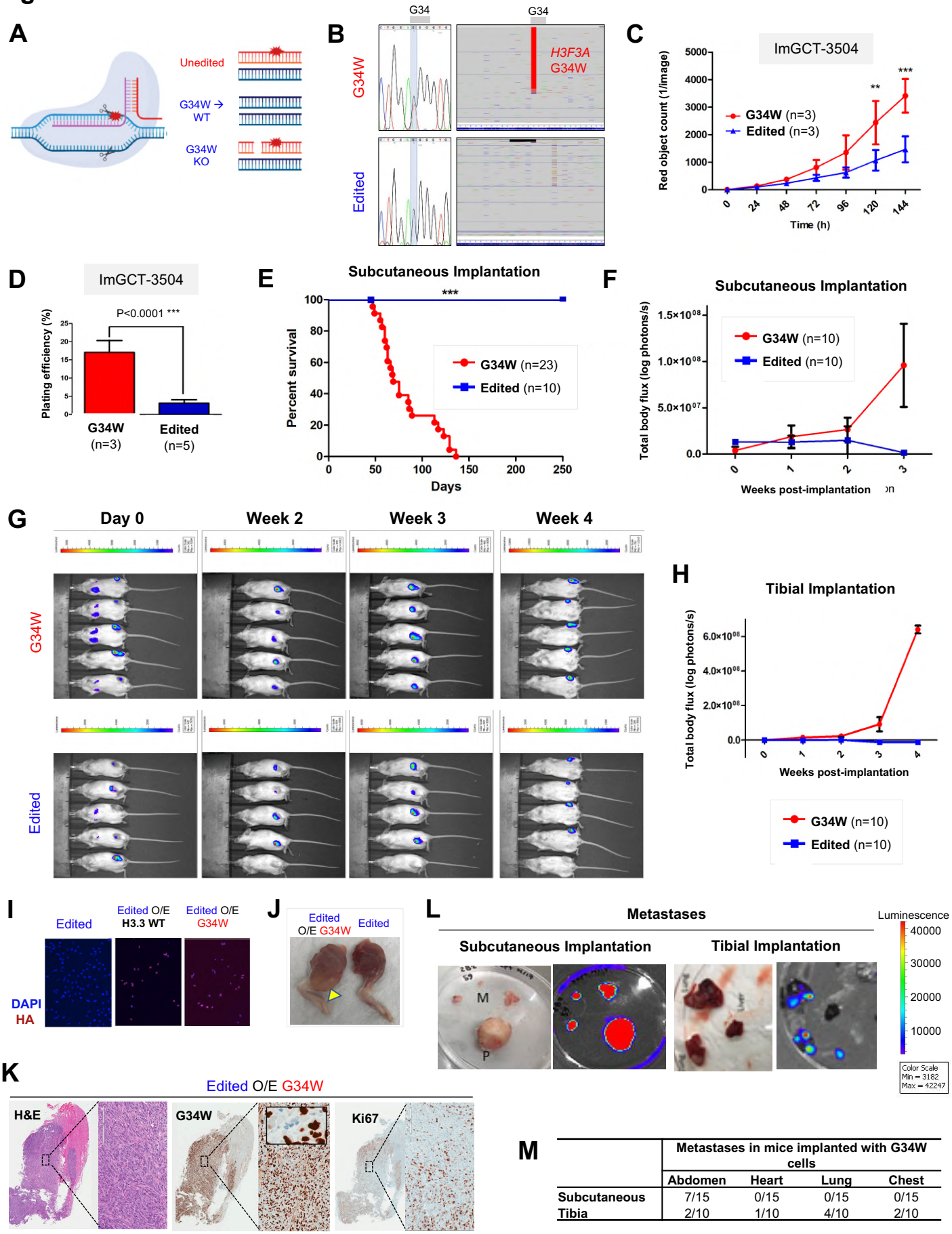
Figure S1



M

	Metastases in mice implanted with G34W cells			
	Abdomen	Heart	Lung	Chest
Subcutaneous	7/15	0/15	0/15	0/15
Tibia	2/10	1/10	4/10	2/10

Figure S1. Related to Figure 1.

(A) Schematic illustrating CRISPR/Cas9 strategy of targeting the *H3F3A* G34W mutation in GCT lines to derive edited clones.

(B) Sanger and MiSeq sequencing of a representative Im-GCT-4072 G34W (top) and Repair to WT (bottom) clone.

(C) G34W lines (Parent: *n*=1; Clone: *n*=1) of Im-GCT-3504 proliferate faster than edited lines (Repair to WT: *n*=3), as measured using the IncuCyte live-cell analysis system for 5 consecutive days. Five technical replicates were measured per line. Data are presented as mean red object count ± SD from biological replicates (independent CRISPR clones). Statistical significance assessed using Student's *t*-test.

(D) G34W lines (Parent: *n*=1; Clone: *n*=1) of Im-GCT-3504 exhibit increased colony formation relative to edited lines (Repair to WT: *n*=5), as measured by manual counting of colonies stained with crystal violet after 3 weeks. Two technical replicates were counted per line, and data are presented as an average of biological replicates (independent CRISPR clones). Statistical significance assessed using Student's *t*-test.

(E) Kaplan-Meier survival curve for subcutaneous implantation of Im-GCT-4072 G34W (Parent: *n*=1; Clone: *n*=2, total 23 mice), and edited lines (Repair to WT: *n*=2, total 10 mice), illustrates the dependence of tumour formation on the presence of G34W mutation.

(F) Quantitative analysis of *in vivo* Xenogen IVIS 200 bioluminescence signal intensity over time, showing increasing signal in mice with subcutaneous implantation of Im-GCT-4072 G34W (Parent: *n*=1; Clone: *n*=1) relative to edited (Repair to WT: *n*=2) lines.

(G) Representative bioluminescence Xenogen IVIS 200 imaging of mice implanted in the tibia with Im-GCT-4072 G34W and edited luciferase-tagged lines, over the time course indicated.

(H) Quantitative analysis of *in vivo* Xenogen IVIS 200 bioluminescence signal intensity over time, showing increasing signal in mice with tibial implantation of Im-GCT-4072 G34W (Parent: *n*=1; Clone: *n*=1) relative to edited (Repair to WT: *n*=2) lines.

(I) Immunofluorescence confirmation of overexpressed HA-tagged H3.3WT and H3.3G34W constructs in Im-GCT-4072 edited clones.

(J) Representative images of mice tibias implanted with edited cells O/E G34W and edited cells after skin removal at the time of sacrifice.

(K) Representative H&E and G34W and Ki67 IHC of decalcified legs derived from tibial implantation of Im-GCT-4072 edited cells O/E G34W. Inset features G34W-negative osteoclasts.

(L) Representative bioluminescence Xenogen IVIS 200 imaging of organs with metastases from subcutaneous and orthotopic tibial implantation of Im-GCT-4072 G34W lines.

(M) Table listing the number of mice implanted with Im-GCT-4072 G34W cells with metastases to listed organs.

