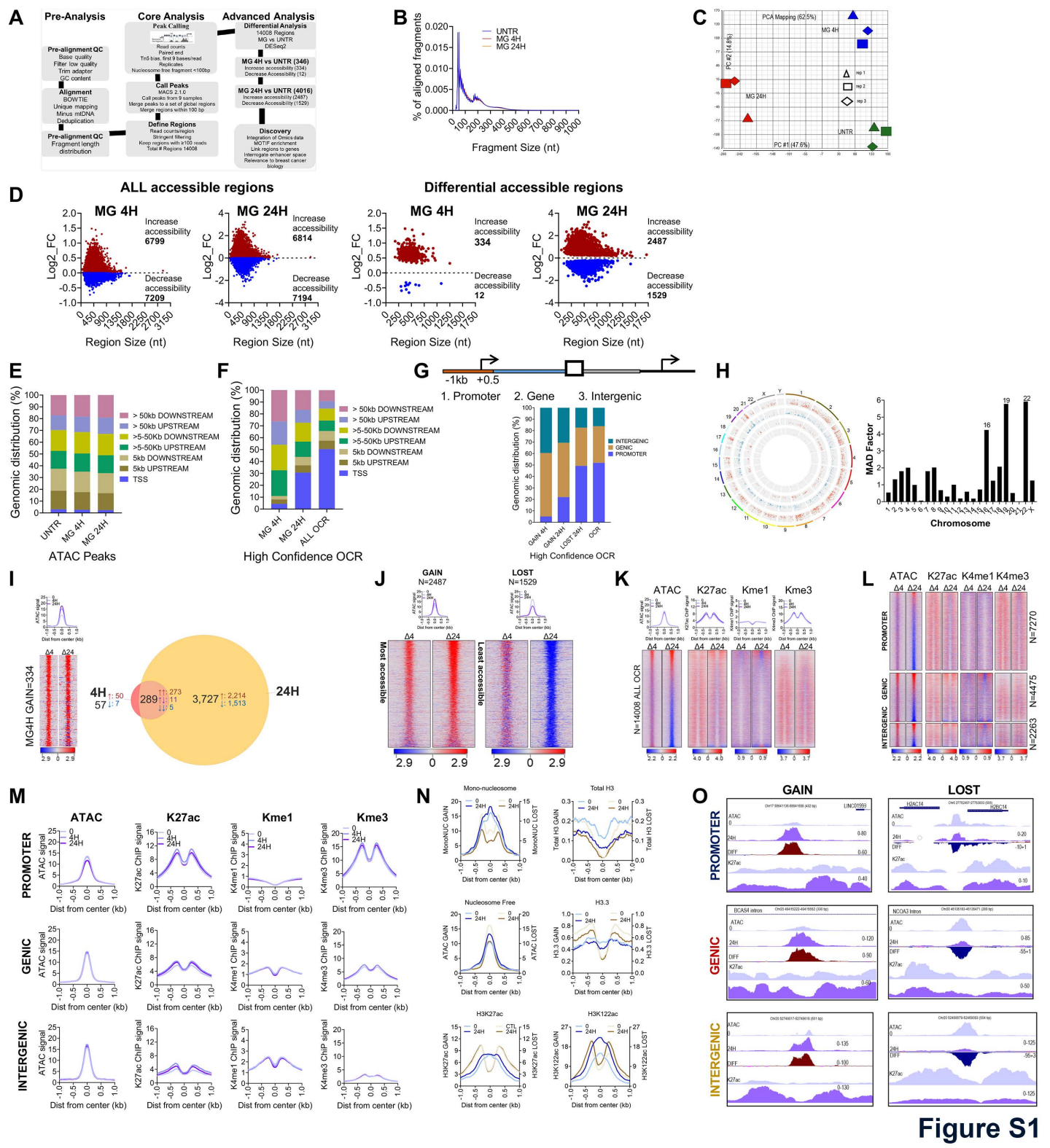
A
Pre-Analysis
Core Analysis
Advanced Analysis
Pre-alignment QC
Base quality
Filter low quality
Trim adapter
GC content
Alignment
BOWTIE
Unique mapping
Minus mtDNA
Deduplication
Pre-alignment QC
Fragment length distribution
Peak Calling
Read counts
Paired end
Tn5 bias, first 9 bases/read
Replicates
Nucleosome free fragment <100bp
Call Peaks
MACS 2.1.0
Call peaks from 9 samples
Merge peaks to a set of global regions
Merge regions within 100 bp
Define Regions
Read counts/region
Stringent filtering
Keep regions with ≥100 reads
Total # Regions 14008
Differential Analysis
14008 Regions
MG vs UNTR
DESeq2
MG 4H vs UNTR (346)
Increase accessibility (334)
Decrease Accessibility (12)
MG 24H vs UNTR (4016)
Increase accessibility (2487)
Decrease Accessibility (1529)
Discovery
Integration of Omics data
MOTIF enrichment
Link regions to genes
Interrogate enhancer space
Relevance to breast cancer biology
B
% of aligned fragments
UNTR
MG 4H
MG 24H
Fragment Size (nt)
C
PCA Mapping (62.5%)
PC #2 (14.8%)
PC #1 (47.6%)
MG 4H
MG 24H
UNTR
rep 1
rep 2
rep 3
D
ALL accessible regions
Differential accessible regions
MG 4H
MG 24H
Increase accessibility
6799
Decrease accessibility
7209
Increase accessibility
6814
Decrease accessibility
7194
Increase accessibility
334
Decrease accessibility
12
Increase accessibility
2487
Decrease accessibility
1529
Log2_FC
Region Size (nt)
E
Genomic distribution (%)
> 50kb DOWNSTREAM
> 50kb UPSTREAM
>5-50Kb DOWNSTREAM
>5-50Kb UPSTREAM
5kb DOWNSTREAM
5kb UPSTREAM
TSS
UNTR
MG 4H
MG 24H
ATAC Peaks
F
Genomic distribution (%)
MG 4H
MG 24H
ALL OCR
High Confidence OCR
G
-1kb +0.5
1. Promoter
2. Gene
3. Intergenic
Genomic distribution (%)
INTERGENIC
GENIC
PROMOTER
GAIN 4H
GAIN 24H
LOST 24H
OCR
High Confidence OCR
H
MAD Factor
Chromosome
I
MG4H GAIN=334
Δ4
Δ24
4H
↑ 50
57 ↓ 7
289
↑ 273
↑ 11
↓ 5
3,727
↑ 2,214
↓ 1,513
24H
J
GAIN
N=2487
LOST
N=1529
Most accessible
Least accessible
K
ATAC
K27ac
Kme1
Kme3
N=14008 ALL OCR
L
ATAC
K27ac
K4me1
K4me3
PROMOTER
GENIC
INTERGENIC
N=7270
N=4475
N=2263
M
ATAC
K27ac
Kme1
