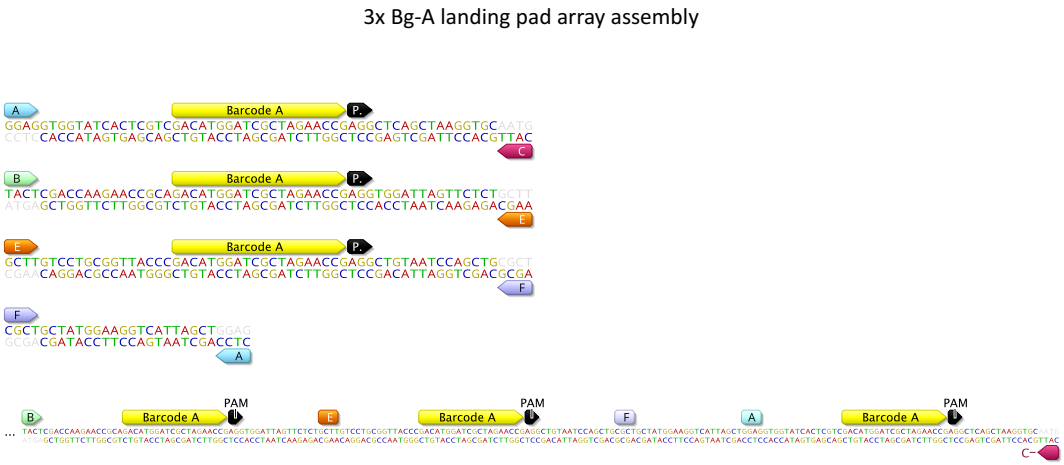
SUPPLEMENTARY INFORMATION

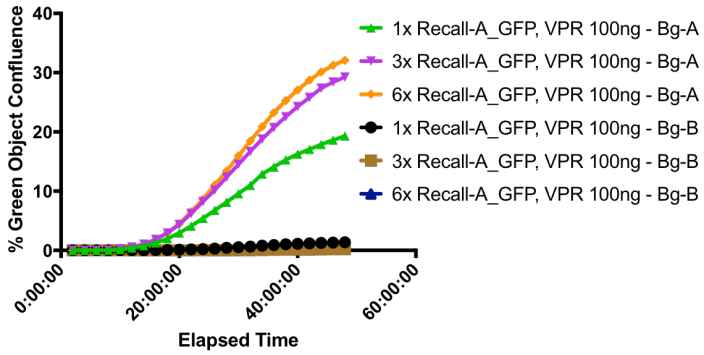
Technical Note: Control of lineage-specific gene expression by functionalized gRNA barcodes



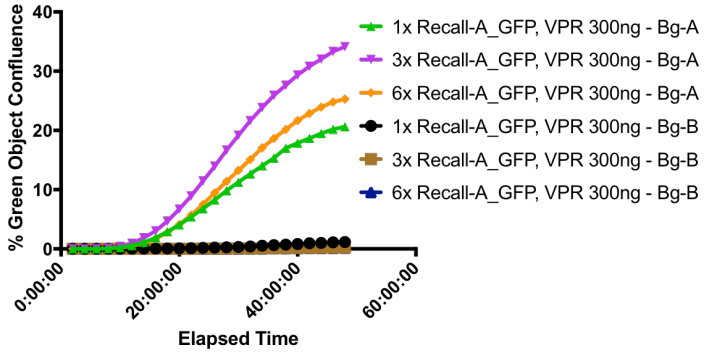
Supplementary Figure S1. Bg-A landing pad array assembly. The 3x barcode landing pad arrays were assembled by first annealing complimentary oligonucleotides containing the barcode of interests and PAM site along with the specified overhangs A-F (a). When combined, these specified overhangs drive assembly of the individual double stranded barcodes to both make the 3x barcode array as well as direct integration into the BbsI digested Recall plasmid (b). Similar schemes were used to assemble larger barcode arrays.

Supplementary Figure S2

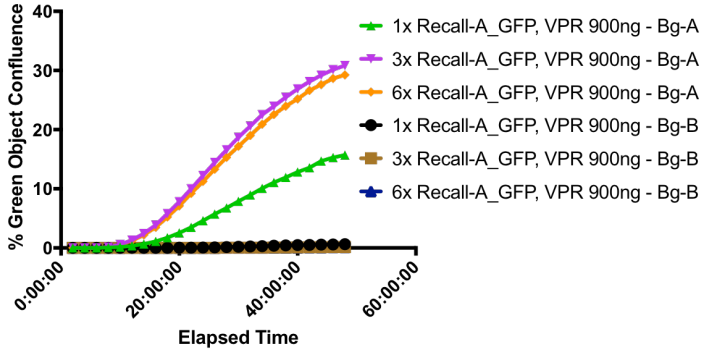
Barcode A 1/3/6x Array with 100ng dCas9-VPR



