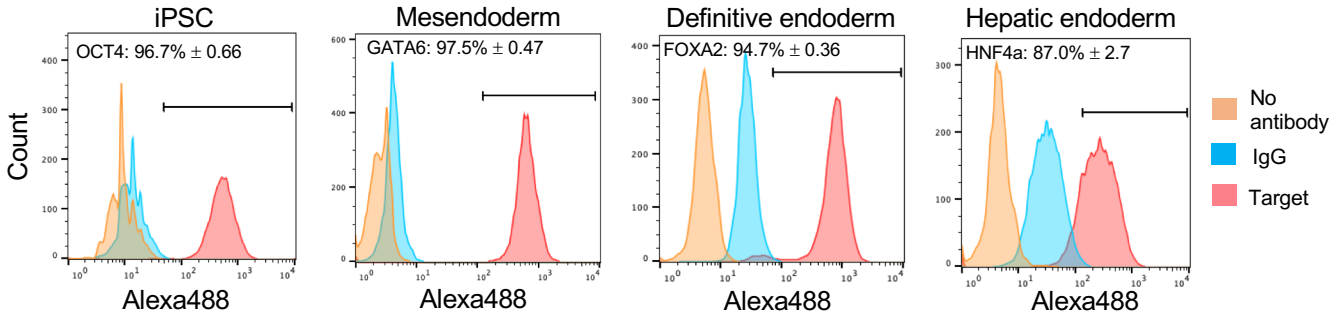
Cell Reports, Volume *35*

Supplemental information

GATA6 defines endoderm fate by controlling chromatin accessibility during differentiation of human-induced pluripotent stem cells

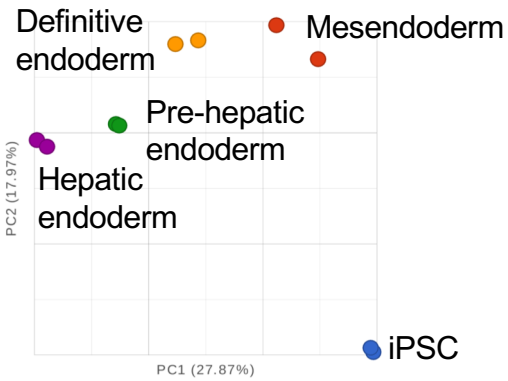
James A. Heslop, Behshad Pournasr, Jui-Tung Liu, and Stephen A. Duncan

A)

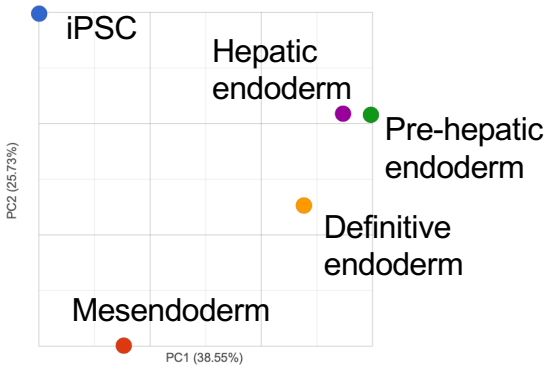


B)

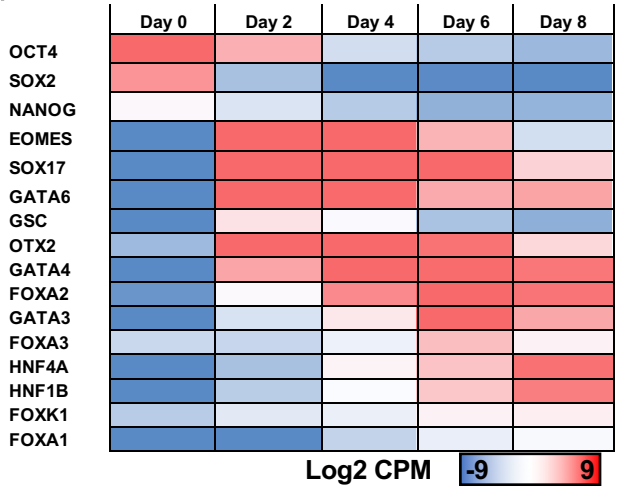
I) **RNA-seq profile**



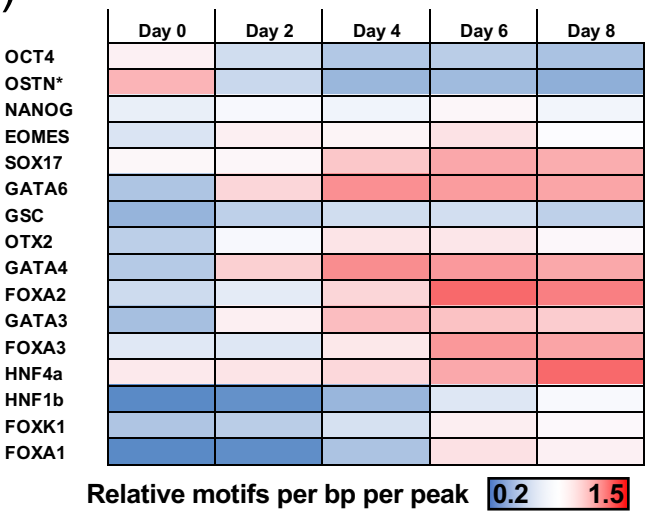
II) **ATAC-seq profile**



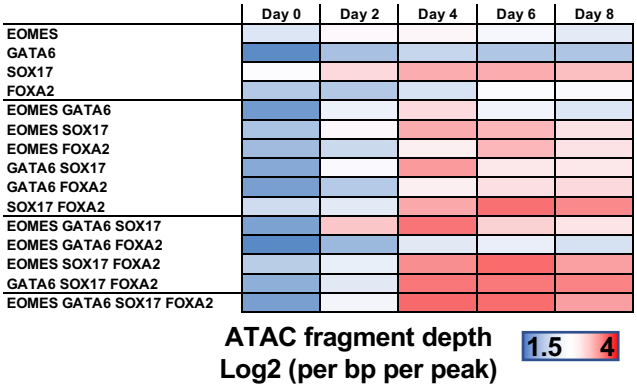
C) I) **RNA-seq**



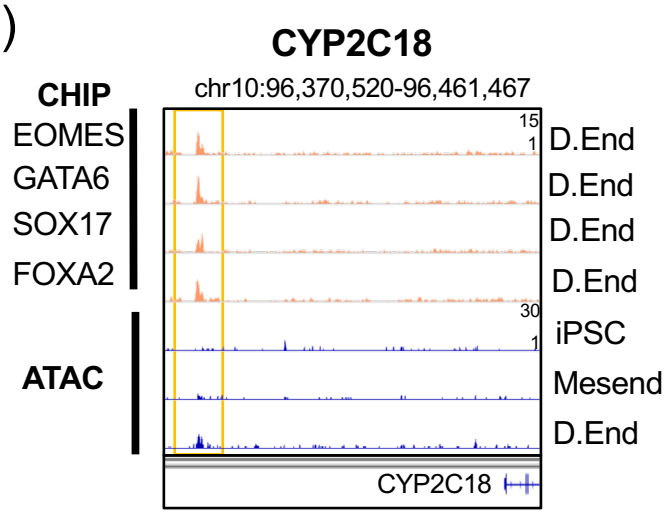
II) **ATAC-seq – motif enrichment**



D)



E)



Day 0: Pluripotent
Day 2: Mesendoderm
Day 4: Definitive endoderm
Day 6: Pre-hepatic endoderm
Day 8: Hepatic endoderm

Figure S1: Further analysis of chromatin accessibility during endoderm formation. Related to Figure 1. A) Flow cytometry analysis of stage-enriched markers during differentiation of pluripotent stem cells to hepatic endoderm, n=3 mean ± SD; B) Principal Component Analysis (PCA) of I) RNA-seq and II) ATAC-seq samples in *GATA6*$^{+/+}$ cells transitioning from pluripotent cells to hepatic endoderm cell populations; C) Comparing RNA-seq and ATAC-seq datasets I) RNA-seq gene expression profile heatmap of transcriptional regulators at 48-hour intervals. Heatmap values normalized by counts per million (CPM; mean, n=2) and shown as log2 values. Regulators listed in order of when peak gene expression is achieved; II) Motif enrichment heatmap of identified transcriptional regulators at ATAC peak center (within 200bp). Values shown relative to motif enrichment 2kb downstream from peak center. Regulators listed according to peak gene expression. *OSTN: OCT4, SOX2, TCF4 and NANOG motif; D) ATAC-seq fragment depth during endoderm formation and hepatic specification at combinations of endoderm-enriched transcription factor co-binding (Figure 1C), displayed as log2 values; E) Genome viewer representation of *CYP2C18* genomic loci using published CHIP-seq EOMES, GATA6, SOX17 and FOXA2 datasets (Tsankov et al., 2015) overlapped with ATAC-seq data during endoderm formation.

