Please cite this article in press as: Yang et al., *GATA6-AS1* Regulates GATA6 Expression to Modulate Human Endoderm Differentiation, Stem Cell Reports (2020), https://doi.org/10.1016/j.stemcr.2020.07.014

Stem Cell Reports
Article
ISSCR
OPEN ACCESS

GATA6-AS1 Regulates GATA6 Expression to Modulate Human Endoderm Differentiation

Jie Yang,[1,4] Pei Lu,[1,4] Mao Li,[1] Chenchao Yan,[1] Tianzhe Zhang,[1] and Wei Jiang[1,2,3,*]
[1]Department of Biological Repositories, Frontier Science Center for Immunology and Metabolism, Medical Research Institute, Zhongnan Hospital of Wuhan University, Wuhan University, Wuhan 430071, China
[2]Hubei Provincial Key Laboratory of Developmentally Originated Disease, Wuhan 430071, China
[3]Human Genetics Resource Preservation Center of Wuhan University, Wuhan 430071, China
[4]Co-first author
*Correspondence: jiangw.mri@whu.edu.cn
https://doi.org/10.1016/j.stemcr.2020.07.014

SUMMARY

Transcriptome analysis has uncovered a series of long noncoding RNAs (lncRNAs) transcribed during cell differentiation, but how lncRNA is integrated with known transcriptional regulatory network is poorly understood. Here, we utilize human definitive endoderm differentiation as a model system and decipher the functional interaction between lncRNA and key transcriptional factor. We have identified *GATA6-AS1*, an lncRNA divergently transcribed from the GATA6 locus, is highly expressed during endoderm differentiation. Knockdown of *GATA6-AS1* in human pluripotent stem cells has no influence on morphology and pluripotency; however, *GATA6-AS1* depletion causes the deficiency of definitive endoderm differentiation. *GATA6-AS1* positively regulates the expression of endoderm key factor GATA6. Further investigation shows *GATA6-AS1* interacts with SMAD2/3 and activates the transcription of GATA6. In addition, overexpression of GATA6 is able to rescue the defect of endoderm differentiation due to the absence of *GATA6-AS1*, suggesting that GATA6 is the functional target of *GATA6-AS1* during endoderm differentiation. Ultimately, our study reveals that *GATA6-AS1* is necessary for human endoderm specification and reveals the underlying mechanism between *GATA6-AS1* and GATA6.