Barcode A 1/3/6x Array with 300ng dCas9-VPR

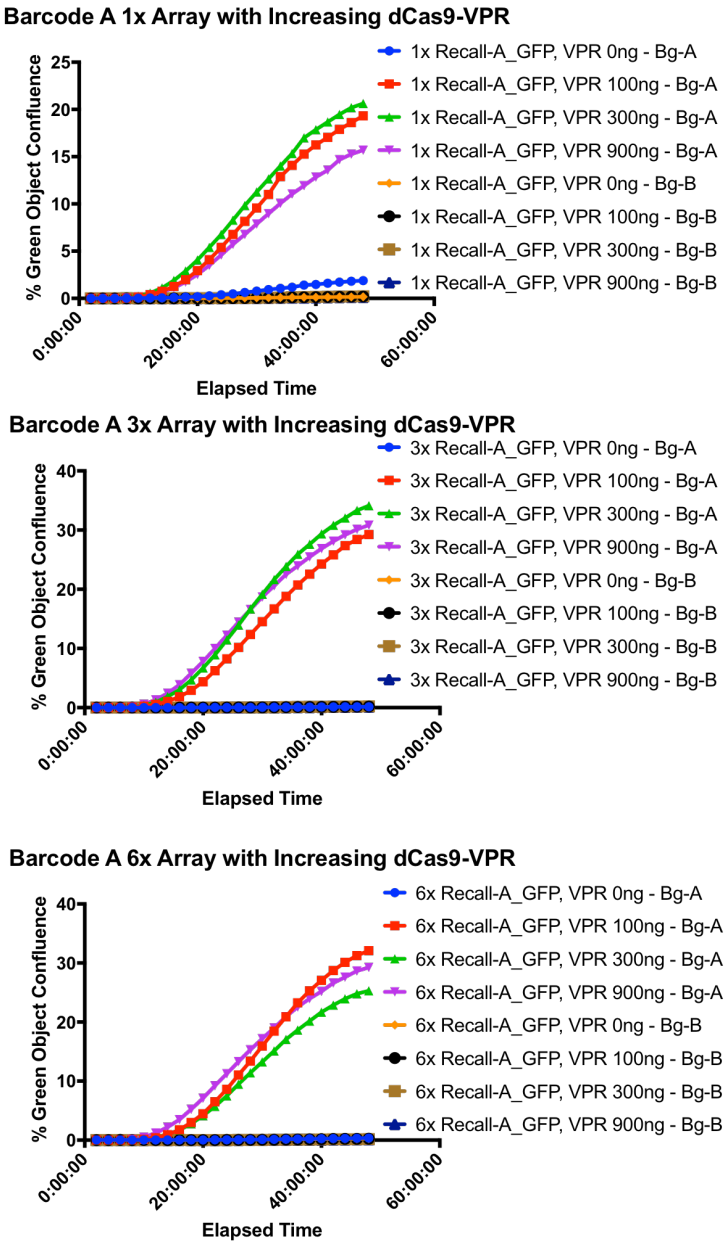


Barcode A 1/3/6x Array with 900ng dCas9-VPR



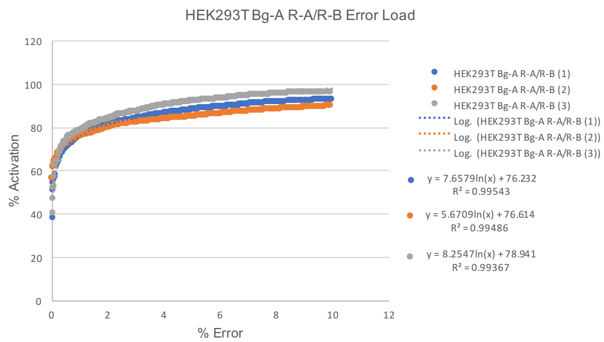
Supplementary Figure S2. Lineage specific gene activation efficiency of 1x, 3x, 6x barcode landing pads at different concentrations of dCas9-VPR. Time lapse fluorescent analysis of percent green object confluence of HEK293Ts Bg-A and Bg-B populations co-transfected with dCas9-VPR and 80ng of Recall-A_GFP plasmids with a 1x, 3x, or 6x barcode array in a 24 well plate. These graphs compare recall activation efficiency between Recall-A_GFP plasmids with a 1x, 3x, or 6x barcode array at given dCas9-VPR amounts.