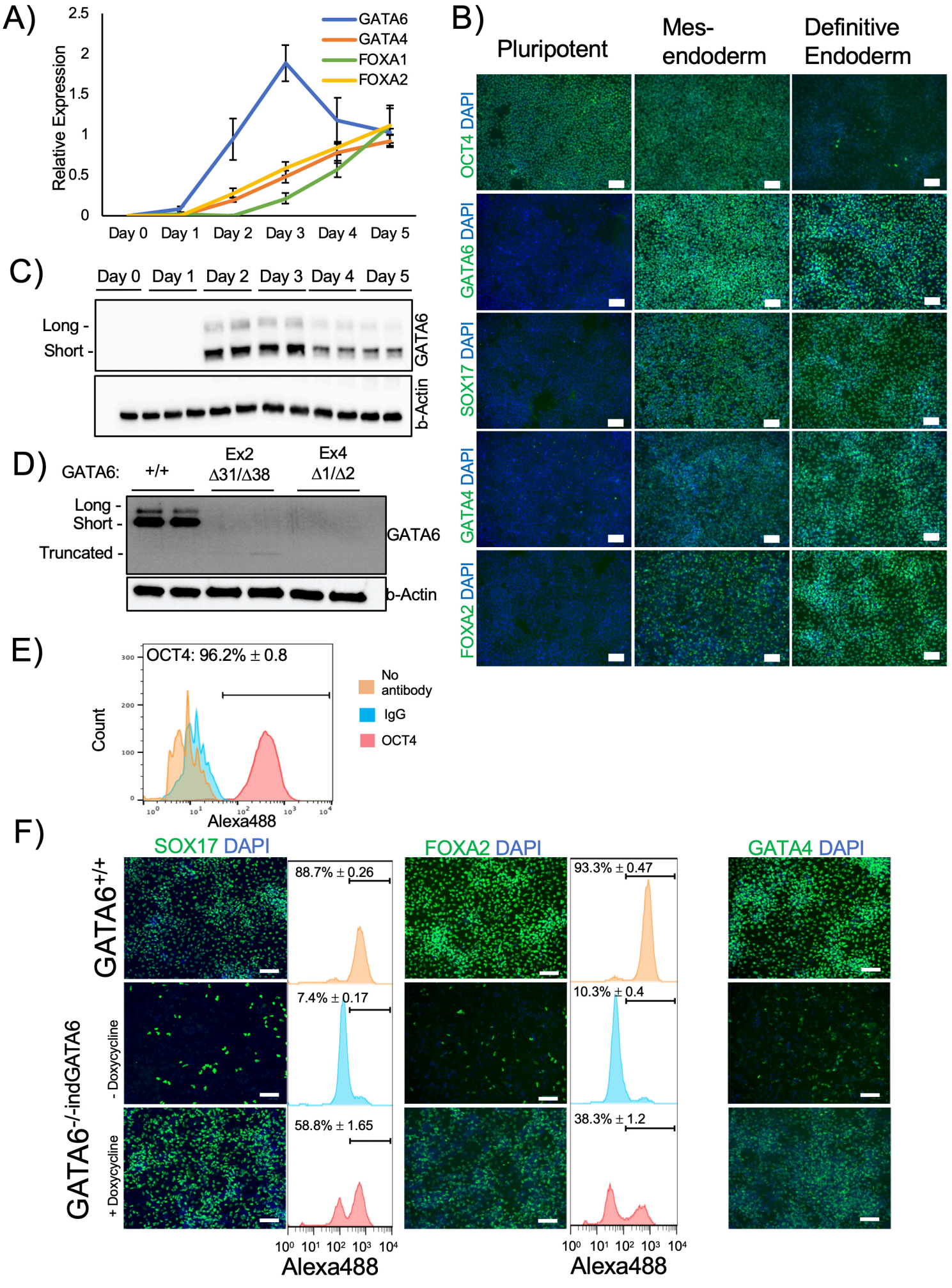
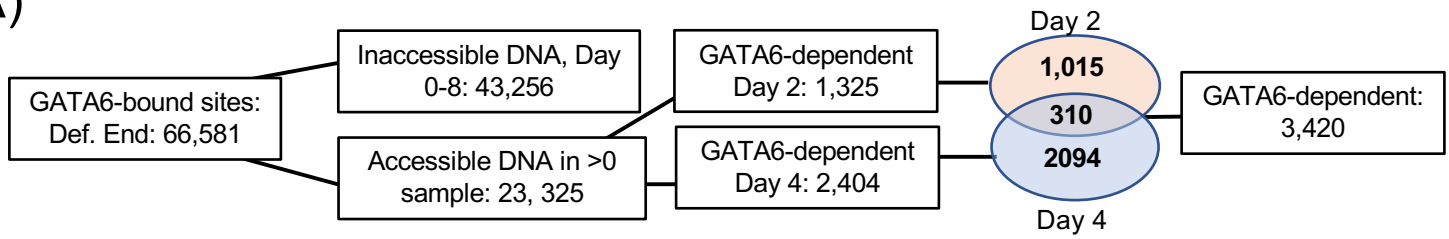
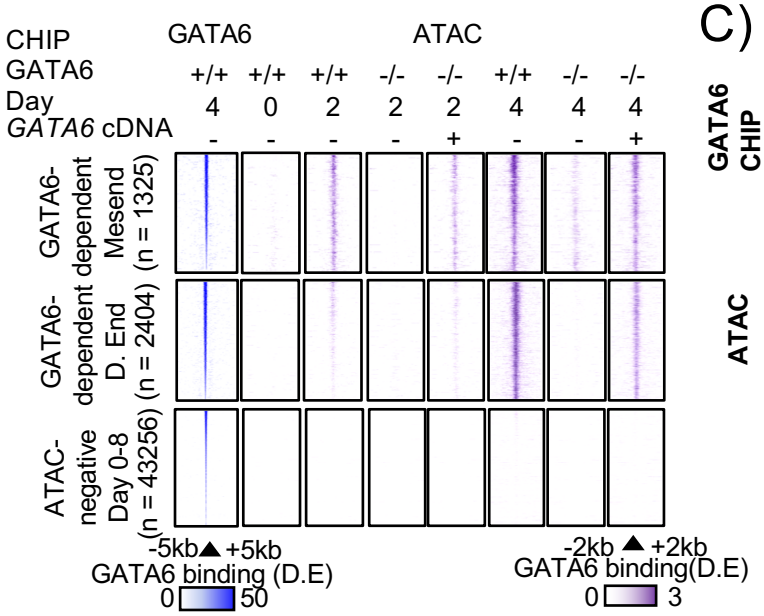

A)
Relative Expression
2.5
2
1.5
1
0.5
0
Day 0
Day 1
Day 2
Day 3
Day 4
Day 5
GATA6
GATA4
FOXA1
FOXA2
B)
Pluripotent
Mes-endoderm
Definitive Endoderm
OCT4 DAPI
GATA6 DAPI
SOX17 DAPI
GATA4 DAPI
FOXA2 DAPI
C)
Day 0
Day 1
Day 2
Day 3
Day 4
Day 5
Long -
Short -
GATA6
b-Actin
D)
GATA6:
+/+
Ex2 Δ31/Δ38
Ex4 Δ1/Δ2
Long -
Short -
Truncated -
GATA6
b-Actin
E)
OCT4: 96.2% ± 0.8
Count
Alexa488
No antibody
IgG
OCT4
F)
SOX17 DAPI
FOXA2 DAPI
GATA4 DAPI
GATA6+/+
GATA6-/-indGATA6
- Doxycycline
+ Doxycycline
88.7% ± 0.26
7.4% ± 0.17
58.8% ± 1.65
93.3% ± 0.47
10.3% ± 0.4
38.3% ± 1.2
Alexa488
Alexa488

