Repair (DDR) foci marker per nuclei is significantly higher than random chance colocalization of DDR foci markers per nuclei. Two-sided paired student's t-test, *p $<$ 0.05, **p $<$ 0.01, ***p $<$ 0.001, *****p $<$ 0.0001. Boxes indicate the median and interquartile range (IQR) with whiskers extending 1.5× IQR past the upper and lower quartiles. The p-values (from left to right) are 1.61e-277, 9.83e-272, 0.00e+00, 1.05e-291, 0.00e+00 (n=10000 iterations for each condition). **(B)** Quantification of the number of specific colocalized DDR foci per cell across cell cycle in irradiated (IR) cells. Two-sided independent student's t-test, *p $<$ 0.05, **p $<$ 0.01, ***p $<$ 0.001, *****p $<$ 0.0001. (Outliers are omitted in the plot.) The p-values (from left to right) are 0.00e+00, 2.33e-269, 7.15e-245, 1.79e-234, 1.18e-25. n = 1 with 106,116 cells. **(C)** Proportion of DDR foci marker colocalized with another DDR foci marker for S/G2 phase shows strong colocalization of RAD51-BRCA1, RAD51-γH2AX, BRCA1-γH2AX, 53BP1-γH2AX highlighting the prevalence of Non-Homologous End Joining and Homologous Recombination. **(D)** Proportion of DDR foci marker colocalized with another DDR foci marker for G1 phase shows strong colocalization of 53BP1-γH2AX highlighting the prevalence of Non-Homologous End Joining in G1 phase.