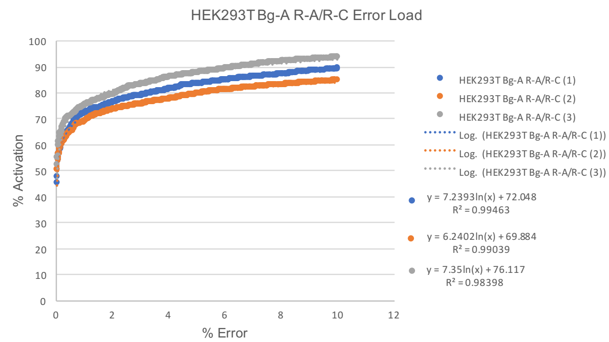
Supplement Figure S3

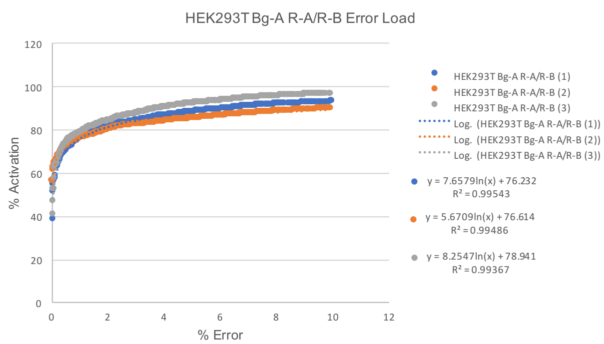


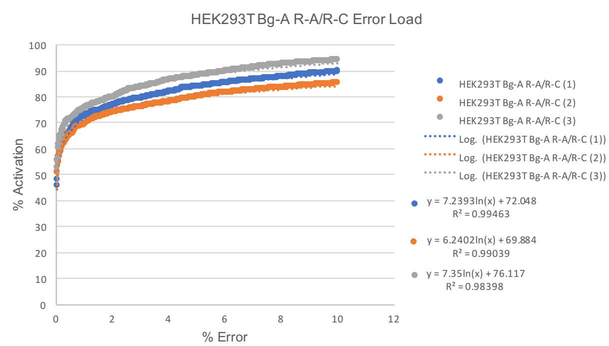
Supplementary Figure S3. Lineage specific gene activation efficiency with increase concentrations of dCas9-VPR in coordination with 1x, 3x, or 6x barcode landing pads. Time lapse fluorescent analysis of percent green object confluence of HEK293Ts Bg-A and Bg-B populations co-transfected with 0, 100, 300, and 900ng of dCas9-VPR and 80ng of Recall-A_GFP plasmids with a 1x, 3x, or 6x barcode array in a 24 well plate. These graphs compare recall activation efficiency of increasing amounts dCas9-VPR when co-transfected with 80ng Recall-A_GFP plasmids with a 1x, 3x, or 6x barcode array.

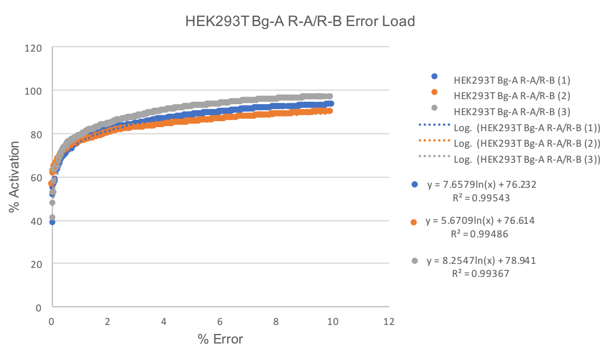
Supplement Figure S4

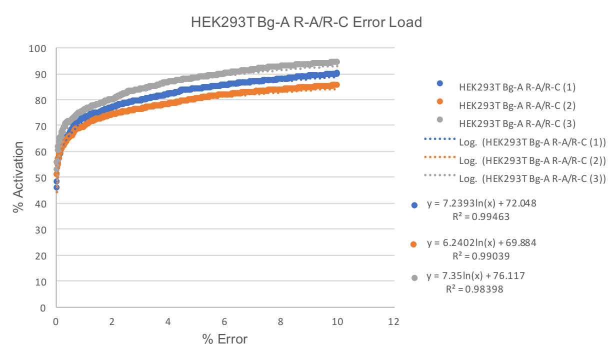




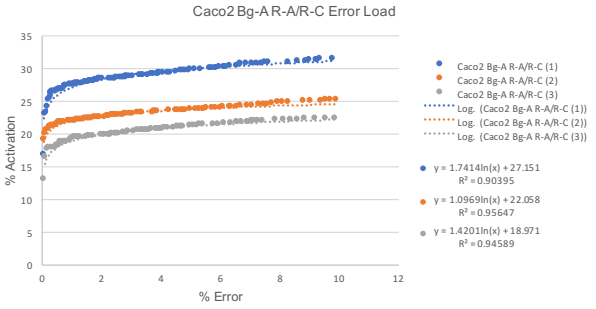


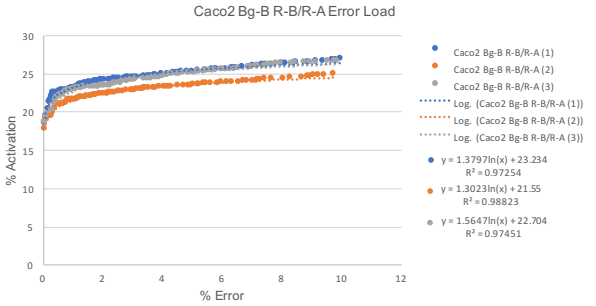


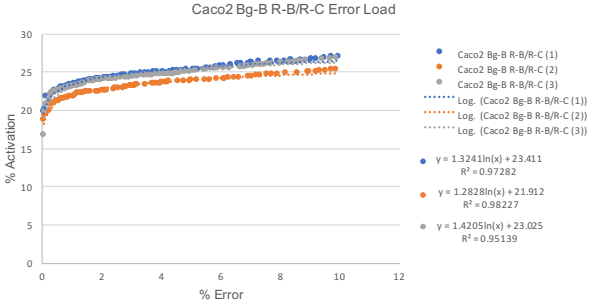


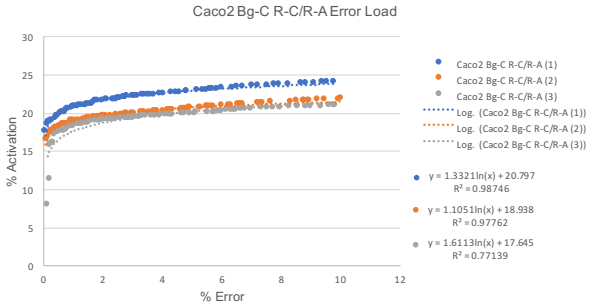


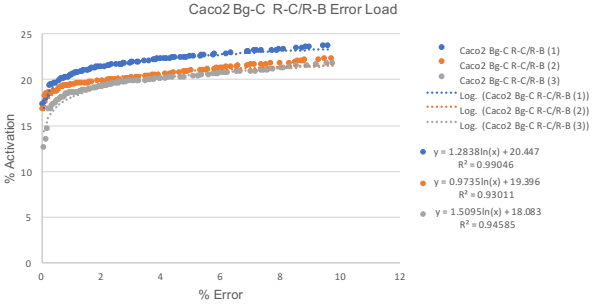

Caco2 Bg-A R-A/R-B Error Load
% Activation
% Error
Caco2 Bg-A R-A/R-B (1)
Caco2 Bg-A R-A/R-B (2)
Caco2 Bg-A R-A/R-B (3)
Log. (Caco2 Bg-A R-A/R-B (1))
Log. (Caco2 Bg-A R-A/R-B (2))
Log. (Caco2 Bg-A R-A/R-B (3))
y = 1.709ln(x) + 26.734
R² = 0.97446
y = 1.1991ln(x) + 22.031
R² = 0.89462
y = 1.3363ln(x) + 18.864
R² = 0.9674