Figure S2: Extended characterization of *GATA6*$^{-/-}$ cells during endoderm formation. Related to Figure 2. A) Graph showing RT-qPCR analysis of *GATA4, GATA6, FOXA1* and *FOXA2* mRNAs during endoderm formation. mRNA levels normalized to the housekeeping mRNA *RPL13a* (n=4 mean ± SD); B) Immunofluorescence analysis of OCT4, GATA6, GATA4, SOX17 and FOXA2 with DAPI nuclei staining in *GATA6*$^{+/+}$ cells during endoderm formation. Scale bar: 100μM; C) Western blot analysis of GATA6 and β-actin protein expression during endoderm formation in *GATA6*$^{+/+}$ cells; D) Extended exposure western blot analysis of GATA6 and β-actin in *GATA6*$^{+/+}$, *GATA6*$^{Ex2\Delta31/\Delta38}$ and *GATA6*$^{Ex4\Delta1/\Delta2}$ cell lines at day 2 of differentiation showing truncated protein in *GATA6*$^{Ex2\Delta31/\Delta38}$ cells; E) Flow cytometry analysis for OCT4 positive cells within the *GATA6*$^{Ex4\Delta1/\Delta2;ind\ GATA6}$ pluripotent cell population, n=3 mean ± SD; F) Immunofluorescence analysis of SOX17, FOXA2 and GATA4 in *GATA6*$^{+/+}$ and *GATA6*$^{Ex4\Delta1/\Delta2;ind\ GATA6}$ cells with and without doxycycline at day 4 of differentiation. Scale bar: 100μM. SOX17 and FOXA2 positive cells within the population determined by flow cytometry, n=3 mean ± SD.

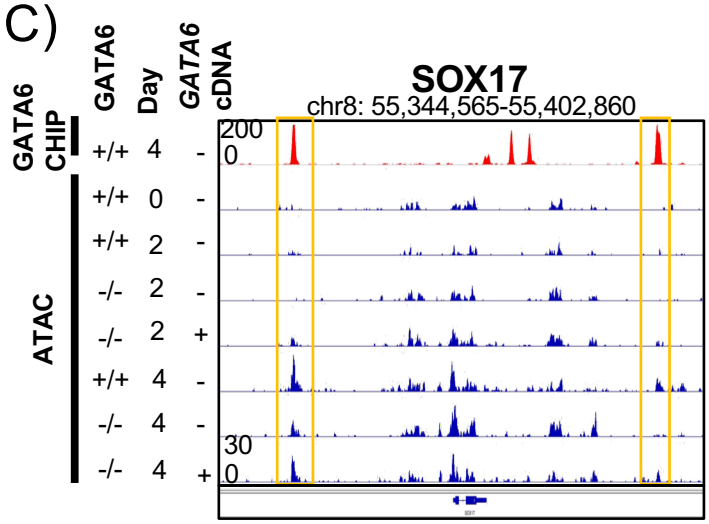
A)



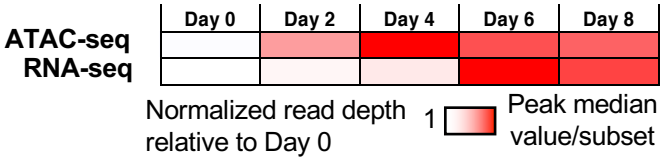
B)



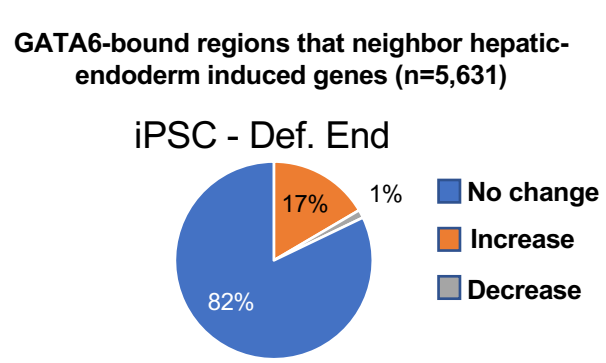
C)



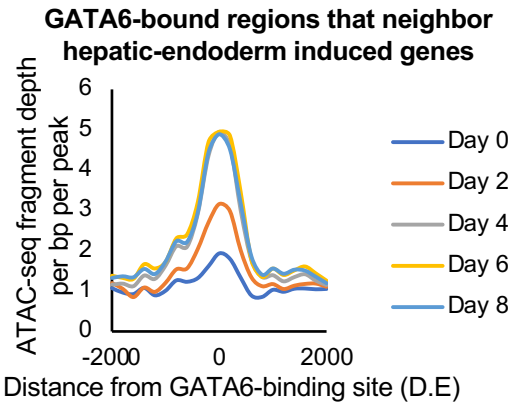
D)



E)



F)



G)

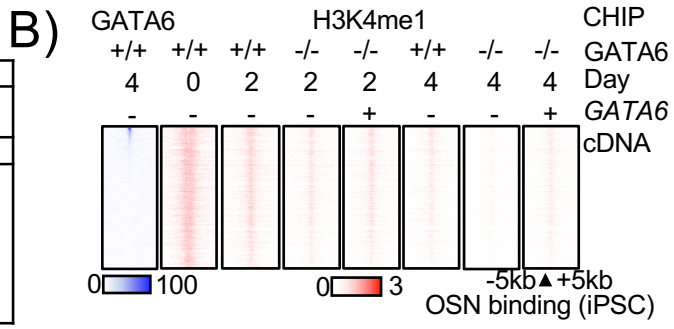
Mesendoderm			
GATA6-dependent n=2206		GATA6$^{Ex4\Delta1/\Delta2;ind\ GATA6}$ enriched n=1242	
GATA4	1.00E-78	EOMES	1.00E-52
GATA6	1.00E-77	TBOX:SMAD	1.00E-18
GATA3	1.00E-77	ZIC3	1.00E-14
SOX17	1.00E-75	LHX1	1.00E-09
EOMES	1.00E-69	TGIF1	1.00E-09
ZIC3	1.00E-12	ISL1	1.00E-07
GSC	1.00E-06	TGIF2	1.00E-07
OTX2	1.00E-05	MEIS1	1.00E-05

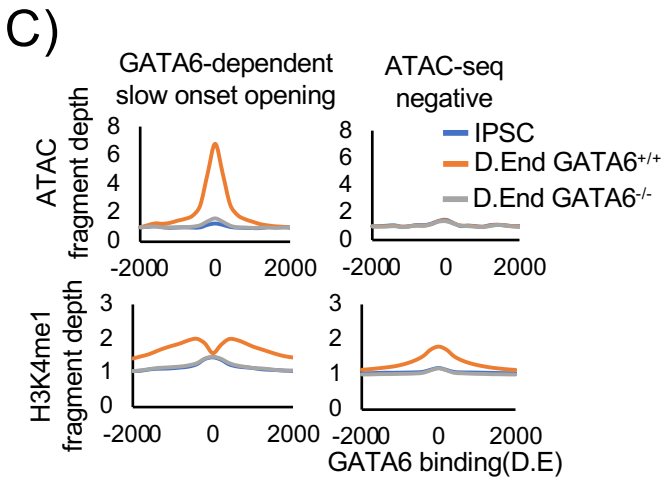
Definitive endoderm			
GATA6-dependent n=3342		GATA6$^{Ex4\Delta1/\Delta2;ind\ GATA6}$ enriched n=899	
GATA4	1.00E-149	OSTN*	1.00E-06
GATA6	1.00E-147	OCT4	1.00E-05
GATA3	1.00E-134	GATA6	1.00E-04
GSC	1.00E-28	GATA4	1.00E-04
OTX2	1.00E-26	KLF4	1.00E-04
SOX17	1.00E-23	GATA3	1.00E-04
FOXA1	1.00E-20	TCF3	1.00E-03
EOMES	1.00E-18	SP5	1.00E-03