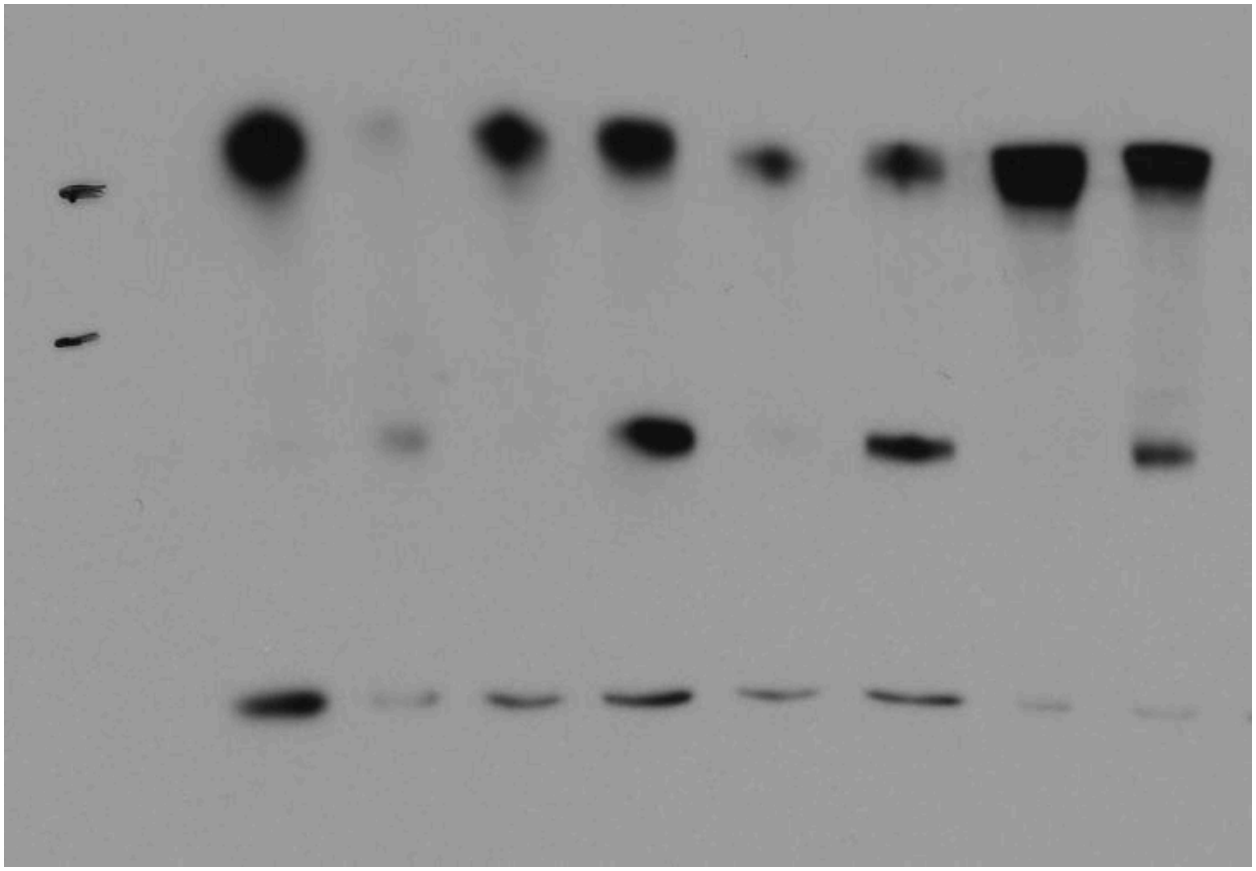

A
Guide_ID
Guide_ID
4
2
0
−2
−4
B
Enrichment Z-Score of interactions
Percentage of sequencing reads / Cells
5%
4%
3%
2%
1%
0%
1%
Guide_ID
Plasmid library
Tissue sample 1
Tissue sample 2
Tissue sample 3
C
% barcoded cells in tissue
% reads in plasmid library
sample 1 (r=0.12)
sample 2 (r=0.13)
sample 3 (r=0.14)
D
Estimated tissue area (mm^2)
Target cells passed QC per guide
50 guides
100 guides
150 guides
E
Estimated tissue area (mm^2)
Target clonal cells passed QC per guide
50 guides
100 guides
150 guides
F
E-Cadherin
CD31
cleaved PARP
overlap
G
Round 1 Only
Overlap
Round 12 Only

Supplementary Figure 17. Quantitative analysis on CRISPRmap barcode readout in tissue. (A) Interaction matrix among guides detected on tissue samples, showing mostly homotypic interactions. All guides with at least 10 cells are shown, and enrichment z scores of interactions are displayed. **(B)** Visualization of the relative abundance of guides in the DDR364 library based on the percentage of sequencing reads in plasmid library (bottom) and the stacked percentage of barcode-assigned cells in three tissue samples (top). **(C)** Correlation of the relative abundance of guides in the DDR364 library based on sequencing reads in plasmid library and barcode-assigned cells in tissue samples. Line, linear regression fit; shaded area, 95% confidence interval of the fit. The Pearson correlation (r) is 0.12, 0.13, 0.14, respectively. **(D)** Estimated tissue area required for a CRISPRmap screen of varying library sizes (50, 100, 150 guides) at different targets of cells passed Quality Check (QC) per guide. Three estimations were made for each target cell number per guide for a given library size based on three tissue samples. Line, linear regression fit; shaded area, 95% confidence interval of the fit. **(E)** As in D) for different targets of clonal cells passed QC per guide. **(F)** Visualization of the void regions on the tissue section, showing regions associated with cell death marked by cleaved PARP (displayed in yellow) and regions associated with mouse vasculature marked by CD31 (shown in cyan). Green regions are identified as an overlap of vasculature and cell death. Unmarked regions are not classified into either category. Raw cleaved PARP, CD31 and E-cadherin fluorescence are shown in yellow, cyan and red, respectively. Scale bar, 200 μm. **(G)** Visualization of the overlay of DAPI fluorescence between the first (Round 1) and the last (Round 12) imaging round on a tissue section, showing minimal tissue loss through the 12 rounds of cyclic CRISPRmap barcode readout and antibody staining. Round 1 and Round 12 DAPI signal is displayed in cyan and red, respectively. The overlap between Round 1 and Round 12 DAPI signal is displayed in white. Scale bar, 200 μm.