Figure S2

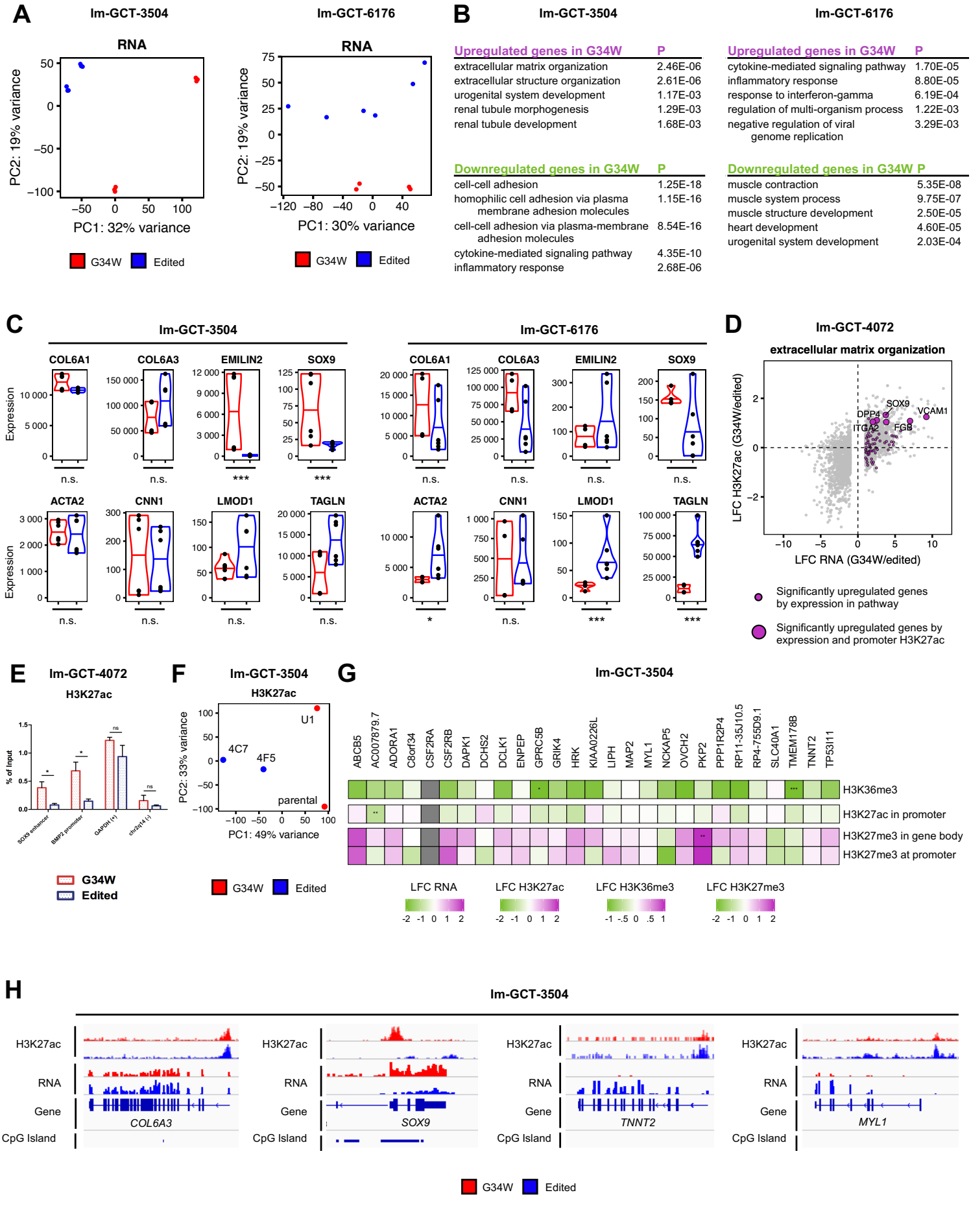

A
lm-GCT-3504
lm-GCT-6176
RNA
PC2: 19% variance
PC1: 32% variance
PC1: 30% variance
G34W
Edited
B
lm-GCT-3504
lm-GCT-6176
Upregulated genes in G34W P
extracellular matrix organization 2.46E-06
extracellular structure organization 2.61E-06
urogenital system development 1.17E-03
renal tubule morphogenesis 1.29E-03
renal tubule development 1.68E-03
Downregulated genes in G34W P
cell-cell adhesion 1.25E-18
homophilic cell adhesion via plasma membrane adhesion molecules 1.15E-16
cell-cell adhesion via plasma-membrane adhesion molecules 8.54E-16
cytokine-mediated signaling pathway 4.35E-10
inflammatory response 2.68E-06
Upregulated genes in G34W P
cytokine-mediated signaling pathway 1.70E-05
inflammatory response 8.80E-05
response to interferon-gamma 6.19E-04
regulation of multi-organism process 1.22E-03
negative regulation of viral genome replication 3.29E-03
Downregulated genes in G34W P
muscle contraction 5.35E-08
muscle system process 9.75E-07
muscle structure development 2.50E-05
heart development 4.60E-05
urogenital system development 2.03E-04
C
lm-GCT-3504
lm-GCT-6176
COL6A1
COL6A3
EMILIN2
SOX9
ACTA2
CNN1
LMOD1
TAGLN
Expression
n.s.
D
lm-GCT-4072
extracellular matrix organization
LFC H3K27ac (G34W/edited)
LFC RNA (G34W/edited)
SOX9
VCAM1
DPP4
ITGA2
FGB
Significantly upregulated genes by expression in pathway
Significantly upregulated genes by expression and promoter H3K27ac
E
lm-GCT-4072
H3K27ac
% of Input
SOX9 enhancer
BMP2 promoter
GAPDH (+)
chr2q14 (-)
G34W
Edited
F
lm-GCT-3504
H3K27ac
PC2: 33% variance
PC1: 49% variance
U1
4C7
4F5
parental
G34W
Edited
G
lm-GCT-3504
ABCB5
AC007879.7
ADORA1
C8orf34
CSF2RA
CSF2RB
DAPK1
DCHS2
DCLK1
ENPEP
GPRC5B
GRIK4
HRK
KIAA0226L
LIPH
MAP2
MYL1
NCKAP5
OVCH2
PKP2
PPP1R2P4
RP11-35J10.5
RP4-755D9.1
SLC40A1
TMEM178B
TNNT2
TP53I11
H3K36me3
H3K27ac in promoter
H3K27me3 in gene body
H3K27me3 at promoter
LFC RNA
LFC H3K27ac
LFC H3K36me3
LFC H3K27me3
H
lm-GCT-3504
H3K27ac
RNA
Gene
CpG Island
COL6A3
SOX9
TNNT2
MYL1
G34W
Edited

Figure S2. Related to Figure 2.

(A) Left: PCA reveals distinct transcriptomic profiles between G34W (red; *n*=6) and edited lines (blue; *n*=6) from Im-GCT-3504. Read counts were counted over Ensembl genes, normalized using the median-of-ratios procedure and transformed using the variance-stabilizing transformation of DESeq2. Refer to **Table S1** for more details on clones.

Right: PCA reveals distinct transcriptomic profiles between G34W (red; *n*=4) and edited lines (blue; *n*=6) from Im-GCT-6176.

(B) Pathway enrichment analysis of statistically significantly up- (purple) and down-regulated (green) genes in G34W compared to edited lines from Im-GCT-3504 (left) and Im-GCT-6176 (right). Pathway enrichment analysis was performed using g:Profiler. Top 5 statistically significantly enriched terms (GO: BP, term size<1000, $P<0.05$) are shown. The complete table can be found in **Table S3**.

(C) Violin plots depicting expression levels of extracellular matrix genes *COL6A1, COL6A3, EMILIN2* and *SOX9* as well as muscle contraction genes *ACTA2, CNN1, LMOD1* and *TAGLN* in G34W (red) and edited (blue) lines from Im-GCT-3504 (left) and Im-GCT-4072 (right). *: $P<0.05$, **: $P<0.01$, ***: $P<0.001$, n.s.: non-significant. Gene expression levels reported in median-of-ratios normalized read counts. Significance was assessed using DESeq2. The complete table can be found in **Table S2**.

