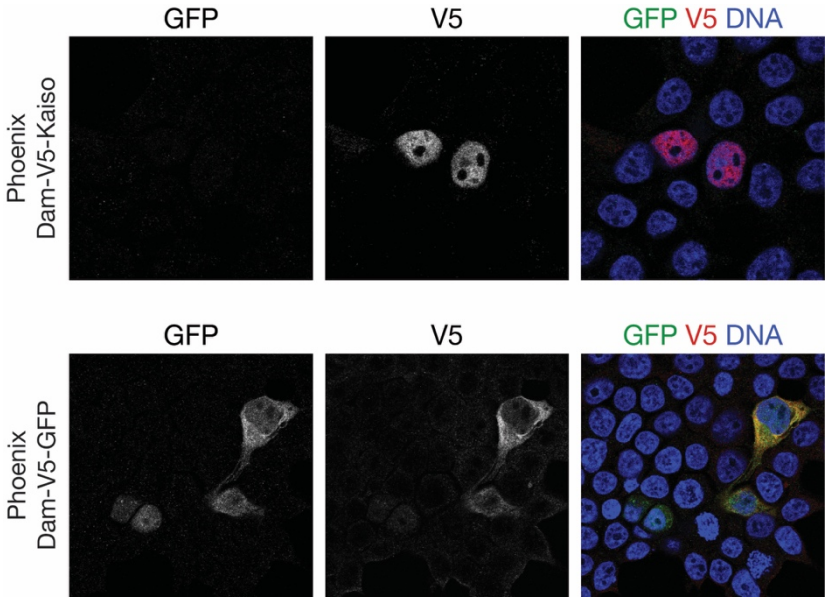
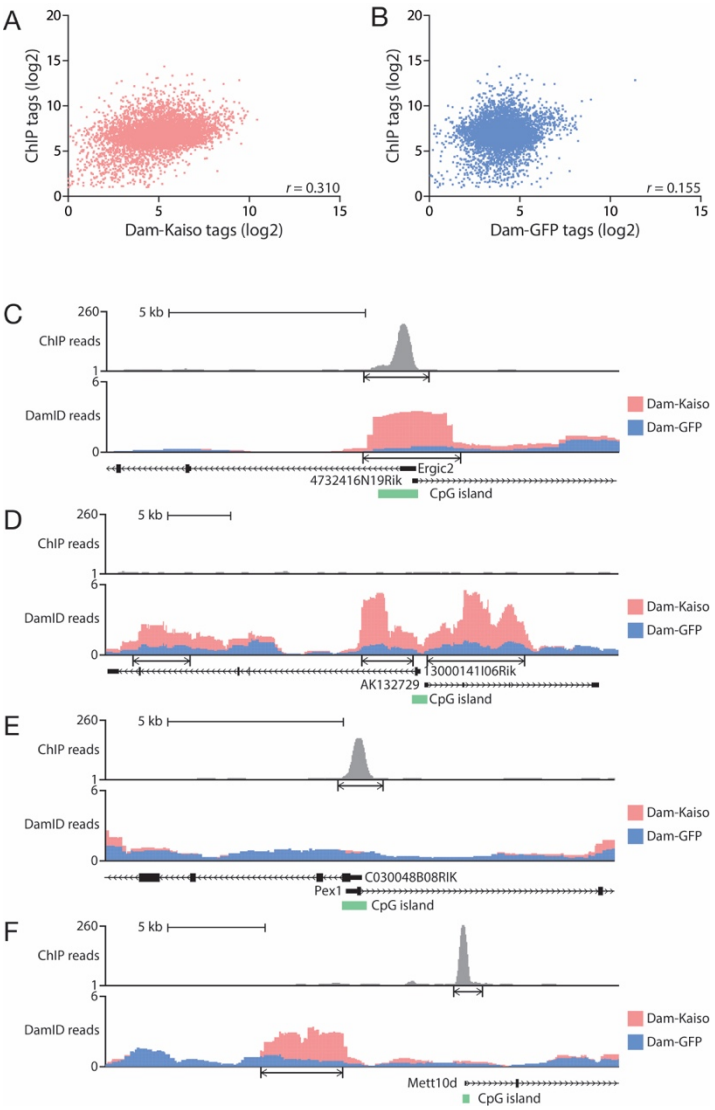
Supplemental Figures and Supplemental Tables



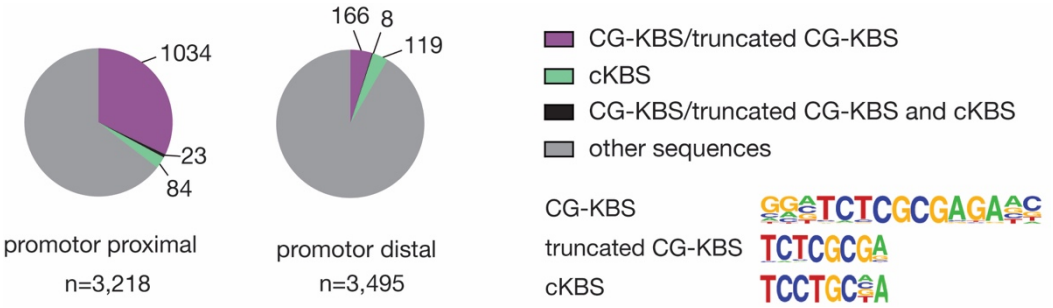
Supplemental Figure 1.