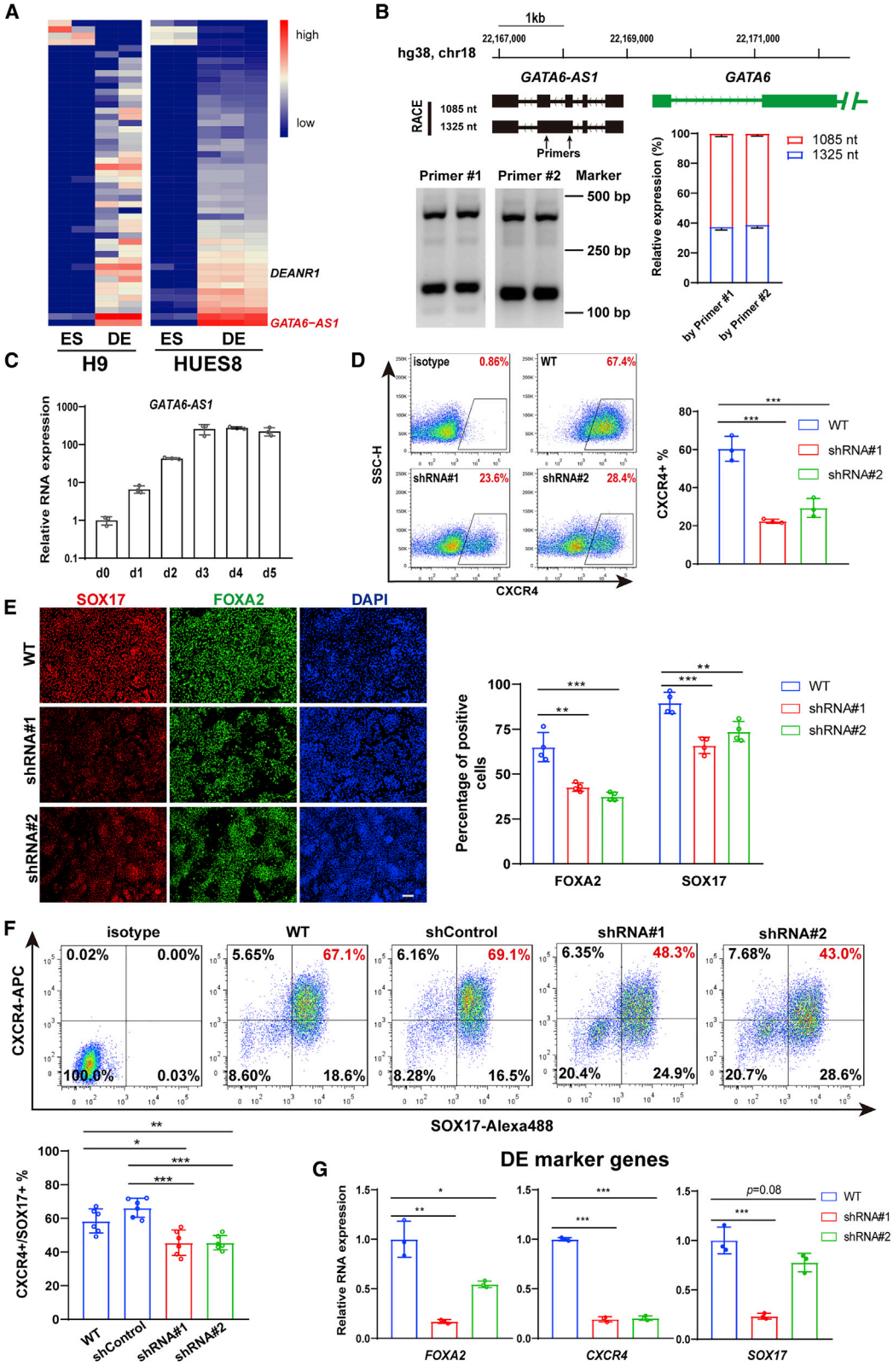
INTRODUCTION

Long noncoding RNAs (LncRNAs) are longer than 200 nucleotides and with low protein-coding potential, and they are expressed in a more cell-type-specific manner during development and in certain cancers (Delas and Hannon, 2017; Rinn and Chang, 2012). More than 77,000 lncRNA transcripts have been identified according to the RefLnc database (Jiang et al., 2019), but the majority have not been functionally dissected yet. During the last decade, lncRNAs have been uncovered as an important functional player in tissue development and cell differentiation (Fatica and Bozzoni, 2014). Among them, many are *cis*-acting lncRNAs that regulate local gene expression in a manner dependent on the location from where they originate (Gil and Ulitsky, 2019). Recently, *yylncT*, a lncRNA localizing to the BRACHYURY (T) locus, regulates the expression of the key mesoderm specifier T and thus contributes to mesoderm commitment (Frank et al., 2019); *CCR5-AS*, the antisense lncRNA of the CCR5 locus, protects *CCR5* mRNA from degradation and facilitates the HIV infection of CD4-positive T cells (Kulkarni et al., 2019). Mechanistically, *cis*-acting lncRNAs could promote or repress local transcription by recruiting transcription regulatory protein or bridging proximal enhancers into the promoter of the target gene or competitively binding transcription factors or epigenetic modifiers (Gil and Ulitsky, 2019).

