Figure S1

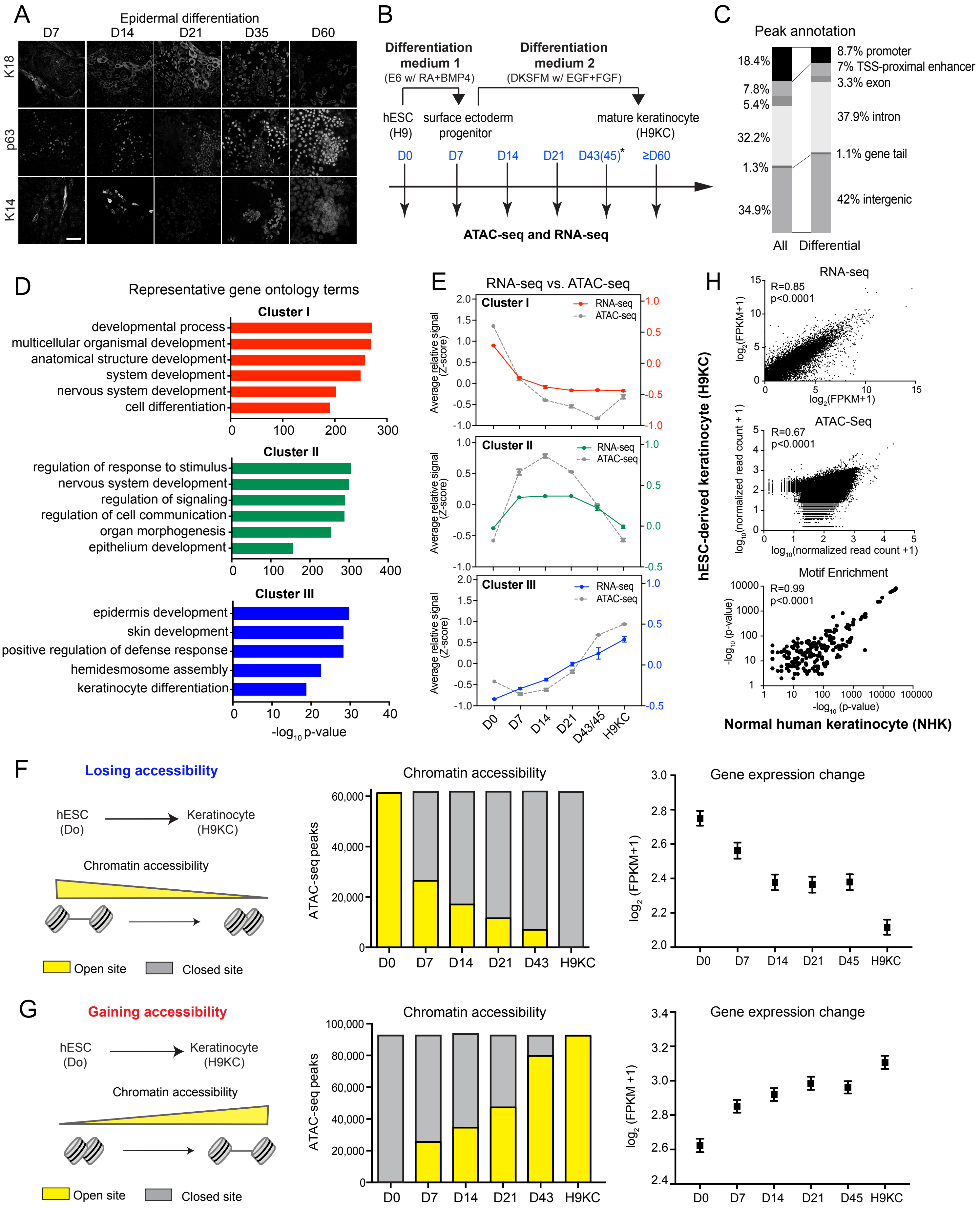

A
Epidermal differentiation
D7
D14
D21
D35
D60
K18
p63
K14
B
Differentiation medium 1
(E6 w/ RA+BMP4)
Differentiation medium 2
(DKSFM w/ EGF+FGF)
hESC (H9)
surface ectoderm progenitor
mature keratinocyte (H9KC)
D0
D7
D14
D21
D43(45)*
≥D60
ATAC-seq and RNA-seq
C
Peak annotation
18.4%
7.8%
5.4%
32.2%
1.3%
34.9%
8.7% promoter
7% TSS-proximal enhancer
3.3% exon
37.9% intron
1.1% gene tail
42% intergenic
All
Differential
D
Representative gene ontology terms
Cluster I
developmental process
multicellular organismal development
anatomical structure development
system development
nervous system development
cell differentiation
0 100 200 300
Cluster II
regulation of response to stimulus
nervous system development
regulation of signaling
regulation of cell communication
organ morphogenesis
epithelium development
0 100 200 300 400
Cluster III
epidermis development
skin development
positive regulation of defense response
hemidesmosome assembly
keratinocyte differentiation
0 10 20 30 40
-log10 p-value
E
RNA-seq vs. ATAC-seq
Cluster I
Cluster II
Cluster III
RNA-seq
ATAC-seq
Average relative signal (Z-score)
2.0 1.5 1.0 0.5 0.0 -0.5 -1.0
1.0 0.5 0.0 -0.5 -1.0
D0 D7 D14 D21 D43/45 H9KC
F
Losing accessibility
hESC (Do)
Keratinocyte (H9KC)
Chromatin accessibility
Open site
Closed site
ATAC-seq peaks
60,000 40,000 20,000 0
D0 D7 D14 D21 D43 H9KC
Gene expression change
log2 (FPKM+1)
3.0 2.8 2.6 2.4 2.2 2.0
D0 D7 D14 D21 D45 H9KC
G
Gaining accessibility
hESC (Do)
Keratinocyte (H9KC)
Chromatin accessibility
Open site
Closed site
ATAC-seq peaks
100,000 80,000 60,000 40,000 20,000 0
D0 D7 D14 D21 D43 H9KC
Gene expression change
log2 (FPKM +1)
3.4 3.2 3.0 2.8 2.6 2.4
D0 D7 D14 D21 D45 H9KC
H
hESC-derived keratinocyte (H9KC)
RNA-seq
R=0.85
p<0.0001
log2(FPKM+1)
0 5 10 15
ATAC-Seq
R=0.67
p<0.0001
log10(normalized read count +1)
0 1 2 3 4 5
Motif Enrichment
R=0.99
p<0.0001
-log10 (p-value)
1 10 100 1000 10000 100000
Normal human keratinocyte (NHK)