Immunofluorescent analysis of Amphotropic 293T (Phoenix) cells transfected with Dam-V5-Kaiso (Dam::Kaiso; top panels) and Dam-V5-GFP (Dam::GFP; bottom panels) for GFP (green) and V5 (red). Note the nuclear Dam::Kaiso expression *versus* diffuse expression of Dam::GFP.



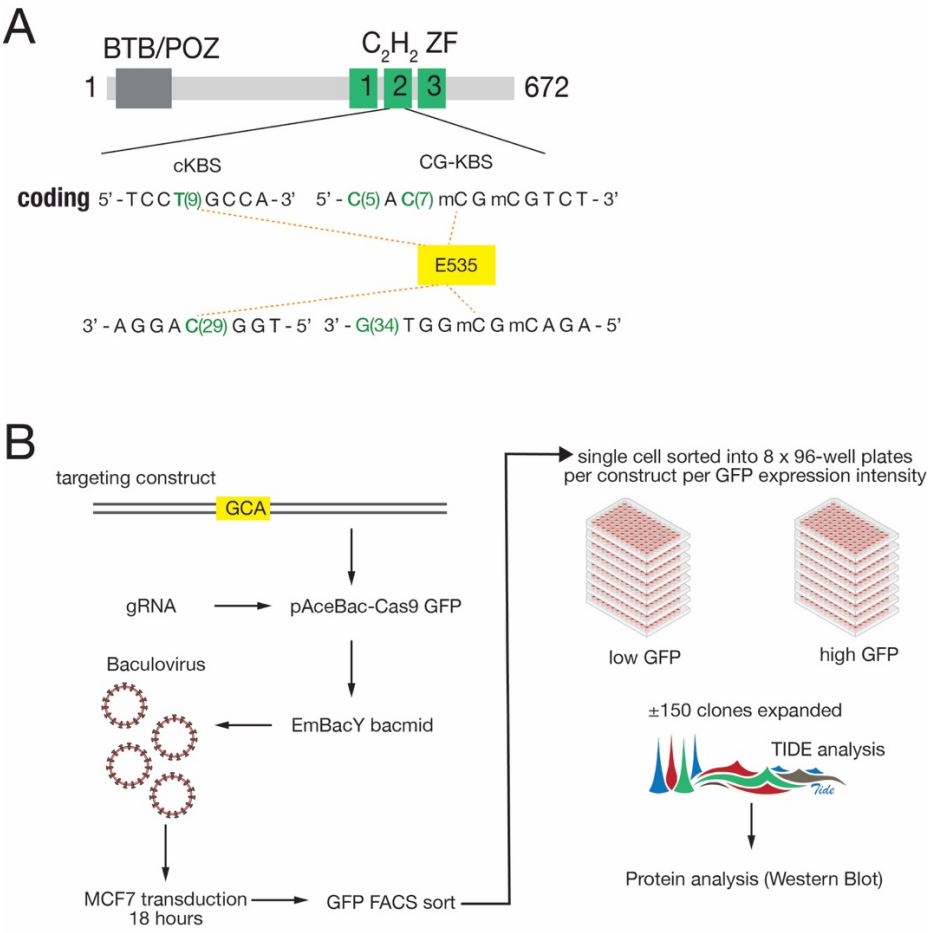
Supplemental Figure 2.

A and **B.** Overlap between Kaiso binding sites identified by ChIP and Dam-ID. Bivariate scatterplot showing the normalized tag count of Kaiso ChIP-seq within a 5 kb window surrounding ChIP-peak-centers versus the presence of Dam-Kaiso (A) and Dam-GFP tags (B). r = Pearson's correlation coefficient. **C-F.** Illustrative examples of genomic locations showing the identified Kaiso binding sites detected by both ChIP and DamID (C), exclusively by DamID (D), exclusively by ChIP (E) or by discontinuous ChIP and Dam-ID signals (F). ChIP reads are presented in black, Dam-GFP is depicted in blue and Dam-Kaiso is depicted in pink. Location of RefSeq genes and CpG islands according to the UCSC genome browser are depicted schematically at the bottom. Locations of the CG islands are indicated with green bars.



Supplemental Figure 3.