Germ layer specification is one of the most critical events during early development. Sperm-egg fusion generates the zygote and through multiple rounds of mitotic division the embryo goes through two-cell, four-cell, morula stage, and subsequently enters blastocyst stage. The inner cell mass in blastocyst undergoes differentiation and generates three germ layers: ectoderm, mesoderm, and definitive endoderm (Kiecker et al., 2016; Okabe, 2013; Wamaitha and Niakan, 2018). The definitive endoderm contributes to the respiratory and gastrointestinal tracts and their derived tissues and organs, and endoderm dysplasia is associated with numerous human diseases that torment millions of people every year (Zorn and Wells, 2009). Due to the limitation to obtain clinical embryonic materials, very few studies on human endoderm development are performed. Embryonic stem cells (ESCs) are established from human blastocysts and provide a suitable model to study early development, disease modeling, and drug screening, as well as cell-based therapy (Thomson et al., 1998; Yiangou et al., 2018).

Transcription factors regulation network and epigenetic modifications have been extensively reported to contribute to definitive endoderm differentiation (Li et al., 2019; Tsankov et al., 2015). However, whether and how lncRNAs could function in endoderm differentiation was largely unknown for a long time. In 2013 Young group reported that more than 60% of lncRNA species expressed in human and murine ESCs were divergently transcribed from active protein-coding genes (Sigova et al., 2013). A following study together revealed that these divergently transcribed lncRNA/mRNA gene pairs exhibited coordinated

Please cite this article in press as: Yang et al., *GATA6-AS1* Regulates GATA6 Expression to Modulate Human Endoderm Differentiation, Stem Cell Reports (2020), https://doi.org/10.1016/j.stemcr.2020.07.014



(legend on next page)

transcriptional changes during differentiation into mesoderm and endoderm lineage through genome-wide analysis (Luo et al., 2016; Sigova et al., 2013). For example, lncRNA *Evx1as* bound to regulatory sites on chromatin and interacted with Mediator to facilitate EVX1 transcription and eventually contributed to mouse mesoderm differentiation (Luo et al., 2016). In 2015, lncRNA *DEANR1* was first identified to contribute to human endoderm differentiation by facilitating FOXA2 transcriptional activation (Jiang et al., 2015). Subsequently, Daneshvar et al. (2016) uncovered that *DIGIT*, an lncRNA transcribed from the Goosecoid (GSC) locus, regulated GSC expression to manipulate endoderm differentiation.

GATA6-AS1 is an lncRNA divergently transcribed from the GATA6 locus. GATA6 is an essential transcription factor during definitive endoderm and pancreas development and β cell function (Chia et al., 2019; Fisher et al., 2017; Schrode et al., 2014; Shi et al., 2017; Tiyaboonchai et al., 2017). However, the function of *GATA6-AS1* is unknown in endoderm and derived organs. Here, using a definitive endoderm differentiation system, in which both *GATA6-AS1* and GATA6 were highly expressed, we dissected the biological role of *GATA6-AS1* and further investigated the functional link between *GATA6-AS1* and GATA6.

RESULTS

GATA6-AS1 Is an lncRNA Highly Expressed in Definitive Endoderm

First, we performed transcriptomic analysis during definitive endoderm differentiation using human ESC line HUES8. By combining the data with previous published data based on H9 cells (GSE44875) (Jiang et al., 2015), we screened a number of endoderm-specific lncRNAs (Table S1). Among these, *GATA6-AS1*, a lncRNA divergently transcribed from GATA6 locus, exhibited the highest expression in definitive endoderm (Figure 1A). Since the current annotation of lncRNAs is not as accurate as coding genes, we performed 5′ and 3′ RACE experiments and identified two isoforms of *GATA6-AS1* with one region difference: one was 1,085 nucleotides and the other was 1,325 nucleotides (Supplemental Information). We further determined the expression levels of both isoforms and found that the longer form is approximately 40%, while the shorter form is about 60% (Figure 1B). Because both isoforms were largely similar in sequence and no isoform preference was obviously observed, we decided to not make a distinction between the isoform types in following studies. In addition, both isoforms were analyzed with low protein-coding potential by Coding Potential Calculator 2 (CPC2) (Kang et al., 2017) and Coding Potential Assessment Tool (CPAT) (Wang et al., 2013), respectively (Figure S1A).

