SUPPLEMENTAL FIGURE LEGENDS

Figure S1. Related to Figures 1&2. GRHL2 binding correlates with nucleosome removal and full enhancer activation. A) Principal component analysis of Combat batch-normalized RNA-seq data for ESCs, EpiLCs, inner cell mass (ICM), and pre- and post-implantation epiblast. B) Quantitative PCR of the Grainyhead family of TFs (*Grhl1*, *2*, and *3*) in EpiLCs vs ESCs. Error bars indicate standard deviation for n=3 biological replicates. $p<0.05$ for GRHL2 and GRHL3. C) A pair of guide RNAs against exon 2 of *Grhl2* were designed (cut sites indicated by black triangles) and introduced along with Cas9 nickase to generate protein truncations. Primers surrounding the target site (green arrows) were designed to amplify the region by PCR for sequencing verification. Positive clones were verified by western blot using a GRHL2 antibody (arrow indicates specific band). TBP = TATA binding protein loading control D) Metagene analysis for K4me3 at GRHL2 sites and all active transcriptional start sites in EpiLCs (having at least 1 count per million). E) Heatmaps for enhancer marks in ESCs and EpiLCs at GRHL2 sites that were identified relative to IgG and that contain a GRHL2 motif. All plots are shown for an 8kb window centered at GRHL2 binding sites. Sites are shown in descending order based on mean GRHL2 ChIP signal intensity. F) Heatmaps for enhancer marks in ESCs and EpiLCs at GRHL2 sites that were identified relative to IgG but lack a GRHL2 motif. All plots are shown for an 8kb window centered at GRHL2 binding sites. Sites are shown in descending order based on mean GRHL2 ChIP signal intensity. G) Metagene analysis at GRHL2 sites for K4me1 signal in ESCs and EpiLCs (Buecker et al. 2014). H)

Metagene analysis at GRHL2 sites for K27ac signal in ESCs and EpiLCs (Buecker et al. 2014). I) Metagene analysis of average signal across SMC1 sites common to both ESCs and EpiLCs in wildtype ESCs and EpiLCs. All plots are shown for an 8000bp window centered at SMC1 binding sites.

Figure S2. Related to Figure 3. GRHL2 is necessary for full enhancer activation. A) Heatmaps for enhancer marks in WT and GRHL2 KO EpiLCs at GRHL2 sites that were identified relative to IgG and that contain a GRHL2 motif. All plots are shown for an 8kb window centered at GRHL2 binding sites. Sites are shown in descending order based on mean GRHL2 ChIP signal intensity. B) Heatmaps for enhancer marks in WT and GRHL2 KO EpiLCs at GRHL2 sites that were identified relative to IgG but lack a GRHL2 motif. All plots are shown for an 8kb window centered at GRHL2 binding sites. Sites are shown in descending order based on mean GRHL2 ChIP signal intensity. C) Metagene analysis of average signal across EpiLC-specific SMC1 sites (minus sites co-bound by GRHL2) in GRHL2 knockout and wildtype EpiLCs. All plots are shown for an 8000bp window centered at SMC1 binding sites. D) Quantitative RT PCR in wildtype (V6.5) and GRHL2 knockout ESCs and EpiLCs was performed for naïve (*Klf2*, *Klf4*, *Rex1*) and formative (*Fgf5*, *Otx2*, *Dnmt3b*) markers, and a fold change was calculated in EpiLCs vs ESCs for each line. Error bars indicate standard deviation for n=4 biological replicates. $p<0.05$ for *Dnmt3b* only in WT vs KO.

Figure S3. Related to Figure 3. GRHL2 is sufficient for full enhancer activation. A) V6.5 ESCs containing *Rosa26*-M2rtTA and a modified *ColA1* locus containing a neomycin resistance gene flanked by FRT sites and a

promoterless, ATG-less hygromycin resistance cassette were targeted with a plasmid containing a dox-inducible HA-GRHL2 insert. The entire plasmid is inserted downstream of the *ColA1* locus via co-electroporation of a Flp recombinase plasmid, replacing the neomycin gene. Correct insertion of the targeting construct places a PGK promoter followed by an ATG start codon right next to the hygromycin sequence, allowing positive clones to be selected with hygromycin treatment. B) Western blot for GRHL2 of hygromycin resistant clones plus or minus 24 hours of doxycycline treatment. TBP = TATA binding protein loading control. C) Quantitative RT-PCR for GRHL2 expression in dox-inducible HA-GRHL2 ESCs treated with dox for 24 hours relative to wildtype EpiLCs. Error bars indicate standard deviation for n = 3 biological replicates. D) Heatmaps for enhancer marks in WT and GRHL2 overexpression ESCs at all ectopic GRHL2 sites that are not bound by endogenous GRHL2 in EpiLCs. All plots are shown for an 8kb window centered at GRHL2 binding sites. Sites are shown in descending order based on mean GRHL2 ChIP signal intensity. E) Fold change in expression of naïve (*Klf2*, *Klf4*, *Rex1*) and formative (*Fgf5*, *Otx2*, *Dnmt3b*) markers in WT and dox-inducible HA-GRHL2 ESCs after 24 hours of doxycycline treatment. $p < 0.05$ for *Klf4*, *Rex1*, and *Dnmt3b* by Students t-test. n = 3 biological replicates.