A. B. Distribution of CG-KBS, incomplete CG-KBS motifs and the cKBS in ChIP-peaks within promoters (promoter proximal) and outside promoters (promoter distal). Promoters were defined as localized 1 kb up- and downstream of the TSS.



Supplemental Figure 4.

Kaiso KBS loss of function approach. **A.** Shown are schematic representation of the Kaiso (ZBTB33) protein (grey bar) including the BTB/POZ domain (grey box) and three zinc fingers (green boxes). The minimal binding sequences for the canonical (cKBS) and CG-containing (CG-KBS) DNA motifs are depicted. Potential cytosine methylation sites (mC) are indicated. E535 represents the KBS interacting residue of Kaiso (yellow box) Cytosine are potentially methylated (based on[68]). Dashed lines indicate crucial hydrogen bonds between the depicted amino acid and DNA sequence. **B.** Baculoviral-based targeting approach. Shown are the targeting construct for mutating the E535 residue (yellow box), and the flow for the subsequent experiments to determine the genetic consequences and impact on protein expression.

35 **Supplemental Table 1. Characterised Kaiso MCF7 clones**

36 Protein expression was determined by western blotting using a Kaiso monoclonal antibody.

Clone 1#	Clone Alias	Tide Sequencing	Kaiso protein expression
1	1H23	*GCA* (E535) + indels	knock-out
2	3H14	*GCA* (E535) + indels	knock-out
3	1M19	indels	knock-out
4	1H26	indels	knock-out
5	2M314	indels	knock-out
6	14H	deletion	residual weak expression
7	12H	ND	knock-out
8	13H	ND	knock-out
9	27L	ND	knock-out

37

Supplemental Table 2. List of DNA Damage repair genes identified in the GO-term analysis. Genes with a bold name were confirmed to have Kaiso binding in the promoter region in the Encode ChIP-seq database for MCF7 cells [35].

Supplemental Table 3: DDR reverse transcriptase qPCR primers. All primer pairs designed for quantitative PCR to determine expression levels of the 11 genes written in bold in Suppl. Table 1.

Murine Gene Symbol	Human Gene Symbol	entrez gene #	Full name
Chaf1a	**CHAF1A**	27221	chromatin assembly factor 1, subunit A (p150)
Endov	**ENDOV**	338371	endonuclease V
Taok1	**TAOK1**	216965	TAO kinase 1
Epc2	**EPC2**	227867	enhancer of polycomb homolog 2
Stub1	**STUB1**	56424	STIP1 homology and U-Box containing protein 1
Usp47	**USP47**	74996	ubiquitin specific peptidase 47
Xrcc6	**XRCC6**	14375	X-ray repair complementing defective repair in Chinese hamster cells 6
Npm1	**NPM1**	18148	nucleophosmin 1
Pole	**POLE**	18973	polymerase (DNA directed), epsilon
Setx	**SETX**	269254	senataxin
Sprtn	**SPRTN**	244666	SprT-like N-terminal domain
Ino80	INO80	68142	INO80 complex subunit
Supt16	SUPT16	114741	SPT16, facilitates chromatin remodeling subunit
Exo5	EXO5	73172	exonuclease 5
Pola1	POLA1	18968	polymerase (DNA directed), alpha 1
Huwe1	HUWE1	59026	HECT, UBA and WWE domain containing 1
Smarca5	SMARCA5	93762	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 5
Smc5	SMC5	226026	structural maintenance of chromosomes 5
Xpc	XPC	22591	xeroderma pigmentosum, complementation group C
Yy1	YY1	22632	YY1 transcription factor
Mre11a	MRE11A	17535	MRE11A homolog A, double strand break repair nuclease
Fancc	FANCC	14088	Fanconi anemia, complementation group C
Smc1b	SMC1B	140557	structural maintenance of chromosomes 1B
Ercc6l2	ERCC6L2	76251	excision repair cross-complementing rodent repair deficiency, complementation group 6 like 2
Mms22l	MMS22L	212377	MMS22-like, DNA repair protein
Hmgb2	HMGB2	97165	high mobility group box 2