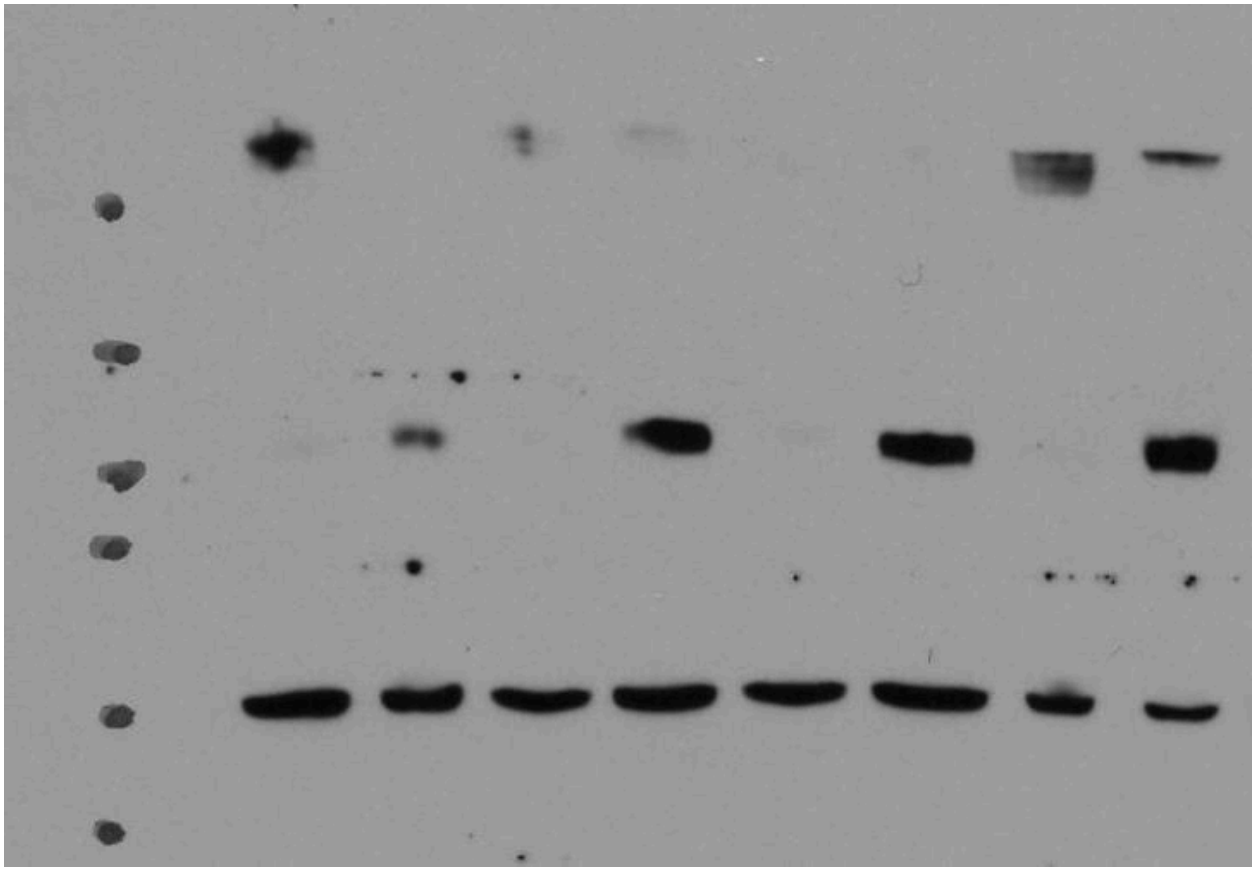
Source Data

Supplementary Figure 12B

BRCA1