Figure S4. Related to Figure 4. Identification of GRHL2 target genes and alternative ESC enhancers. A) Histogram of distances in kb between all GRHL2 sites and the nearest gene promoter within the same TAD. B) Ward hierarchical clustering of 4 independent clones of WT V6.5 and GRHL2 KO EpiLCs based on

expression profiling data. C) Quantitative RT-PCR verification of microarray profiling data for significantly down-regulated GRHL2 targets in WT and GRHL2 KO EpiLCs. $P < 0.05$ by Students t-test for each gene, except for *Elf3* ($p = 0.051$). Error bars indicate standard deviation for n = 3-4 biological replicates. D) Quantitative RT-PCR for GRHL2 targets in (C), but in WT and GRHL2 KO ESCs. No significant differences were observed. Error bars indicate standard deviation for n = 3 biological replicates. E) Schematic for identifying alternative ESC enhancers. We identified likely GRHL2 targets in EpiLCs by identifying the gene nearest to GRHL2 binding sites within the same TAD (based on Dixon et al. 2012). We then identified genes that are likely regulated by enhancers in ESCs by identifying genes within the same TAD that are nearest to active ESC enhancers, based on presence of a significant ATAC peak (FDR < 0.05) and both K4me1 and K27ac signal. To identify GRHL2 targets that are regulated by an enhancer in ESCs, we overlapped the ESC genes nearest to an active enhancer with GRHL2 targets in EpiLCs. To find the alternative enhancer associated with these genes in ESCs, we then matched these genes back to the corresponding enhancer sites. F) Histogram of distances in kb between GRHL2 target promoters and their corresponding alternative ESC enhancer sites.

Figure S5. Related to Figure 5. Identification of candidate transcription factors driving GRHL2 targets in ESCs. A) Histogram of distances in kb between GRHL2 target promoters and their corresponding KLF4 enhancer sites in ESCs. B) Western blot for KLF2 and KLF4 in Klf2 and Klf4 double knockout (DKO) ESCs. TBP = TATA-binding protein as a loading control. C) Number of

genes that are significantly up- or down-regulated with KLF2/4 DKO for all genes and predicted GRHL2 targets.

Figure S6. Related to Figure 6. GRHL2-regulated genes undergo an enhancer switch during the ESC to EpiLC transition. A) Levels of enhancer marks (ATAC signal, H3K4me1, H3K27ac, and SMC1) at the *Jam1* locus in ESCs and EpiLCs. KLF4 and GRHL2 ChIP-seq tracks and significant binding sites are indicated by purple and blue bars, respectively. KLF4 and GRHL2 sites deleted using CRISPR are indicated. B) Fold change in *Jam1* expression with deletion of the nearby KLF4-bound enhancer (left) or GRHL2-bound enhancer (right) in ESCs and EpiLCs. ** indicates $p < 0.01$ and *** indicates $p< 0.005$ by Student's t-test.

Figure S7. Related to Figure 7. GRHL2 maintains expression of epithelial and cell adhesion genes in EpiLCs. A) Brightfield images of wildtype and GRHL2 knockout ESCs and EpiLCs. B) Gene ontology of genes predicted to be regulated by GRHL2 enhancers by GREAT analysis. C) Gene ontology analysis using Enrichr for genes nearest to GRHL2 enhancers within the same TAD. D) Expression fold change of epithelial and mesenchymal markers in wildtype EpiLCs vs ESCs measured by quantitative RT-PCR. Error bars indicate standard deviation for n=3 biological replicates. $p<0.05$ for *Vim* only; $p=0.09$ for *Cldn6*. E) Quantitative RT-PCR for EMT markers in WT and GRHL2 KO ESCs. No significant differences were observed for n = 3 biological replicates. Error bars indicate standard deviation. F) Spearman correlations in E6.5 single cell RNA-

seq data for expression of GRHL2, anterior epiblast markers at E6.5 (*Cdh1*, *Otx2*, *Sox2*) and primitive streak markers (*Evx1*, *Fgf8*, *Snai1*, *T*).