(D) Scatterplot illustrating changes at promoter H3K27ac for significantly upregulated genes in extracellular matrix organization pathway in G34W lines compared to edited lines from Im-GCT-4072. X-axis: LFC of expression. Y-axis: $\log_2$ fold-change (LFC) of promoter H3K27ac. Grey: significant differentially expressed genes; purple: transcriptionally upregulated genes in pathway; big circle: significant changes in promoter H3K27ac and gene expression. Replicates: H3K27ac (G34W *n*=3; edited *n*=4), RNA (G34W *n*=7; edited *n*=9). The complete table can be found in **Table S3**.

(E) H3K27ac ChIP-qPCR for *SOX9* enhancer and *BMP2* promoter loci comparing Im-GCT-4072 G34W (red; *n=2*) and edited (blue; *n=3*) lines. *GAPDH* promoter and chr2q14 are positive and negative control regions respectively.

(F) PCA reveals distinct H3K27ac profiles between G34W (red; *n*=3) and edited lines (blue; *n*=4) from Im-GCT-4072. Read counts were counted over 10kb genomic bins and normalized to RPKM. The effect of the genotype is captured in PC1 (49% of the variance). The labels correspond to clones.

(G) Heatmap illustrating epigenetic changes (H3K27ac, H3K36me3, H3K27me3) at genes that were consistently transcriptionally deregulated across the three isogenic cell models. Differential histone mark enrichment is reported in $\log_2$ fold-change (LFC) in G34W lines over edited lines from Im-GCT-3054. Significance was assessed using DESeq2. *: $P<0.05$, **: $P<0.01$, ***: $P<0.001$. The complete table can be found in **Table S2**.

(H) Genomic tracks highlighting changes in H3K27ac and gene expression at *COL6A3, SOX9, TNNT2* and *MYL1* loci between G34W (red) and edited (blue) lines from Im-GCT-3054. Signals overlaid by replicates, reported in RPKM, and group auto-scaled by genomic assay.

Figure S3

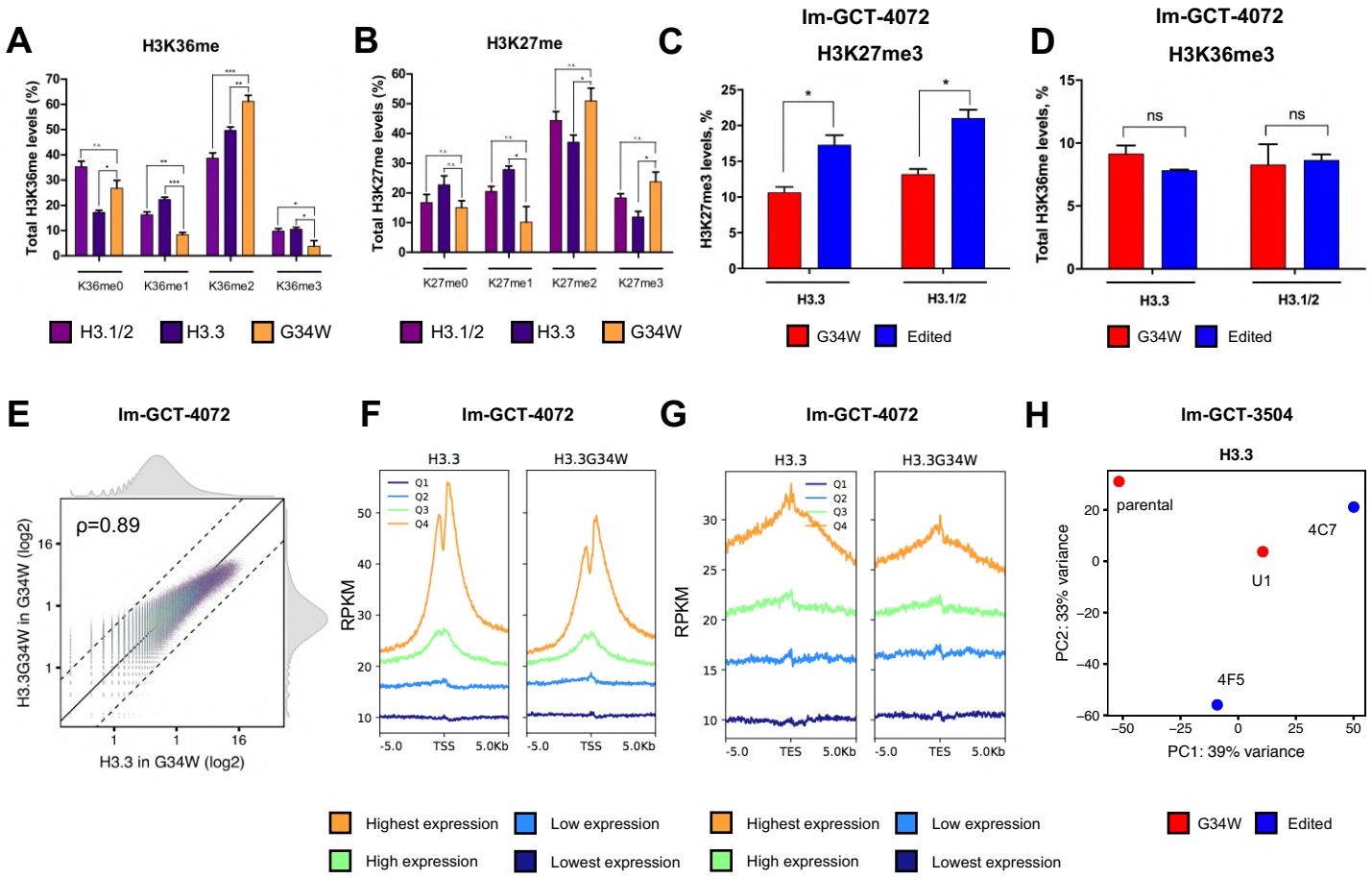

A
H3K36me
Total H3K36me levels (%)
K36me0
K36me1
K36me2
K36me3
H3.1/2
H3.3
G34W
B
H3K27me
Total H3K27me levels (%)
K27me0
K27me1
K27me2
K27me3
H3.1/2
H3.3
G34W
C
Im-GCT-4072
H3K27me3
H3K27me3 levels, %
H3.3
H3.1/2
G34W
Edited
D
Im-GCT-4072
H3K36me3
Total H3K36me levels, %
ns
ns
H3.3
H3.1/2
G34W
Edited
E
Im-GCT-4072
ρ=0.89
H3.3G34W in G34W (log2)
H3.3 in G34W (log2)
F
Im-GCT-4072
H3.3
H3.3G34W
Q1
Q2
Q3
Q4
RPKM
-5.0
TSS
5.0Kb
Highest expression
Low expression
High expression
Lowest expression
G
Im-GCT-4072
H3.3
H3.3G34W
RPKM
-5.0
TES
5.0Kb
Highest expression
Low expression
High expression
Lowest expression
H
Im-GCT-3504
H3.3
parental
4C7
U1
4F5
PC2: 33% variance
PC1: 39% variance
G34W
Edited

Figure S3. Related to Figure 3.

(A-B) Histone mass spectrometry reveals *in cis* changes in H3K36 **(A)** and H3K27 **(B)** methylation on G34W (yellow) compared to WT H3.1/2 (light purple) or WT H3.3 (purple) peptides in primary G34W-mutant GCT cell lines (n=4). *: P<0.05; **: P<0.01, ***: P<0.001, n.s.: non-significant. Significance was assessed using Student's t-test.

(C-D) Histone mass spectrometry quantification of global H3K27me3 **(C)** and H3K36me3 **(D)** in G34W lines (red; n=2) compared to edited lines (blue; n=3) from Im-GCT-4072. *: P<0.05; **: P<0.01, ***: P<0.001, n.s.: non-significant. Significance was assessed using Student's t-test.