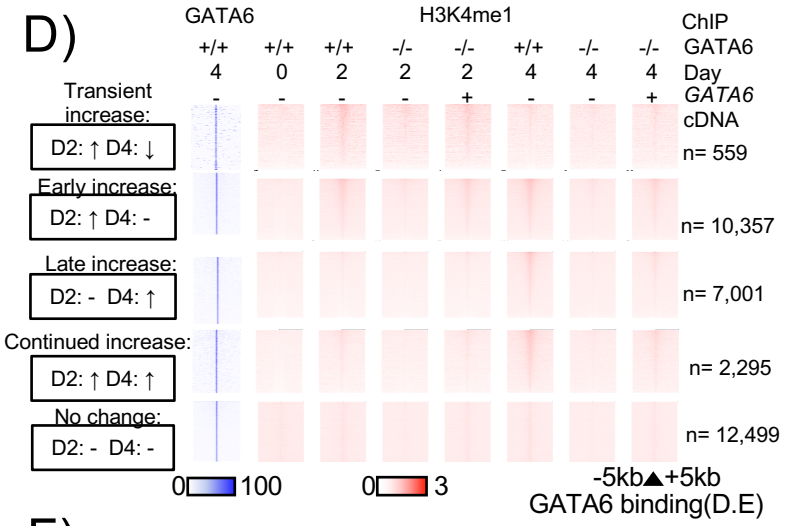
Figure S3: GATA6-dependent changes in chromatin accessibility. Related to Figure 3. A) The alignment of ATAC-seq datasets to GATA6-bound sites in definitive endoderm (Fisher et al., 2017); B) Heatmap representation of GATA6 (blue) and ATAC-seq (purple) at GATA6 binding sites during endoderm formation in *GATA6*$^{+/+}$ and *GATA6*$^{Ex4\Delta1/\Delta2;ind\ GATA6}$ pluripotent cells +/- doxycycline. Signal intensity represents read density. Heatmaps are split based on chromatin accessibility profile during endoderm formation determined in Figure S3A. *GATA6*$^{-/-}$ refers to *GATA6*$^{Ex4\Delta1/\Delta2;ind\ GATA6}$ cells; C) Genome viewer representation of ATAC-seq and GATA6-CHIP-seq datasets demonstrating GATA6-dependent chromatin remodeling at presumptive transcriptional enhancers of SOX17. *GATA6*$^{-/-}$ refers to *GATA6*$^{Ex4\Delta1/\Delta2;ind\ GATA6}$ cells; D) Heatmap representation of chromatin accessibility at GATA6-bound enhancers with GATA6-dependent accessibility and the mRNA expression of genes that neighbor the selected enhancers during the hepatic differentiation protocol. Selected sites have GATA6-dependent chromatin accessibility in definitive endoderm, and mRNA levels that are significantly induced post-definitive endoderm. Median values of the ATAC-seq and RNA-seq (ATAC seq: Tag density. RNA-seq: counts per million) datasets displayed relative to day 0 values, n=180 regions relating to Figure 3C; E) Pie chart depicting all GATA6-bound regions that neighbor genes that are significantly induced during hepatic specification. Regions with 2-fold change difference in ATAC-seq fragment between iPSC (Day 0) and definitive endoderm (Day 4) populations identified as regions that increased or decreased in accessibility; F) Graph showing ATAC-seq fragment depth at GATA6-bound regions that neighbor genes induced during hepatic specification between Day 0 (iPSCs) and Day 8 (hepatic endoderm). Graph centered on GATA6 binding sites in definitive endoderm, n=5,631 regions; G) Motifs enriched at all sites of GATA6-dependent chromatin accessibility and at sites that are increased in the absence of GATA6 in *GATA6*$^{Ex4\Delta1/\Delta2;ind\ GATA6}$ cells in mesendoderm and definitive endoderm. P values: hypergeometric enrichment test using HOMER motif analysis with default background. *OSTN: OCT4, SOX2, TCF4, NANOG motif

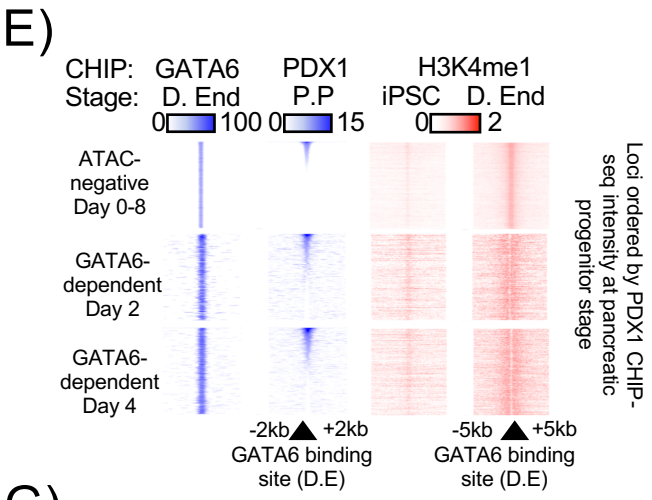
A)

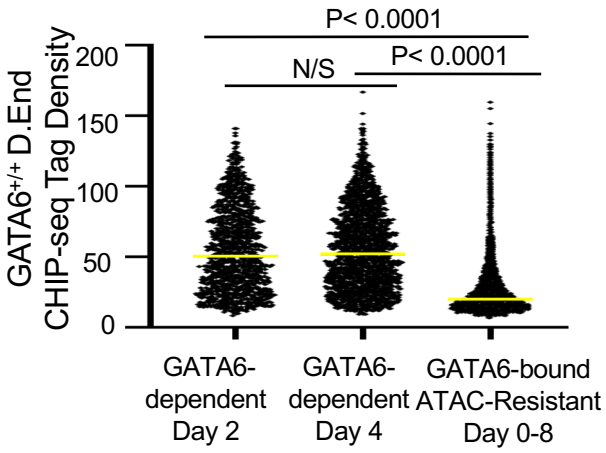
IPSC (Day 0) – Definitive Endoderm (Day 4)			
Increased H3K4me1 n=3728		Decreased H3K4me1 n=3085	
Motif	**P-value**	**Motif**	**P-value**
GATA6	1e-35	OCT4-SOX2-TCF-NANOG	1e-67
GATA4	1e-35	OCT4	1e-27
GATA3	1e-28	BRN1	1e-18
FOXP1	1e-8	FOSL2	1e-8
SOX6	1e-7	FRA1	1e-8