Figure S2

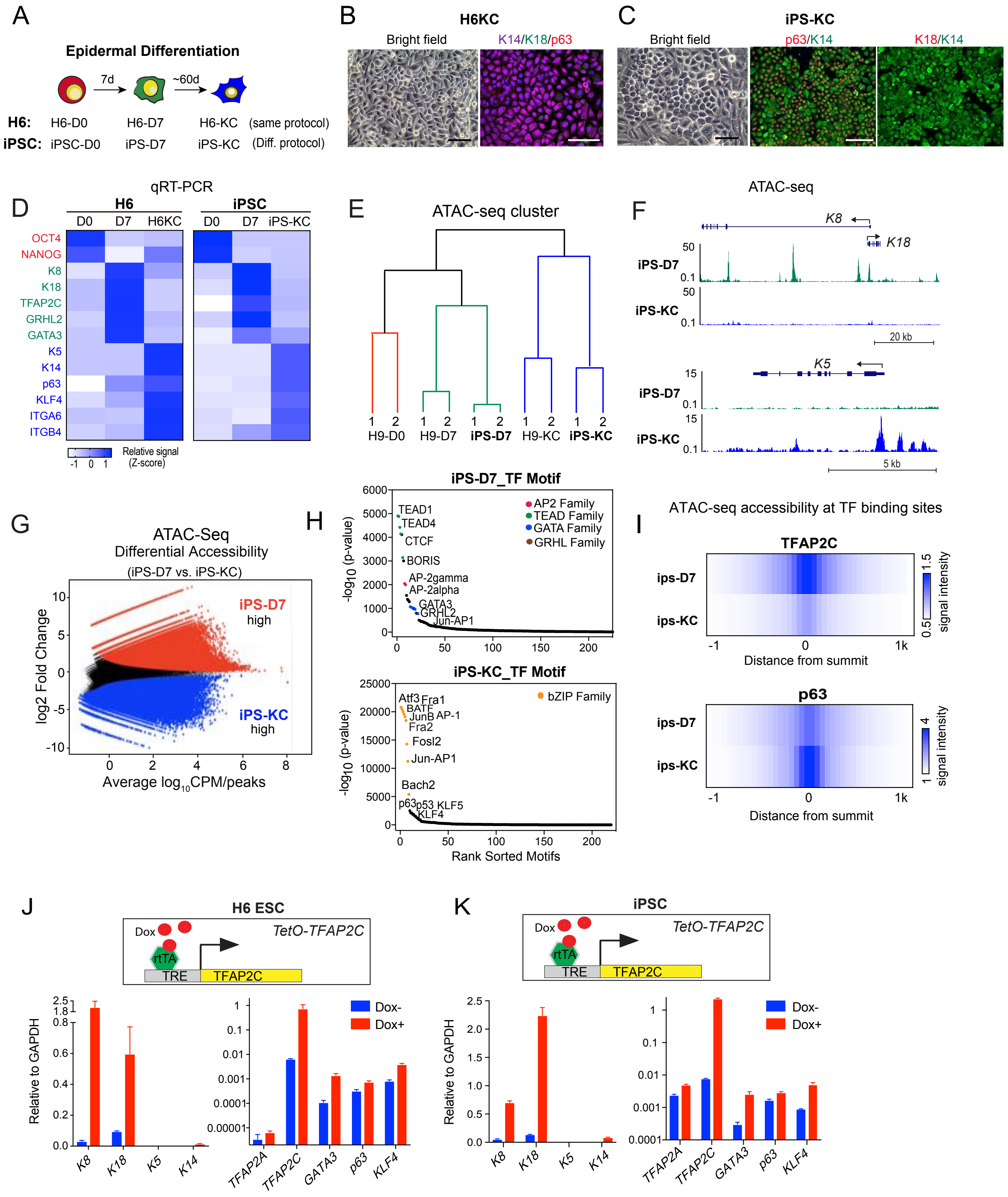
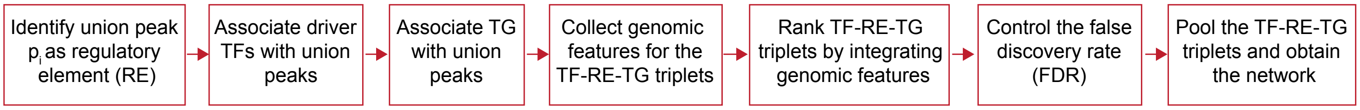