(E) Scatterplot depicting changes in G34W (x-axis) and H3.3 (y-axis) ChIP abundance in the parental line Im-GCT-4072. Color: point density. Solid line: no change in abundance. Dotted lines: two-fold change in abundance. Reads counted over 10kb bins, averaged per condition, normalized to RPKM and reported in $\log_2$ scale.

(F-G) H3.3 (left) and H3.3G34W (right) abundances in the ±5kb window around the transcription start site (TSS) **(F)** and transcription end site (TES) **(G)** correlate with gene expression. Genes are stratified into four quartiles of expression in the parental line Im-GCT-4072. Histone abundances reported in RPKM.

(H) PCA illustrating genome-wide deposition patterns of H3.3 between G34W (red; n=2) and edited lines (blue; n=2) from Im-GCT-3504. PCA was performed on H3.3 abundance at consensus H3.3 peaks. The labels correspond to clones. Refer to **Table S1** for more details on clones.

Figure S4

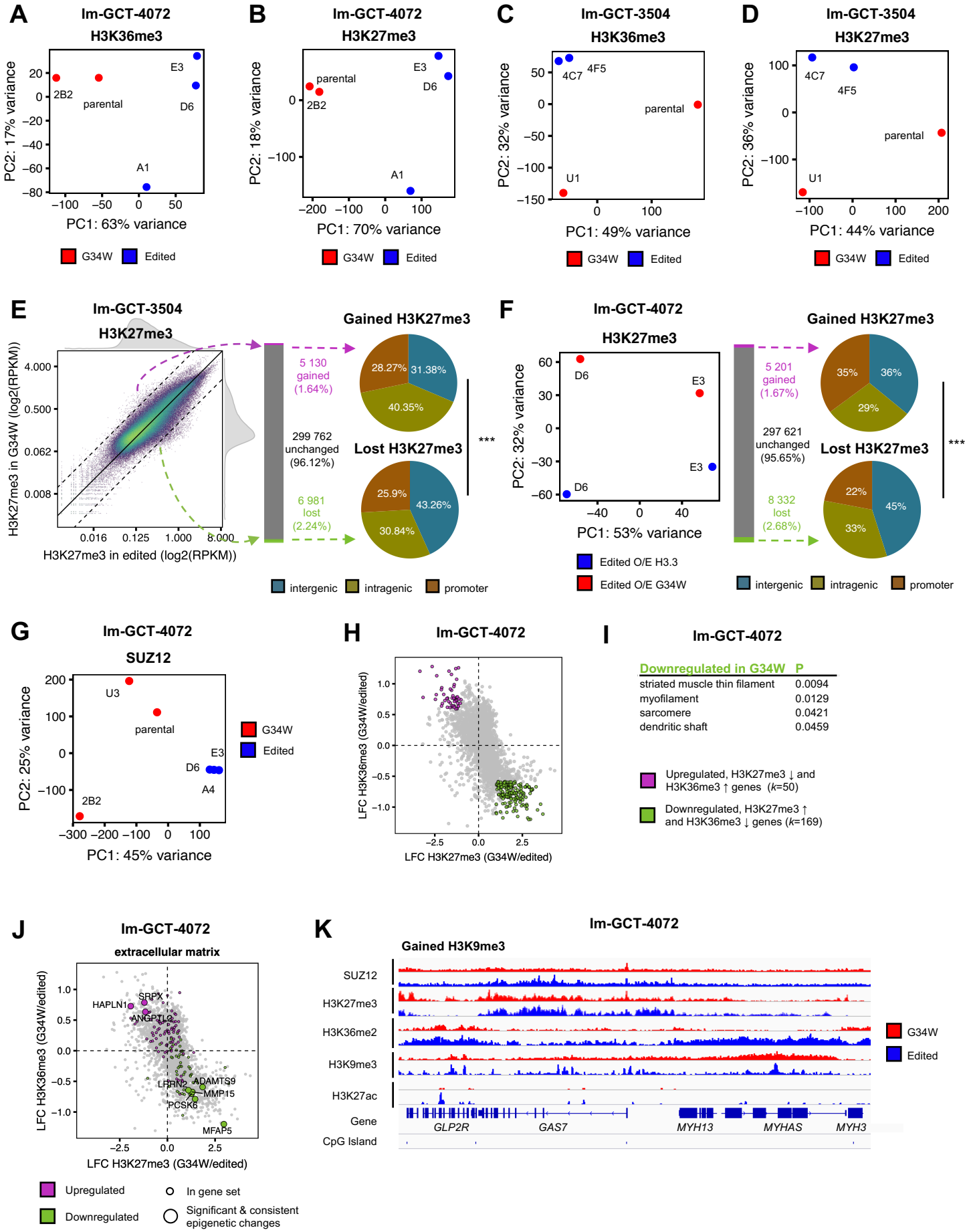

A
lm-GCT-4072
H3K36me3
PC2: 17% variance
PC1: 63% variance
2B2
parental
E3
D6
A1
G34W
Edited
B
lm-GCT-4072
H3K27me3
PC2: 18% variance
PC1: 70% variance
C
lm-GCT-3504
H3K36me3
PC2: 32% variance
PC1: 49% variance
4C7
4F5
U1
D
lm-GCT-3504
H3K27me3
PC2: 36% variance
PC1: 44% variance
E
lm-GCT-3504
H3K27me3
H3K27me3 in G34W (log2(RPKM))
H3K27me3 in edited (log2(RPKM))
5 130 gained (1.64%)
299 762 unchanged (96.12%)
6 981 lost (2.24%)
Gained H3K27me3
28.27%
31.38%
40.35%
Lost H3K27me3
25.9%
43.26%
30.84%
***
intergenic
intragenic
promoter
F
lm-GCT-4072
H3K27me3
PC2: 32% variance
PC1: 53% variance
Edited O/E H3.3
Edited O/E G34W
5 201 gained (1.67%)
297 621 unchanged (95.65%)
8 332 lost (2.68%)
Gained H3K27me3
35%
36%
29%
Lost H3K27me3
22%
45%
33%
***
intergenic
intragenic
promoter
G
lm-GCT-4072
SUZ12
PC2: 25% variance
PC1: 45% variance
U3
A4
G34W
Edited
H
lm-GCT-4072
LFC H3K36me3 (G34W/edited)
LFC H3K27me3 (G34W/edited)
I
lm-GCT-4072
Downregulated in G34W P
striated muscle thin filament 0.0094
myofilament 0.0129
sarcomere 0.0421
dendritic shaft 0.0459
Upregulated, H3K27me3 ↓ and H3K36me3 ↑ genes (k=50)
Downregulated, H3K27me3 ↑ and H3K36me3 ↓ genes (k=169)
J
lm-GCT-4072
extracellular matrix
HAPLN1
SRPX
ANGPTL6
LRRN2
ADAMTS9
MMP15
PCSK6
MFAP5
LFC H3K36me3 (G34W/edited)
LFC H3K27me3 (G34W/edited)
Upregulated
Downregulated
In gene set
Significant & consistent epigenetic changes
K
lm-GCT-4072
Gained H3K9me3
SUZ12
H3K27me3
H3K36me2
H3K9me3
H3K27ac
Gene
CpG Island
GLP2R
GAS7
MYH13
MYHAS
MYH3
G34W
Edited

Figure S4. Related to Figure 4.

(A-B) PCA reveals distinct H3K36me3 **(A)** and H3K27me3 **(B)** profiles between G34W (red; *n*=2) and edited lines (blue; *n*=3) from Im-GCT-4072. Read counts were counted over 10kb genomic bins and normalized to RPKM. The effect of the genotype is captured in PC1 (63-70% of the variance). The labels correspond to clones. Refer to **Table S1** for more details.

