Genomic dissection of enhancers uncovers principles of combinatorial regulation and cell type-specific wiring of enhancer-promoter contacts

Verena Thormann[1], Maika C. Rothkegel[1], Robert Schöpflin[1], Laura V. Glaser[1] Petar Djuric[1],Na Li[1], Ho-Ryun Chung[1], Kevin Schwahn[1], Martin Vingron[1], Sebastiaan H. Meijsing[1§]

[1]Max Planck Institute for Molecular Genetics, Ihnestraße 63-67, 14195 Berlin, Germany

[§] Corresponding author

Contact information:

Sebastiaan H. Meijsing

Max Planck Institute for Molecular Genetics

Ihnestrasse 63-73

14195, Berlin, Germany

Tel: +49-30-84131176

Email: meijsing@molgen.mpg.de

This PDF includes:

- **Supplemental Figures**
- **Supplemental Methods**
- **Supplemental Tables**
- **Supplemental reference**

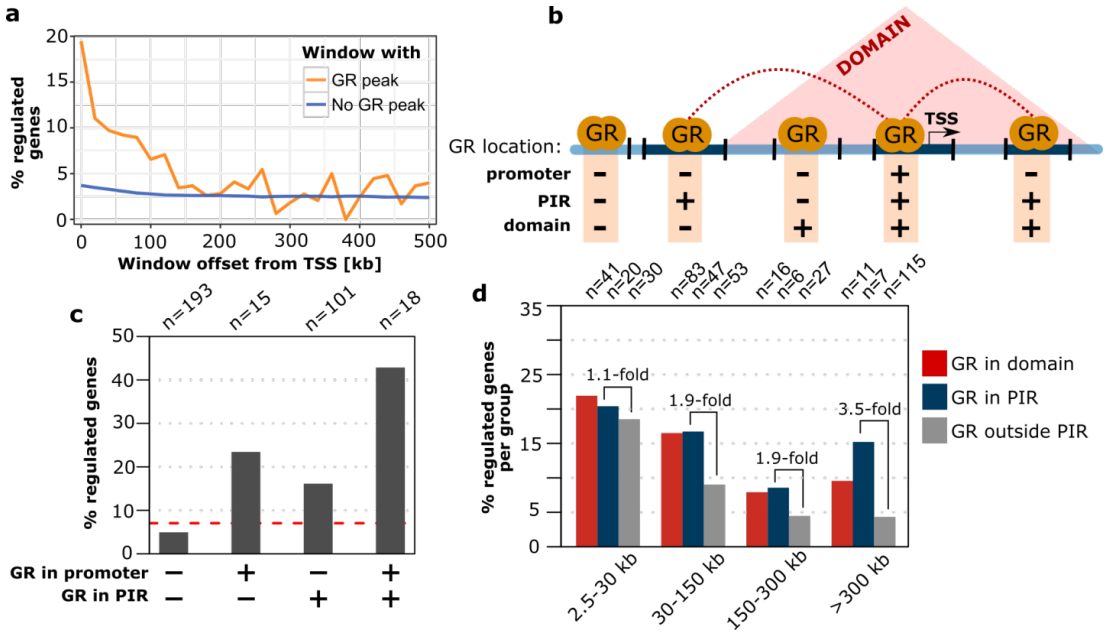
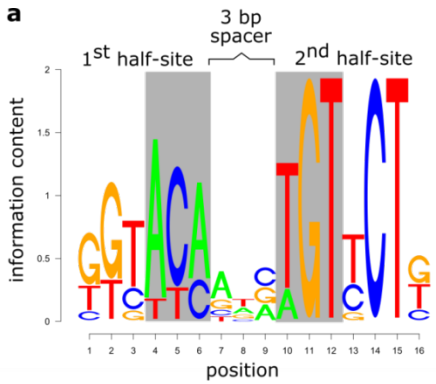


Figure S1. Linking GR binding to the GR-dependent regulation of genes. (a) Percentage of genes regulated by GR in IMR90 cells (absolute log2 fold change (|log2FC|) upon dexamethasone treatment > 0.5) vs. distance between the TSS and the nearest active GR peak (orange curve). Genes were grouped in distance intervals of 20 kb. Genes without an active GR peak between the TSS and the end of the distance interval are shown as a blue curve. (b) Cartoon depicting the different regions that were used to test if the link between GR binding and gene regulation benefits from incorporating information regarding the three dimensional organization of the genome in the nucleus. Promoter regions: Restriction fragments with the TSS of genes; PIR: promoter interacting region with increased contact frequencies with promoter regions based on Hi-C data. Domain: contact domains based on (1) (c) Percentage of genes regulated by GR (|log2FC| > 0.5) for different GR binding configurations: (i) without an active GR peak in the promoter region and all promoter interacting regions (shown as red dashed line), (ii) active GR peak in the promoter region only, (iii) active GR peak in PIR only, and (iv) Active GR peak in the promoter region and at least one PIR. (d) Percentage of genes regulated by GR (|log2FC| > 0.5) for genes grouped by the location of the nearest active GR peak: (i) in a PIR (blue bars), (ii) in the same chromatin domain (red bars), and (iii) outside of PIRs (gray bars).

a



b

locus	clone	editing	sequence	Matrix ID	sequence p-value	difference to wt log(p)
GILZ	wt	-	AGAACATTGGGTTCC	M00955 M00192 M00205	1.6E-03 3.8E-03 1.3E-02	-
	G1-4	1 bp insertion	AGAACCATTGGGTTCC	M00955 M00192 M00205	3.7E-01 1.5E-01 1.9E-01	2.35 1.59 1.17
	G5	12 bp deletion	AGAAC----------CACATATACAA	M00955 M00192 M00205	2.3E-01 1.4E-01 1.4E-01	2.16 1.57 1.04
	G6	23 bp deletion	ACCA----------CATATACAACAG	M00955 M00192 M00205	3.5E-01 1.4E-01 1.3E-01	2.33 1.57 1.00
DUSP1	wt	-	GGAACATTCTGCGG	M00955 M00205 M00192	4.6E-01 5.3E-01 1.3E-02	-
	D1	1 bp insertion	GGAACATTTCTGCGG	M00955 M00205 M00192	7.8E-01 5.6E-01 1.3E-02	0.23 0.02 0.02
	D2	allele 1: 1 bp insertion	GGAACATTTCTGCGG	M00955 M00205 M00192	7.8E-01 5.6E-01 1.3E-02	0.23 0.02 0.02
		allele 2: 1 bp insertion	GGAACATTCCTGCGG	M00955 M00205 M00192	7.8E-01 5.9E-01 1.3E-02	0.23 0.04 0.02
	D3	1 bp deletion	GGAACA-TCTGCGGT	M00955 M00205 M00192	7.8E-01 6.0E-01 1.3E-02	0.23 0.05 0.01
	D4	allele 1: 5 bp deletion	GGAACATT-----GTTTCCT	M00955 M00205 M00192	6.1E-02 1.4E-01 8.9E-03	-0.87 -0.59 -0.16
		allele 2: 1 bp deletion	GGAACA-TCTGCGGT	M00955 M00205 M00192	7.8E-01 6.0E-01 1.3E-02	0.23 0.05 0.01
	D5	allele 1: 1 bp deletion	GGAACA-TCTGCGGT	M00955 M00205 M00192	7.8E-01 6.0E-01 1.3E-02	0.23 0.05 0.01
		allele 2: 216 bp insertion	NA	M00955 M00205 M00192	NA	NA

Figure S2. Effect of CRISPR/Cas9-induced indels on the *GILZ* and *DUSP1* GBS1 GR consensus motif-matches. (a) Position weight matrix of a GR consensus motif (Matrix-ID: M00205). For orientation, the first 3 bp of each GR half-site are highlighted in grey. (b) Table summarizing how indels for clonal lines and loci as indicated influence the positioning of GR half sites and how the changes influence the predicted p-value of the motif-match as calculated using the sTRAP tool using standard settings (matrix file: transfac 12.1 metazoans, background model: chordate_conserved_elements, multiple test correction: Benjamini-Hochberg).