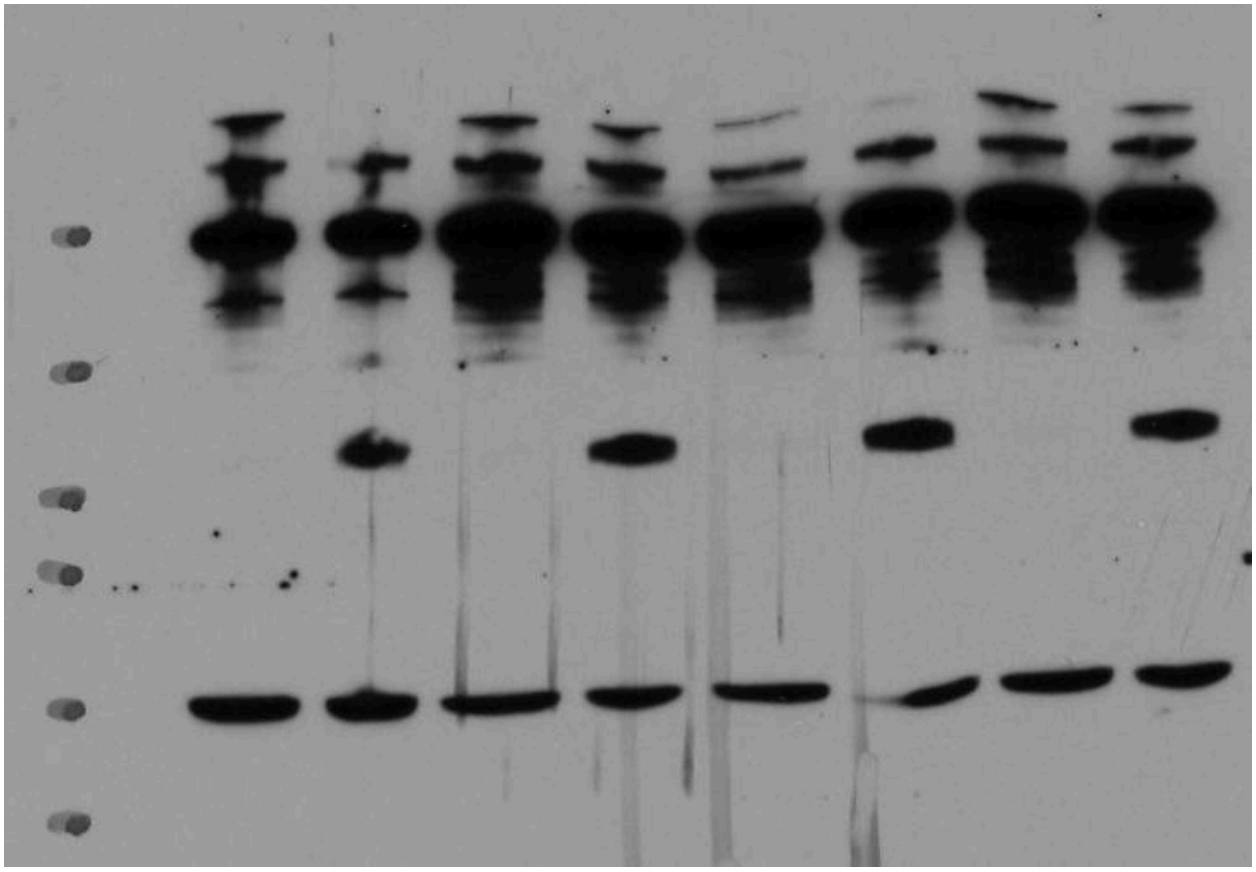


pKAP1 and Tubulin (Different exposure time on the same gel)