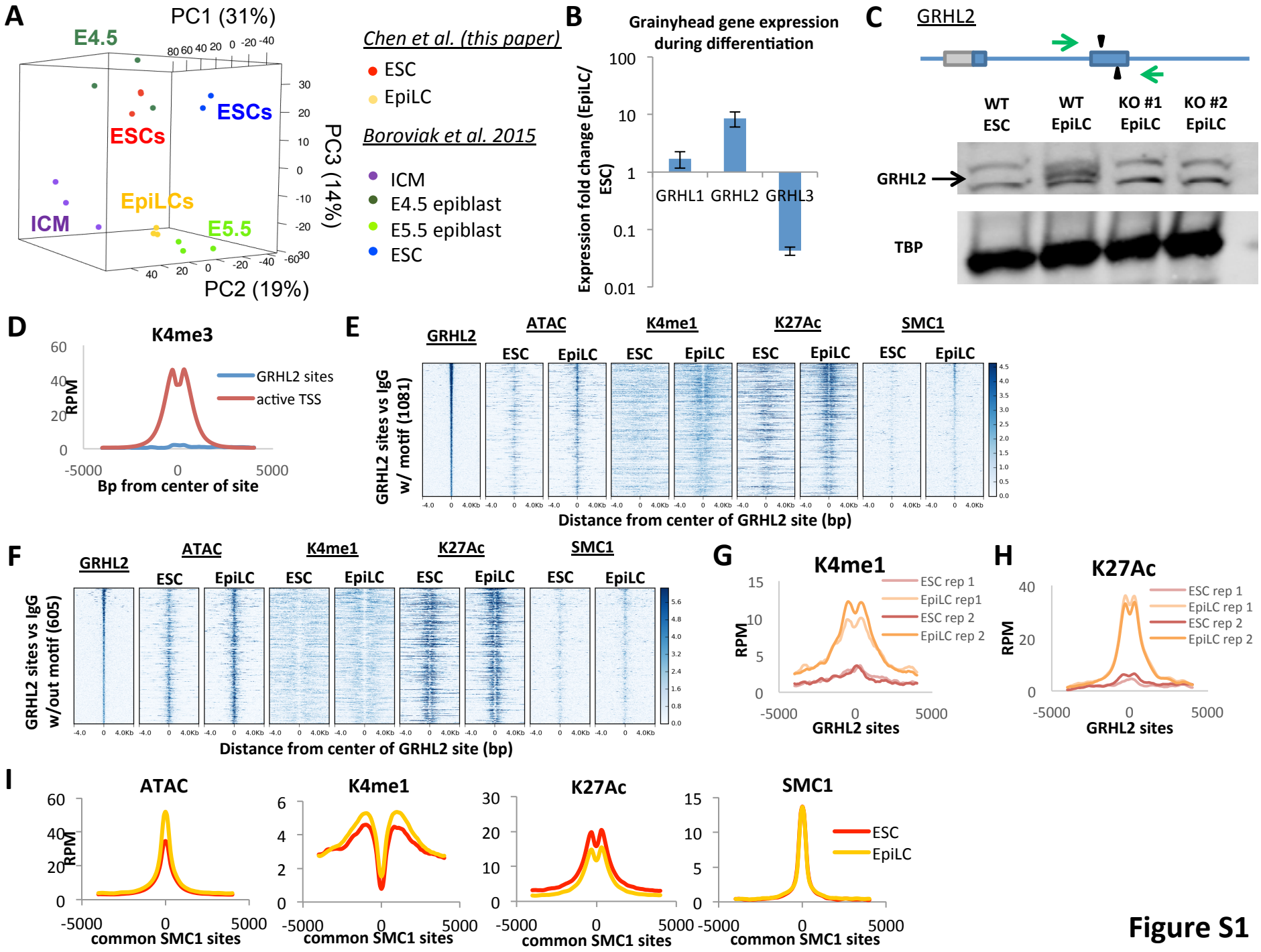

A
PC1 (31%)
E4.5
ESCs
ESCs
EpiLCs
ICM
E5.5
PC3 (14%)
PC2 (19%)
Chen et al. (this paper)
ESC
EpiLC
Boroviak et al. 2015
ICM
E4.5 epiblast
E5.5 epiblast
ESC
B
Grainyhead gene expression during differentiation
Expression fold change (EpiLC/ ESC)
GRHL1 GRHL2 GRHL3
C
GRHL2
WT ESC
WT EpiLC
KO #1 EpiLC
KO #2 EpiLC
GRHL2
TBP
D
K4me3
RPM
GRHL2 sites
active TSS
Bp from center of site
E
GRHL2
ATAC
K4me1
K27Ac
SMC1
ESC
EpiLC
GRHL2 sites vs IgG w/ motif (1081)
Distance from center of GRHL2 site (bp)
F
GRHL2 sites vs IgG w/out motif (605)
Distance from center of GRHL2 site (bp)
G
K4me1
ESC rep 1
EpiLC rep1
ESC rep 2
EpiLC rep 2
RPM
GRHL2 sites
H
K27Ac
ESC rep 1
EpiLC rep 1
ESC rep 2
EpiLC rep 2
RPM
GRHL2 sites
I
ATAC
K4me1
K27Ac
SMC1
RPM
ESC
EpiLC
common SMC1 sites