A
Epidermal Differentiation
7d
~60d
H6: H6-D0 H6-D7 H6-KC (same protocol)
iPSC: iPSC-D0 iPS-D7 iPS-KC (Diff. protocol)
B
H6KC
Bright field
K14/K18/p63
C
iPS-KC
Bright field
p63/K14
K18/K14
qRT-PCR
D
H6
iPSC
D0 D7 H6KC
D0 D7 iPS-KC
OCT4
NANOG
K8
K18
TFAP2C
GRHL2
GATA3
K5
K14
p63
KLF4
ITGA6
ITGB4
Relative signal (Z-score)
-1 0 1
E
ATAC-seq cluster
1 2 1 2 1 2 1 2 1 2
H9-D0 H9-D7 iPS-D7 H9-KC iPS-KC
ATAC-seq
F
K8
K18
iPS-D7
iPS-KC
50
0.1
20 kb
K5
15
5 kb
G
ATAC-Seq
Differential Accessibility
(iPS-D7 vs. iPS-KC)
iPS-D7 high
iPS-KC high
log2 Fold Change
10 5 0 -5 -10
0 2 4 6 8
Average log10CPM/peaks
H
iPS-D7_TF Motif
-log10 (p-value)
6000 5000 4000 3000 2000 1000 0
TEAD1
TEAD4
CTCF
BORIS
AP-2gamma
AP-2alpha
GATA3
GRHL2
Jun-AP1
AP2 Family
TEAD Family
GATA Family
GRHL Family
0 50 100 150 200
iPS-KC_TF Motif
25000 20000 15000 10000 5000 0
Atf3
Fra1
BATF
JunB
AP-1
Fra2
Fosl2
Jun-AP1
Bach2
p63
p53
KLF5
KLF4
bZIP Family
Rank Sorted Motifs
I
ATAC-seq accessibility at TF binding sites
TFAP2C
ips-D7
ips-KC
1.5
0.5
signal intensity
-1 0 1k
Distance from summit
p63
4
1
J
H6 ESC
Dox
rtTA
TRE
TFAP2C
TetO-TFAP2C
Relative to GAPDH
2.5 1.8 0.8 0.6 0.4 0.2 0.0
K8 K18 K5 K14
1 0.1 0.01 0.001 0.0001 0.00001
TFAP2A TFAP2C GATA3 p63 KLF4
Dox-
Dox+
K
iPSC
2.5 2.0 1.5 1.0 0.5 0.0
1 0.1 0.01 0.001 0.0001

Figure S3

A

TF-Accessible Regulatory Element-TG based network reconstruction pipeline



Identify union peak p_i as regulatory element (RE) → Associate driver TFs with union peaks → Associate TG with union peaks → Collect genomic features for the TF-RE-TG triplets → Rank TF-RE-TG triplets by integrating genomic features → Control the false discovery rate (FDR) → Pool the TF-RE-TG triplets and obtain the network

B

	Transcription Factor	**Accessible Regulatory Element (time course ATAC-seq union peaks)**	**Target Gene**
Investigated features	Expression level (X_{TF})** TF motif enrichment (E_{TF})*** Binding strength (B_i)*	Motif sequence conservation (C_i)* Openness level of peak region (O_i)***	Expression level (X_{TG})** Promoter openness (O_{TG})***
Public chromatin connection data	Chromatin contacts in keratinocyte, Hi-C in NHEK (Rao et al., Cell 2014) Cross-cell-type correlation of DHS (DNaseI hypersensitive sites) signal at distal enhancer and promoter (Thurman et al., Nature 2012)		

* Context non-specific features
** Context specific features from RNA-seq
*** Context specific features from ATACseq

C

Transcription factors shown in the regulatory network

Initiation Process	Maturation Process
TF associated with gaining accessibility **TFAP2C TFAP2A GRHL2 GATA3 JUN ATF3** FOS TEAD1 TEAD2 TEAD4 ELF1 CEBPB ETS1 ETS2 GABPA OCT4 SOX2	TF associated with gaining accessibility **p63 p53 KLF4 KLF5 BATF JUN FOS**
TF associated with losing accessibility **OCT4 NANOG SOX2** JUN ATF3 FOS TEAD1 TEAD4	TF associated with losing accessibility **TFAP2C GATA3 TEAD1 TEAD2 TEAD4** JUND EWSR1 PBX1 TCF3 GABPA RFX3 HOXB4 NF1 NR4A1