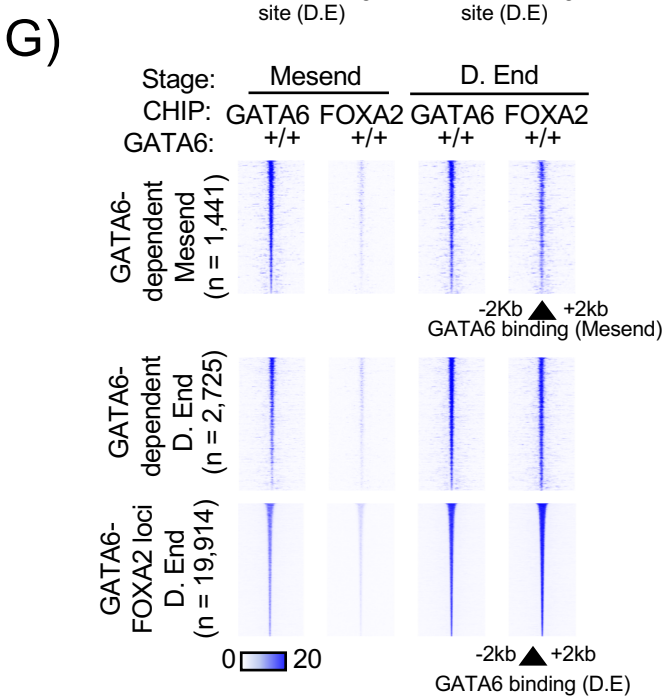












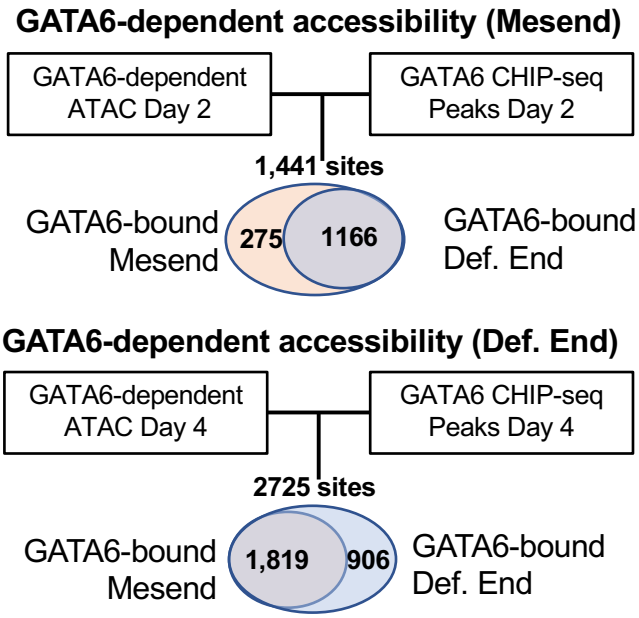
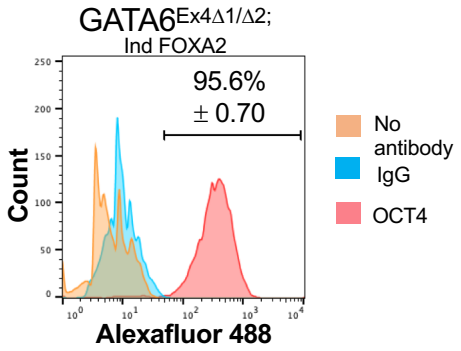
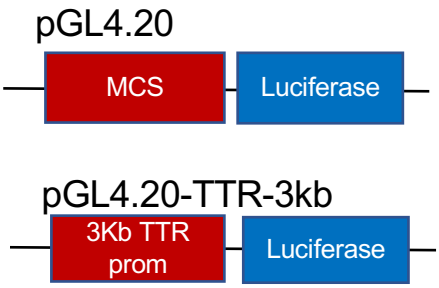


Figure S4: GATA6 binding induces H3K4me1 accumulation and most commonly binds in mesendoderm at GATA6-dependent chromatin remodeled loci. Related to Figures 3 and 4. A) Transcription factor binding motifs enriched at regions that increase or decrease in H3K4me1 accumulation during definitive endoderm formation. Differential H3K4me1 peaks determined as FDR<0.01 and fold change >2; P values: hypergeometric enrichment test using HOMER motif analysis with default background; B) Heatmap depicting GATA6-independent loss of H3K4me1 (red) at regions of OCT4-SOX2-NANOG binding during definitive endoderm formation. OCT4, SOX2 NANOG binding sites determined from previously published CHIP-seq datasets (Tsankov et al., 2015); C) Histogram depicting ATAC-seq and H3K4me1 fragment depths at GATA6 binding sites (definitive endoderm) with GATA6-dependent chromatin accessibility (n=2,404) or that remain ATAC-negative (n=43,256) in *GATA6*$^{+/+}$ and *GATA6*$^{Ex4\Delta1/\Delta2;ind\ GATA6}$ cells; D) Temporal profile of H3K4me1 (red) during definitive endoderm formation at GATA6 binding loci (blue) that remain ATAC-negative between day 0 and day 8 in *GATA6*$^{+/+}$ and *GATA6*$^{Ex4\Delta1/\Delta2;ind\ GATA6}$ cells +/- doxycycline. Differential H3K4me1 tag density at GATA6 bound loci determined as being >1.5 log2 fold change; E) Heatmap depicting published PDX1 CHIP-seq density at pancreatic progenitor (P.P) stage (Lee et al., 2019) aligned with GATA6-bound loci in definitive endoderm demonstrating PDX1 engagement at subsets of GATA6-dependent patterning during early pancreatic specification and that these regions increase in H3K4me1 deposition (red). See Figure S3A for no. of sites in each subset; F) GATA6-CHIP-seq tag density scatter graph at regions of GATA6-dependent chromatin remodeling and ATAC-negative sites. Mann-Whitney U Test, P-values displayed on graph, median value shown as yellow line. See Figure S3A for no. of sites in each subset; G) Heatmaps depicting GATA6 and FOXA2 CHIP-seq intensity at loci of GATA6-FOXA2 co-binding, and GATA6-dependent chromatin remodeling at day 2 (mesendoderm) and day 4 (definitive endoderm) in *GATA6*$^{+/+}$ definitive endoderm; H) Number of alignments between GATA6-dependently accessible chromatin loci in mesendoderm (upper panel) and definitive endoderm (lower panel) with the corresponding GATA6 CHIP-seq peaks from *GATA6*$^{+/+}$ mesendoderm and definitive endoderm cell populations. Venn diagram depicting the overlap of mesendoderm and definitive endoderm stage GATA6 CHIP-seq peaks with regions of rapid and slow onset GATA6-dependent chromatin remodeling.

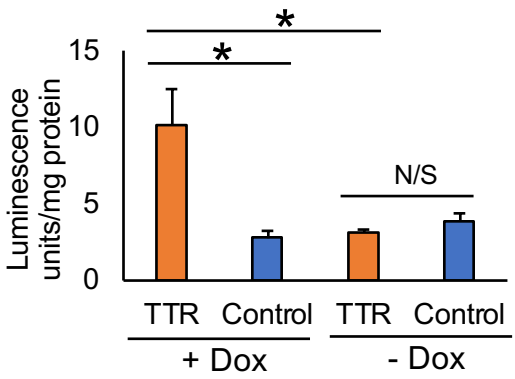
A)



B)



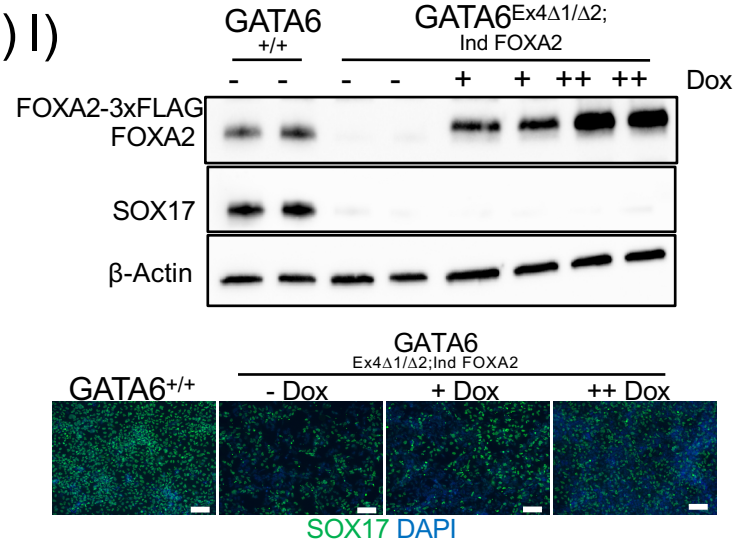
C)



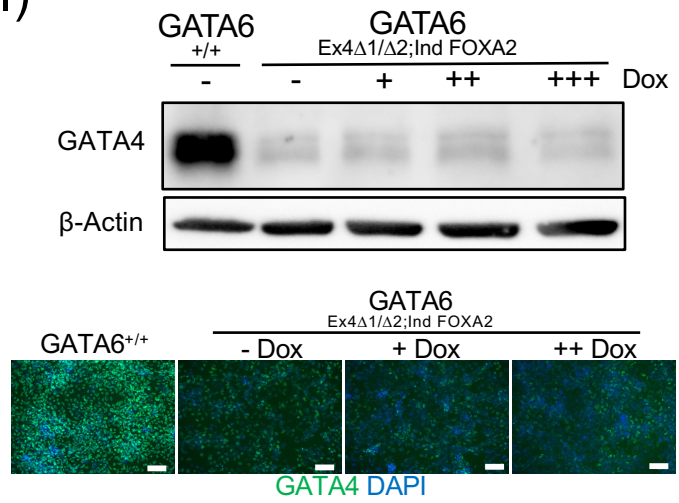
D)

Motif Name	P-value
ZIC	1e-5
NF1-halfsite	1e-4
MAFA	1e-4
OCT4	1e-4
GATA3	1e-4
JUNB/AP1	1e-3

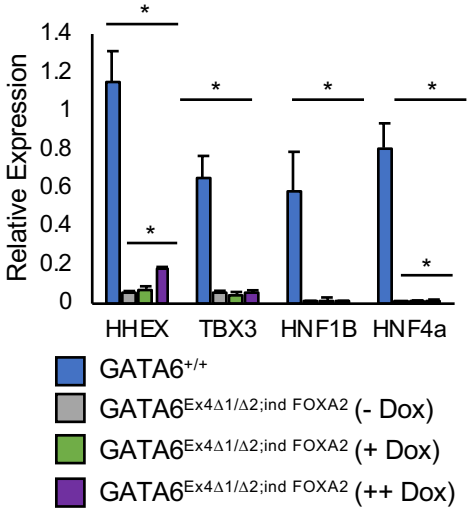
E) I)