By qRT-PCR, we validated the expression pattern of *GATA6-AS1* during endoderm differentiation, and the results showed that *GATA6-AS1* was gradually increased and achieved the highest expression level at day 3 (Figure 1C), when the endoderm differentiation was considered almost completed. Consistent with the endoderm-specific expression pattern, we surveyed available databases including most human tissues (Fagerberg et al., 2014) and found *GATA6-AS1* was highly expressed in endoderm-derived tissues and organs, such as stomach, small intestine, and duodenum (Figure S1B). These results indicated that *GATA6-AS1* is an lncRNA highly expressed in definitive endoderm.

GATA6-AS1 Is Required for Definitive Endoderm Differentiation

To further determine the biological function of *GATA6-AS1* in definitive endoderm differentiation, we utilized a quick definitive endoderm differentiation system. Time course experiment showed that most endoderm-specific genes achieved the expression peaks and both pluripotency and mesoderm-associated genes quickly decreased at day 3 (Figure S1C). Immunofluorescence assay of 3-day differentiated cells showed there were about 74% of FOXA2-positive

Figure 1. *GATA6-AS1* Is Highly Expressed in and Required for Human Endoderm Differentiation

(A) Heatmap showing the differentially expressed lncRNAs between ESCs and definitive endoderm (DE) cells (abs(log_2(fold-change)) > 5, and $p < 0.05$). Red represents higher expression and blue represents lower expression. The list of differentially expressed lncRNAs is shown in Table S1.

(B) Diagram of the *GATA6-AS1* and GATA6 gene locus. Two isoforms of *GATA6-AS1* were identified by 5′ and 3′ RACE and validated by RT-PCR. The relative expression levels of two isoforms were calculated according to PCR analysis (n = 3).

(C) Time course (day 0 to 5) expression pattern of *GATA6-AS1* during endoderm differentiation (n = 3). All data were normalized to day 0.

(D) Flow cytometric analysis showed the percentage of CXCR4-positive cells in differentiated HUES8 and *GATA6-AS1* knockdown cells (n = 3).

(E) Immunofluorescence of endoderm markers SOX17 and FOXA2 in differentiated HUES8 and *GATA6-AS1* knockdown cells (n = 4 for quantification). Scale bar, 100 μm.

(F) Co-flow cytometric analysis of SOX17 and CXCR4 in differentiated HUES8, shControl, and *GATA6-AS1* knockdown cells (n = 6).

(G) mRNA levels of endoderm marker genes, including *FOXA2*, *SOX17*, and *CXCR4* were detected by qRT-PCR (n = 3). All data were normalized to control.

and 82% of SOX17-positive cells (Figure S1D). These results together with our previous studies (Jiang et al., 2013a, 2013b, 2015) showed that the differentiation protocol was a simple, fast, and effective system to further investigate lncRNA function in definitive endoderm. In addition, CXCR4 was a widely used surface marker to evaluating the differentiation efficiency of definitive endoderm (Chu et al., 2016; D'Amour et al., 2005; Drukker et al., 2012; Jiang et al., 2015; Jiang et al., 2013a; Jiang et al., 2013b). We also confirmed this notion in our differentiation system by immunostaining with both CXCR4 and SOX17 (Figure S1E), and intracellular flow cytometric analysis using fluorescence-labeled antibody against SOX17 and CXCR4 (Figure S1F). Next, we designed five shRNAs targeting *GATA6-AS1* and two control shRNAs targeting unrelated sequences in our initial two rounds of screening (Figure S2A). The results showed three shRNAs effectively interfered with endoderm differentiation, while the other two shRNAs with ineffective knockdown efficiency and the control shRNAs did not (Figure S2A). We also compared, side-by-side, the control shRNA and wild-type cells, and the results showed no difference in terms of the differentiation capability toward definitive endoderm (Figure S2B), suggesting the lentiviral vector has no effect on the phenotype. Next, we evaluated the knockdown efficiency of shRNAs to target the common regions of both isoforms of *GATA6-AS1* in the endoderm and 293T cells and obtained two efficient shRNAs (Figure S2C). Then, accordingly, we established two human ESC lines with stable *GATA6-AS1* knockdown (Figure S2D). *GATA6-AS1* knockdown ESCs had no influence on colony morphology or expression of pluripotent genes (Figures S2D–S2F), which is consistent with the notion that *GATA6-AS1* is not expressed in undifferentiated pluripotent stem cells.