D

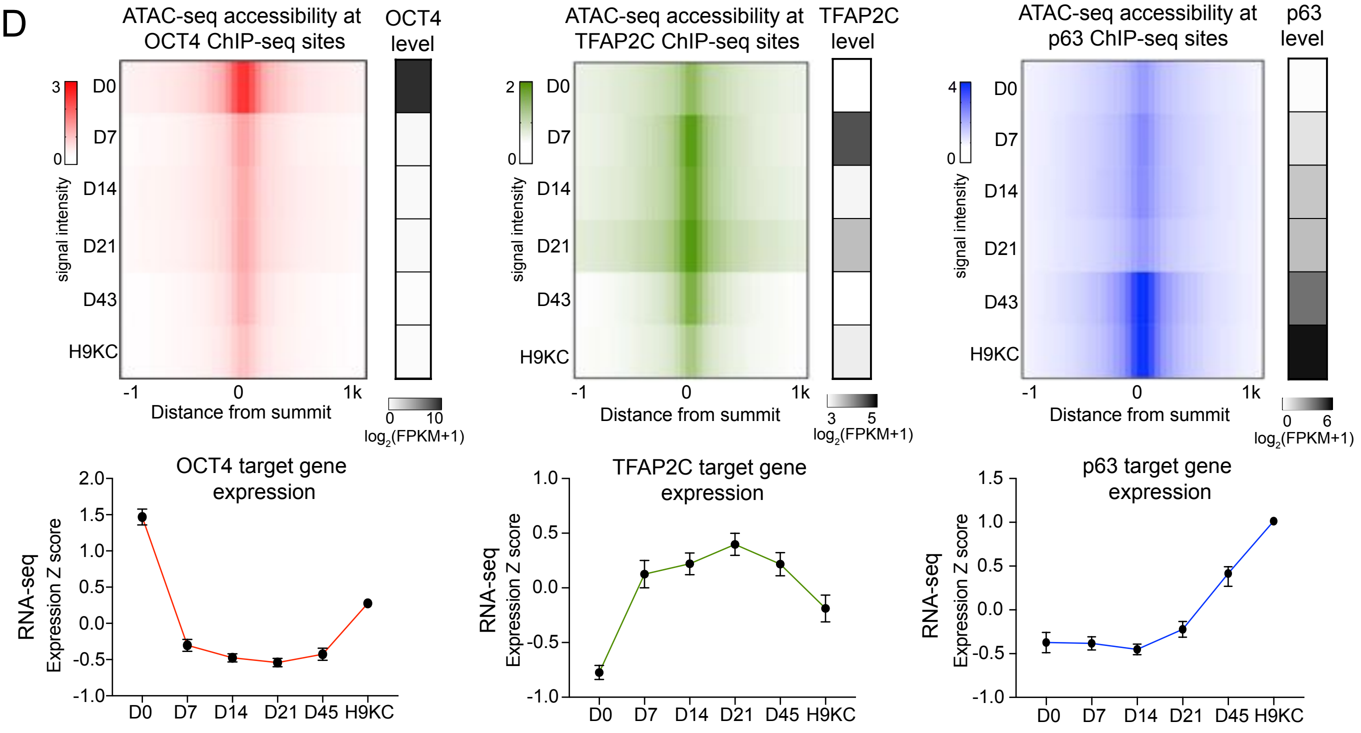


Figure S4

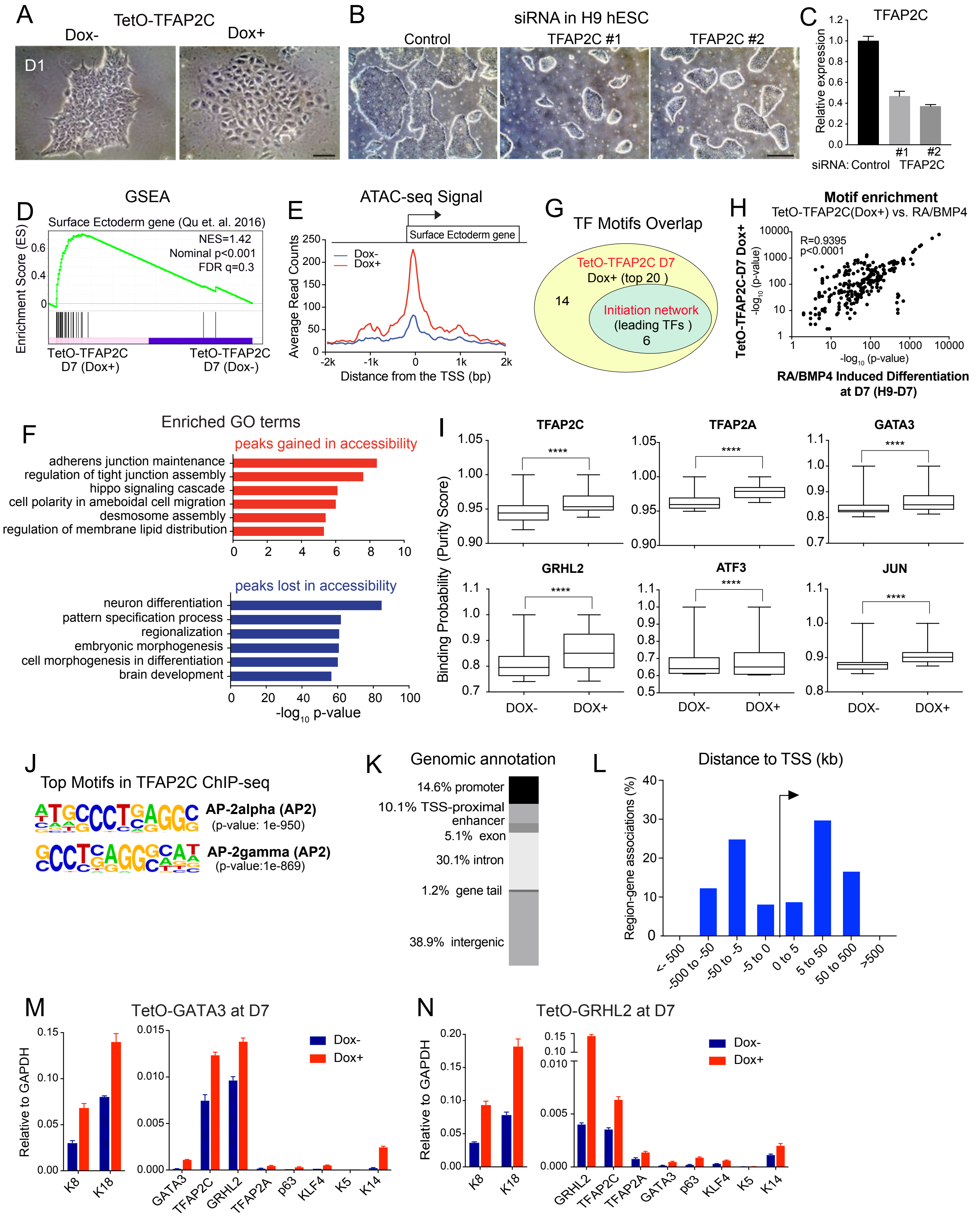
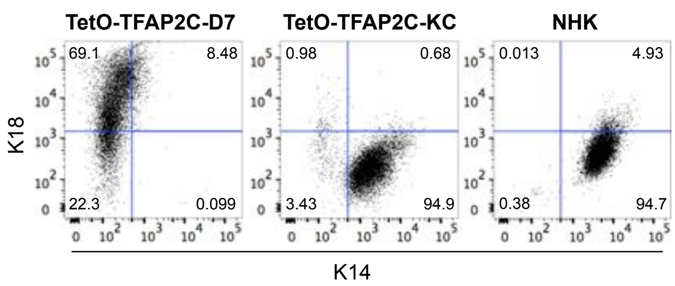