(C-D) PCA reveals distinct H3K36me3 **(C)** and H3K27me3 **(D)** profiles between G34W (red; *n*=2) and edited lines (blue; *n*=2) from Im-GCT-3504. The effect of the genotype is captured in PC2 (32-36% of the variance), while some variability between parental and clonal lines is also captured in the PCA.

(E) Left: Scatterplot depicting changes in H3K27me3 in G34W (y-axis; *n*=2) and edited (x-axis; *n*=3) lines from Im-GCT-3504. Color: point density. Solid line: no change in abundance. Dotted lines: two-fold change in abundance. Reads counted over 10kb bins, averaged per condition, normalized to RPKM and reported in $\log_2$ scale.

Middle: Bar plot quantifying the number of 10kb bins with gained (purple), unchanged (grey) and lost (green) H3K27me3 in G34W relative to edited lines. Bins with above-median average H3K27me3 in either condition and with an absolute $\log_2$ fold-change (LFC) of H3K27me3 exceeding 0.58 were called as gains and losses.

Right: H3K27me3 is redistributed from intergenic to genic regions in G34W lines. Pie charts illustrating proportion of 10kb bins gaining or losing H3K27me3 that overlap promoters, gene bodies and intergenic regions. ***: $P<0.001$. Significance was assessed using $\chi 2$ test.

(F) Left: PCA reveals changes in H3K27me3 in edited lines from Im-GCT-4072 upon re-expression of H3.3G34W (red; *n*=2) compared to H3.3WT (blue; *n*=2). Clone-specific effects are captured in PC1 (53% of variance), whereas G34W-dependent effects are captured by PC2 (32% of the variance).

Middle: Bar plot quantifying the number of 10kb bins with gained (purple), unchanged (grey) and lost (green) H3K27me3 in edited lines re-expressing H3.3G34W compared to H3.3WT. Bins with above-median average H3K27me3 in either condition and with an absolute $\log_2$ fold-change (LFC) of H3K27me3 exceeding 0.32 (1.25-fold change) were called as gains and losses.

Right: Pie charts illustrating proportion of 10kb bins gaining or losing H3K27me3 that overlap promoters, gene bodies and intergenic regions. ***: $P<0.001$.

(G) PCA reveals distinct SUZ12 profiles between G34W (red; *n*=3) and edited (blue; *n*=3) lines from Im-GCT-4072. The effect of the genotype is captured in PC1 (45% of the variance).

(H) Scatterplot illustrating genes with significant changes in expression, genic H3K36me3 and H3K27me3 in G34W lines compared to edited lines from Im-GCT-4072. X-axis: $\log_2$ fold-change (LFC) of H3K27me3. Y-axis: LFC of H3K36me3. Grey: significantly differentially expressed genes; purple: upregulated genes with significant loss of H3K27me3 and gain of H3K36me3; green: downregulated genes with significant gain of H3K27me3 and gain of H3K36me3. Replicates: H3K27me3 (G34W *n*=2; edited *n*=3), H3K36me3 (G34W *n*=2; edited *n*=3), RNA (G34W *n*=7; edited *n*=9). The complete table can be found in **Table S2**.

(I) Pathway enrichment analysis of up- (purple) and downregulated (green) genes with consistent changes in H3K36me3 and H3K27me3 between G34W and edited lines from Im-GCT-4072. Pathway enrichment analysis was performed using g:Profiler. All statistically significantly enriched terms ($P<0.05$) are shown. No pathways were enriched among the upregulated genes. The complete table can be found in **Table S3**.

(J) Scatterplot illustrating epigenetic changes at significantly deregulated genes in extracellular matrix pathway in G34W lines compared to edited lines from Im-GCT-4072, indicating comparatively fewer consistent epigenetic changes relative to actin filament-based process pathway genes. X-axis: $\log_2$ fold-change (LFC) of H3K27me3. Y-axis: LFC of H3K36me3. Grey: significant differentially expressed genes; purple/green: up- and downregulated genes in pathway; big circle: significant changes in genic H3K36me3 and H3K27me3. Replicates: H3K27me3 (G34W *n*=2; edited *n*=3), H3K36me3 (G34W *n*=2; edited *n*=3), RNA (G34W *n*=7; edited *n*=9).

(K) Representative track at the *MYH* cluster illustrating changes in SUZ12, H3K27me3, H3K36me2, H3K9me3, and H3K27ac between G34W (red) and edited lines (blue) from Im-GCT-4072. Signals overlaid by replicates, reported in RPKM, and group auto-scaled by chromatin mark.