Next, we subjected the *GATA6-AS1* knockdown ESCs to endoderm differentiation assay using our differentiation scheme described above. As CXCR4-positive cells could represent human definitive endoderm cells in such a system (Figures S1E and S1F) and be easily detected, we first performed flow cytometric analysis of CXCR4 and the results showed a significant decrease in the percentage of CXCR4-positive cells in *GATA6-AS1* knockdown cells compared with wild-type cells (Figure 1D). Consistently, immunofluorescence assay showed significantly decreased endoderm markers, including SOX17 and FOXA2, in knockdown cells (Figure 1E). In addition, the intracellular flow cytometry analysis of CXCR4 and SOX17 showed that CXCR4-positive/SOX17-positive cells were significantly reduced (Figure 1F), which was further supported by qRT-PCR analysis of endoderm marker genes (Figure 1G). Meanwhile, both pluripotent genes (*SOX2* and *OCT4*) and mesoderm marker genes (*T* and *MIXL1*) showed increased expression in a certain extent in differentiated *GATA6-AS1* knockdown cells (Figure S2G). To further confirm the *GATA6-AS1* function in definitive endoderm, we performed the definitive endoderm differentiation in another differentiation system supplemented with 1 μM WNT inhibitor XAV-939 at day 3, which was a more efficient definitive endoderm differentiation protocol (Jiang et al., 2013b). The results of flow cytometry showed that *GATA6-AS1* knockdown cells exhibited a significant decrease of CXCR4-positive/SOX17-positive cells, similar to the results described above (Figure S2H). These results suggested that *GATA6-AS1* was a functional lncRNA in endoderm differentiation, and that depletion of *GATA6-AS1* blocked human endoderm differentiation from ESCs.

GATA6-AS1 Influences the Transcriptome Profile of Definitive Endoderm

To further confirm the important role of *GATA6-AS1* and investigate the downstream targets in endoderm differentiation, we performed RNA sequencing (RNA-seq) experiment using differentiated endoderm cells from *GATA6-AS1* knockdown cells and wild-type cells. Transcriptome analysis indicated that approximately 8% of coding genes exhibited differential expression levels, including pluripotent genes and endoderm marker genes (Figure 2A). Moreover, the expression of ESC-specific genes was still maintained in differentiated *GATA6-AS1* knockdown cells, but the expression levels of endoderm-specific genes were lower compared with differentiated wild-type cells (Figure 2B; Table S2). Consistently, genes with higher expression in wild-type cells were significantly enriched in regulation of gastrulation by gene set enrichment analysis (GSEA) (Figure 2C; Table S3). In addition, gene ontology (GO) terms of downregulated genes in *GATA6-AS1* knockdown endoderm cells included pattern specification process, embryonic organ development, cell fate commitment, and stem cell differentiation (Figure 2D; Table S3), supporting that *GATA6-AS1* played an important role in early development. Further analysis of differently expressed transcription factors showed pluripotent genes and endoderm-specific genes, especially *GATA6*, were interacting with other proteins or transcription factors (Figure 2E). Taken together, these results provide transcriptome evidence that *GATA6-AS1* is a functional lncRNA in human early endoderm differentiation.