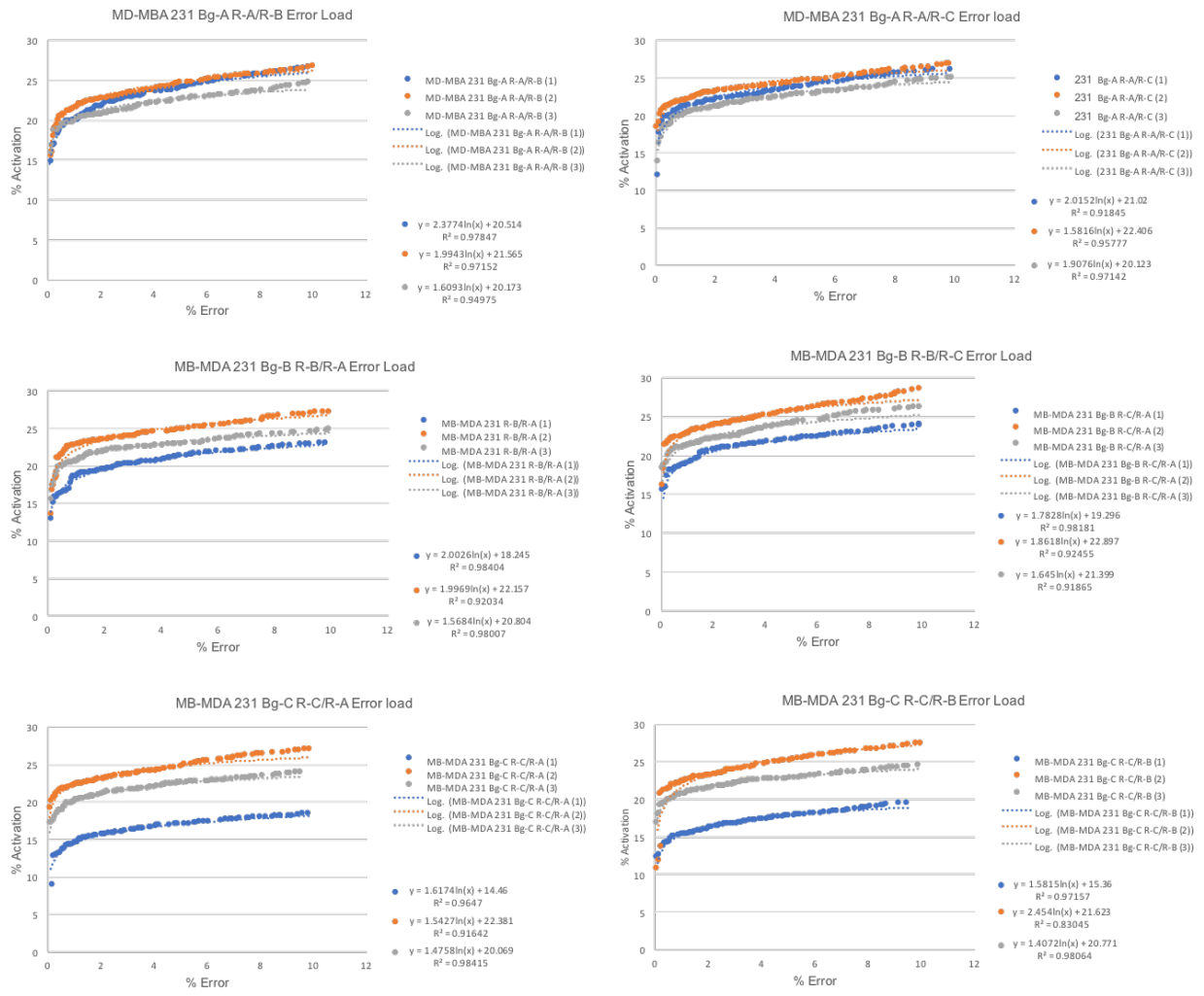

Caco2 Bg-A R-A/R-C Error Load
% Activation
% Error
Caco2 Bg-A R-A/R-C (1)
Caco2 Bg-A R-A/R-C (2)
Caco2 Bg-A R-A/R-C (3)
Log. (Caco2 Bg-A R-A/R-C (1))
Log. (Caco2 Bg-A R-A/R-C (2))
Log. (Caco2 Bg-A R-A/R-C (3))
y = 1.7414ln(x) + 27.151
R² = 0.90395
y = 1.0969ln(x) + 22.058
R² = 0.95647
y = 1.4201ln(x) + 18.971
R² = 0.94589