II)



F)



G)

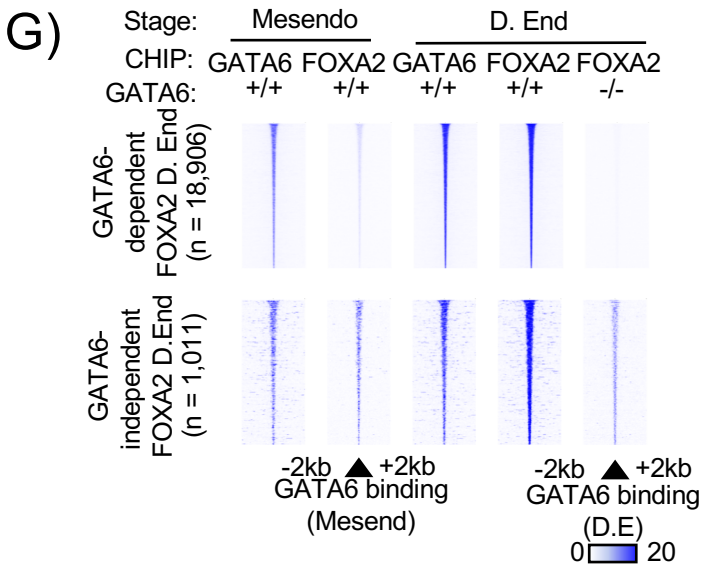


Figure S5: Expression of doxycycline inducible FOXA2-3xflag cDNA in GATA6$^{-/-}$ cells. Related to Figure 5. A) Flow cytometry analysis of *GATA6*$^{Ex4\Delta1/\Delta2;ind\ FOXA2}$ pluripotent cells positive for OCT4 expression, n=3 mean ± SD; B) Plasmid map of pGL4.20 Promega luciferase vector with and without inserted 3kb TTR promoter region; C) Graph showing analysis of luminescence from the pGL4.20-TTR-3kb or pGL4.20 luciferase vectors. Luminescence units normalized with total protein content, n=3 mean ± SD. Student's t-test: * = p<0.05; D) Transcription factor motifs enriched at sites of increased chromatin accessibility in *GATA6*$^{Ex4\Delta1/\Delta2;ind\ FOXA2}$ (++dox) cells when compared to *GATA6*$^{Ex4\Delta1/\Delta2;ind\ FOXA2}$(-dox) cells. P values: hypergeometric enrichment test using HOMER motif analysis with default background, n=1,927 regions; E) Western blot and immunofluorescence analysis of *GATA6*$^{+/+}$ and *GATA6*$^{Ex4\Delta1/\Delta2;ind\ FOXA2}$ cells with increasing doxycycline supplementation at day 4 of differentiation for I) FOXA2, SOX17, β-actin and II) GATA4 and β-actin protein expression. Immunofluorescence scale bar: 100μM; F) Graph showing RT-qPCR analysis of foregut-enriched mRNAs at day 4 of differentiation in *GATA6*$^{+/+}$ and *GATA6*$^{Ex4\Delta1/\Delta2;ind\ FOXA2}$ cells with increasing doxycycline supplementation. Levels of mRNA were normalized to the housekeeping mRNA *RPL13a*. n=3, mean ± SD. Student's t-test: * = p<0.05; G) Heatmap depicting FOXA2 binding at regions of co-binding with GATA6 in *GATA6*$^{+/+}$ definitive endoderm that are GATA6-dependent and GATA6-independent. GATA6 and FOXA2 CHIP-seq intensity displayed from *GATA6*$^{+/+}$ and *GATA6*$^{Ex4\Delta1/\Delta2;indFOXA2}$ (++dox) cells during definitive endoderm formation.