Figure S5

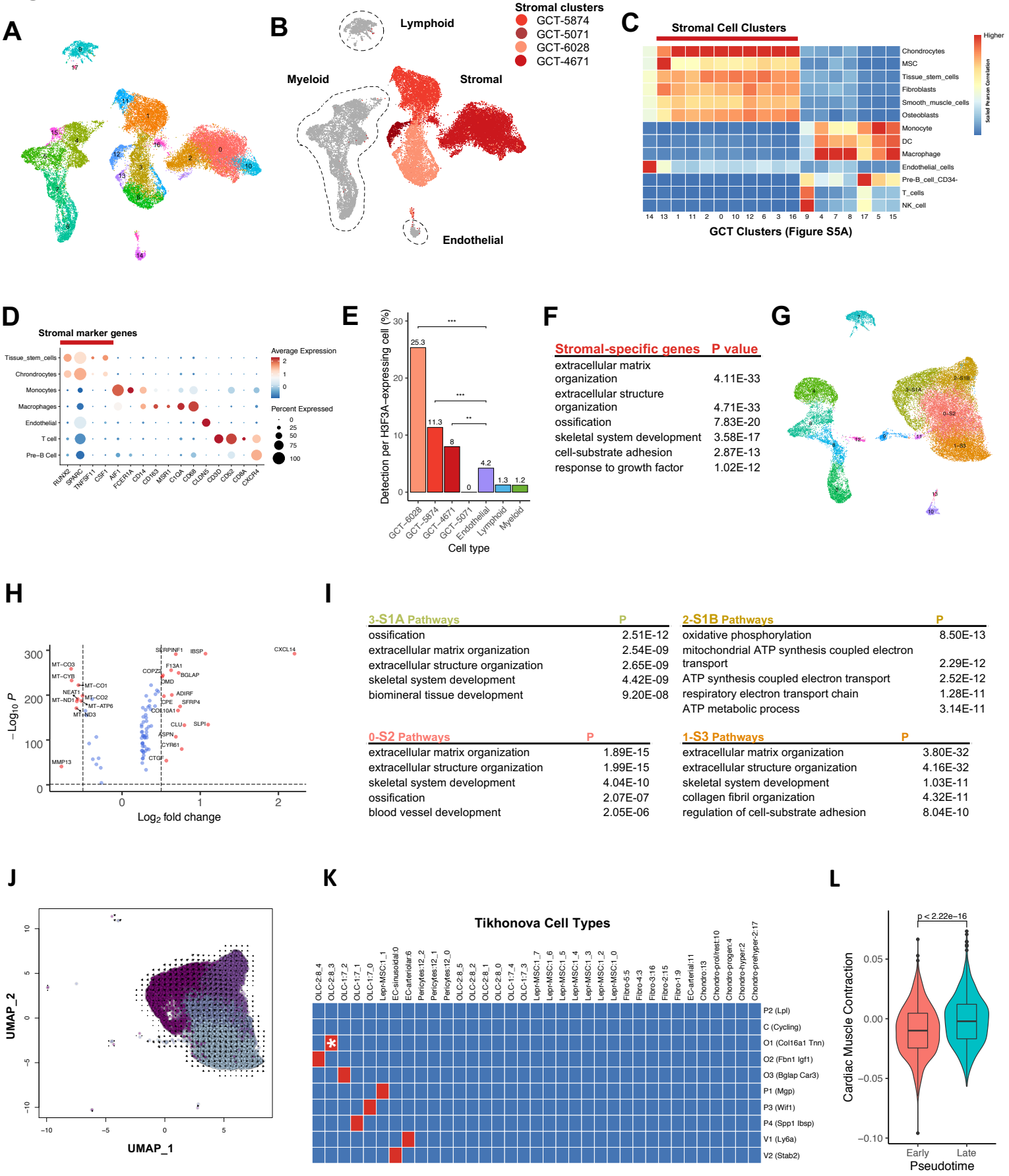

A
B
Lymphoid
Myeloid
Stromal
Endothelial
Stromal clusters
GCT-5874
GCT-5071
GCT-6028
GCT-4671
C
Stromal Cell Clusters
Chondrocytes
MSC
Tissue_stem_cells
Fibroblasts
Smooth_muscle_cells
Osteoblasts
Monocyte
DC
Macrophage
Endothelial_cells
Pre-B_cell_CD34-
T_cells
NK_cell
Higher
Scaled Pearson Correlation
14 13 1 11 2 0 10 12 6 3 16 9 4 7 8 17 5 15
GCT Clusters (Figure S5A)
D
Stromal marker genes
Tissue_stem_cells
Chrondrocytes
Monocytes
Macrophages
Endothelial
T cell
Pre–B Cell
RUNX2 SPARC TNFSF11 CSF1 AIF1 FCER1A CD14 CD163 MSR1 C1QA CD68 CLDN5 CD3D CD82 CD8A CXCR4
Average Expression
Percent Expressed
E
Detection per H3F3A–expressing cell (%)
25.3
11.3
8
0
4.2
1.3
1.2
GCT–6028 GCT–5874 GCT–4671 GCT–5071 Endothelial Lymphoid Myeloid
Cell type
F
Stromal-specific genes P value
extracellular matrix organization 4.11E-33
extracellular structure organization 4.71E-33
ossification 7.83E-20
skeletal system development 3.58E-17
cell-substrate adhesion 2.87E-13
response to growth factor 1.02E-12
G
H
−Log10 P
Log2 fold change
SERPINF1 IBSP CXCL14 F13A1 COPZ2 BGLAP OMD ADIRF CPE SFRP4 COL10A1 CLU SLPI ASPN CYR61 CTGF MMP13 MT−CO3 MT−CYB MT−CO1 NEAT1 MT−ND1 MT−CO2 MT−ATP6 MT−ND3
I
3-S1A Pathways P
ossification 2.51E-12
extracellular matrix organization 2.54E-09
extracellular structure organization 2.65E-09
skeletal system development 4.42E-09
biomineral tissue development 9.20E-08
2-S1B Pathways P
oxidative phosphorylation 8.50E-13
mitochondrial ATP synthesis coupled electron transport 2.29E-12
ATP synthesis coupled electron transport 2.52E-12
respiratory electron transport chain 1.28E-11
ATP metabolic process 3.14E-11
0-S2 Pathways P
extracellular matrix organization 1.89E-15
extracellular structure organization 1.99E-15
skeletal system development 4.04E-10
ossification 2.07E-07
blood vessel development 2.05E-06
1-S3 Pathways P
extracellular matrix organization 3.80E-32
extracellular structure organization 4.16E-32
skeletal system development 1.03E-11
collagen fibril organization 4.32E-11
regulation of cell-substrate adhesion 8.04E-10
J
UMAP_2
UMAP_1
K
Tikhonova Cell Types
OLC-2:8_4 OLC-2:8_3 OLC-1:7_2 OLC-1:7_1 OLC-1:7_0 Lepr-MSC:1_1 EC-sinusoidal:0 EC-arteriolar:6 Pericytes:12_2 Pericytes:12_1 Pericytes:12_0 OLC-2:8_5 OLC-2:8_2 OLC-2:8_1 OLC-2:8_0 OLC-1:7_4 OLC-1:7_3 Lepr-MSC:1_7 Lepr-MSC:1_6 Lepr-MSC:1_5 Lepr-MSC:1_4 Lepr-MSC:1_3 Lepr-MSC:1_2 Lepr-MSC:1_0 Fibro-5:5 Fibro-4:3 Fibro-3:16 Fibro-2:15 Fibro-1:9 EC-arterial:11 Chondro:13 Chondro-prol/rest:10 Chondro-progen:4 Chondro-hyper:2 Chondro-prehyper-2:17
P2 (Lpl)
C (Cycling)
O1 (Col16a1 Tnn)
O2 (Fbn1 Igf1)
O3 (Bglap Car3)
P1 (Mgp)
P3 (Wif1)
P4 (Spp1 Ibsp)
V1 (Ly6a)
V2 (Stab2)
Baryawno Cell Types
L
p < 2.22e−16
Cardiac Muscle Contraction
Early
Late
Pseudotime

Figure S5. Related to Figure 5.

(A) Unbiased clustering of single cells from GCT (*n*=4 tumours) on UMAP plot illustrating 17 distinct clusters.

(B) UMAP plot displaying cell cluster annotation of GCT (*n*=4 tumours) scRNA-seq data. Myeloid, lymphoid and endothelial cell clusters are colored in grey, whereas stromal clusters are shown in shades of red based on the tumour of origin.

(C) SingleR classification of 17 GCT clusters (from **Fig. S5A**) to Human Primary Cell Atlas (HPCA) cell types. Stromal clusters map variously to chondrocytes, osteoblasts, fibroblasts and smooth muscle cells.

(D) Dotplot illustrating average expression (color) and percent expression (dot size) of marker genes for various cell type clusters. Stromal cell specific marker genes are highlighted.

(E) Bar graph showing G34W mutation detection in cells expressing at least one count of *H3F3A*. ***: $P<0.0001$, **: $P<0.001$. Significance was assessed using $\chi 2$ test. Detected G34W is significantly higher in the putative stromal cells of 3 out of 4 GCT samples. G34W failed to be detected in any cells from one sample (GCT-5071) known to be G34W positive, likely due to the technical limitations of single-cell variant detection.

(F) Pathway enrichment analysis of stromal-specific genes performed using g:Profiler shows similar enriched pathways as **Fig. 2B, S2B**. Top 6 statistically significant enriched GO terms (GO:BP, term size<1000, $P<0.05$) are shown.

(G) Unbiased clustering of GCT cells on UMAP plot following Harmony data integration reveals 13 clusters, four of which represent integrated stromal cells: S1A, S1B, S2, S3.

(H) Volcano plot showing genes significantly upregulated (positive LFC) and downregulated (negative LFC) in S1A stromal cells compared to S1B. Horizontal dashed line indicates a p = 0.05 significance cut-off and vertical lines indicate a 0.5 LFC cut-off. Labelled red points are genes reaching both cut-off's, blue points indicate genes failing to reach either of the cut-offs.

(I) Pathway enrichment analysis for stromal clusters S1-S3 from Harmony integrated data using g:Profiler. Top 5 statistically enriched GO terms (GO: BP, term size<1000, $P<0.05$) are shown.

(J) Lineage inference analysis using Velocyto RNA velocity overlaid on UMAP embedding of stromal cells from **Fig. S5G** illustrates two trajectories; from S1A to S1B, and from S1A to S3.