To exclude the effect of *GATA6-AS1* in endoderm differentiation was not due to the genetic background of individual cell lines, we re-constructed another stable cell line with one shRNA achieving effective *GATA6-AS1* knockdown in H9 ESCs (Figure S3A), while the *GATA6-AS1* knockdown cells maintained typical colony morphology (Figure S3B). After being subjected to endoderm differentiation, we observed a reduced percentage of CXCR4-positive cells in *GATA6-AS1* knockdown cells (Figure S3C), along with

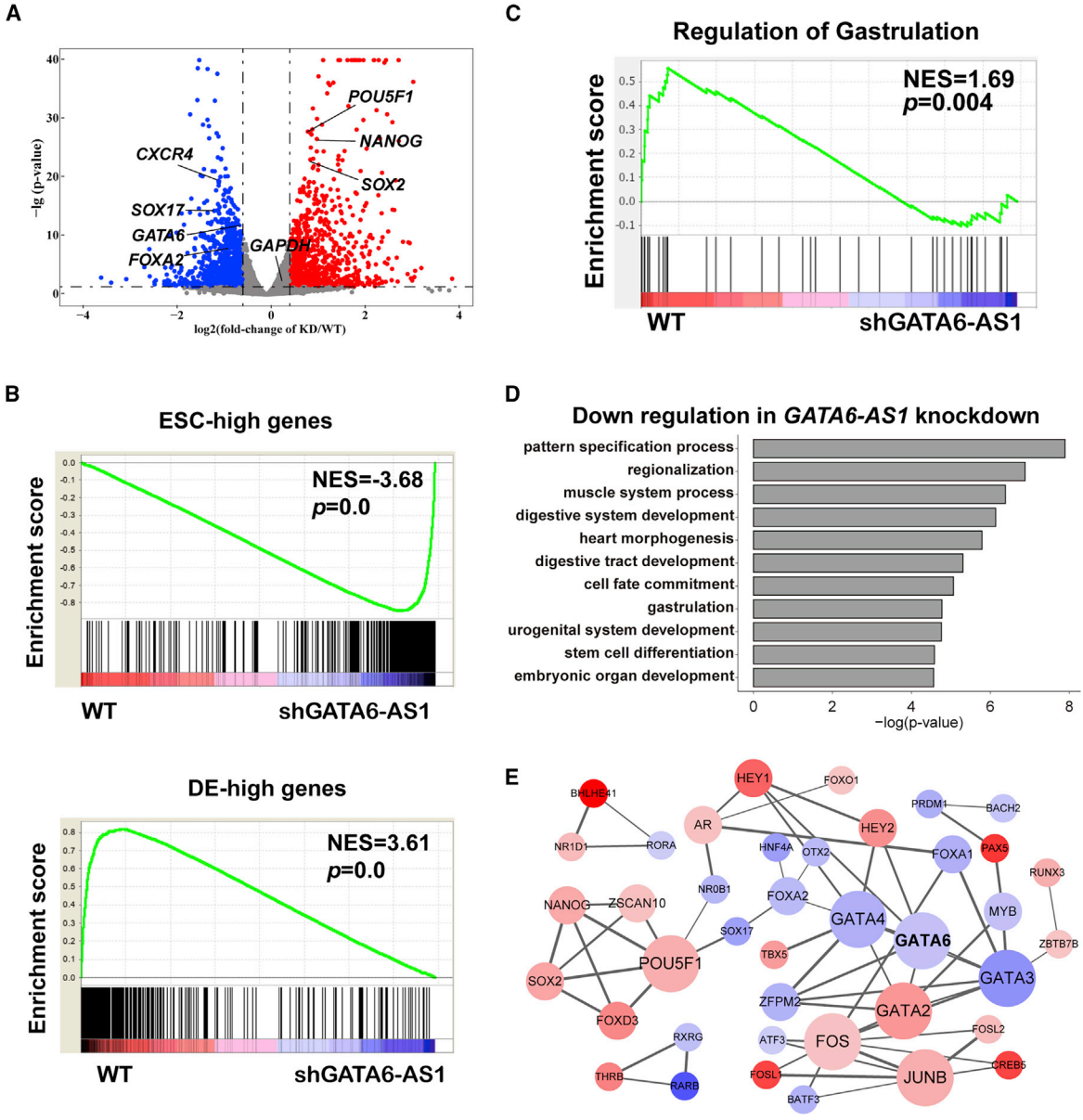


Figure 2. *GATA6-AS1* Influences the Transcriptome Profile of Human Endoderm

(A) Scatterplot showing differentially expressed genes ($p < 0.05$ and abs($\log_2$(fold-change)) > 0.5) between differentiated control cells and *GATA6-AS1* knockdown cells. Up- and downregulated genes in *GATA6-AS1* knockdown cells are shown in red and blue, respectively.