Figure S1

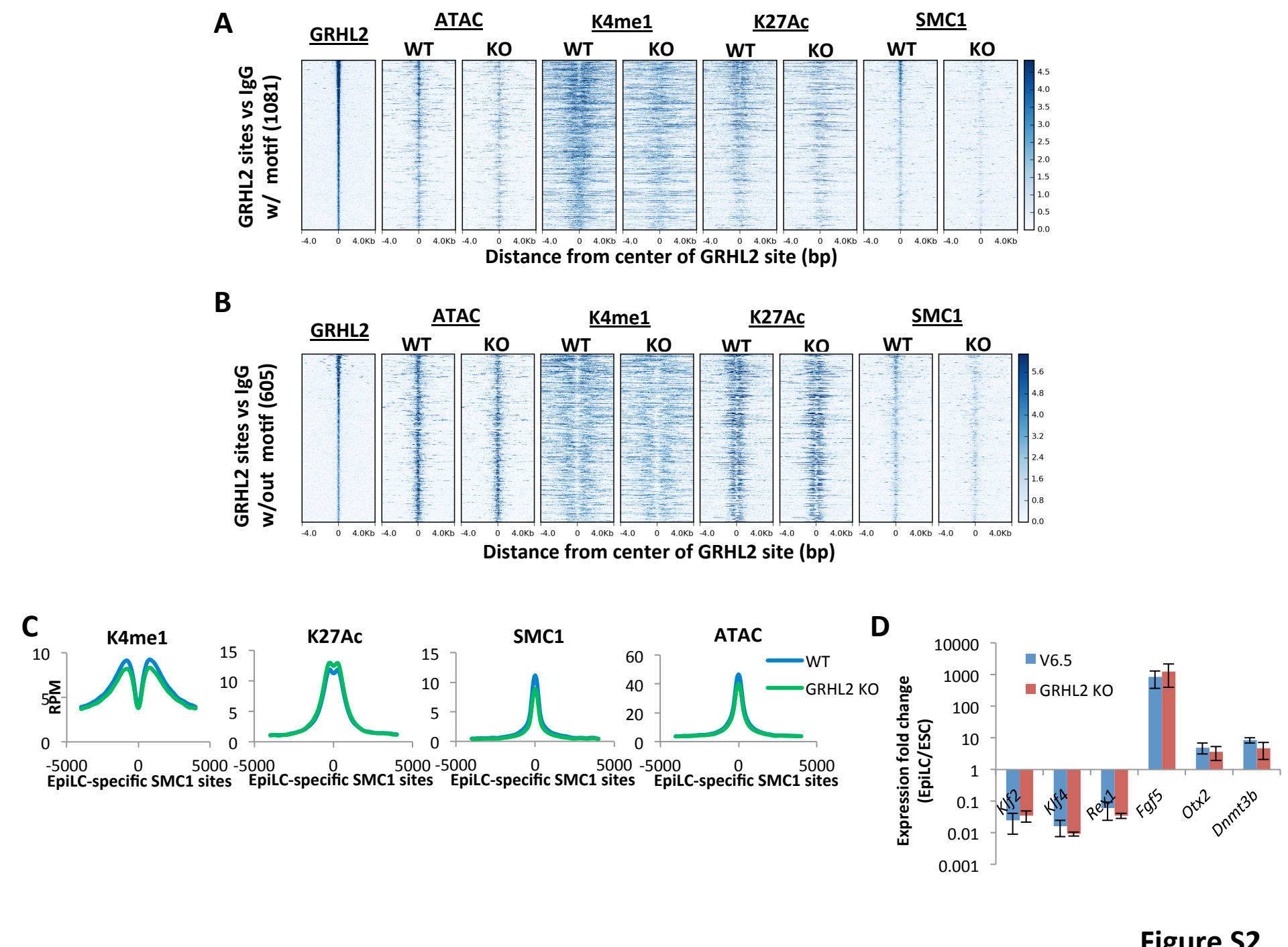
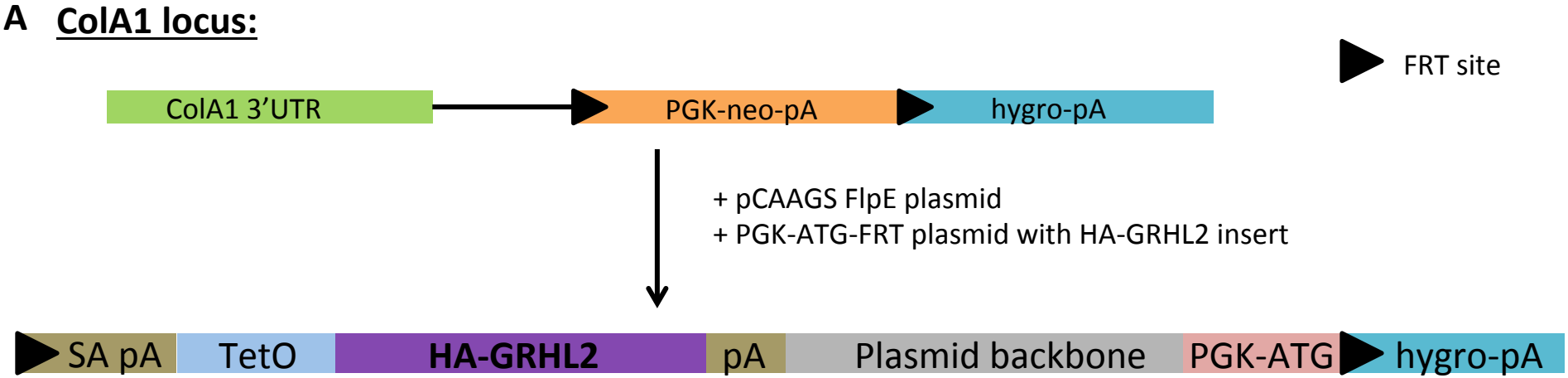
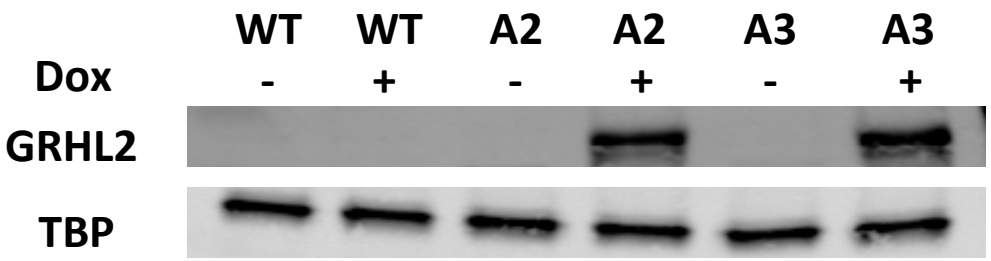
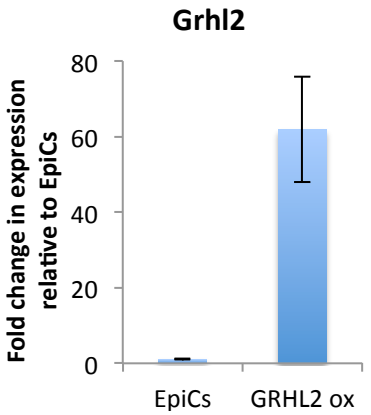
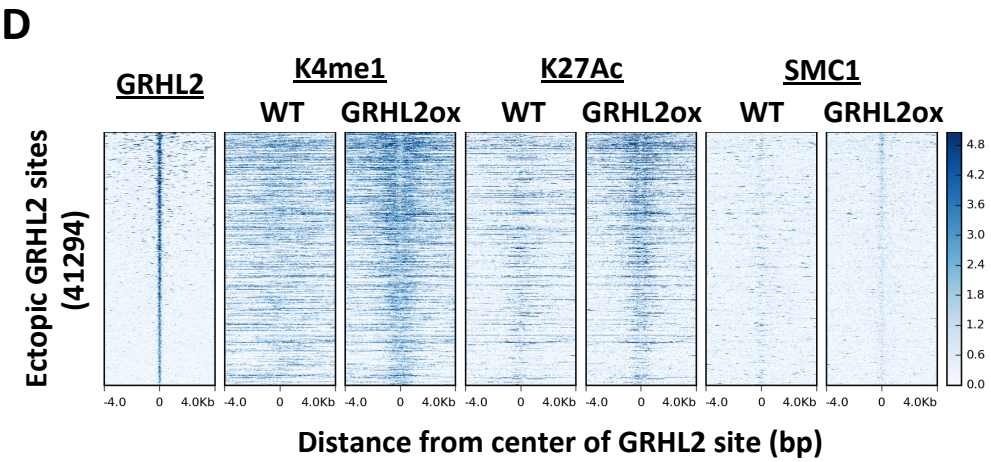