Caco2 Bg-B R-B/R-A Error Load
% Activation
% Error
Caco2 Bg-B R-B/R-A (1)
Caco2 Bg-B R-B/R-A (2)
Caco2 Bg-B R-B/R-A (3)
Log. (Caco2 Bg-B R-B/R-A (1))
Log. (Caco2 Bg-B R-B/R-A (2))
Log. (Caco2 Bg-B R-B/R-A (3))
y = 1.3797ln(x) + 23.234
R² = 0.97254
y = 1.3023ln(x) + 21.55
R² = 0.98823
y = 1.5647ln(x) + 22.704
R² = 0.97451


Caco2 Bg-B R-B/R-C Error Load
% Activation
% Error
Caco2 Bg-B R-B/R-C (1)
Caco2 Bg-B R-B/R-C (2)
Caco2 Bg-B R-B/R-C (3)
Log. (Caco2 Bg-B R-B/R-C (1))
Log. (Caco2 Bg-B R-B/R-C (2))
Log. (Caco2 Bg-B R-B/R-C (3))
y = 1.3241ln(x) + 23.411
R² = 0.97282
y = 1.2828ln(x) + 21.912
R² = 0.98227
y = 1.4205ln(x) + 23.025
R² = 0.95139


Caco2 Bg-C R-C/R-A Error Load
% Activation
% Error
Caco2 Bg-C R-C/R-A (1)
Caco2 Bg-C R-C/R-A (2)
Caco2 Bg-C R-C/R-A (3)
Log. (Caco2 Bg-C R-C/R-A (1))
Log. (Caco2 Bg-C R-C/R-A (2))
Log. (Caco2 Bg-C R-C/R-A (3))
y = 1.3321ln(x) + 20.797
R² = 0.98746
y = 1.1051ln(x) + 18.938
R² = 0.97762
y = 1.6113ln(x) + 17.645
R² = 0.77139


Caco2 Bg-C R-C/R-B Error Load
% Activation
% Error
Caco2 Bg-C R-C/R-B (1)
Caco2 Bg-C R-C/R-B (2)
Caco2 Bg-C R-C/R-B (3)
Log. (Caco2 Bg-C R-C/R-B (1))
Log. (Caco2 Bg-C R-C/R-B (2))
Log. (Caco2 Bg-C R-C/R-B (3))
y = 1.2838ln(x) + 20.447
R² = 0.99046
y = 0.9735ln(x) + 19.396
R² = 0.93011
y = 1.5095ln(x) + 18.083
R² = 0.94585



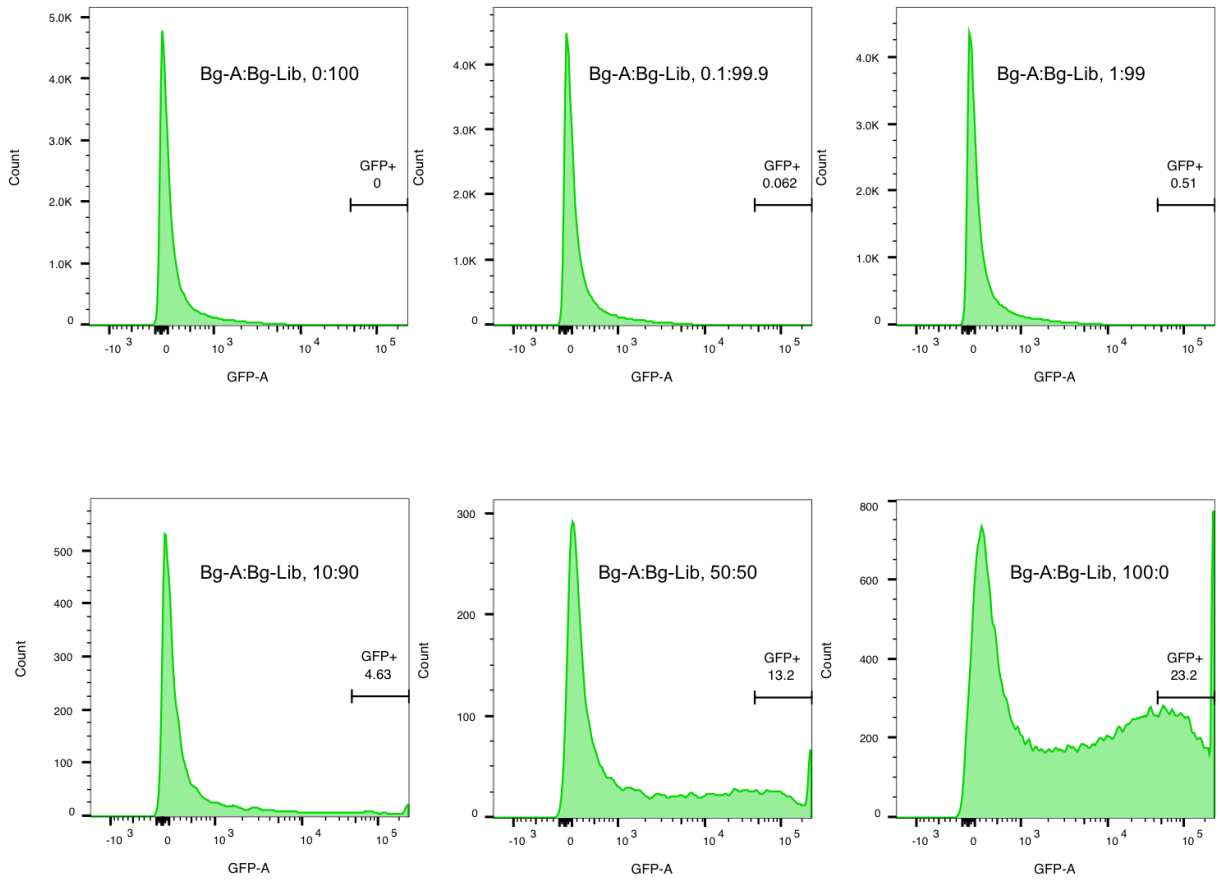
Supplementary Figure S4. Error load quantification.