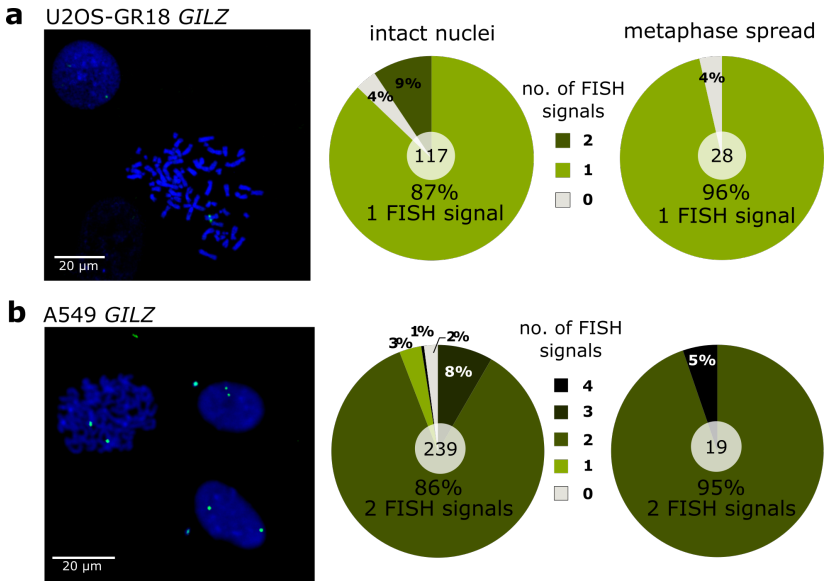


Figure S3. DNA FISH analysis of the *GILZ* locus Xq22.3 in U2OS-GR18 and A549 cells. (a) Representative image of a DNA FISH experiment targeting the *GILZ* locus (green) for U2OS-GR18 cells stably expressing GR. DNA was counterstained by Hoechst 33342 (blue). Pie charts show the quantification of the number of FISH signals per nucleus for either metaphase spreads or intact nuclei. The total number of analyzed metaphase spreads or intact nuclei is depicted in the inner circle of each pie chart. (b) Same as for (a) except that the cell line analyzed was A549.

a TRAP analysis *GILZ* GBS1-4

GBS	sequence	Matrix ID	TRAP score
GBS1	AGAACACATTGGGTTCC	M00955	10.4
		M00192	8.6
		M00205	8.5
GBS2	AGGACACATTCTGTTAA	M00955	3.0
		M00192	5.7
		M00205	8.5
GBS3	TGAACACACGGTGTTTG	M00955	3.3
		M00192	6.2
		M00205	1.2
GBS4	AGAACACAAGTGCAGG	M00955	7.4
		M00192	6.7
		M00205	3.0

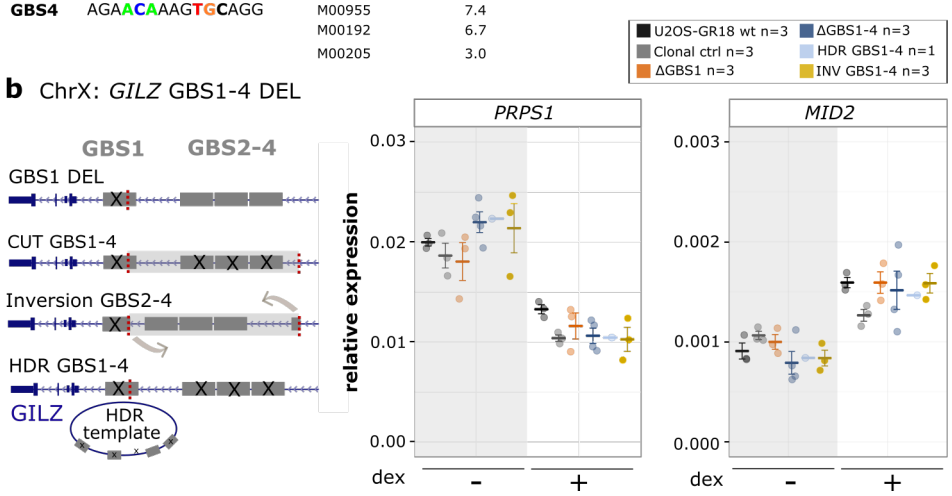


Figure S4. TRAP motif score of GILZ GBS1-4 and effect of *GILZ* GBS1 deletion on neighboring GR-regulated genes *PRPS1* and *MID2*. (a) Table summarizing the weight score as determined by TRAP using standard settings (matrix file: transfac_2010.1 vertebrates, background model: human promoters, multiple test correction: Benjamini-Hochberg). (b) Left: Schematics of the *GILZ* GBS1-4 enhancer and the various clonal lines that were generated to assess the role of the GBSs found at this locus, targets for gRNAs highlighted in red. Clonal cell lines with GBS1-4 deletion were either generated by cutting with two gRNAs or by using a homology directed repair (HDR) template to mutate each of the 4 GBSs by HDR. Right: Relative *PRPS1* and *MID2* expression levels as determined by qPCR for clonal lines with GBS1 deletion (n=3), cut GBS1-4 deletion (n=3), inversion GBS1-4 (n=3) and HDR GBS1-4 deletion (n=1). Circles indicate values for each individual clonal line. Horizontal lines and error bars: Averages ± SEM for cells treated with vehicle or 1 µM dex overnight.

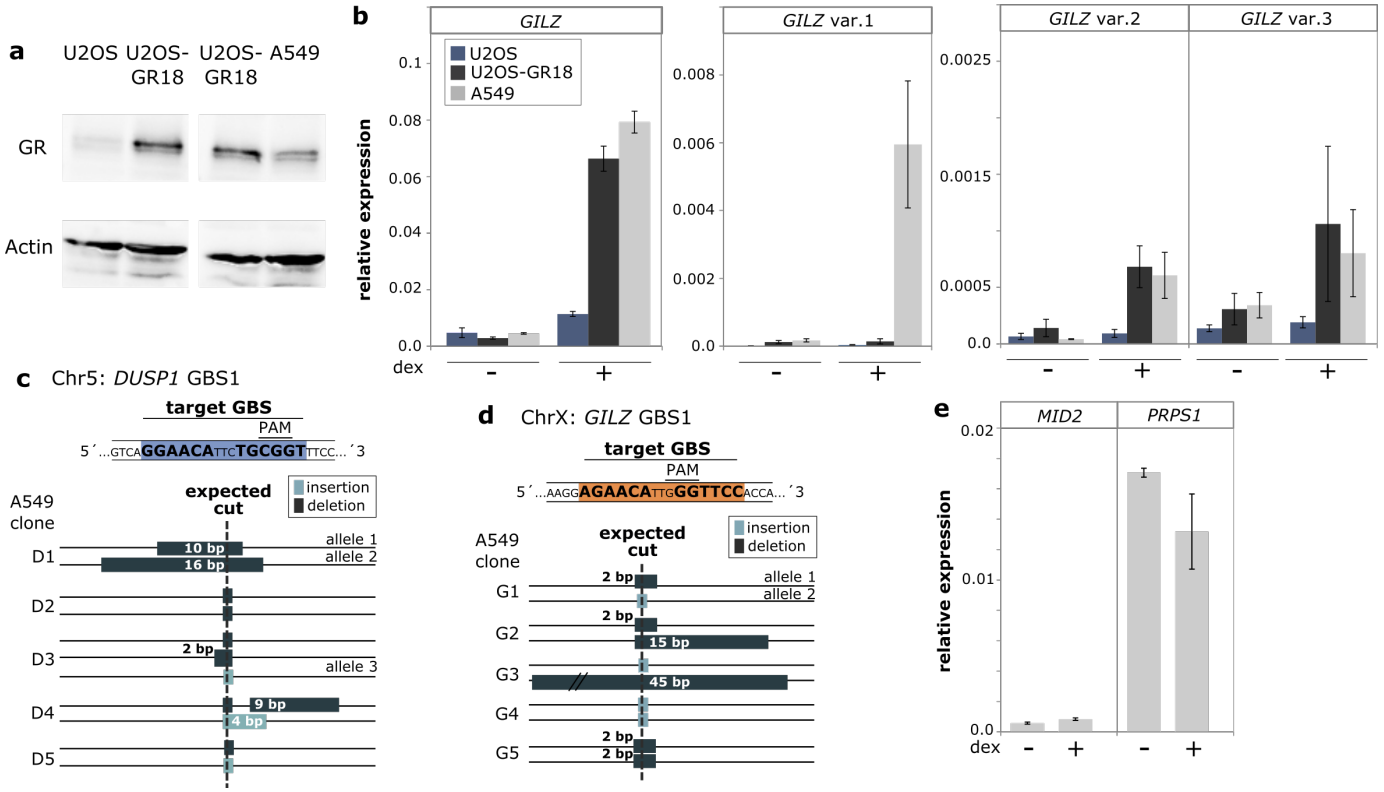


Figure S5. Genotyping results of A549-derived clonal cell lines successfully edited at the *GILZ* or *DUSP1* locus. (a) Expression levels of GR in U2OS, U2OS-GR18 and A549 cells was analyzed by immunoblotting with antibodies against GR and actin as a control for loading. (b) Relative mRNA expression levels as determined by qPCR for *GILZ* transcript variants as shown for U2OS, U2OS-GR18 and A549 cells. Averages ± SEM (n≥3) for cells treated with vehicle or 1 µM dex overnight. (c) Schematic representation of the CRISPR/Cas9-induced insertions (light grey) or deletions (dark gray) for each of the alleles of the *GILZ* GBS1 locus in A549 cells. (d) Same as for (c) expect for the *DUSP1* GBS1 locus. The gRNAs used for CRISPR/Cas9-editing and the corresponding expected location of the Cas9-induced DNA double-strand break (dashed line 3 bp upstream of the respective protospacer adjacent motif (PAM)), are indicated on the top of each figure panel. (e) Relative mRNA expression of the *MID2* and *PRPS1* genes upon overnight treatment with 1µM dex or ethanol as vehicle, was determined by qPCR in A549 cells. Averages ± standard error of mean (SEM) are shown (n=3).