(B) Gene set enrichment analysis (GSEA) profile of ESC-high and DE-high genes. The gene list is shown in Table S2.

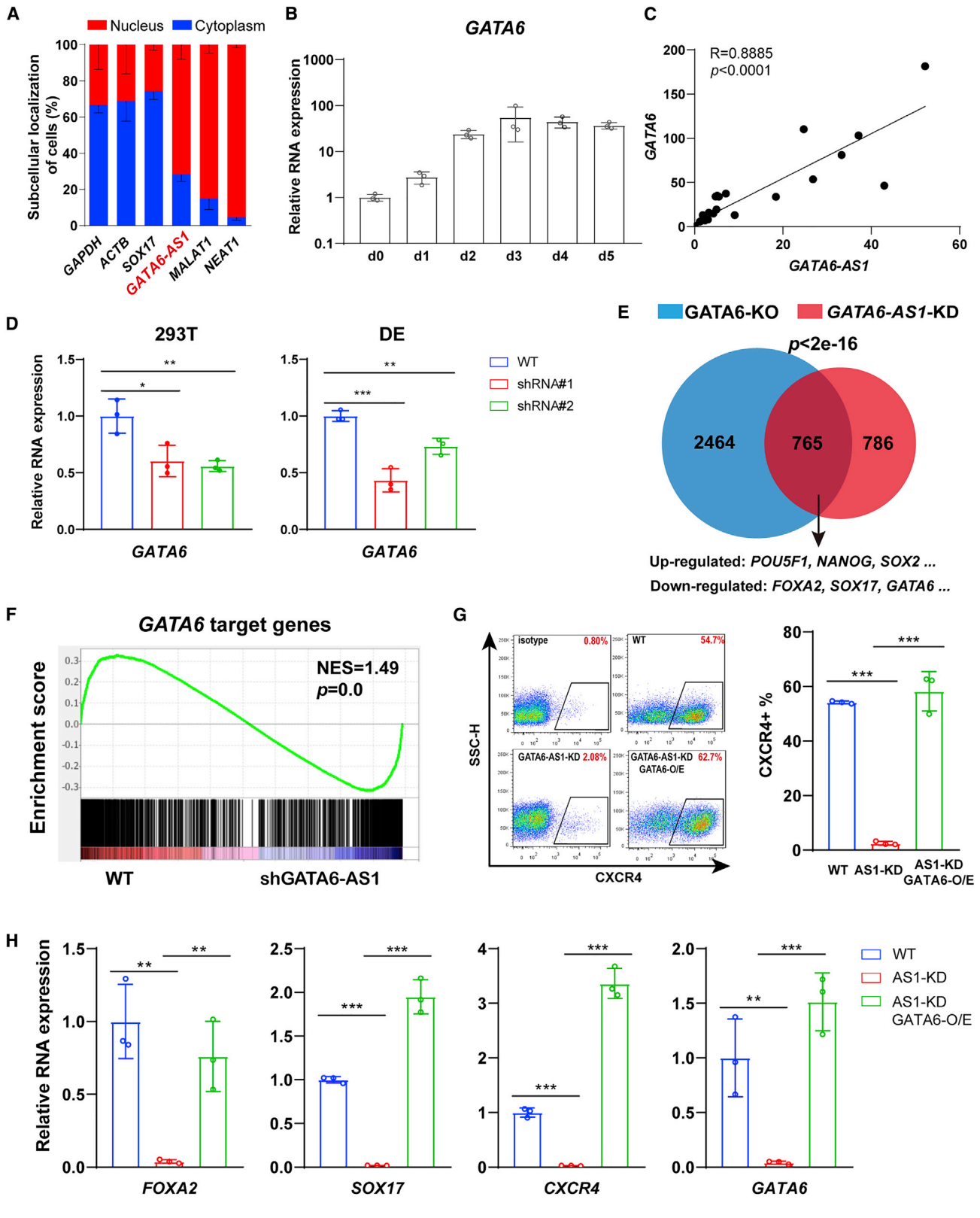
(C) GSEA profile of "regulation of gastrulation" set. The gene list is shown in Table S3.