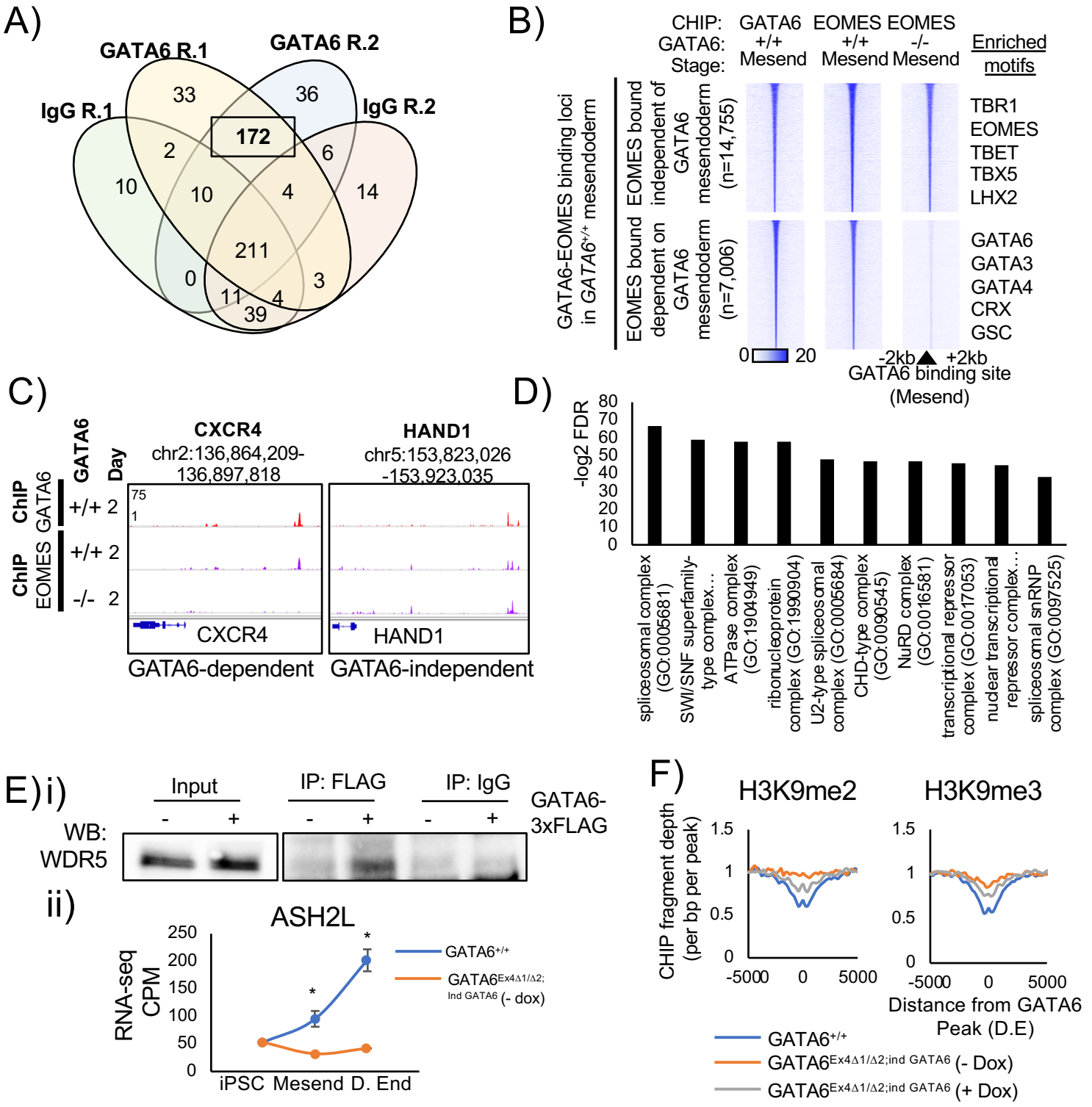

A)
GATA6 R.1
GATA6 R.2
IgG R.1
IgG R.2
33
36
172
2
6
10
10
4
14
211
0
3
11
4
39
B)
CHIP:
GATA6:
Stage:
GATA6 +/+ Mesend
EOMES +/+ Mesend
EOMES -/- Mesend
Enriched motifs
GATA6-EOMES binding loci in GATA6+/+ mesendoderm
EOMES bound independent of GATA6 mesendoderm (n=14,755)
EOMES bound dependent on GATA6 mesendoderm (n=7,006)
TBR1
EOMES
TBET
TBX5
LHX2
GATA6
GATA3
GATA4
CRX
GSC
0
20
-2kb
+2kb
GATA6 binding site (Mesend)
C)
ChIP GATA6
ChIP EOMES
GATA6
Day
+/+ 2
+/+ 2
-/- 2
CXCR4
chr2:136,864,209-136,897,818
HAND1
chr5:153,823,026-153,923,035
75
1
CXCR4
HAND1
GATA6-dependent
GATA6-independent
D)
-log2 FDR
80
70
60
50
40
30
20
10
0
spliceosomal complex (GO:0005681)
SWI/SNF superfamily-type complex... ATPase complex (GO:1904949)
ribonucleoprotein complex (GO:1990904)
U2-type spliceosomal complex (GO:0005684)
CHD-type complex (GO:0090545)
NuRD complex (GO:0016581)
transcriptional repressor complex (GO:0017053)
nuclear transcriptional repressor complex...
spliceosomal snRNP complex (GO:0097525)
E) i)
Input
IP: FLAG
IP: IgG
- + - + - +
GATA6-3xFLAG
WB: WDR5
ii)
ASH2L
RNA-seq CPM
250
200
150
100
50
0
GATA6+/+
GATA6Ex4Δ1/Δ2; Ind GATA6 (- dox)
iPSC Mesend D. End
F)
H3K9me2
H3K9me3
CHIP fragment depth (per bp per peak)
1.5
1
0.5
0
-5000 0 5000
Distance from GATA6 Peak (D.E)
GATA6+/+
GATA6Ex4Δ1/Δ2;ind GATA6 (- Dox)
GATA6Ex4Δ1/Δ2;ind GATA6 (+ Dox)
G)
H3K4me2
H3K27ac
CHIP fragment depth (per bp per peak)
6
4
2
0
10
5
0
-2000 0 2000
Distance from GATA6 Peak (D.E)
IPSC
Mesend
Def. End
Hep. End
H)
GATA6 bound, transiently accessible loci in definitive endoderm vs loci that retain accessibility
All loci (n=691)
GATA6-dependent loci (n=228)
Name motif P-value
BORIS 1e-7
TBET 1e-5
E2A 1e-5
ZEB1 1e-4
Name motif P-value
NKX2.1 1e-3
PRDM14 1e-2
CDX2 1e-2
PAX8 1e-2

Figure S6: Extended investigation of GATA6-interacting proteins. Related to Figure 6. A) Venn diagram depicting the proteins which were identified GATA6 and IgG immunoprecipitations, n=2 replicates; B) Heatmap depicting GATA6 and EOMES CHIP-seq density at regions of GATA6-dependent and GATA6-independent EOMES binding in *GATA6*$^{+/+}$ and *GATA6*$^{Ex4\Delta1/\Delta2;ind\ GATA6}$ (-dox) mesendoderm cell populations (day 2). Top 5 motifs uniquely overrepresented in each subset of GATA6-EOMES peaks. P values: hypergeometric enrichment test using HOMER motif analysis. The alternate peak subset used as the background dataset to remove motifs equally enriched within both datasets; C) Genome viewer representation of GATA6 and EOMES CHIP-seq data in mesendoderm cell populations at the *CXCR4* and *HAND1* gene loci in *GATA6*$^{+/+}$ and *GATA6*$^{Ex4\Delta1/\Delta2;ind\ GATA6}$ cells without doxycycline; *GATA6*$^{-/-}$ refers to *GATA6*$^{Ex4\Delta1/\Delta2;ind\ GATA6}$ cells; D) GO analysis for the complexes enriched in the GATA6 interactome; E) Investigation of GATA6-H3K4 methyltransferase interactions I) co-immunoprecipitation validation for WDR5 protein II) RNA-seq gene expression data of *ASH2L,* a H3K4 methyltransferase-associated factor, normalized by counts per million, n=2 mean ± SD. *:p<0.05 using GSA analysis; F) CHIP-seq fragment depth of H3K9me2 (KDM1A target) and H3K9me3 (KDM1A refractive) fragment depth profiles at regions of slow onset (day 4) GATA6 dependent chromatin remodeling (n=2404 sites) reveal similar profiles. Graphs centered on GATA6 binding site in definitive endoderm; G) Temporal profile of equivalent stage H3K27ac and H3K4me2 published CHIP-seq (Loh et al., 2014) fragment depth during the early stages of liver development at sites of GATA6-dependent chromatin remodeling (n=2,404 sites); H) Motifs enriched at GATA6-bound loci that become accessible during endoderm formation, then reduce in accessibility during hepatic specification. P values: hypergeometric enrichment test using HOMER motif analysis. Regions that maintain accessibility during hepatic specification used as the background dataset to remove equally enriched motifs from the analysis.

Day 0-2			
Increased		Decreased	
No. of Peaks	Transcription factor	No. of Peaks	Transcription factor
243	EOMES FOXA2 GATA6 SOX17	210	NANOG
126	GATA6	103	NANOG OCT4 SALL4 SOX2
99	EOMES GATA6 SOX17	83	NANOG OCT4
99	GATA6 SOX17	63	NANOG OCT4 SOX2
94	EOMES	47	OCT4
69	EOMES FOXA2 GATA6	40	NANOG SALL4
56	FOXA2	34	FOXA2
52	FOXA2 GATA6 SOX17	32	FOXA2 NANOG
49	EOMES GATA6	27	NANOG OCT4 SALL4
47	EOMES FOXA2 GATA6 NANOG SOX17	25	HNF4A
Day 2-4			
Increased		Decreased	
No. of Peaks	Transcription factor	No. of Peaks	Transcription factor
293	EOMES FOXA2 GATA6 SOX17	34	NANOG
263	FOXA2	16	NANOG OCT4 SALL4 SOX2
140	FOXA2 GATA6 SOX17	15	OCT4 SALL4
103	EOMES FOXA2	13	FOXA2
83	FOXA2 GATA6	12	NANOG OCT4
76	EOMES FOXA2 GATA6	11	NANOG OCT4 SALL4 SOX2 T
75	FOXA2 SOX17	8	FOXA2 NANOG
74	EOMES FOXA2 SOX17	6	NANOG OCT4 SOX2
73	SOX17	6	SMAD2/3
72	GATA6 SOX17	5	EOMES
Day 4-6			
Increased		Decreased	
No. of Peaks	Transcription factor	No. of Peaks	Transcription factor
205	FOXA2	64	EOMES FOXA2 GATA6 SOX17
36	EOMES FOXA2	43	GATA6
28	HNF4A	41	GATA6 SOX17
23	FOXA2 HNF4A	37	EOMES GATA6 SOX17
20	FOXA2 SOX17	25	NANOG
17	FOXA2 NANOG	23	FOXA2 GATA6 SOX17
16	HNF1B	20	EOMES FOXA2 GATA6
16	NANOG	19	FOXA2
14	EOMES FOXA2 GATA6 SOX17	17	EOMES GATA6
14	EOMES FOXA2 SOX17	16	SOX17
Day 6-8			
Increased		Decreased	
No. of Peaks	Transcription factor	No. of Peaks	Transcription factor
77	HNF4A	40	FOXA2
54	FOXA2	33	EOMES FOXA2 GATA6 SOX17
26	NANOG	15	EOMES FOXA2
11	FOXA2 HNF4A	14	EOMES FOXA2 GATA6
8	HNF1B	13	EOMES FOXA2 SOX17
7	NANOG OCT4	9	EOMES FOXA2 GATA6 NANOG SOX17
6	EOMES FOXA2	8	SOX17
6	FOXA2 NANOG	7	NANOG
6	GATA6	6	EOMES
5	FOXA2 GATA6	6	EOMES GATA6 SOX17

Table S2: Overlapping chromatin accessibility with transcription factor co-binding.
Overlap of transcription factor binding sites identified by CHIP-seq with differential chromatin accessibility at 48-hour intervals during endoderm formation and specification (Fisher et al., 2017; Tsankov et al., 2015). Related to Figure 1.