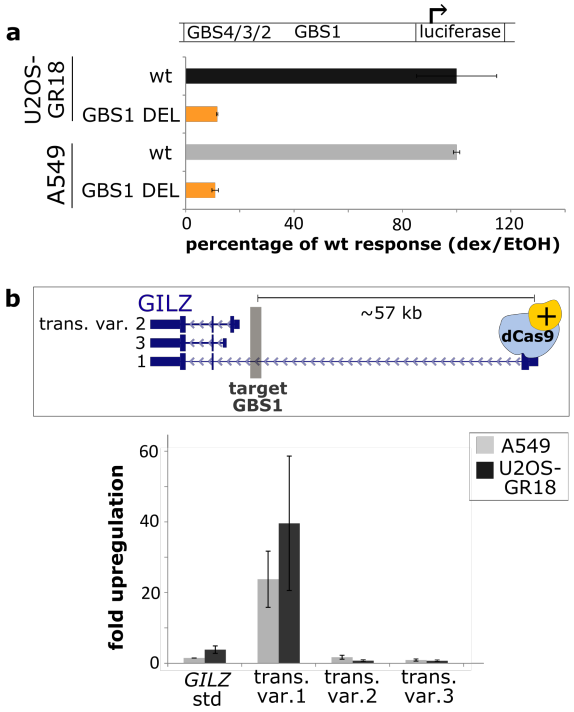


Figure S6. Luciferase reporter analysis of the *GILZ* GBS1-4 enhancer and activation of the endogenous *GILZ* transcript variant 1 by targeted recruitment of dCas9-SAM. (a) Transcriptional activity of luciferase reporters containing a minimal promoter and either the wildtype *GILZ* GBS enhancer region or one with a GBS1 deletion. Activity is shown as percentage of the fold induction upon overnight treatment with 1 µM dexamethasone observed for wildtype reporter ± SEM (n = 3). (b) Comparison between U2OS-GR18 and A549 cells of the transcriptional induction of the endogenous *GILZ* transcript variants as indicated by dCAS9-SAM. Average fold induction by targeted recruitment to the *GILZ* TSS1 relative to gRNAs targeting other regions ± SEM (n=3) is shown.

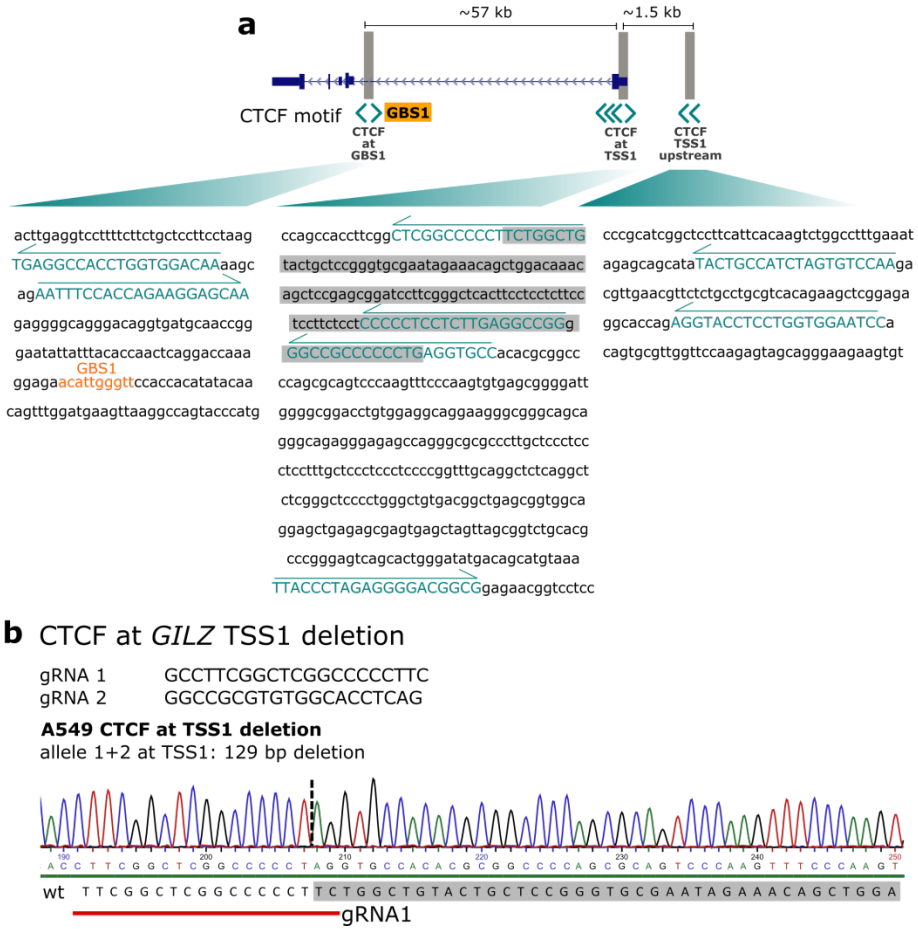


Figure S7. CTCF motif-matches at the *GILZ* locus and genotyping results for an A549 clonal line where we deleted a region with multiple CTCF binding sites at the *GILZ* TSS1. (a) CTCF motif-matches (highlighted in turquoise) and their orientation (indicated with arrows) as well as GBS1 (highlighted in orange) at: Left: GBS1; Middle: TSS1 of *GILZ* transcript variant 1 and Right: CTCF ChIP-seq peak approximately 1.5 kb upstream of TSS1. Nucleotides deleted for the A549 CTCF deletion clone are highlighted in grey. (b) Sanger sequencing for the A549 CTCF deletion clone. In addition, the gRNA used for CRISPR/Cas9-editing and the expected location of the Cas9-induced DNA double-strand break (dashed black line) are indicated.