A
TetO-TFAP2C
Dox-
Dox+
D1
B
siRNA in H9 hESC
Control
TFAP2C #1
TFAP2C #2
C
TFAP2C
Relative expression
1.2
1.0
0.8
0.6
0.4
0.2
0.0
siRNA: Control
#1
#2
TFAP2C
D
GSEA
Surface Ectoderm gene (Qu et. al. 2016)
Enrichment Score (ES)
0.6
0.4
0
NES=1.42
Nominal p<0.001
FDR q=0.3
TetO-TFAP2C D7 (Dox+)
TetO-TFAP2C D7 (Dox-)
E
ATAC-seq Signal
Surface Ectoderm gene
Average Read Counts
250
200
150
100
50
0
Dox-
Dox+
-2k
-1k
0
1k
2k
Distance from the TSS (bp)
G
TF Motifs Overlap
TetO-TFAP2C D7
Dox+ (top 20)
14
Initiation network
(leading TFs)
6
H
Motif enrichment
TetO-TFAP2C(Dox+) vs. RA/BMP4
TetO-TFAP2C-D7 Dox+
-log10 (p-value)
R=0.9395
p<0.0001
10000
1000
100
10
1
RA/BMP4 Induced Differentiation at D7 (H9-D7)
F
Enriched GO terms
peaks gained in accessibility
adherens junction maintenance
regulation of tight junction assembly
hippo signaling cascade
cell polarity in ameboidal cell migration
desmosome assembly
regulation of membrane lipid distribution
0
2
4
6
8
10
peaks lost in accessibility
neuron differentiation
pattern specification process
regionalization
embryonic morphogenesis
cell morphogenesis in differentiation
brain development
0
20
40
60
80
100
-log10 p-value
I
Binding Probability (Purity Score)
TFAP2C
TFAP2A
GATA3
GRHL2
ATF3
JUN
****
1.05
1.00
0.95
0.90
1.1
1.0
0.9
0.8
0.7
0.6
0.5
DOX-
DOX+
J
Top Motifs in TFAP2C ChIP-seq
AP-2alpha (AP2)
(p-value: 1e-950)
AP-2gamma (AP2)
(p-value:1e-869)
K
Genomic annotation
14.6% promoter
10.1% TSS-proximal enhancer
5.1% exon
30.1% intron
1.2% gene tail
38.9% intergenic
L
Distance to TSS (kb)
Region-gene associations (%)
40
30
20
10
0
< -500
-500 to -50
-50 to -5
-5 to 0
0 to 5
5 to 50
50 to 500
>500
M
TetO-GATA3 at D7
Relative to GAPDH
0.15
0.10
0.05
0.00
0.015
0.010
0.005
0.000
K8
K18
GATA3
TFAP2C
GRHL2
TFAP2A
p63
KLF4
K5
K14
Dox-
Dox+
N
TetO-GRHL2 at D7
Relative to GAPDH
0.20
0.15
0.10
0.05
0.00
0.15
0.10
0.010
0.005
0.000
K8
K18
GRHL2
TFAP2C
TFAP2A
GATA3
p63
KLF4
K5
K14
Dox-
Dox+