Supplementary Figure 12C

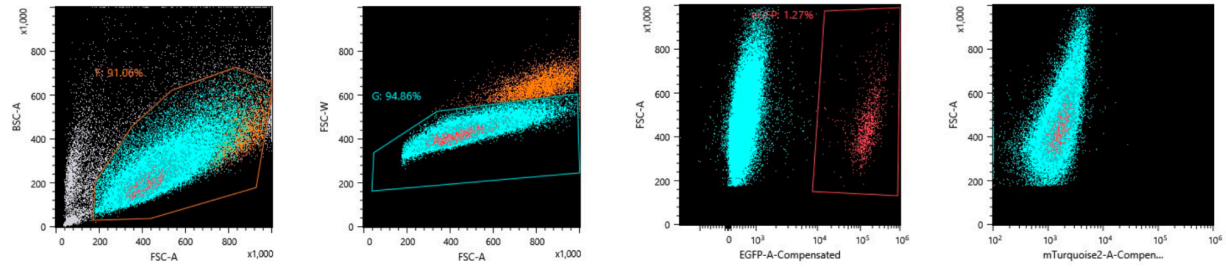
BRCA2, pKAP1 and Tubulin



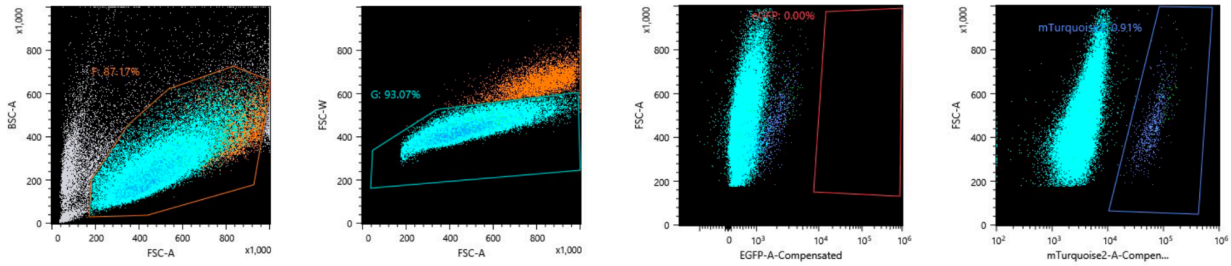
Flow cytometry plots

For CRISPRmap assay on the GFP/mTurquoise2-reporting cells

GFP+ cells



mTurquoise2+ cells



Supplementary Table Information

Supplementary Table 1. Cost estimation of the CRISPRmap assay. The cost of probes and fluorophores, key chemicals and reagents, and enzymatic reactions are listed. The calculation of cost per million cells is based on the base-editing screening on MCF7 cells using the DDR364 library. The actual cost may differ according to price changes on the vendors' side. The total setup cost mainly results from the purchase of Padlock, primer, splint and readout probes, which require a minimum amount to be ordered. Cost of Conventional OPS (Feldman et al., 2019) is included for side-by-side comparison with CRISPRmap on a per six-well plate basis.

Supplementary Table 2. GFP-pilot analysis. This table shows the single-cell level data for cells analyzed in the CRISPRmap knockout screening targeting GFP, related to Figure 2 and Supplementary Figure 1. The number of amplicons detected as well as the average nuclear GFP intensity in each cell are listed. Single-cell level data for the non-infection control in the same experiment is included for the evaluation of non-specific amplicon generation. Cells that passed QC are listed with their unique guide identities. The pairwise Mann-Whitney U test p values of each GFP-NTC guide pair are listed, related to Supplementary Figure 1A. Sensitivity, specificity, precision and the ratio of cells passed quality check (QC) on HT1080s in different experimental settings are listed, related to Supplementary Figure 1B. And the sensitivity, specificity, precision and ratio of cells passedQC under 25 QC metrics are listed, related to Supplementary Figure 1C-E. Number of reads classified into each category in the recombination analysis are listed for the GFPpilot and DDR364 library at the Opool, Plasmid, and gDNA level, respectively, related to Supplementary Figure 1J. The ratio of cells with unique or double barcode detection are listed, side-by-side with the Poisson expected double transduction rate, for the cells transduced with the GFP-pilot library at different multiplicity of infection (MOI), related to Supplementary Figure 1L. The total number of cells profiled and the fraction of cells with barcode detection in the CRISPRmap and Conventional OPS assays performed on different cell types are listed, related to Figure 2F.