Independent barcoded populations, Bg-A, Bg-B, and Bg-C, were generated in cell lines: HEK293T, MB-MDA 231, and Caco2. The barcoded populations were co-transfected with each of the Recall plasmids (R-A, R-B, R-C) containing GFP + dCas9-VPR plasmid independently, causing instances of either match or mismatch with regards to the barcoded gRNA and Recall plasmid. GFP expression was measured 48 hours after transfection. Error load was quantified by comparatively tallying the values of the GFP-histogram, from high to low GFP intensity, of matching and mismatching recall samples. Specifically, sum totals of matching recall events were tabulated with respect to and along with each new accruing mismatch recall event. From these tabulations, % activation at a given error was calculated using the formula:

$$\frac{\Sigma\, matching\ recall\ events\ at\ given\ error}{total\ matching\ events\ analyzed}.$$

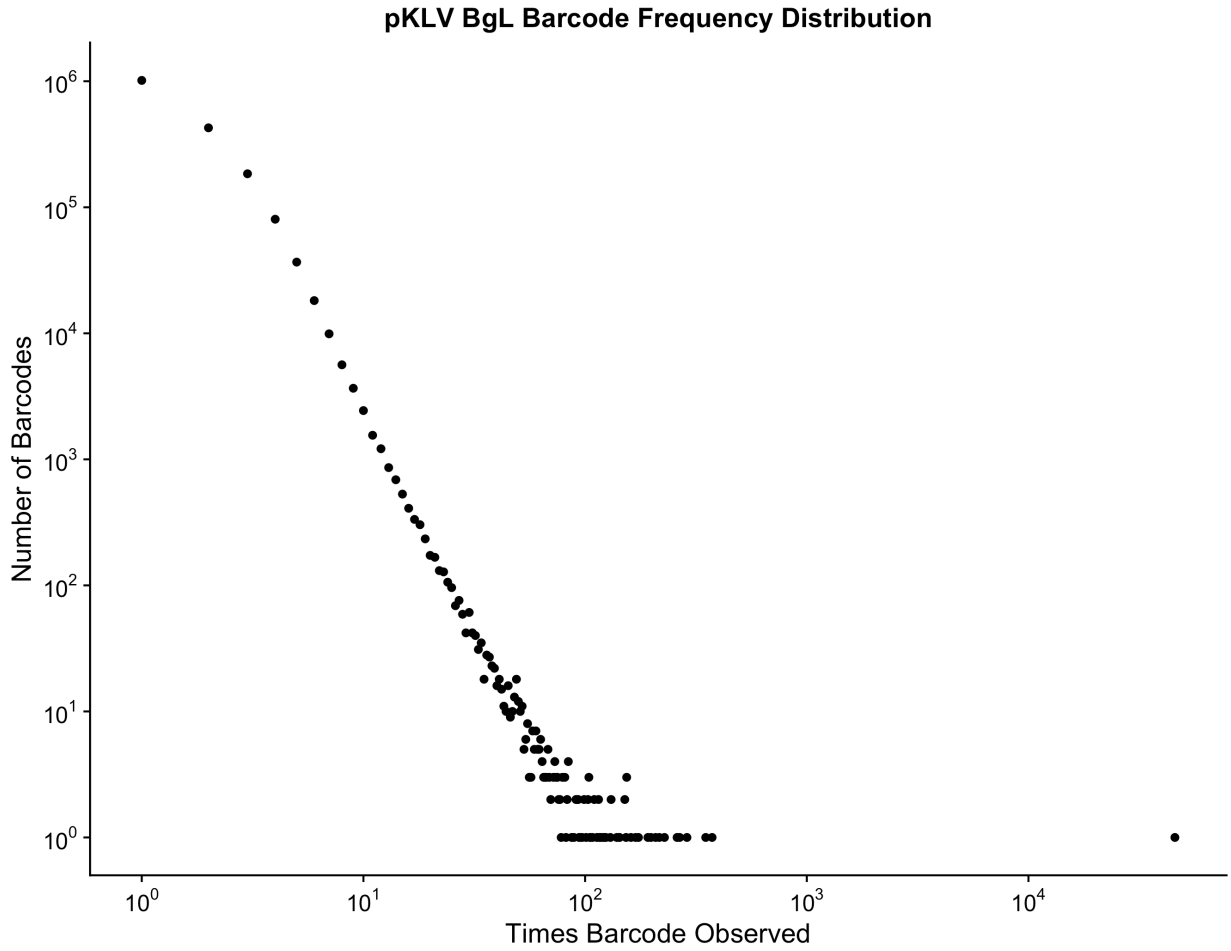
% Error was calculated using the formula: $\frac{\Sigma\, mismatching\ recall\ events}{\Sigma\, matching\ recall\ events}$. The following data was charted onto a scatter plot and, using least squares fitting, generated exponential equations to model the data. (**Suppl Fig S4**). Fits for a vast majority of samples had an R^2 value of > 0.95. Experiments were performed in biological triplicate.