Figure S5

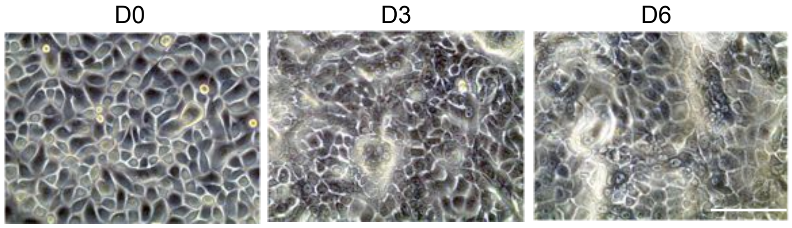
A

FACS

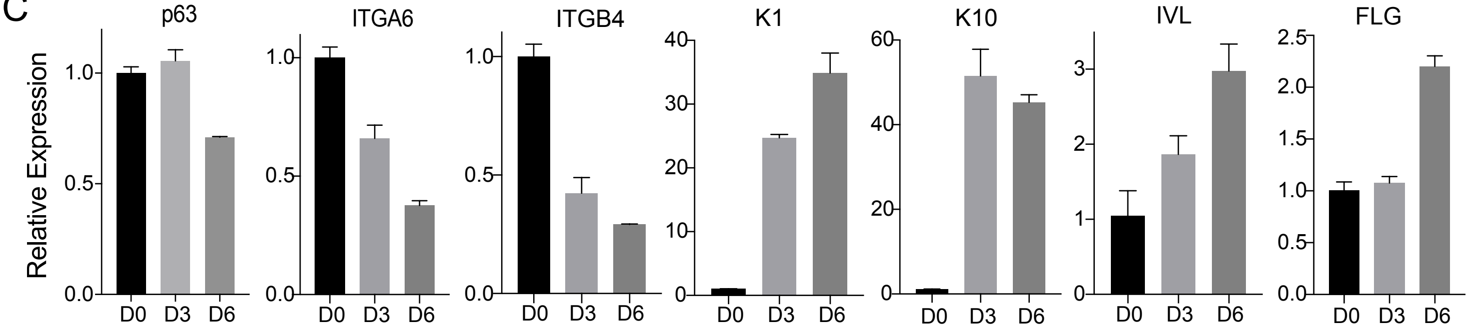


B

CaCl2 Treatment

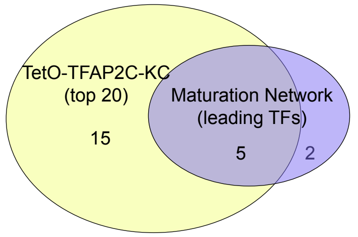


C



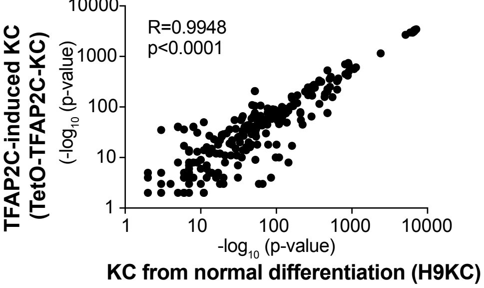