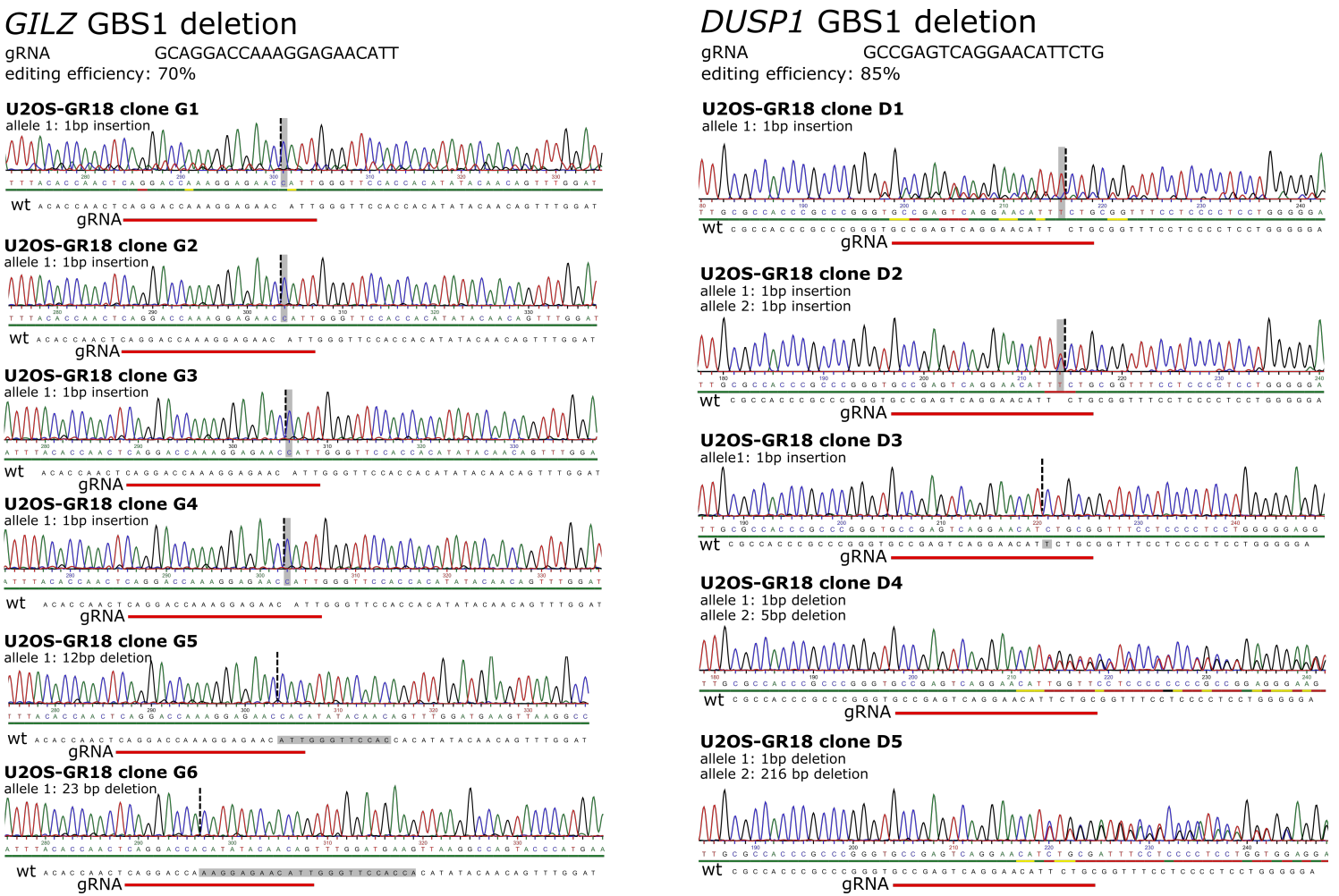


Figure S8. Genotyping results for U2OS-GR18 *GILZ* GBS1 and *DUSP1* GBS1 deletion cell lines. Sanger sequencing for U2OS-GR18 *GILZ* GBS1 and *DUSP1* GBS1 deletion cell lines. Deleted nucleotides are highlighted in grey. In addition, the gRNA used for CRISPR/Cas9-editing, the editing efficiency and the expected location of the Cas9-induced DNA double-strand break (dashed black line) are indicated.

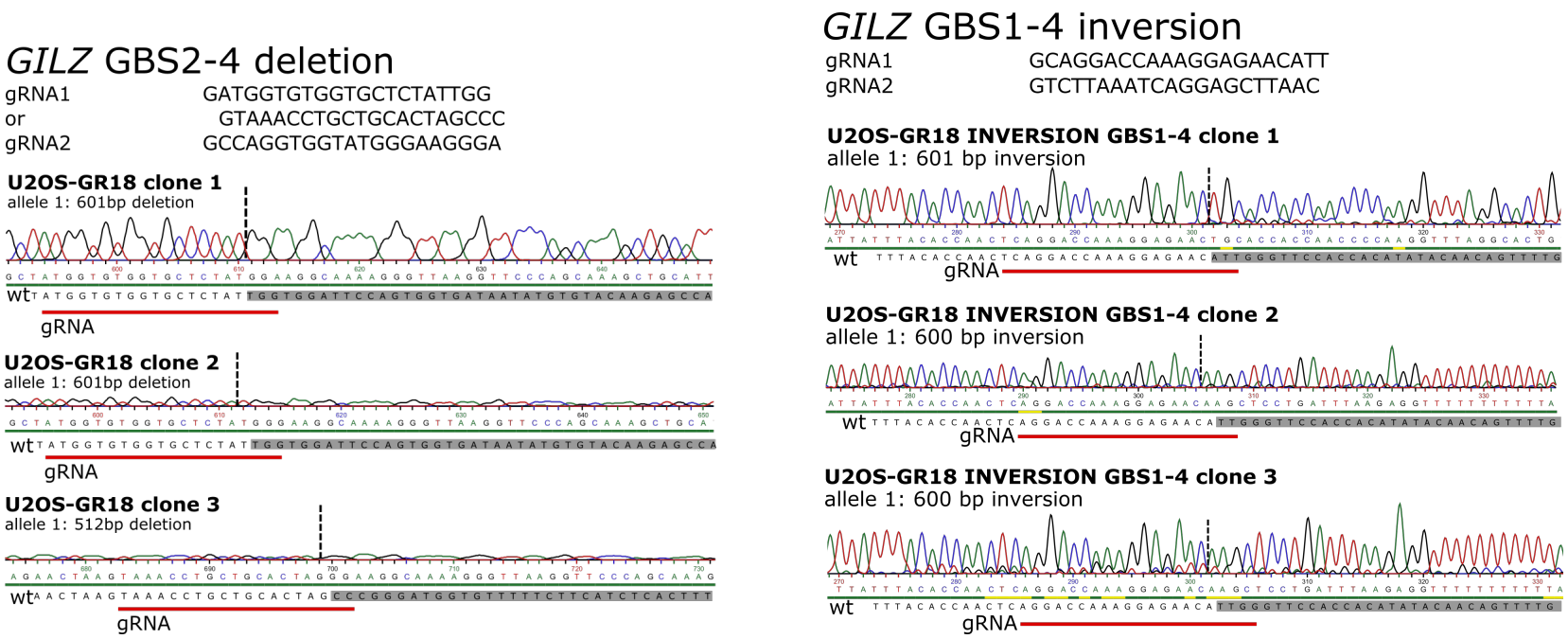


Figure S9. Genotyping results for U2OS-GR18 *GILZ* GBS2-4 deletion and *GILZ* GBS1-4 inversion cell lines. Sanger sequencing for U2OS-GR18 *GILZ* GBS2-4 deletion and *GILZ* GBS1-4 inversion cell lines. Deleted nucleotides are highlighted in grey. In addition, the gRNA used for CRISPR/Cas9-editing and the expected location of the Cas9-induced DNA double-strand break (dashed black line) are indicated.

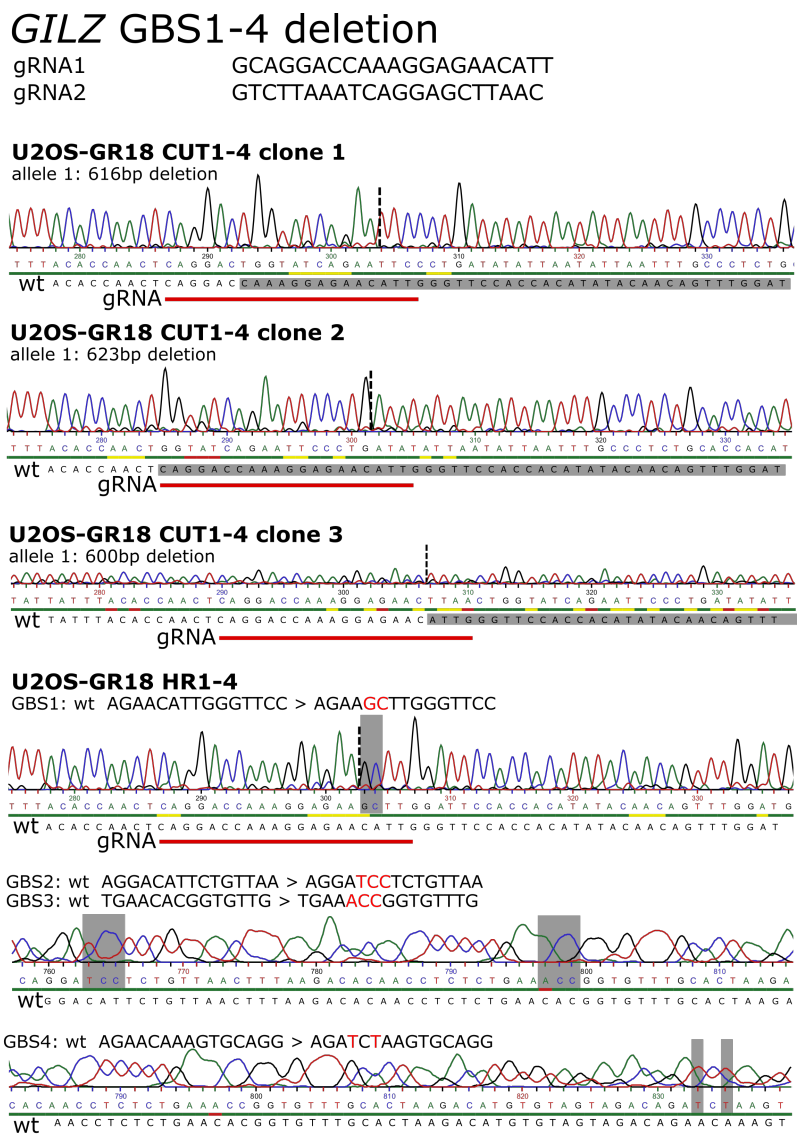


Figure S10. Genotyping results for U2OS-GR18 *GILZ* GBS1-4 deletion cell lines. Sanger sequencing for U2OS-GR18 *GILZ* GBS1-4 deletion cell lines. Deleted or mutated nucleotides are highlighted in grey. In addition, the gRNA used for CRISPR/Cas9-editing and the expected location of the Cas9-induced DNA double-strand break (dashed black line) are indicated.