(K) Matched reference comparison showing the probability of mutual assignment between the humanized Baryawno and Tikhonova references, illustrating that there are certain shared populations between the datasets. White asterisk indicates that OLC-2:8_3 and O1 (Col16a1 Tnn) (subsets that GCT stromal cells most closely resemble) are each other's exclusive match, indicating that they represent a similar cell type.

(L) Boxplot displaying higher average expression of genes from the GO:BP geneset "Cardiac muscle contraction" (GO:0060048) in cells in "Late" pseudotime compared to "Early" pseudotime. Early and Late pseudotime cells were defined as the first and final 1000 cells along the S1A to S3 lineage by ranked pseudotime. Significance was assessed using a linear model, correlating expression of the geneset with increasing pseudotime of all cells along the S1A to S3 trajectory.

Figure S6

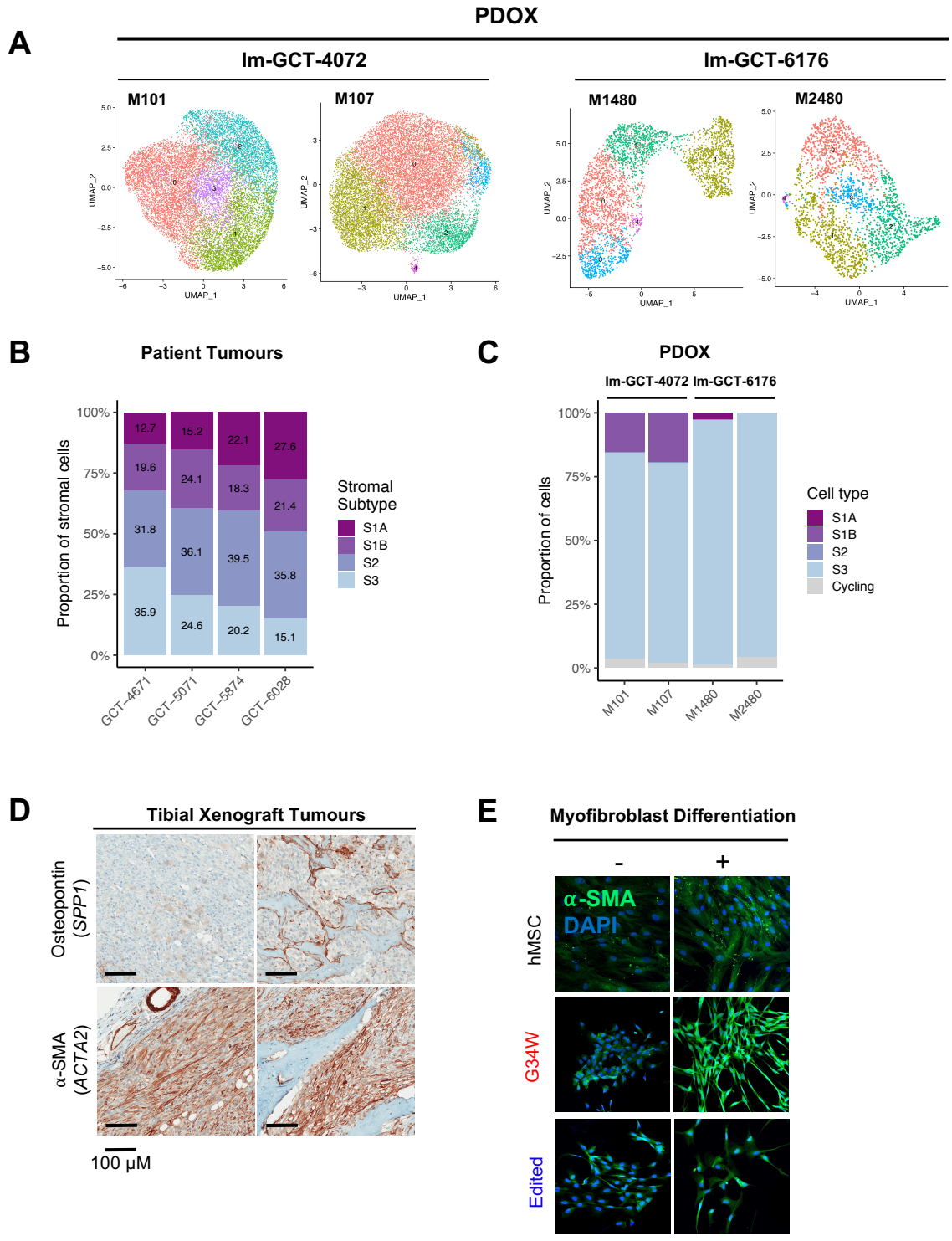

A
PDOX
Im-GCT-4072
Im-GCT-6176
M101
M107
M1480
M2480
UMAP_1
UMAP_2
B
Patient Tumours
Proportion of stromal cells
100%
75%
50%
25%
0%
12.7
19.6
31.8
35.9
15.2
24.1
36.1
24.6
22.1
18.3
39.5
20.2
27.6
21.4
35.8
15.1
GCT-4671
GCT-5071
GCT-5874
GCT-6028
Stromal Subtype
S1A
S1B
S2
S3
C
PDOX
Im-GCT-4072 Im-GCT-6176
Proportion of cells
M101
M107
M1480
M2480
Cell type
S1A
S1B
S2
S3
Cycling
D
Tibial Xenograft Tumours
Osteopontin (SPP1)
α-SMA (ACTA2)
100 μM
E
Myofibroblast Differentiation
-
+
hMSC
G34W
Edited
α-SMA
DAPI

Figure S6. Related to Figure 5.

(A) UMAP plots displaying unbiased clustering of snRNA-seq data from n=4 PDOX tumours from 2 patient cell lines, Im-GCT-4072 and Im-GCT-6176. Refer to **Table S1** for more details.

(B) Stacked bar graph showing the percent of stromal cells from each patient tumour sample that belong to each stromal cell subtype.

(C) Stacked bar graph showing the percent of stromal cells from each PDOX sample that were classified using the patient tumour cells (**Fig. 5B**) as a SingleR reference.

(D) Representative IHC for SPP1 (osteopontin) and ACTA2 (α-SMA) in two PDOX tumours, showing comparatively more ACTA2+ cells than SPP1+ cells, consistent with an elevated proportion of S3 cells in PDOX tumours.

(E) Representative immunofluorescence images for alpha smooth muscle actin (α-SMA) staining for hMSCs and isogenic Im-GCT-4072 cells maintained in non-induced (-) or myofibroblast differentiation media (+) for 2 weeks.