A
GRHL2
ATAC
K4me1
K27Ac
SMC1
WT
KO
GRHL2 sites vs IgG w/ motif (1081)
Distance from center of GRHL2 site (bp)
B
GRHL2 sites vs IgG w/out motif (605)
Distance from center of GRHL2 site (bp)
C
K4me1
K27Ac
SMC1
ATAC
RPM
WT
GRHL2 KO
EpiLC-specific SMC1 sites
D
Expression fold change (EpiLC/ESC)
V6.5
GRHL2 KO
Klf2
Klf4
Rex1
Fgf5
Otx2
Dnmt3b

Figure S2


A
ColA1 locus:
FRT site
ColA1 3'UTR
PGK-neo-pA
hygro-pA
+ pCAAGS FlpE plasmid
+ PGK-ATG-FRT plasmid with HA-GRHL2 insert
SA pA
TetO
HA-GRHL2
pA
Plasmid backbone
PGK-ATG
hygro-pA


B
WT
WT
A2
A2
A3
A3
Dox
-
+
-
+
-
+
GRHL2
TBP


C
Grhl2
Fold change in expression relative to EpiCs
80
60
40
20
0
EpiCs
GRHL2 ox


D
GRHL2
K4me1
K27Ac
SMC1
WT
GRHL2ox
WT
GRHL2ox
WT
GRHL2ox
Ectopic GRHL2 sites (41294)
-4.0
0
4.0Kb
4.8
4.2
3.6
3.0
2.4
1.8
1.2
0.6
0.0
Distance from center of GRHL2 site (bp)

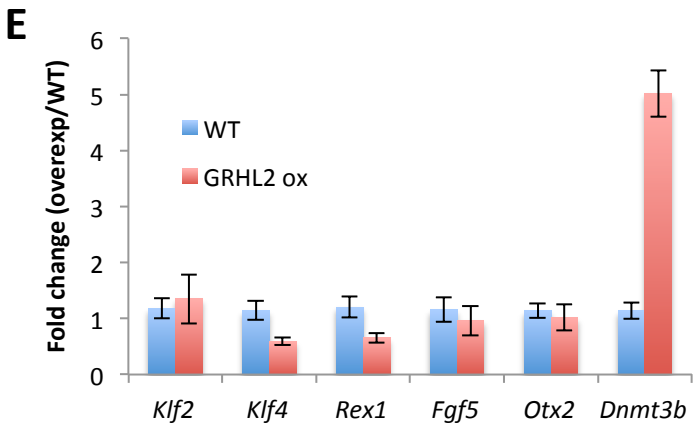
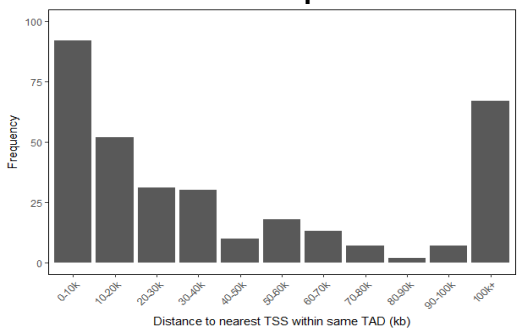
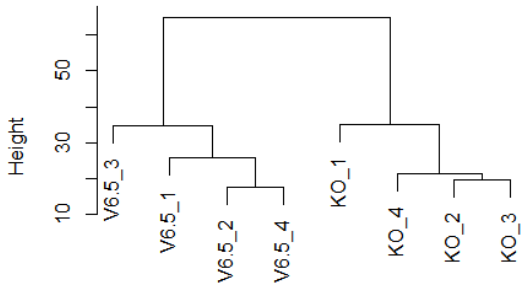

E
Fold change (overexp/WT)
6
5
4
3
2
1
0
WT
GRHL2 ox
Klf2
Klf4
Rex1
Fgf5
Otx2
Dnmt3b