D

TF Motifs Overlap



E

Motif enrichment
TetO-TFAP2C-KC vs. H9KC



F

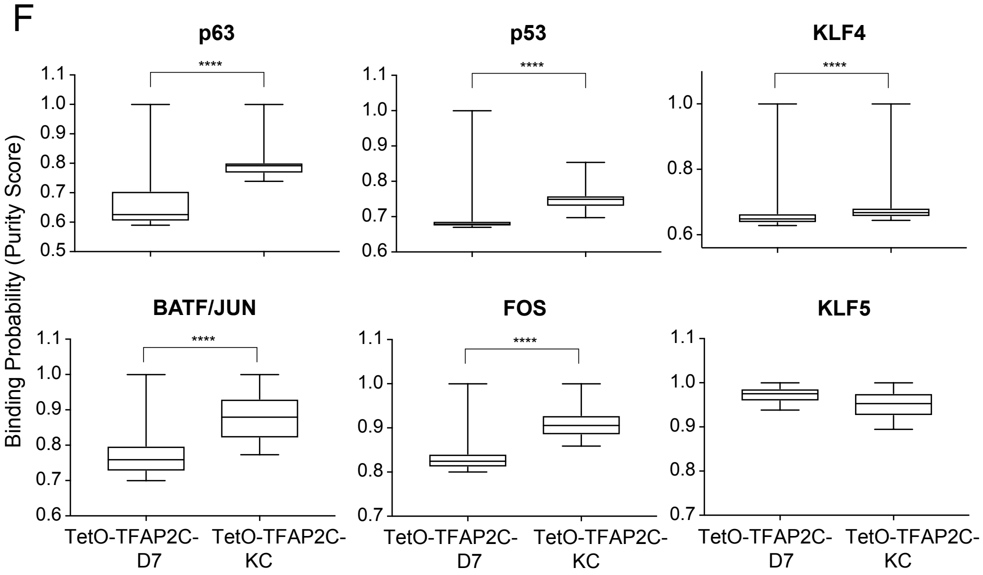


Figure S6

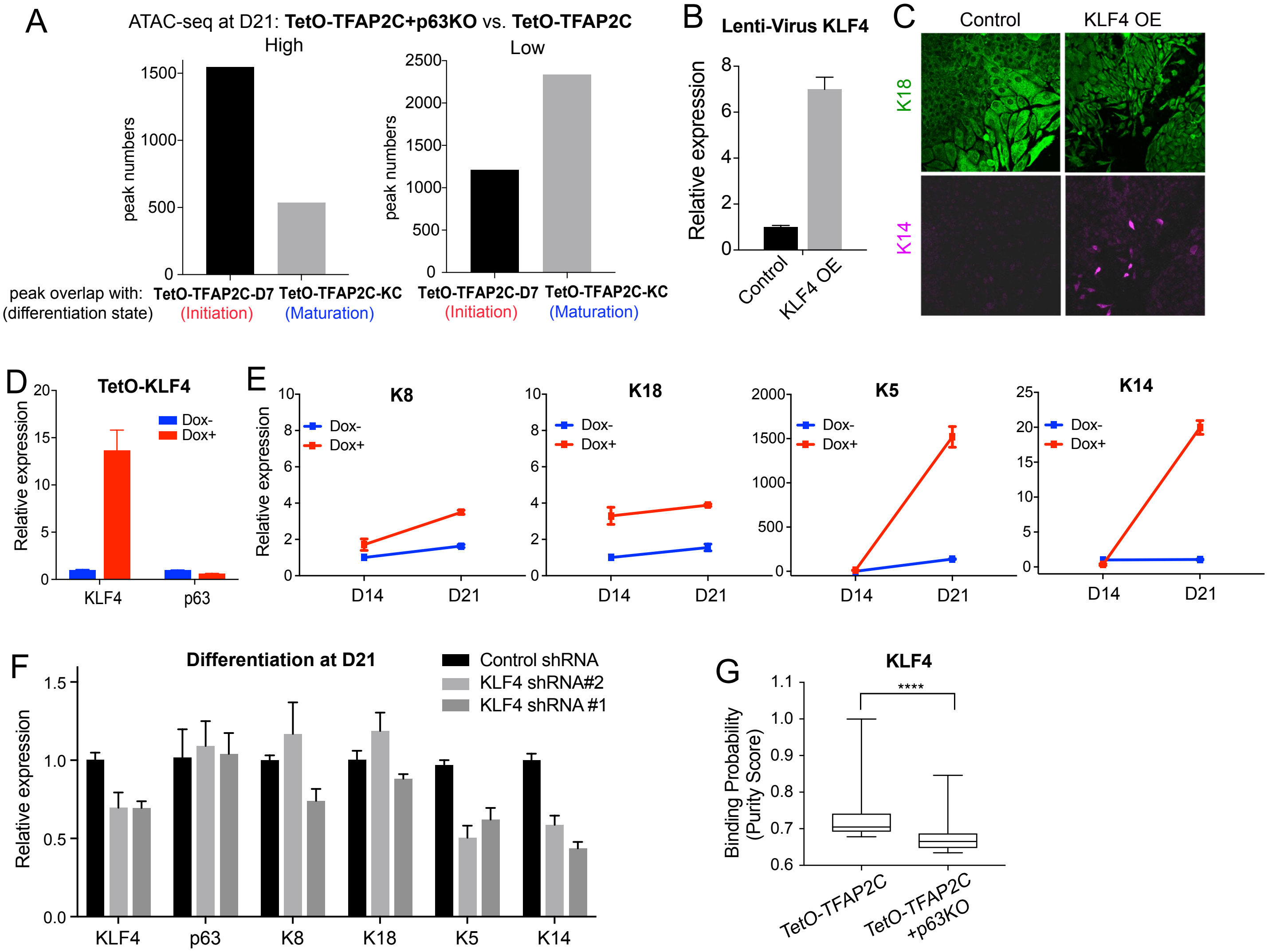
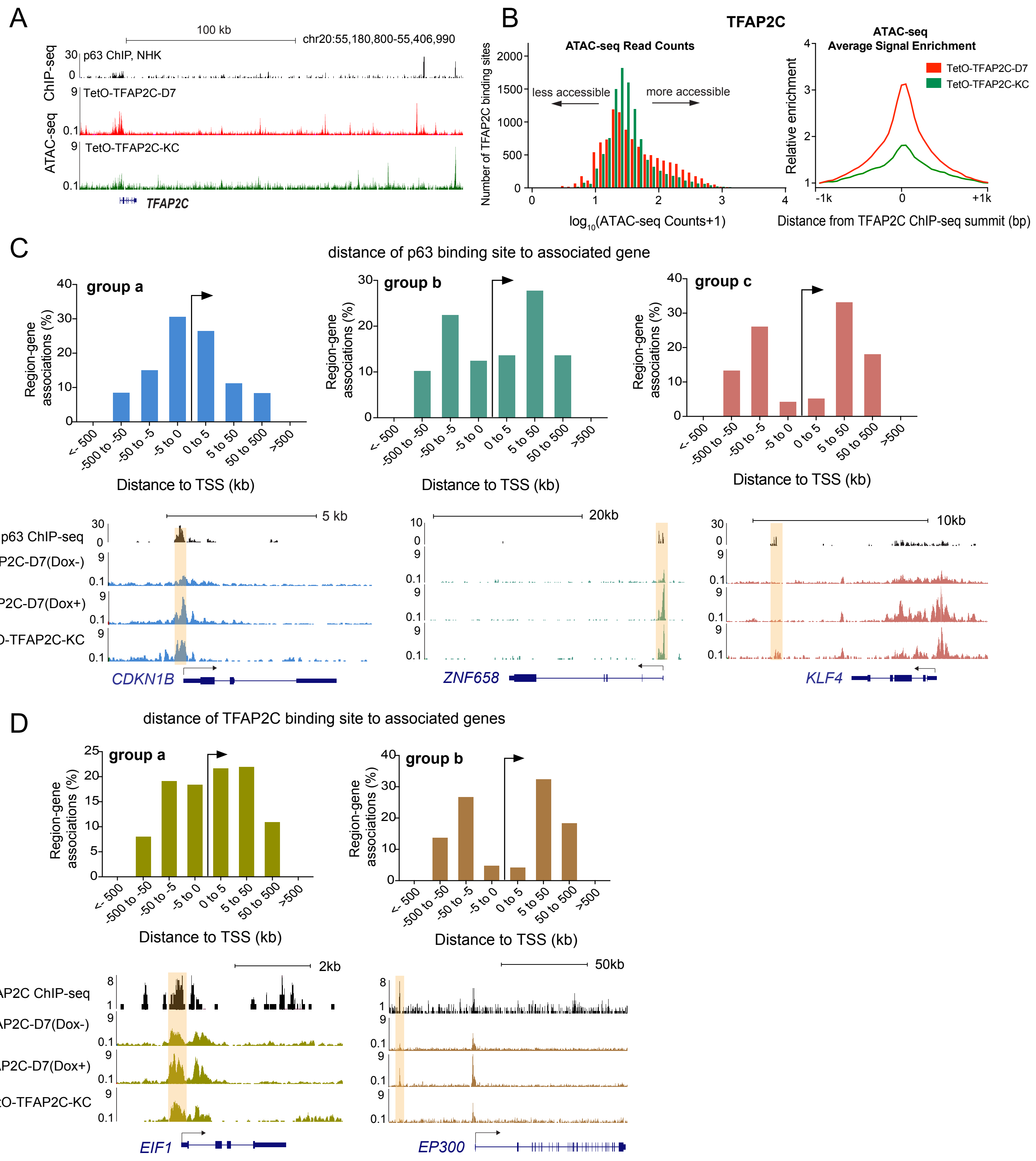