Supplemental Table 3: Reverse Transcriptase qPCR primers

Primer name	Sequence
CHAF1A_qPCR#1 Fw	5' - TTAGACCGAAACTTGTCAACGG- 3'
CHAF1A_qPCR#1 Rv	5' - GTCTGGCTGCTCATTCGAGT- 3'
CHAF1A_qPCR#2 Fw	5' - CAGCCAGACAGTCTTGTGGAC- 3'
CHAF1A_qPCR#2 Rv	5' - GTCGTTCTGAATGGCCTTCAA- 3'
ENDOV_qPCR#1 Fw	5' - CACTGTGGAAACGGGAGCAA- 3'
ENDOV_qPCR#1 Rv	5' - TTCACGAAGGACACGTCAACG- 3'
TAOK1_qPCR#1 Fw	5' - ATGCCATCAACTAACAGAGCAG- 3'
TAOK1_qPCR#1 Rv	5' - CGCACATCTCGTGCAAAATAC- 3'
TAOK1_qPCR#2 Fw	5' - TGCACGAGATGTGCGTACC- 3'
TAOK1_qPCR#2 Rv	5' - TGTGTTCACGTAAATAACAGCCT- 3'
TAOK1_qPCR#3 Fw	5' - ACATGGTGCTCTTCAGGGATT- 3'
TAOK1_qPCR#3 Rv	5' - GGCGTTCCCACAAAGGAATTG- 3'
EPC2_qPCR#1 Fw	5' - CAAGCCGCTGCCTATCTACC- 3'
EPC2_qPCR#1 Rv	5' - AAATGATGTTCCGATTCCTCCTC- 3'
EPC2_qPCR#2 Fw	5' - AGACAACGAGCAACCAGATTATG- 3'
EPC2_qPCR#2 Rv	5' - GCAGTTTTTACGTTTTCTCACCC- 3'
EPC2_qPCR#3 Fw	5' - TGACCCTTATGTTGCCTTTCG- 3'
EPC2_qPCR#3 Rv	5' - TCACCACCATAGTCTCCCAAAT- 3'
STUB1_qPCR#1 Fw	5' - AGCAGGGCAATCGTCTGTTC- 3'
STUB1_qPCR#1 Rv	5' - CAAGGCCCGGTTGGTGTAATA- 3'
USP47_qPCR#1 Fw	5' - CTCGACGCTAATTTTGAGCCA- 3'
USP47_qPCR#1 Rv	5' - CTCTTGGAAGCGGACCTATAAAC- 3'
USP47_qPCR#2 Fw	5' - AACCACAGATGTTACAAGGAGC- 3'
USP47_qPCR#2 Rv	5' - AAGATATGTGTCGATTCGCCAG- 3'
XRCC6_qPCR#1 Fw	5' - GTTGATGCCTCCAAGGCTATG- 3'
XRCC6_qPCR#1 Rv	5' - CCCCTTAAACTGGTCAAGCTCTA- 3'
XRCC6_qPCR#2 Fw	5' - TTGCTTCTGCCTAGCGATACC- 3'
XRCC6_qPCR#2 Rv	5' - AAACCTGGATCATCAAACCGTT- 3'
NPM1_qPCR#1 Fw	5' - GGAGGTGGTAGCAAGGTTCC- 3'
NPM1_qPCR#1 Rv	5' - TTCACTGGCGCTTTTTCTTCA- 3'
NPM1_qPCR#2 Fw	5' - GGTGGTAGCAAGGTTCCACAG- 3'
NPM1_qPCR#2 Rv	5' - TTCTTCACTGGCGCTTTTTCT- 3'
NPM1_qPCR#3 Fw	5' - GCACTTAGTAGCTGTGGAGGA- 3'
NPM1_qPCR#3 Rv	5' - TTCACTGGCGCTTTTTCTTCA- 3'
POLE_qPCR#1 Fw	5' - ATTCAAGATGACGGAAGCAGATT- 3'
POLE_qPCR#1 Rv	5' - AACCCTTTCTGGTCGCAATGT- 3'
POLE_qPCR#2 Fw	5' - TTGCGACCAGAAAGGGTTGT- 3'
POLE_qPCR#2 Rv	5' - TGATTTGGCAAGTCCAGATCCT- 3'
SETX_qPCR#1 Fw	5' - TGACATCACTGGGCAAGACTT- 3'
SETX_qPCR#1 Rv	5' - CCAACGCCGAACCATTTCATT- 3'
SETX_qPCR#2 Fw	5' - GCCAGCGTTACACTGTTTTATG- 3'
SETX_qPCR#2 Rv	5' - GTAGCTTGCGTTGTTGATAATGG- 3'
SPRTN_qPCR#1 Fw	5' - GGTGATGGGGAACCTTGAGAT- 3'
SPRTN_qPCR#1 Rv	5' - CTGTCACTTCTCGAATCCACTG- 3'
SPRTN_qPCR#2 Fw	5' - GACCCAGGTCTACGATGGGAA- 3'
SPRTN_qPCR#2 Rv	5' - GTGAGCTGAGTCAAGCGGAG- 3'
GAPDH_qPCR #1 FW	5' - CATTTCCTGGTATGACAACG- 3'
GAPDH_qPCR #1 Rv	5' - CTCTTCCTCTTGTGCTCTTG- 3'
ACTIN_qPCR #1 Fw	5' - CGTCACCAACTGGGACGACA- 3'
ACTIN_qPCR #1 Rv	5' - CTTCTCGCGGTTGGCCTTGG- 3'