Kme3
PROMOTER
GENIC
INTERGENIC
0
4H
24H
ATAC signal
K27ac ChIP signal
K4me1 ChIP signal
K4me3 ChIP signal
Dist from center (kb)
N
Mono-nucleosome
Total H3
Nucleosome Free
H3.3
H3K27ac
H3K122ac
MonoNUC GAIN
MonoNUC LOST
Total H3 GAIN
Total H3 LOST
ATAC GAIN
ATAC LOST
H3.3 GAIN
H3.3 LOST
H3K27ac GAIN
H3K27ac LOST
H3K122ac GAIN
H3K122ac LOST
CTL
24H
Dist from center (kb)
O
GAIN
LOST
PROMOTER
GENIC
INTERGENIC
ATAC
24H
DIFF
K27ac
HZAC14
KDEC14
LINC01999
BCAS4 intron
NCOA3 Intron

Figure S1

Figure S1 . Systematic analysis of MG132 Differential Open Chromatin Regions (DOCRs) A) Flowchart describing the computational analysis pipeline used to define DOCRs from ATAC-seq. Major steps of analysis are highlighted. **B)** Graph showing insert size distribution of ATAC-seq fragments. **C)** Principal component analysis of ATAC-seq reads showing clear separation among treatment conditions, UNTR (Control, green), MG4H (blue), MG24H (red) are represented on two principal components with PC1 on x-axis and PC2 on y-axis. Three biological replicates per treatment condition are represented (rep1, rep2, rep3). PCA is based on ATAC signal, Tn5 transposase sensitive reads derived from 78, 380 high-confidence regions (FDR 0.05) with at least 10 reads in the 9 samples. **D)** Scatter plots showing the relationship between genomic size of accessible region (nt) and change in chromatin accessibility in MG132 treated samples (log2FC). ALL accessible regions (OCR) 4H (left) and 24H (right) and DOCRs that significantly increase or decrease chromatin accessibility at 4H (left) and 24H (right). **E)** Graph showing genomic distribution of accessible (ATAC) peaks in each treatment condition. **F)** Graph showing genomic distribution of MG132 high confidence open chromatin regions (OCRs) used in the analysis **G)** Graph showing distribution of high confidence OCRs by genomic category. TOP: Schematic defining genomic categories. **H)** Circos plot showing chromosome distribution of MG132 DOCRs. The outer ring represents reference chromosomes 1 through 22 in clockwise