A
ATAC-seq at D21: **TetO-TFAP2C+p63KO** vs. **TetO-TFAP2C**
High
peak numbers
0
500
1000
1500
Low
0
500
1000
1500
2000
2500
peak overlap with:
(differentiation state)
TetO-TFAP2C-D7
(Initiation)
TetO-TFAP2C-KC
(Maturation)
TetO-TFAP2C-D7
(Initiation)
TetO-TFAP2C-KC
(Maturation)
B
Lenti-Virus KLF4
Relative expression
0
2
4
6
8
Control
KLF4 OE
C
Control
KLF4 OE
K18
K14
D
TetO-KLF4
Relative expression
0
5
10
15
20
Dox-
Dox+
KLF4
p63
E
Relative expression
K8
0
2
4
6
8
10
Dox-
Dox+
D14
D21
K18
0
2
4
6
8
10
Dox-
Dox+
D14
D21
K5
0
500
1000
1500
2000
Dox-
Dox+
D14
D21
K14
0
5
10
15
20
25
Dox-
Dox+
D14
D21
F
Differentiation at D21
Relative expression
0.0
0.5
1.0
1.5
Control shRNA
KLF4 shRNA#2
KLF4 shRNA #1
KLF4
p63
K8
K18
K5
K14
G
KLF4
Binding Probability
(Purity Score)
0.6
0.7
0.8
0.9
1.0
1.1
****
TetO-TFAP2C
TetO-TFAP2C
+p63KO

Figure S7


A
100 kb
chr20:55,180,800-55,406,990
ChIP-seq
30
0
p63 ChIP, NHK
ATAC-seq
9
0.1
TetO-TFAP2C-D7
9
0.1
TetO-TFAP2C-KC
TFAP2C
B
TFAP2C
ATAC-seq Read Counts
Number of TFAP2C binding sites
2000
1500
1000
500
0
less accessible
more accessible
0
1
2
3
4
log10(ATAC-seq Counts+1)
ATAC-seq
Average Signal Enrichment
TetO-TFAP2C-D7
TetO-TFAP2C-KC
Relative enrichment
4
3
2
1
-1k
0
+1k
Distance from TFAP2C ChIP-seq summit (bp)
C
distance of p63 binding site to associated gene
group a
group b
group c
Region-gene associations (%)
40
30
20
10
0
< -500
-500 to -50
-50 to -5
-5 to 0
0 to 5
5 to 50
50 to 500
>500
Distance to TSS (kb)
p63 ChIP-seq
TetO-TFAP2C-D7(Dox-)
TetO-TFAP2C-D7(Dox+)
TetO-TFAP2C-KC
5 kb
20kb
10kb
30
0
10
0
9
0.1
CDKN1B
ZNF658
KLF4
D
distance of TFAP2C binding site to associated genes
group a
group b
Region-gene associations (%)
25
20
15
10
5
0
40
30
20
10
0
< -500
-500 to -50
-50 to -5
-5 to 0
0 to 5
5 to 50
50 to 500
>500
Distance to TSS (kb)
TFAP2C ChIP-seq
TetO-TFAP2C-D7(Dox-)
TetO-TFAP2C-D7(Dox+)
TetO-TFAP2C-KC
2kb
50kb
8
1
9
0.1
EIF1
EP300