Figure S3

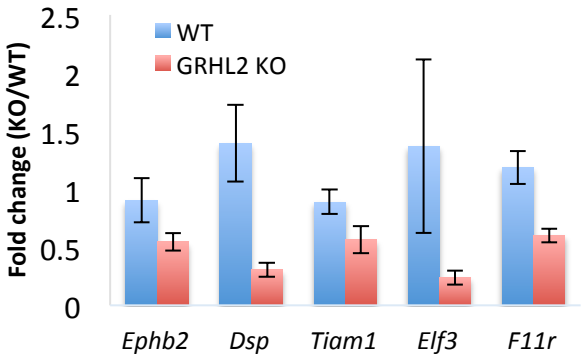
A



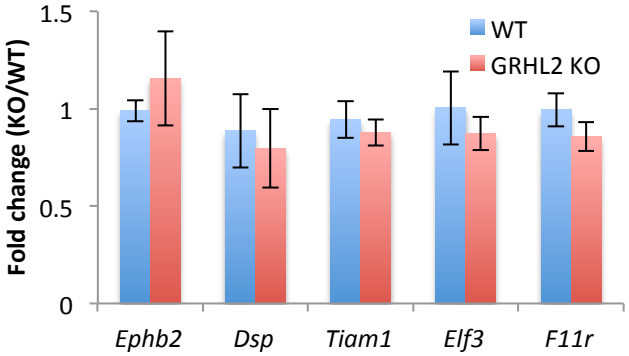
B



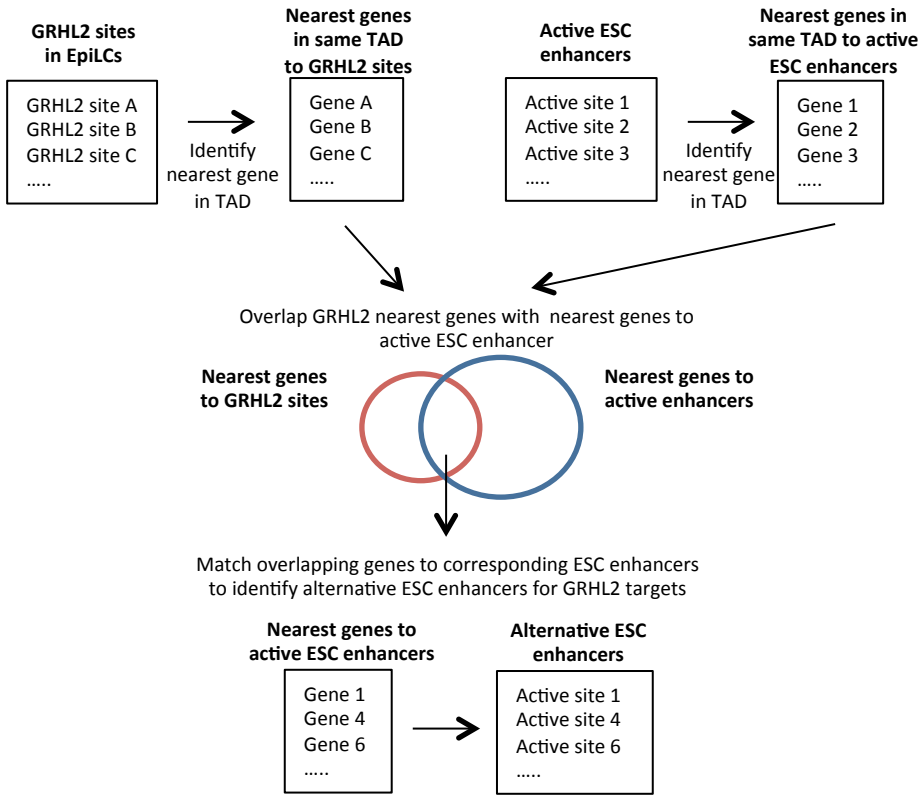
C



D



E



F

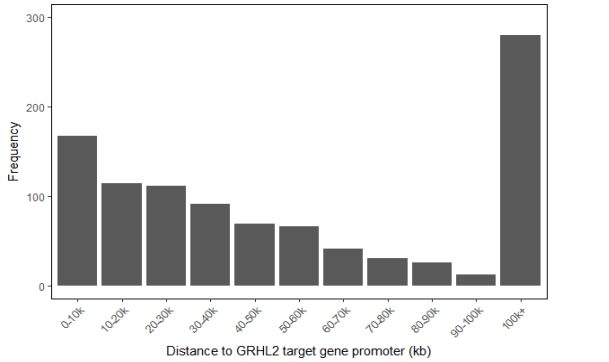
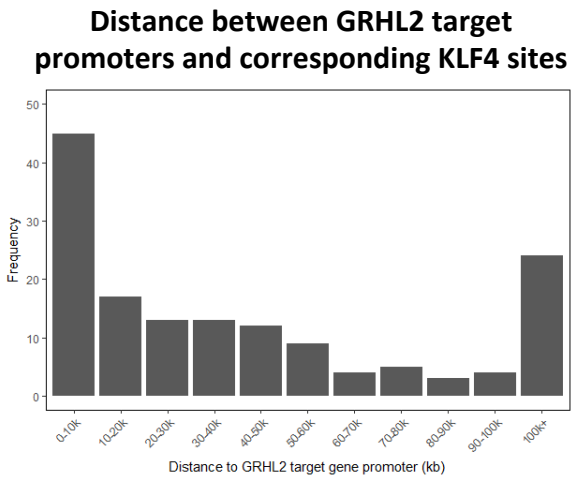
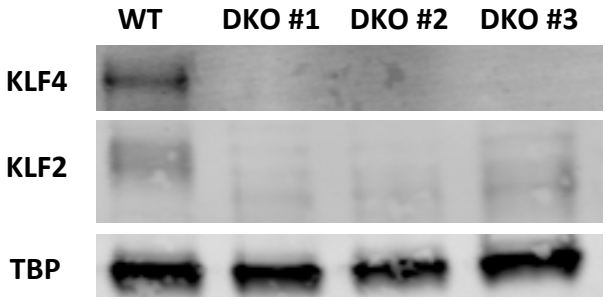