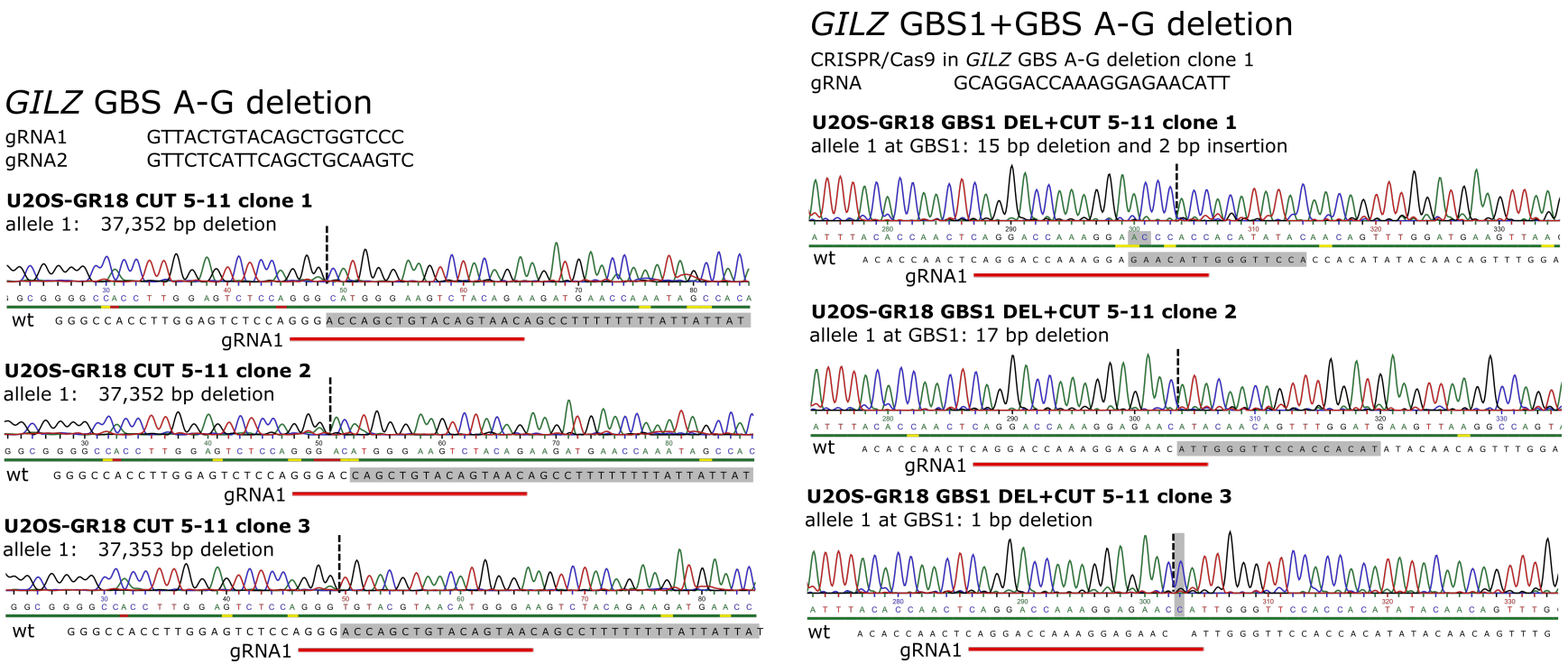


Figure S11. Genotyping results for U2OS-GR18 *GILZ* GBS A-G and GBS1 + GBS A-G deletion cell lines. Sanger sequencing for U2OS-GR18 *GILZ* GBS A-G and GBS1 + GBS A-G deletion cell lines. Deleted nucleotides are highlighted in grey. In addition, the gRNA used for CRISPR/Cas9-editing and the expected location of the Cas9-induced DNA double-strand break (dashed black line) are indicated.

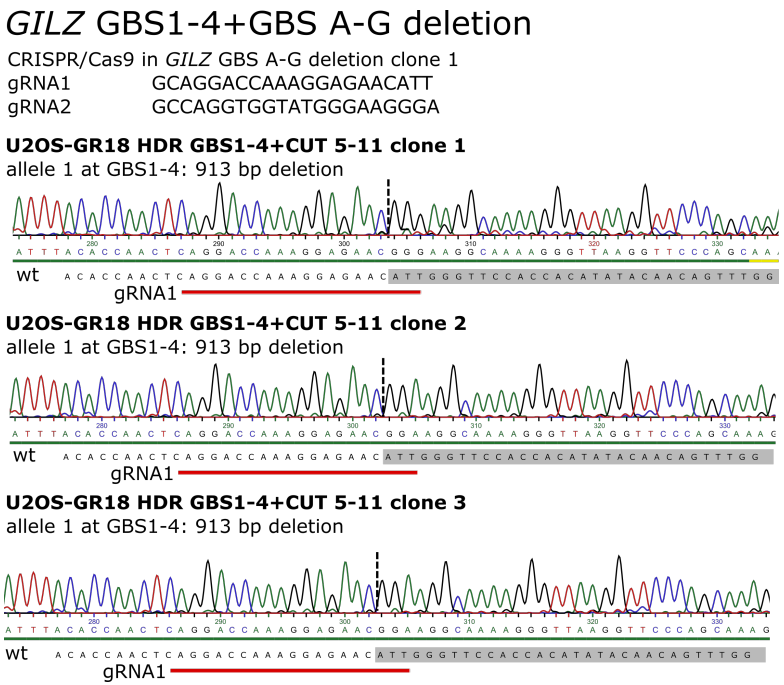


Figure S12. Genotyping results for U2OS-GR18 *GILZ* GBS1-4 + GBS A-G deletion cell lines. Sanger sequencing for U2OS-GR18 GBS1-4 + GBS A-G deletion cell lines. Deleted nucleotides are highlighted in grey. In addition, the gRNA used for CRISPR/Cas9-editing and the expected location of the Cas9-induced DNA double-strand break (dashed black line) are indicated.

Figure S13. Genotyping results for A549 *GILZ* GBS1 and *DUSP1* GBS1 deletion cell lines. Sanger sequencing for A549 *GILZ* GBS1 and *DUSP1* GBS1 deletion cell lines. In addition, the gRNA used for CRISPR/Cas9-editing, the editing efficiency and the expected location of the Cas9-induced DNA double-strand break (dashed black line) are indicated.

Supplemental Methods

Cell lines, plasmids, transient transfections luciferase assays and immunoblotting

A549 (ATCC CCL-185), U2OS and U2OS cells stably expressing rat GRα (U2OS-GR18) (2) were cultured in DMEM supplemented with 5% FBS. IMR90 cells (ATCC CCL-186) were cultured in EMEM supplemented with 10% FBS. The pGL3-promoter *GILZ* reporter and variants with mutated GBSs have been described previously (3). The luciferase reporter construct containing the endogenous *GILZ* variant 2 promoter and approximately 100 bp downstream of the transcriptional start site driving the expression of the luciferase gene (Fig. 6a,b) was generated by amplifying the region (GRCh37/hg19 ChrX:106960191-106962152) from genomic DNA and cloning it into pGL3-basic (Promega). The GBS1 of this construct was mutated by site directed mutagenesis using the primers listed in Table S1. dCas9-SAM plasmids have been described elsewhere (4). Modified gRNAs including 2 MS2 stem loops (4) were designed using the CRISPR Design tool (5) (http://crispr.mit.edu, sequences listed in Table S2), ordered as gBlocks from IDT and cloned into the pCR-Blunt vector (Invitrogen). Transient transfections and immunoblotting were done essentially as described (6). Luciferase activity was measured using the dual luciferase assays kit (Promega).

Genome editing using CRISPR/Cas9

gRNAs for the deletion of GBSs (Table S3) were designed using the CRISPR Design tool (5) (http://crispr.mit.edu, sequences listed in Table S2). gRNAs were ordered as gBlocks from IDT and cloned into the pCR-Blunt vector (Invitrogen). Cells were transfected using 600 ng each of the gRNA construct, hCas9 expression construct (Addgene #41815) (7) and pPUR (Clontech) by nucleofection (U2OS-GR18: kit V; A549 kit T, Lonza) according to the manufacturer's instructions. One day after nucleofection, we added puromycin (U2OS-GR18: 10 µg/ml, A549: 2 µg/ml) for 24h to select for transfected cells. To genotype single cell-derived clonal lines, genomic DNA was isolated using the Blood and Tissue kit (Qiagen), the targeted region was amplified by PCR using primers spanning the target (Table S4) and PCR products were analyzed by sequencing. For gene editing using homology directed repair (HDR), we added 3 µg of a HDR template to the transfection mix. Furthermore, to increase gene editing by HDR, we treated transfected cells for 24 h with 10 µM SCR7 prior to puromycin selection. The HDR template for the *GILZ* GBS1 region (GRCh37/hg19 ChrX:106960177-106962953) was PCR amplified from genomic DNA and cloned into the pCR-Blunt vector (Invitrogen). GBSs1-4 within the HDR template constructs were mutated by site directed mutagenesis using the primers listed in Table S1.