Figure S7

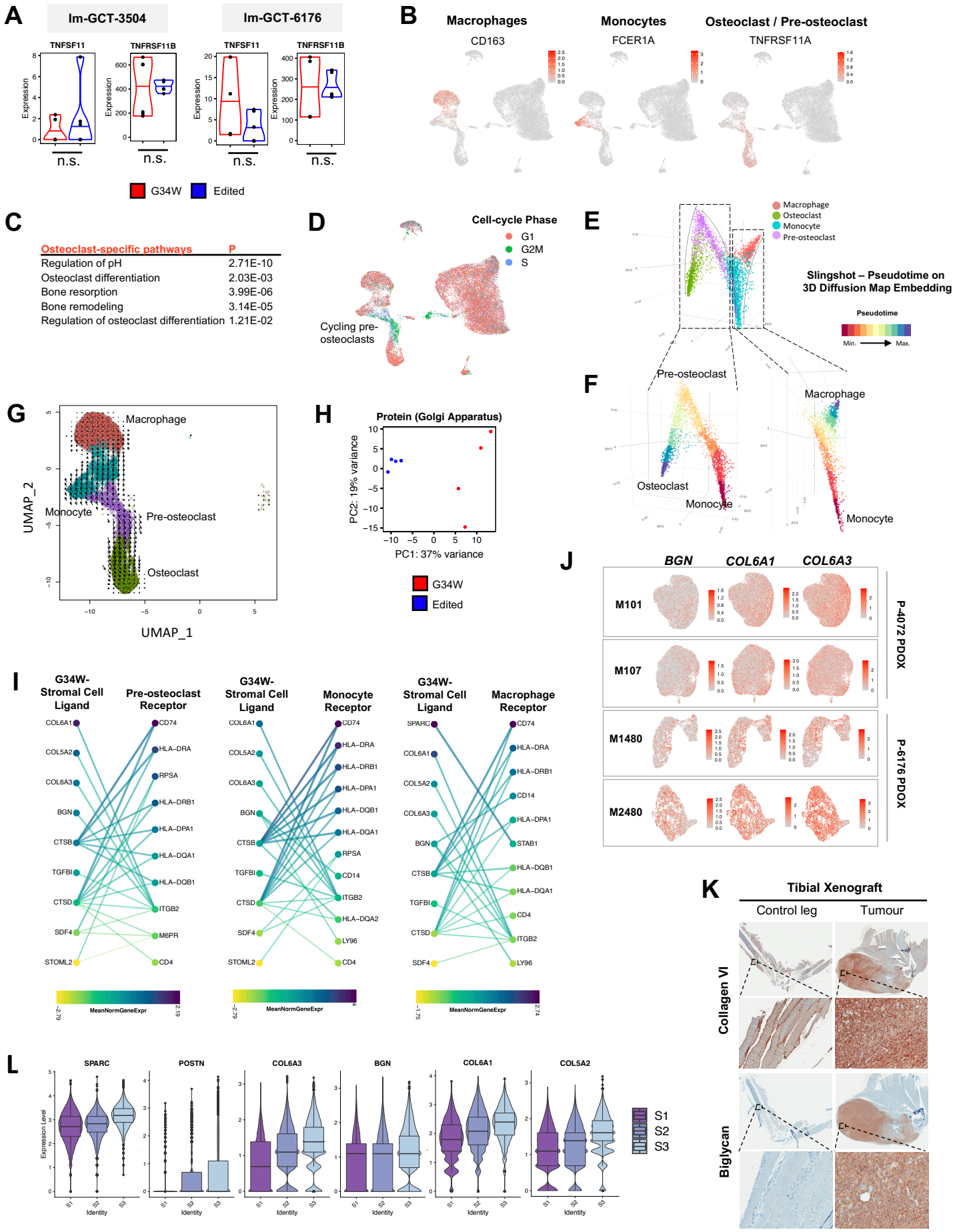

A
Im-GCT-3504
Im-GCT-6176
TNFSF11
TNFRSF11B
Expression
n.s.
G34W
Edited
B
Macrophages
CD163
Monocytes
FCER1A
Osteoclast / Pre-osteoclast
TNFRSF11A
C
Osteoclast-specific pathways P
Regulation of pH 2.71E-10
Osteoclast differentiation 2.03E-03
Bone resorption 3.99E-06
Bone remodeling 3.14E-05
Regulation of osteoclast differentiation 1.21E-02
D
Cell-cycle Phase
G1
G2M
S
Cycling pre-osteoclasts
E
Macrophage
Osteoclast
Monocyte
Pre-osteoclast
Slingshot – Pseudotime on 3D Diffusion Map Embedding
Pseudotime
Min.
Max.
F
Pre-osteoclast
Macrophage
Osteoclast
Monocyte
Monocyte
G
Macrophage
Monocyte
Pre-osteoclast
Osteoclast
UMAP_2
UMAP_1
H
Protein (Golgi Apparatus)
PC2: 19% variance
PC1: 37% variance
G34W
Edited
I
G34W-Stromal Cell Ligand
Pre-osteoclast Receptor
Monocyte Receptor
Macrophage Receptor
COL6A1
COL5A2
COL6A3
BGN
CTSB
TGFBI
CTSD
SDF4
STOML2
SPARC
CD74
HLA-DRA
HLA-DRB1
HLA-DPA1
HLA-DQB1
HLA-DQA1
RPSA
CD14
ITGB2
HLA-DQA2
LY96
M6PR
CD4
STAB1
MeanNormGeneExpr
J
BGN
COL6A1
COL6A3
M101
M107
M1480
M2480
P-4072 PDOX
P-6176 PDOX
K
Tibial Xenograft
Control leg
Tumour
Collagen VI
Biglycan
L
SPARC
POSTN
COL6A3
BGN
COL6A1
COL5A2
Expression Level
Identity
S1
S2
S3

Figure S7. Related to Figure 6.

(A) Violin plots depicting expression levels of *TNFSF11* (RANKL) and *TNFRSF11B* (OPG) in G34W (red) and edited (blue) lines from Im-GCT-3504 (left) and Im-GCT-6176 (right). *: $P<0.05$, ***: $P<0.001$, n.s.: non-significant. Gene expression levels reported in median-of-ratios normalized read counts. Significance assessed using DESeq2.

(B) UMAP plot illustrating expression of the tumour associated macrophage marker *CD163,* monocyte marker *FCER1A,* and osteoclast marker *TNFRSF11A* (RANK).

(C) Pathway enrichment using g:Profiler (GO:BP, term size<1000, $P<0.05$) for putative osteoclast-specific genes, showing enrichment for terms relevant to osteoclast differentiation and function.

(D) Average expression of cell-cycle associated genes plotted on UMAP illustrates that pre-osteoclasts are highly cycling cells relative to other myeloid cell types.

(E) 3D diffusion map using the lineage inference method Slingshot to visualize the complete myeloid trajectory, showing a bifurcating fate for monocytes and some pre-osteoclasts.

(F) 3D diffusion map from **Fig. S7E**, segregated by Slingshot predicted lineages and coloured by pseudotime (red: start, blue: end). Left: the osteoclast lineage trajectory shown in **Fig. 6C** in 3D. Right: the monocyte to macrophage lineage trajectory.

(G) Lineage inference analysis using Velocyto RNA velocity overlaid on UMAP embedding of myeloid cells suggests two trajectories with monocytes at the origin: one towards macrophages and the other towards the osteoclast lineage, with some pre-osteoclasts following a macrophage trajectory.

(H) PCA reveals distinct Golgi protein abundance profiles in G34W (red) and edited (blue) lines from Im-GCT-4072. Protein abundance profiles were derived from intensities and normalized to log2(1+x). $n=4$ technical replicates were used per condition. Refer to **Table S1** for more details.

(I) CCInx predicted ligand-receptor interactions between GCT stromal cells (left) and pre-osteoclasts, monocytes or macrophages. Colors represent the mean normalized gene expression in each cell type. Only interactions between genes specific to stromal or highly expressed in pre-osteoclasts/monocyte/macrophage cells are shown.

(J) UMAP plot illustrating expression of osteoclast-interacting ligands identified in **Fig. 6E** in 4 PDOX tumour samples, validating the importance of these ligands in both patient and PDOX tumours.

(K) Representative IHC for collagen type VI and biglycan in a tibial xenograft tumour from implantation of Im-GCT-4072 G34W cells shown in **Fig. 1F**. Stronger extracellular collagen VI and BGN staining is observed in G34W xenograft relative to the bone matrix of a control contralateral mouse leg.

(L) Box plots illustrating gene expression for genes found to have significantly enriched expression (Seurat Wilcox test, $P<0.05$, FDR corrected) in the S3 stromal subtype (*SPARC, POSTN, COL6A3, BGN, COL6A1, COL5A2*) and significantly enriched secretion in G34W stromal cells ($P<0.05$).