Figure S4

A



B



C

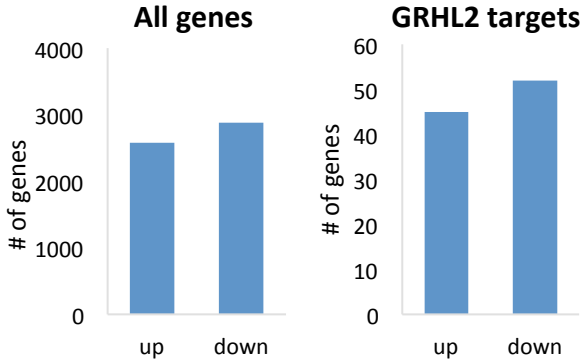
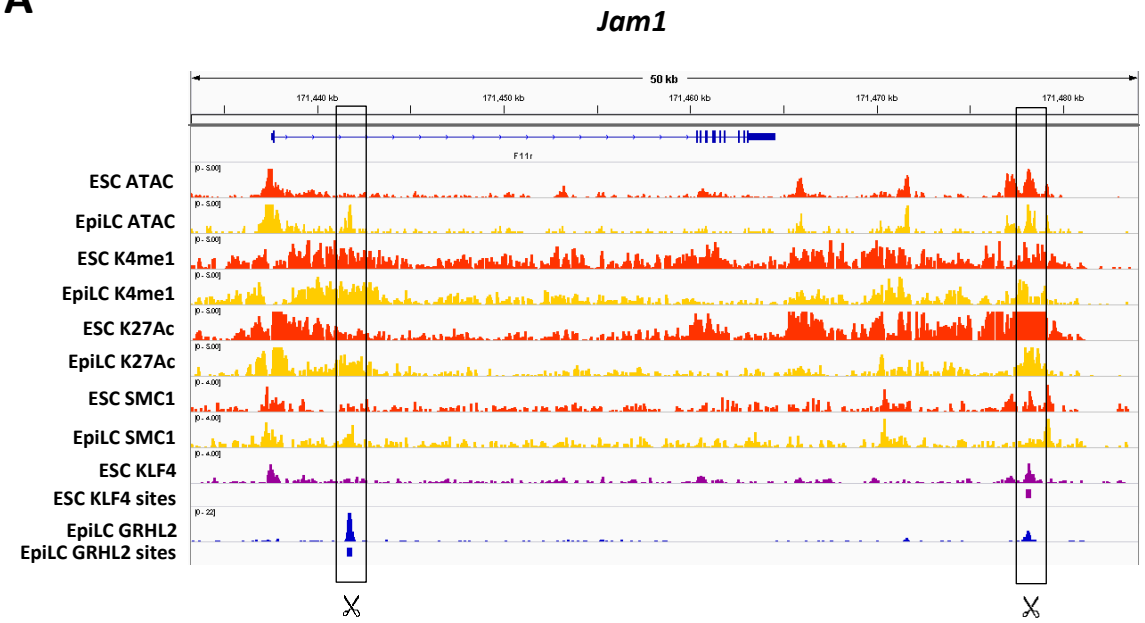


Figure S5

A


Jam1
50 kb
171,440 kb
171,450 kb
171,460 kb
171,470 kb
171,480 kb
ESC ATAC
EpiLC ATAC
ESC K4me1
EpiLC K4me1
ESC K27Ac
EpiLC K27Ac
ESC SMC1
EpiLC SMC1
ESC KLF4
ESC KLF4 sites
EpiLC GRHL2
EpiLC GRHL2 sites

B

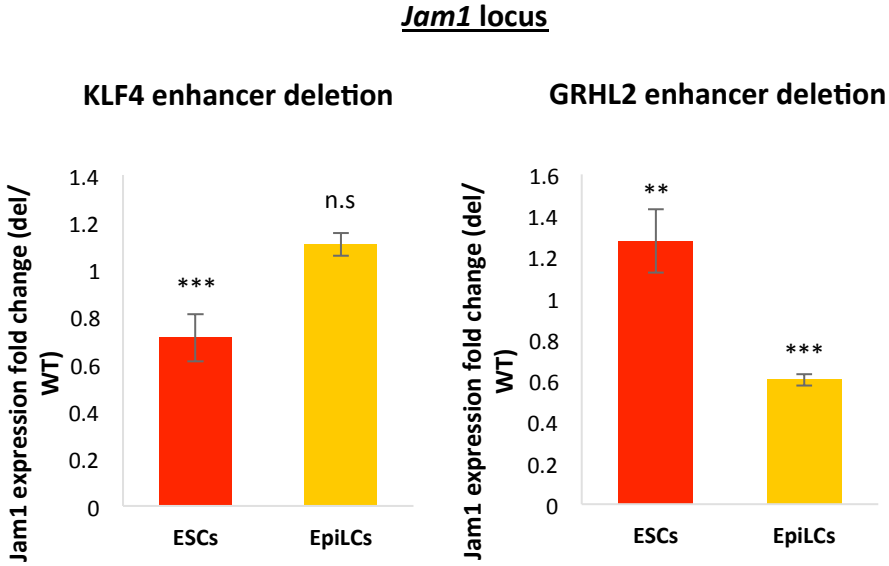

Jam1 locus
KLF4 enhancer deletion
GRHL2 enhancer deletion
Jam1 expression fold change (del/ WT)
1.4
1.2
1
0.8
0.6
0.4
0.2
0
***
n.s
ESCs
EpiLCs
Jam1 expression fold change (del/ WT)
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
**
***
ESCs
EpiLCs