RNA preparation and analysis

Cells were cultured to confluency and treated overnight with 1 µM dexamethasone or vehicle control (ethanol). The next day, RNA was extracted, reverse transcribed and analyzed by qPCR as described previously (6) using the primer pairs listed in Table S5.

dCas9-SAM activation of endogenous target genes

The ability of dCas9-SAM to activate endogenous target genes was tested by transfecting 600 ng each of dCas9-VP64, MS2-p65-HSF1 activator complex, the modified gRNA including 2 MS2 stem

loops and a GFP expression construct by nucleofection (Lonza) according to the manufacturer's instructions. 24 h post-transfection, GFP-positive cells were isolated by FACS sorting, RNA was isolated and analyzed by qPCR as described above.

Chromatin Immunoprecipitation (ChIP and ChIP-seq)

ChIP assays targeting GR were performed as described (6). In brief, *GILZ* GBS1 deletion clone G2 and *DUSP1* GBS1 deletion clone D2 were cultured to confluency and treated for 90 min with 1 μM dexamethasone or vehicle control (ethanol). Next, chromatin was fixed at room temperature for 3 min with 1% formaldehyde. After quenching with 200mM glycine, the cells were sonicated using a Bioruptur (Diagenode). GR-bound regions were immunoprecipitated using the N499 GR-antibody and pulled down by protein A/G agarose beads (Santa Cruz Biotechnology). ChIP-assays targeting CTCF were essentially done as described above using a polyclonal CTCF-antibody (Active Motif, Cat. No. 61311) except that a modified RIPA wash buffer (50 mM HEPES-KOH, 1 mM EDTA, 1% NP40, 0.7% Na-deoxycholat, 500 mM LiCl, pH 7.5) was used. For ChIP assays targeting H3K27ac, three million fixed IMR90 cells (with formaldehyde (final concentration 1%) for 3 minutes) were lysed in 300 μl chromatin lysis buffer (50 mM Tris-HCl pH 8.1, 100 mM NaCl, 1% SDS, 3% Triton X-100, 5 mM EDTA, 0.2% NaN_3) supplemented with 3x protease inhibitors (Roche complete protease inhibitor cocktail) on ice for 10 minutes. Lysates were then diluted 3x with dilution buffer (50 mM Tris-HCl pH 8.6, 100 mM NaCl, 5 mM EDTA, 0.2% NaN_3) and homogenized ten times with a syringe (271/2 gauge). The lysate was then aliquoted (200 μl) into 1.5 ml TPX polymethylpentene tubes (Diagenode) and sheared at 4 °C in a Bioruptor Pico for 2x10 cycles. Chromatin Immunoprecipitation: The ChIP was performed using the Diagenode Auto Histone ChIPseq kit on an IPstar SX-8G compact automated system (Diagenode) using their indirect method (Ag + Ab → Beads). 1 μg of Diagenode ChIPseq grade rabbit polyclonal antibody (H3K27Ac; pAb-196-050) were used per ChIP. After de-crosslinking, RNaseA and proteinase K digestion, the DNA was isolated using ChIP DNA concentrator columns (Zymo Research, D5205) according to the manufacturer's instructions. Subsequently, DNA was either quantified by qPCR using primer pairs as listed in Table S6, or 10 ng was used to prepare ChIP-seq libraries.

Circularized chromosome conformation capture (4C)

4C experiments were performed as previously described (8), using $5x10^6$ dexamethasone-treated cells (1 μM, 90 minutes). 4-bp cutters were used as primary (CviQI, Thermo Fisher Scientific) and secondary (DpnII, Thermo Fisher Scientific) restriction enzymes. 4C libraries were generated for two biological replicates. Fragments for sequencing were generated for each cell line by inverse-PCR on the 4C library using viewpoint-specific primers (Table S7), which include adaptors for subsequent high throughput sequencing (Illumina HiSeq2500). Reproducibility between replicates was evaluated by correlation between smoothed profiles in a range of ± 500 kb around the viewpoint (excluding the viewpoint region ±5 kb): Spearman's rho A549: 0.74, U2OS-GR18: 0.51; Pearson's r: A549: 0.91, U2OS-GR18: 0.92. A representative profile is shown for each cell line in Fig. 7.

Fluorescent in situ hybridization (FISH)

To arrest cells in metaphase, A549 and U2OS-GR18 cells were cultivated in DMEM supplemented

with 5% FBS and 0.1 ug/ml colcemid (Thermo Fisher Scientific) for 3 hours. Arrested cells were transferred into a hypotonic salt solution of 0.56% KCl and incubated for 10 min at 37°C. Subsequently, the cells were fixed (75% methanol+ 25% acetic acid (v/v)), transferred to microscope slides and stored in 100% ethanol at -20°C for at least two days. To prepare slides for hybridization with FISH probes, slides were rinsed in saline sodium citrate buffer (SSC) and subsequently incubated for 10 min at 37°C in a pepsin solution (1% HCl + 0.007% pepsin). Next, the slides underwent a series of washing steps (2x in PBS for 5 min each, 1x PBS-$MgCl_2$ (5% 1M Magnesium chloride in PBS) for 3 min, 1x PBS-$MgCL_2$ containing 1% formaldehyde for 10 min, 1x 70% ethanol for 3 min, 1x 85% ethanol for 3 min, 1x 100% ethanol for 3 min). FISH-probes (labeled with Green 5-Fluorescein) targeting the *GILZ* locus (BAC, RP11-81I3) were purchased from Empire Genomics and slides were hybridized with FISH probes according to the manufacturer's instructions. Images of metaphase spreads and intact FISH-labeled cells were captured using a fluorescence microscope (Zeiss LSM700).

Computational analysis

Correlating genomic GR binding with gene regulation

Data A549: All data sets for A549 cells were taken from an ENCODE time series (9). Counts per gene from RNA-seq experiments after 0h and 4h treatment with 100 nM dexamethasone were obtained from the GEO database (GEO:GSE91305, GEO:GSE91243). Differentially expressed genes were determined from replicates using DESeq2 (10). Genes were considered to be regulated, when the absolute log2 fold change was > 0.5 and the adjusted p-value < 0.01. Genes were considered to be non-regulated, when the absolute value of the log2 fold change was <0.1 or not available. Genes with intermediate log2 fold change values were not considered. Transcription start site (TSS) annotations for genome build GRCh38 were taken from the 'hsapiens_gene_ensembl' dataset accessed via the biomaRt package (11,12). For genes with multiple transcripts the most 5' TSS was used. Promoter regions were defined as 10 kb window (-5000bp, +4999bp) around the TSS. ChIP-seq data of GR binding after 1h treatment with 100nM dexamethasone was obtained from (GEO:GSE91285, GEO:GSE91357) and ChIP-seq data for H3K27ac after 4h treatment with 100nM dexamethasone from (GEO:GSE91347, GEO:GSE91282). Short reads were mapped to reference genome GRCh38 using Bowtie2 v2.1.0 (--end-to-end --sensitive) (13). Reads with mapping quality <10 and duplicate reads were removed. Replicates were pooled for further analysis. Peaks of GR binding were called using MACS2 v2.1.1.20160309 (14) using a qvalue of 0.01. To determine active GR peaks, the number of H3K27ac ChIP-seq reads was counted in a 1000 bp window (-500, +499bp) around each peak center and divided by the number of input reads +1 in the same window. GR peaks were labeled active when their level of H3K27ac was within the top 0.25 quartile of all peaks. Chromosome contact domains as well as chromatin contact lists for A549 cells after 4h of treatment with 100nM dexamethasone were obtained from (GEO:GSE92804). Interacting regions in chromatin contacts were adjusted to be 10kb around the center of the region, independent of their initial size which was 5 kb or 10 kb. A TSS is part of a chromatin contact, when it is located within an interacting region. The second region being part of this contact is termed promoter interacting region (PIR). Data IMR90: Processed data regarding changes in gene expression upon treatment of IMR90 cells with 1 μM dexamethasone for 4 h was downloaded from EBI ArrayExpress (accession number: E-MTAB-2954) (15). Genes were considered regulated, when the absolute value of the log2 fold change was > 0.5 and non-regulated when it was < 0.1, genes with a log2 fold change in

between were not considered. GR ChIP-seq data for IMR90 cells are from (15) and are available from EBI ArrayExpress under accession number E-MTAB-2955. Coordinates of ChIP-seq peaks were transferred to hg18 using the liftOver tool from UCSC Genome Browser (16). To call active GR peaks, we counted the number of H3K27ac ChIP-seq reads from hormone-treated IMR90 cells in a 1000 bp window (-500bp, +499bp) around each GR peak and divided it by the number of input reads +1 in the same window. Peaks were labeled active when their H3K27ac level was within the top 0.1 quantile of all peaks. Hi-C-based long-range chromatin interactions are from (17). Chromosome contact domain coordinates are from (1) and were transferred to hg18 using the liftOver tool from UCSC Genome Browser. TSS annotations are from Bioconductor package TxDb.Hsapiens.UCSC.hg18.knownGene: Annotation package for TxDb object(s) (Carlson M and Maintainer BP (2015). R package version 3.2.2.). For genes with multiple transcripts, we used the most 5' TSS. Genes with transcripts on different chromosomes or strands were removed.