Supplement Figure S5



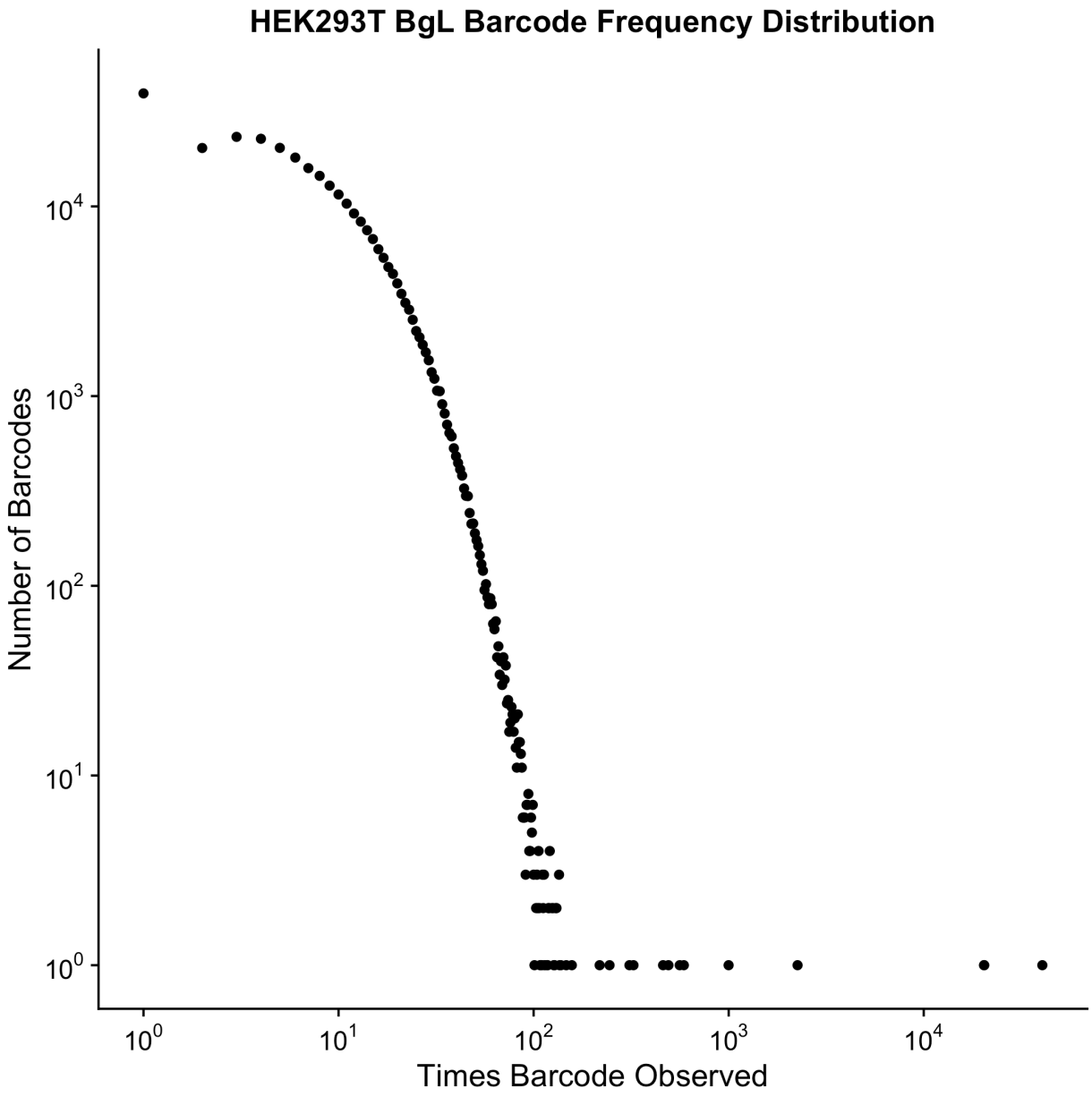
Supplementary Figure S5. HEK293T Bg-A/Bg-Library lineage dilution standard.
To initialize a standard and generate the gates for lineage isolation, a range of HEK293T Bg-A/ Bg-library dilutions were plated in a 6 well plate: 0%Bg-A, 0.1%Bg-A, 1%Bg-A, 10%Bg-A, 50%Bg-A, and 100%Bg-A. The barcoded cell populations were co-transfected with Recall-A_GFP and dCas9-VPR. Cells were analyzed via flow cytometry 48 hours post transfection. Sorting gates were set based off of the 0%Bg-A to both maximize capture of GFP positive cells while minimizing the capture noisy non-lineages of interest.