Supplementary Table 3. Base-editing screening data analysis. Information of the guides in the DDR364 library are listed in detail, including the sgRNA sequences, type of intended mutations, target genes, intended Amino Acid change, updated Clinvar categories. The number of RNAmap spots and TPM RNA-seq reads for the 12 transcripts are listed, related to Figure 3E. The KS test p values, FDR, and log2-fold change (LFC) of each guide on each foci features in the irradiation (irradiation) and DNA damaging agents (chemo) screening are shown, the sgRNA_IDs, Rule Set 2 on-target scores, sgRNA categories and Clinvar categories are listed alongside, related to Figure 4, 5 and Supplementary Figure 8, 13-15. P values of gene enrichment fisher test are listed, related to Figure 4I&L and Supplementary Figure 8I&L. Percentage library representation of guides based on cells assigned with guide-reporting barcodes in seven conditions in the base-editing screen and based on reads sequenced at the plasmid library level are shown, related to Supplementary Figure 7. The wasserstein distances for all foci features are shown, related to Supplementary Figure 8A&B. The p values of beta binomial tests for all binary features are shown, related to Supplementary Figure 8C.

The FDR and LFC of foci features for the 9 individually transduced guides are shown, related to Supplementary Figure 11. Base-editing efficiency for the 3 hits identified in the IR screen and 2 AAVS1 control guides are listed for in-window and out-of-window C bases, related to Supplementary Figure 12A.

Supplementary Table 4. OE19 tissue analysis. Cells analyzed in the OE19 tissue profiling shown in Figure 6 are listed at the single-cell level for the evaluation of spot count, guide purity, apoptotic marker expression and cell sizes. The library representation in three tissue samples and relative abundance of guides in plasmid library sequencing reads are listed, related to Supplementary Figure 17B&C. The estimated tissue area required for three different library sizes based on three tissue samples are shown, related to Supplementary Figure 17D&E.

Supplementary Table 5. GFP-pilot library design and readout scheme. This table lists the sgRNA_ID, sgRNA sequences and barcode sequences in the GFP-pilot library, as well as the sequences of padlocks, primers, splints and readout probes that were used to readout the barcodes, related to Figure 2 and Supplementary Figure 1. A detailed readout scheme including the order, conjugated fluorophores and imaging setting of the 8 readout probes are listed, as well as the 8-bit codebook encoding the 10 guides in the library. The oligonucleotide sequences used in the GFP-mTurquoise2 barcode detection experiment are also listed.

Supplementary Table 6. DDR364 library design, readout scheme, and antibody staining. This table lists the sgRNA_ID, sgRNA sequences and barcode sequences in the DDR364 library, related to Figure 3, 4, 5 and Supplementary Figure 6-15. The sequences of padlocks, primers, splints and readout probes that were used to readout the barcodes and a detailed readout scheme including the order, conjugated fluorophores and imaging setting of the 24 readout probes are shown, together with the 24-bit codebook encoding the 364 guides in the library. The mutational outcomes and ClinVar annotations of the guides are listed. Antibodies used in the irradiation screen, DNA damaging agents screen (chemo) are listed in detail. The antibodies used in OE19 tissue profiling are also listed in detail, related to Figure 6 and Supplementary Figure 17.

Supplementary Table 7. Oligonucleotide sequences of CRISPRmap probes. This table lists the sequence of all the 54 padlocks and 54 primers used in the DDR364 base-editing screening. All 2,916 possible padlock-primer combinations are listed with their corresponding readout sequences specified. This table also lists the sequences of 12 splints and 24 readout probes that are required for the barcode detection of the 2,916 combinations. Sequences of PCR primers used for library amplification and validation are listed. For RNAmap, the sequences of 36 readout probes and their corresponding 319 36-bit codes are listed, as well as the readout scheme and the codebook for the 12 selected transcripts profiled in the DDR364 irradiation screen. Primer sequences for PCR amplification of the genomic loci of the intended mutations are listed for the five individually transduced guides in the base-editing efficiency validation.

Supplementary Table 8. Cells passed QC in the irradiation base-editing screening and the optical features. This table lists 226,369 cells in the irradiation base-editing screening that

passed the QC criteria, related to Figure 3 and Supplementary Figure 8. The average nuclear intensity of protein stains, cell cycle phase annotation, the raw count for foci features, and the number of RNA-reporting spots measured by RNAmap for the 12 transcripts are listed. Cells are then grouped by sgRNA_ID for FDR and LFC calculation listed in Supplementary Table 3.