Correlating distance between TSS and GR peaks with regulation (Fig. 1a and S1a): To determine the link between gene regulation and GR binding, genes were grouped by the distance between their TSS and the nearest active GR peak in 20kb intervals. For each interval, we computed the percentage of regulated genes or as control this percentage was calculated for genes without any active GR peak between the end of the respective interval and the TSS.

Correlating chromatin looping, GR binding and gene regulation (Fig. 1c and S1c): Genes were grouped based on whether they contained an active GR peak in either (i) the promoter region, (ii) in a distal region looping to the promoter (promoter interacting region, PIR), (iii) both promoter region and PIR or (iv) genes without an active GR peak in either region. For each of the four groups, we calculated the percentage of regulated genes.

Correlating gene regulation with GR binding inside of looping fragments, GR binding outside of looping fragments or binding within contact domains (Fig. 1d and Fig S1d): Percentage of regulated genes was calculated for groups defined as follows: Genes were grouped by the distance between their TSS and the closest active GR peak. Next, genes from each distance group were put into up to two of the following groups: (i) Closest active GR peak maps to a PIR looping to its TSS, (ii) closest active GR peak and TSS within the same contact domain as specified by (1) or (iii) closest active GR peak mapping to a non-PIR fragment. Note, group (i) and (iii) are mutually exclusive, whereas group (i) and (ii), as well as (ii) and (iii) can overlap.

ChIP-seq analysis

H3K27ac ChIP-seq IMR90 (+/-dex): Single-end short reads were mapped to reference genome hg18 using Bowtie2 v2.1.0 (--end-to-end --sensitive) (13). Reads with mapping quality <10 and duplicate reads were removed. CTCF ChIP-seq (A549 & U2OS-GR18): Paired-end short reads were mapped to reference genome hg19 using Bowtie v2.1.0 (--end-to-end --sensitive --maxins 2000). Reads with mapping quality < 10 were filtered out. Duplicates were removed using Picard-tools v2.5.0 (http://broadinstitute.github.io/picard/). Coverage profiles were generated using igvtools v2.3.55 (18,19) (count -z 5 -w 25 --pairs) and the wigToBigWig tool (20).

4C data analysis

Primer sequences were extended to the next 3' restriction site and clipped from short reads

allowing up to three mismatches for the identification of primers. Clipped short reads were mapped in single-end mode to reference genome hg19 using BWA-MEM v0.7.12 (21,22) and sorted by name afterwards. The reference genome was digested virtually according to the first cutter CviQI to obtain restriction fragments. Reads were counted per fragment to obtain interaction profiles. A read was assigned to a fragment only when the first read mapped to a CviQI site (first cutter) and its mate to a DpnII site (second cutter). All primary alignments with a mapping quality ≥ 30 for both mates were considered in the interaction profiles. In case the alignments from the same read pair were separated by less than 9 restriction fragments, only the first one was counted to avoid double counting of reads mapping to the same fragment or coming from undigested DNA fragments. The processing of BAM files to interaction profiles was done with custom Java code using HTSJDK library v1.139 (https://samtools.github.io/htsjdk/). To reduce noise, profiles were smoothed by averaging over a running window of five fragments. For better comparability between cell lines, profiles were normalized as reads per million. The scaling factor was computed from all contacts on the same chromosome excluding the viewpoint region extended by 5 kb on each side. Statistics on the number of processed reads and valid contacts are shown in Table S8.

DNA sequence motif analysis

To identify CTCF binding motif-matches at CTCF-ChIP peaks (Fig. 7c and Fig. 7-figure supplement 1), we analyzed regions of interest using the Transcription Factor Affinity Prediction (TRAP) webtool (23). The CTCF motif-matches (M01259 and M01200) shown have weight scores > 4.5. To determine the individual GR motif weight score for *GILZ* GBS1-4, we used TRAP (23) for a single nucleotide sequence (matrix IDs M00955, M00192 and M00205) using human promoters as a background model and Benjamini-Hochberg correction for multiple testing. To assess the effect of indels at the *DUSP1* GBS1 and *GILZ* GBS1 on the motif score (Fig. 2-figure supplement 1), we used sTRAP (23) (matrix IDs M00955, M00192 and M00205) using chordate conserved elements as background model and Benjamini-Hochberg correction for multiple testing.

Supplemental Tables

Table S1: Primer sequences for site directed mutagenesis (SDM)

Name	Usage	Sequence
SDM *GILZ* GBS1	HDR, luciferase construct	fwd: ACTCAGGACCAAAGGAGAAGCTTGGGTTCCACCACATATACAACAG rev: CTGTTGTATATGTGGTGGAACCCAAGCTTCTCCTTTGGTCCTGAGT
SDM *GILZ* GBS2 HDR	HDR	fwd: TGGGAGACAATAATGATCTCAGGATCCTCTGTTAACTTTAAGACACAACCTCT rev: AGAGGTTGTGTCTTAAAGTTAACAGAGGATCCTGAGATCATTATTGTCTCCCA
SDM *GILZ* GBS2	luciferase construct	fwd: TGGGAGACAATAATGATCTCAGGATTCTCTGTTAACTTTAAGACACAACCTCT rev: AGAGGTTGTGTCTTAAAGTTAACAGAGAATCCTGAGATCATTATTGTCTCCCA
SDM *GILZ* GBS3&4	HDR, luciferase construct	fwd: TCTGAAACCGGTGTTTGCACTAAGACATGTGTAGTAGACAGATCTAAGTGCAGGCT rev: AGCCTGCACTTAGATCTGTCTACTACACATGTCTTAGTGCAAACACCGGTTTCAGA

Table S2: gRNA sequences

Name	Usage	Sequence
gRNA *GILZ* GBS1	GBS1 DEL, HDR GBS1-4, gRNA1 for GBS1-4, dCas9-VP64 SAM	GCAGGACCAAAGGAGAACATT
gRNA *GILZ* GBS4	gRNA2 for CUT GBS1-4 & GBS2-4	GCCAGGTGGTATGGGAAGGGA
gRNA *GILZ* GBS2	gRNA2 for CUT GBS2-4	GATGGTGTGGTGCTCTATTGG GTAAACCTGCTGCACTAGCCC
gRNA *DUSP1* GBS1	*DUSP1* GBS1 DEL	GCCGAGTCAGGAACATTCTG
gRNA *GILZ* GBS A	gRNA1 for CUT GBS A-G	GTTACTGTACAGCTGGTCCC
gRNA *GILZ* GBS G	gRNA1 for CUT GBS A-G	GTTCTCATTCAGCTGCAAGTC
gRNA *GILZ* TSS1	dCas9-VP64 SAM	GAGGGAGCAAGGGCGCGCCC
gRNA CTCF1 at TSS1	gRNA1 for CTCF DEL	GCCTTCGGCTCGGCCCCCTTC
gRNA CTCF3 at TSS1	gRNA2 for CTCF DEL	GGCCGCGTGTGGCACCTCAG

Table S3: Location of target GBS

Name	Location (GRCh37/hg19)
GILZ GBS1	ChrX:106,961,572-106,961,594
GILZ GBS2	ChrX:106,962,033-106,962,053

GILZ GBS3	ChrX:106,962,067-106,962,087
GILZ GBS4	ChrX:106,962,104-106,962,124
DUSP1 GBS1	Chr5:172,199,533-172,199,553

Table S4: Primer sequences for genotyping of CRISPR/Cas9-edited cells

Name	**Sequence**
GILZ GBS1	fwd: GGAAAACACCTGCCCTGTGA rev: CGGGAGGAAATCAAGGCCTT
GILZ GBS1-4	fwd: GGAAAACACCTGCCCTGTGA rev: GTCTGAGTCTGGGCTGAACC
GILZ GBS A-G	fwd: AGTGACACTGTTGGCCTTCC rev: AGCCATTCCTCACCCTGAC
GILZ CTCF TSS1	fwd: TAGCCTGAGTCAGACCTCCC rev: CTAGCTCACTCGCTCTCAGC
DUSP1 GBS1	fwd: CAACCCTCGCTCCCTGTC rev: CCTCTTTGCTGTCCTCGACC