orientation. X axis denotes chromosome demarcation, y axis dots are DOCRs ranked by log2FC from 0 to max 2 change in accessibility. From outer circos ring Chromosome, 24H increase (blue), 24H decrease (red), 4H increase (blue), 4H decrease (red). **Right Graph:** Graph showing the Median Absolute Deviation (MAD Factor) indicating variability in DOCR chromosome distribution **I)** Chromatin accessibility of 334 GAIN DOCRs in cells treated for 4H. Heatmaps representing differential ATAC signal (ATAC) and metaplot (Top) representing ATAC signal coverage at genomic regions that change accessibility at 4H, the lines represent signal in control, 0, 4H and 24H treated cells: **Right Panel:** Venn diagram showing the overlap of genomic regions that increase or decrease chromatin accessibility in cells treated for 4 and 24H. Numbers represent DOCRs in each class, arrows represent direction of change in accessibility (red, GAIN), blue, LOST). **J)** Heatmaps and metaplots (top) representing ATAC signal (ATAC) at DOCRs that GAIN (2487, **Right Panel**) and LOST (1529, **Left Panel**). Heatmaps show differential accessibility (ATAC), metaplots show ATAC signal in each treatment sample. **K)** Heatmaps and metaplots (TOP) showing ATAC, H3K27ac, H3K4me1 and H3K4me3 signal at ALL OCR tested (14008) at 0, 4 and 24H. **L)** Heatmaps showing differential ATAC, H3K27ac, H3K4me1 and H3K4me3 signal at ALL OCRs split in genomic categories, PROMOTER (7270), GENE (4475) and INTERGENIC (2263) regions. **M)** Metaplots showing chromatin state for all OCR PROMOTER (7270), GENE (4475) and INTERGENIC (2263). **N)** Metaplots showing distinct chromatin states of GAIN and LOST promoter DOCR. Unique chromatin architecture reflected by the unimodal and bimodal profiles of mono nucleosome signal, ATAC, Total H3, H3.3, H3K27ac, H3K122ac. **O)** Browser tracks showing examples of GAIN and LOST DOCRs. Tracks show read coverage of chromatin accessibility (ATAC), differential read coverage of accessibility (DIFF, GAIN=red, LOST=blue) and H3K27ac Track colors: Control (0), MG24H (24H).