GATA6-dependent chromatin opening (mesendoderm; n= 1,325) vs ATAC-negative (n=43,246)		GATA6-dependent chromatin opening (definitive endoderm; n=2,404) vs ATAC-negative (n=43,246)	
Motif name	**P-value**	**Motif name**	**P-value**
GATA1	1.00E-05	*FOXA2*	1.00E-12
SOX17	1.00E-05	*FOXA1*	1.00E-12
SOX6	1.00E-04	*SOX17*	1.00E-12
TBOX:SMAD	1.00E-04	*FOXA3*	1.00E-11
GATA2	1.00E-04	*SOX2*	1.00E-09
TBX5	1.00E-03	*FOX:EBOX*	1.00E-09
SOX15	1.00E-03	*FOXM1*	1.00E-09
OCT6	1.00E-03	*SOX15*	1.00E-08
BRN1	1.00E-03	*SOX6*	1.00E-08
EBF	1.00E-03	*SOX3*	1.00E-08
GATA6-bound chromatin opening (hepatic specification; n=259) vs ATAC-negative (n=43,246)		**ATAC-negative (n=43,246) vs All GATA6-dependent chromatin opening (n=3,988)**	
Motif name	**P-value**	**Motif name**	**P-value**
FOX:EBOX	1.00E-08	*GATA3, DR4*	*1.00E-548*
FOXA1	1.00E-06	*PDX1*	*1.00E-141*
FOXM1	1.00E-05	*GATA4*	*1.00E-38*
FOXA3	1.00E-04	*RUNX2*	*1.00E-30*
FOXK2	1.00E-04	*MITF*	*1.00E-30*
FOXA2	1.00E-04	*HOXA2*	*1.00E-28*
HNF1B	1.00E-03	*ZNF675*	*1.00E-25*
FOXF1	1.00E-03	*HOXB4*	*1.00E-25*
HNF4A	1.00E-02	*IRF8*	*1.00E-24*
SOX9	1.00E-02	*LXRE*	*1.00E-23*

Table S4: Motif enrichment in different ATAC-subgroups during endoderm formation at sites of GATA6 binding. Motifs co-enriched at regions GATA6-dependent chromatin accessibility using GATA6-bound ATAC-negative sites as background to remove motifs equally present in both groups. Analysis reversed for ATAC-negative subset. P values: hypergeometric enrichment test. Related to Figure 4.

iPSC- Definitive endoderm		Definitive endoderm – Hepatic endoderm							
Decrease vs Increase		Day 6 Increase vs Decrease		Day 6 Decrease vs Increase		Day 8 Increase vs Down		Day 8 Decrease vs Increase	
Motif Name	**P-value**	**Motif Name**	**P-value**	**Motif Name**	**P-value**	**Motif Name**	**P-value**	**Motif Name**	**P-value**
OCT4-SOX2-TCF-NANOG	1e-5	FOX:EBOX	1e-18	Unknown ESC element	1e-22	HNF1B	1e-7	TBET	1e-10
BORIS	1e-2	FOXM1	1e-17	ZIC3	1e-12	PR	1e-5	PAX6	1e-5
Etv2	1e-2	FOXA1	1e-16	CHOP	1e-9	AP-2alpha	1e-4	EOMES	1e-5
NF1-halfsite	1e-2	FOXA2	1e-12	PITX1:EBOX	1e-9	SOX10	1e-3	GATA:SCL	1e-4
SPDEF	1e-2	FOXL2	1e-11	KLF5	1e-9	HOXB13	1e-3	MAZ	1e-3

Table S6: Motif enrichment at GATA6-bound regions that decrease in accessibility. Motifs enriched at regions of GATA6-bound decreases in accessibility during and post definitive endoderm formation. Sites that increase in accessibility used as background to remove motifs equally present in both groups. Analysis reversed for assessment of GATA6-bound sites that increase in accessibility during the same time period. P values: hypergeometric enrichment test. Related to Figure 6.

CRISPR Ca9 gRNAs and Primers			
GATA6 Exon	Sequence Type	Fwd 5'-3'	Rev 5'-3'
Exon 2	CRISPR/Cas9 gRNA	CACCGAGTGGGCCAGCCAACCACGC	AAACGCGTGGTTGGCTGGCCCACTC
	INDEL PCR primers	TCCATGCTGCCCGGCCTAC	TGCCGTATGGAGGGCTGT
	Sequencing primers	AGATGTACCAGACCCTCGCC	CCCACGTAGGGCGAGTAGG
Exon 4	CRISPR/Cas9 gRNA	CACCGTTATGGCGCAGAAACGCCG	AAACCGGCGTTTCTGCGCCATAAC
	INDEL PCR primers	ACCACCTTATGGCGCAGAAA	GCAGAATACATGGCATACCCC
	Sequencing primers	ACATACTTGTTGATGACAGGGACA	CGTTTGCAATAGTTCAACTGG

qPCR Primers and Probes			
Gene	Fwd (5'-3')	Probe (5'-3')	Rev (5'-3')
GATA6	TTCGTTTCCTGGTTTGAATTCC	TCATAGCAAGTGGTCTGGGCACC	TGCAATGCTTGTGGACTCTAC
GATA4	AGGCGTTGCACAGATAGTG	CATAGCCCCACAGTTGACACACTCT	CGACACCCCAATCTCGAT
FOXA2	GAGCGGTGAAGATGGAAGG	CAGCTACTATGCAGAGCCCGAGG	TGTACGTGTTCATGCCGTT
SOX17	CAACTATCCTGACGTGTGACAG	TGCAGGCCAGAAGCAGTGTTACA	ACCCAGGAGTCTGAGGATTT
HHEX	CATTTAGCGCGTCGATTCTG	CAGCTCAGCGAGAGACAGGTC	GATTCTCCAACGACCAGACC
TBX3	GCGGCCATGTACGTGTAG	TGTCCCCTTTCGGAAGCCTGTTC	TGCCCTTCCACCTCCAG
HNF1B	GATACCAGCAGCATCAGTACA	AACAGTGTCCTCTACAAGCCTGGTG	ACGAAGTAAGTGGTGTGTGG
HNF4A	GGTGGACAAAGACAAGAGGAA	TCTGGACGGCTTCCTTCTTCATGC	CTCATAGCTTGACCTTCGAGTG

Table S7: List of oligonucleotides used. RT-qPCR primers, CRISPR/Cas9 gRNAs, PCR and sequencing primers. CRISPR/Cas9 oligonucleotides were designed based on previous reports of successful GATA6 gene editing (Shi et al., 2017; Tiyaboonchai et al., 2017). Related to STAR methods and Key Resource Table.

Antibodies							
Target	**Company**	**Product Code**	**Western blot**	**Immunofluorescence**	**Flow Cytometry**	**CHIP**	**Co-IP**
GATA6	Cell Signaling	5851	1:1000	1:7500	1:200	1:50	-
GATA4	Santa Cruz	SC-1237	1:200	1:100	-	-	-
FOXA2	RnD	AF2400	1:500	1:200	1:200	5μg	-
SOX17	RnD	AF1924	1:500	1:50	1:50	-	-
OCT4	Santa Cruz	SC-9081	-	1:100	1:100		-
HNF4A	Santa Cruz	SC-6556	-	1:200	1:100	-	-
EOMES	Cell Signaling	66325	-	-	-	1:50	-
FLAG	Sigma-Aldrich	F1804	1:1000	1:1000	-	-	1μg
Normal Mouse IgG	Millipore	12-371	-	-	-	-	1μg
CHD4	Cell Signaling	12011	1:750	-	-	-	-
SMARCA4	Cell Signaling	49360	1:750	-	-	-	-
KDM1A	Cell Signaling	2184	1:750	-	-	-	-
WDR5	Cell Signaling	13105	1:750	-	-	-	-
beta-Actin	Sigma-Aldrich	AC-15	1:5000	-	-	-	-
H3K4me1	Diagenode	C15410194	-	-	-	2μg	-
H3K9me2	Abcam	ab1220	-	-	-	5μg	-
H3K9me3	Abcam	ab8898	-	-	-	2μg	-
Donkey anti-rabbit alexafluor488	ThermoFisher Scientific	A21206	-	1:1000	1:500	-	-
Donkey anti-goat alexafluor488	ThermoFisher Scientific	A11055	-	1:1000	1:500	-	-
Goat anti-mouse alexafluor568	ThermoFisher Scientific	A11031	-	1:1000	-	-	-
Goat anti-Rabbit IgG-HRP	Cell Signaling	7074	1:3000	-	-	-	-
Goat anti-mouse IgG-HRP	ThermoFisher Scientific	31430	1:5000	-	-	-	-
Mouse anti-Goat IgG-HRP	Santa Cruz	SC-2354	1:3000	-	-	-	-

Table S8: List of antibodies used. Antibodies used in the study with the assay-specific dilutions. Related to STAR methods and Key Resource Table.