Figure S6

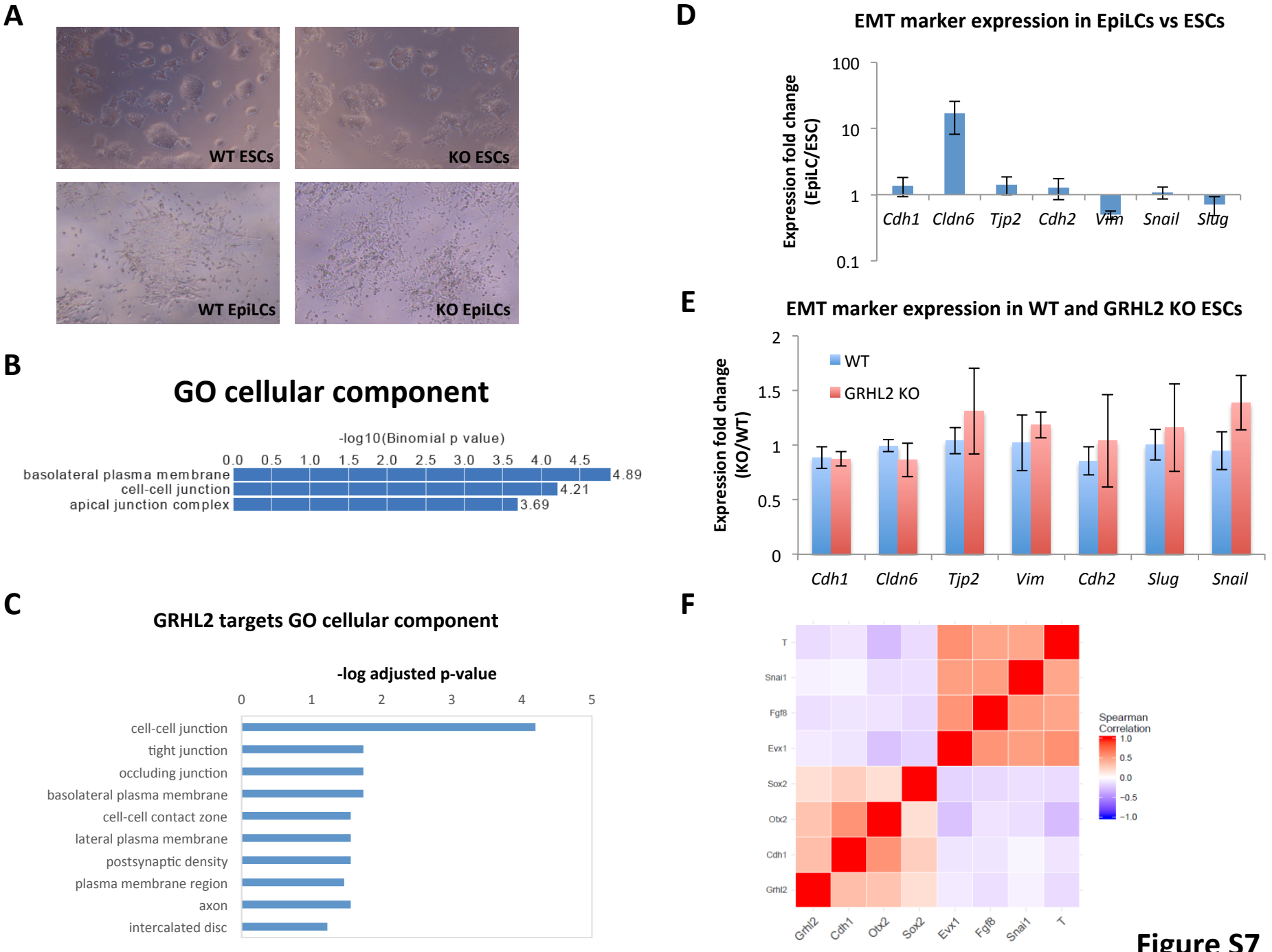

A
WT ESCs
KO ESCs
WT EpiLCs
KO EpiLCs
B
GO cellular component
-log10(Binomial p value)
0.0 0.5 1.0 1.5 2.0 2.5 3.0 3.5 4.0 4.5
basolateral plasma membrane
cell-cell junction
apical junction complex
4.89
4.21
3.69
C
GRHL2 targets GO cellular component
-log adjusted p-value
0 1 2 3 4 5
cell-cell junction
tight junction
occluding junction
basolateral plasma membrane
cell-cell contact zone
lateral plasma membrane
postsynaptic density
plasma membrane region
axon
intercalated disc
D
EMT marker expression in EpiLCs vs ESCs
Expression fold change (EpiLC/ESC)
100
10
1
0.1
Cdh1 Cldn6 Tjp2 Cdh2 Vim Snail Slug
E
EMT marker expression in WT and GRHL2 KO ESCs
Expression fold change (KO/WT)
2
1.5
1
0.5
0
WT
GRHL2 KO
Cdh1 Cldn6 Tjp2 Vim Cdh2 Slug Snail
F
T
Snai1
Fgf8
Evx1
Sox2
Otx2
Cdh1
Grhl2
Spearman Correlation
1.0
0.5
0.0
−0.5
−1.0

Figure S7