Supplement Figure S6



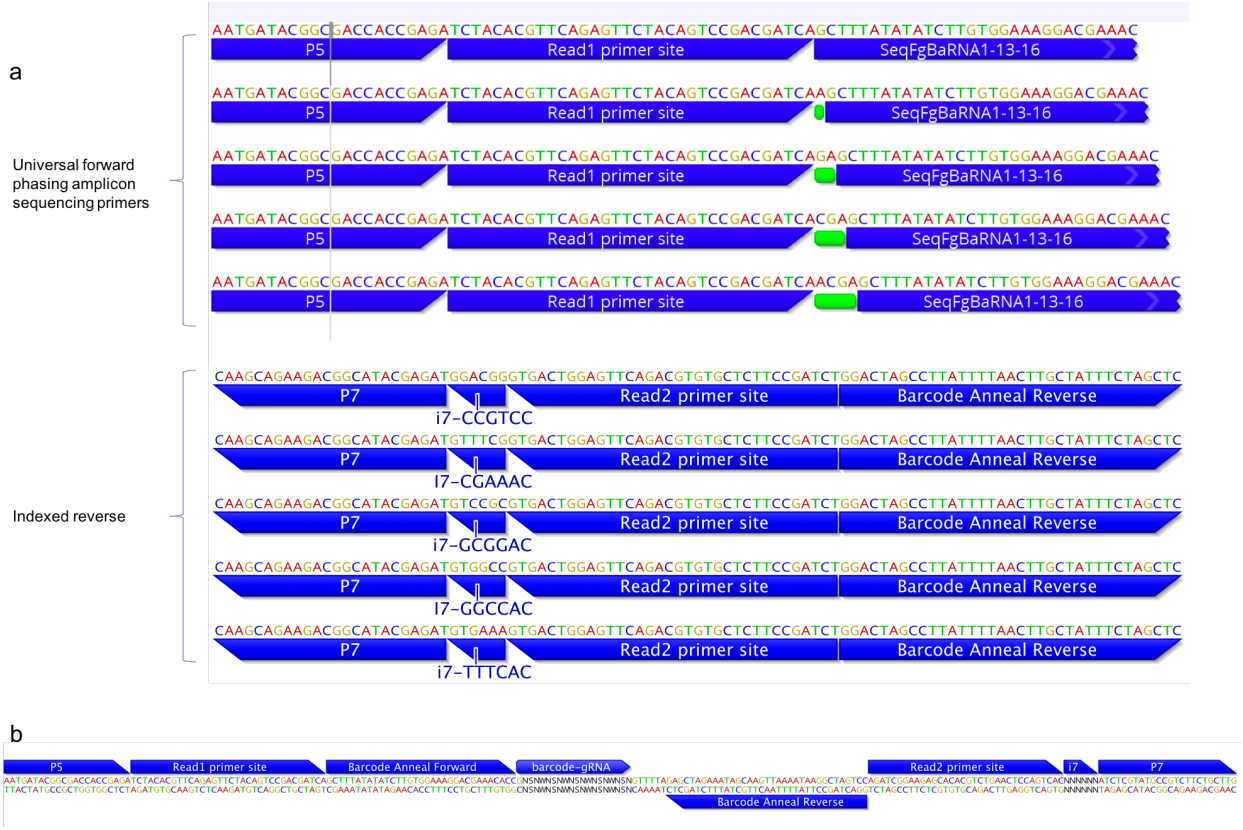
Supplementary Figure S6. Barcode sampling of Lentiviral plasmid Bg-Library.
Barcode sampling of plasmid DNA was performed to confirm the generation of a high diversity barcode-gRNA lentiviral transfer vector library. The barcodes were amplified from 50ng of the Lentiviral plasmid Bg-Library DNA and sequenced. This graph shows the barcode read frequency distribution of the library at a depth of 4.1×10^6 reads. Barcode sampling revealed 1.79×10^6 unique barcodes with >95% of the reads being represented by only 4 or fewer reads.

Supplement Figure S7



Supplementary Figure S7. Barcode sampling of HEK293T Bg-Library.
Barcode sampling of genomic DNA was performed to confirm the generation of a high diversity barcode-gRNA HEK293T library. The barcodes were amplified from 16µg of the HEK293T Bg-Library genomic DNA and sequenced. This graph shows the barcode read frequency distribution of the library at a depth of $3.88x10^6$ reads. Barcode sampling revealed $3.01x10^5$ unique barcodes.

Supplement Figure S8



Supplementary Figure S8. Barcode amplification primers.
(a) Annotated list of both forward and reverse amplification primers. The universal phasing amplicon forward primers and the indexed reverse primers contain both the flanking barcode annealing regions along with the relevant Illumina adaptor/indices, allowing for one-stage amplification of the genomically encoded barcoded gRNA. Utilization of the mixed universal forward phasing primers generates amplicons with staggered read 1 start sites, introducing necessary base diversity during Illumina sequencing of the barcoded amplicons. **(b)** Template of the amplified barcoded gRNA (not including the different phasing variants).

Supplementary Figure S9. Lineage-specific BAX expression time-lapse imaging.
Attached are movies that display lineage-specific activation of BAX and GFP in Caco-2 cells. Images were collected at 1 hour intervals for 24 hours of i) GFP only - R-A_GFP ii) GFP and BAX - R-A_BAX _GFP and iii) GFP and BAX D71A - R-A BAX D71A _GFP. In conditions that simultaneously express GFP and BAX, a signal delay is noted, showing GFP expression with subsequent Annexin V Red staining from BAX induced apoptosis. In the GFP and BAX D71A condition, the hyperactive BAX variant induces apoptosis significantly faster, resulting in contaminant Annexin V red and GFP signal onset.

Filenames of Barcode match movies:
Caco-2 phase_red_green_Bg-A/R-A_GFP
Caco-2 phase_red_green_Bg-A/R-A_BAX _GFP
Caco-2 phase_red_green_Bg-A/R-A BAX D71A _GFP
Caco-2_red_green_Bg-A/R-A_GFP
Caco-2_red_green_Bg-A/R-A_BAX _GFP
Caco-2_red_green_Bg-A/R-A BAX D71A _GFP

Filenames of Barcode mismatch movies:
Caco-2 phase_red_green_Bg-B/R-A_GFP
Caco-2 phase_red_green_Bg-B/R-A_BAX _GFP
Caco-2 phase_red_green_Bg-B/R-A BAX D71A _GFP
Caco-2_red_green_Bg-B/R-A_GFP
Caco-2_red_green_Bg-B/R-A_BAX _GFP
Caco-2_red_green_Bg-B/R-A BAX D71A _GFP