Table S5: qPCR primer sequences for the quantification of gene expression

Name	**Sequence**
GILZ standard	fwd: CCATGGACATCTTCAACAGC rev: TTGGCTCAATCTCTCCCATC
GILZ transcript variant 1	fwd: TACAGTGAGCAACTTTCGGC rev: GTTGATCAGGTAGCAGGGGT
GILZ transcript variant 2	fwd: TGGAGTTTGTGACATACGAGG rev: AGAACGAACCCAAAGCCAAG
GILZ transcript variant 3	fwd: AATTCCTAGCTAGCTTCAGAGC rev: GGCCTGTTCGATCTTGTTGT
DUSP1	fwd: CTGCCTTGATCAACGTCTCA rev: GTCTGCCTTGTGGTTGTCCT
FKBP5	fwd: TGAAGGGTTAGCGGAGCAC rev: CTTGGCACCTTCATCAGTAGTC
RPL19	fwd: ATGTATCACAGCCTGTACCTG rev: TTCTTGGTCTCTTCCTCCTTG
PRPS1	fwd: CGCTTAGTGGAGTGCTTAGG rev: TCACCAATTTCCACACAGGTC
MID2	fwd: AAGTTATTCTCTGGGCCTGC rev: ACAGTTTTGGAGAGCCAAGG

Table S6: qPCR primer sequences for quantification of ChIP experiments

Name	**Sequence**
GILZ distal peak I	fwd: CTTGCTCTGACAGGGAACAA rev: AGATCCCAGAAGAATTGGCAG
GILZ distal peak H	fwd: GATGGAGATAGGAAAAGGGGAG rev: GGAGTACTGCCAAGTGCTTTAT
GILZ distal peak G	fwd: ACTGCCTCTTTTTCTAAGGGC rev: TCTCTCATCTCATCCTCATGGA
GILZ distal peak F	fwd: AACTCAGCAGCTTTTCTTCGT rev: AACCAAGGAATTGGGTCACAT
GILZ distal peak A	fwd: TCAACGTCCAGACATAGCAAG rev: ATAGCTGGGAAATGGTAGCAG
DUSP1 distal peak C	fwd: ATCTTTACAAACAGATCTCCATGC rev: TCACACAATGCTGACTACGG
DUSP1 distal peak B	fwd: AAACCGGATCACACACTGAG rev: TAACTTCACCCGAGTTCCTCT
DUSP1 distal peak A	fwd: TGTCGCTGGTACATTTCCAC rev: CAGCTGGGTTTCCGATTACA
neg. ctrl GR/CTCF-binding	fwd: AATGGCAGCCCCTAGTCATTC rev: AACTGGGAGTGATACTGGTTCC
GILZ GBS1	fwd: GTGAGGCCACCTGGTGG rev: TATATGTGGTGGAACCCAATG
GILZ TSS1	fwd: GAGTGAGCTAGTTAGCGGTC rev: CCGTCCCCTCTAGGGTAATTT
GILZ TSS1 upstream	fwd: GGCCTTTGAAATAGAGCAGC rev: TACTCTTGGAACCAACGCAC
pos. ctrl CTCF-binding	fwd: GTGATCGGTCCAGTGCATAG rev: CTGGCATGTCATGGTAGAGC

Table S7: 4C primer sequences for inverse-PCR

Name	**Sequence**
viewpoint *GILZ* GBS1 adjacent	fwd: CTACACGACGCTCTTCCGATCTAATGTTCAGGTGTGGGAGTAC rev: CAGACGTGTGCTCTTCCGATCTGCCCTCTGCCTCTTGTTAGG

Table S8: 4C quality control statistics: Only read pairs with proper 4C primers and both reads mapping with a MAPQ≥30 to a corresponding restriction site were considered. Both mates can contribute a valid contact, however, the second mate was only considered when it was separated by at least nine restriction fragments from the fragment of the first mate.

	A549 Rep1	**U2OS-GR18 Rep1**
Reads pairs	13,535,799	13,529,198
Read pairs with valid 4C primers	13,105,165	13,069,360
Valid contacts	10,095,475	11,210,416
Cis contacts	9,039,327	10,438,151
Trans contacts	1,056148	772,265

Supplemental References

1. Rao, S.S., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S. *et al.* (2014) A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell*, **159**, 1665-1680.
2. Rogatsky, I., Trowbridge, J.M. and Garabedian, M.J. (1997) Glucocorticoid receptor-mediated cell cycle arrest is achieved through distinct cell-specific transcriptional regulatory mechanisms. *Molecular and cellular biology*, **17**, 3181-3193.
3. Wang, J.C., Derynck, M.K., Nonaka, D.F., Khodabakhsh, D.B., Haqq, C. and Yamamoto, K.R. (2004) Chromatin immunoprecipitation (ChIP) scanning identifies primary glucocorticoid receptor target genes. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 15603-15608.
4. Konermann, S., Brigham, M.D., Trevino, A.E., Joung, J., Abudayyeh, O.O., Barcena, C., Hsu, P.D., Habib, N., Gootenberg, J.S., Nishimasu, H. *et al.* (2015) Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature*, **517**, 583-588.
5. Hsu, P.D., Scott, D.A., Weinstein, J.A., Ran, F.A., Konermann, S., Agarwala, V., Li, Y., Fine, E.J., Wu, X., Shalem, O. *et al.* (2013) DNA targeting specificity of RNA-guided Cas9 nucleases. *Nature biotechnology*, **31**, 827-832.
6. Meijsing, S.H., Pufall, M.A., So, A.Y., Bates, D.L., Chen, L. and Yamamoto, K.R. (2009) DNA binding site sequence directs glucocorticoid receptor structure and activity. *Science*, **324**, 407-410.
7. Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E. and Church, G.M. (2013) RNA-guided human genome engineering via Cas9. *Science*, **339**, 823-826.
8. van de Werken, H.J., de Vree, P.J., Splinter, E., Holwerda, S.J., Klous, P., de Wit, E. and de Laat, W. (2012) 4C technology: protocols and data analysis. *Methods in enzymology*, **513**, 89-112.
9. Consortium, E.P. (2012) An integrated encyclopedia of DNA elements in the human genome. *Nature*, **489**, 57-74.
10. Love, M.I., Huber, W. and Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology*, **15**, 550.
11. Durinck, S., Spellman, P.T., Birney, E. and Huber, W. (2009) Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nature protocols*, **4**, 1184-1191.

12. Durinck, S., Moreau, Y., Kasprzyk, A., Davis, S., De Moor, B., Brazma, A. and Huber, W. (2005) BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. *Bioinformatics (Oxford, England)*, **21**, 3439-3440.
13. Langmead, B. and Salzberg, S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nature methods*, **9**, 357-359.
14. Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W. *et al.* (2008) Model-based analysis of ChIP-Seq (MACS). *Genome biology*, **9**, R137.
15. Starick, S.R., Ibn-Salem, J., Jurk, M., Hernandez, C., Love, M.I., Chung, H.R., Vingron, M., Thomas-Chollier, M. and Meijsing, S.H. (2015) ChIP-exo signal associated with DNA-binding motifs provides insight into the genomic binding of the glucocorticoid receptor and cooperating transcription factors. *Genome research*, **25**, 825-835.
16. Hinrichs, A.S., Karolchik, D., Baertsch, R., Barber, G.P., Bejerano, G., Clawson, H., Diekhans, M., Furey, T.S., Harte, R.A., Hsu, F. *et al.* (2006) The UCSC Genome Browser Database: update 2006. *Nucleic acids research*, **34**, D590-598.
17. Jin, F., Li, Y., Dixon, J.R., Selvaraj, S., Ye, Z., Lee, A.Y., Yen, C.A., Schmitt, A.D., Espinoza, C.A. and Ren, B. (2013) A high-resolution map of the three-dimensional chromatin interactome in human cells. *Nature*, **503**, 290-294.
18. Robinson, J.T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G. and Mesirov, J.P. (2011) Integrative genomics viewer. *Nature biotechnology*, **29**, 24-26.
19. Thorvaldsdottir, H., Robinson, J.T. and Mesirov, J.P. (2013) Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Briefings in bioinformatics*, **14**, 178-192.
20. Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M. and Haussler, D. (2002) The human genome browser at UCSC. *Genome research*, **12**, 996-1006.
21. Li, H. and Durbin, R. (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics (Oxford, England)*, **25**, 1754-1760.
22. Heng, L. (2013) Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *Genomics*, **1303**.
23. Manke, T., Heinig, M. and Vingron, M. (2010) Quantifying the effect of sequence variation on regulatory interactions. *Human mutation*, **31**, 477-483.