(D) Gene ontology (GO) enrichment analysis of downregulated genes upon *GATA6-AS1* knockdown. The full GO terms are listed in Table S3.

(E) Differently expressed transcription factors interaction network. Red and blue represent upregulated and downregulated transcription factors after *GATA6-AS1* knockdown, respectively. The circle size represents the number of interaction proteins and the line width represents combined score of proteins.

significantly decreased mRNA levels of *SOX17*, *GATA6*, *FOXA2*, and *CXCR4* (Figure S3D). The protein levels of GATA6 and SOX17 were obviously reduced after *GATA6-AS1* depletion (Figure S3E). Moreover, when we performed the following differentiation toward pancreatic lineage based on our previously established protocols (Jiang et al., 2015; Tan et al., 2019), we found severe defects in pancreatic lineage marked by PDX1 and hepatic lineage marked

Please cite this article in press as: Yang et al., *GATA6-AS1* Regulates GATA6 Expression to Modulate Human Endoderm Differentiation, Stem Cell Reports (2020), https://doi.org/10.1016/j.stemcr.2020.07.014



(legend on next page)

by AFP in *GATA6-AS1* knockdown cells (Figure S3F), supporting the notion that *GATA6-AS1* is important for endoderm fate. These data together confirmed that *GATA6-AS1* was required for human endoderm differentiation, and that deficiency of *GATA6-AS1* caused the downregulation of definitive endoderm signature genes and blocked stem cell differentiation and cell fate commitment.

GATA6 Is the Target of *GATA6-AS1*

Since the subcellular distribution of lncRNAs is usually correlated with different action models (Chen, 2016), we asked whether *GATA6-AS1* was localized in the nucleus or cytosol. We separated cytoplasmic and nucleus fractions of endoderm cells and determined the expression levels of *GATA6-AS1* by qRT-PCR. The result showed that the *GATA6-AS1* transcript was mainly localized in the nucleus (Figure 3A). Given that *GATA6-AS1* was divergently transcribed from the GATA6 locus, and GATA6 is one of the key nodes of the protein-protein interaction network affected by *GATA6-AS1* (Figure 2E), we proposed that GATA6 might be the target of *GATA6-AS1*. Consistent with this hypothesis, GATA6 was upregulated during endoderm differentiation (Figure 3B) in a similar trend to *GATA6-AS1* (Figure 1C). We also found a significant correlation between *GATA6-AS1* and *GATA6* expression from either different RNA-seq samples (Jiang et al., 2015; Li et al., 2018; Xin et al., 2013) (Figure 3C; Table S4) or our qRT-PCR results (Figure S3G). More importantly, knockdown of *GATA6-AS1* led to a decreased expression of *GATA6* in both 293T cells and differentiated endoderm cells (Figures 3D and S3D), suggesting that *GATA6-AS1* might regulate GATA6.

To further investigate the relationship between *GATA6-AS1* and GATA6, we downloaded and re-analyzed the publicly available RNA-seq data of human endoderm cells with GATA6 depletion (E-MTAB-5958 [Chia et al., 2019]), and identified 3,229 differentially expressed genes due to GATA6-KO (Table S1). We found that there was significant overlap between *GATA6-AS1*-regulated genes (Table S1) and GATA6-regulated genes (Figure 3E). In addition, we identified GATA6 direct targets by combining the RNA-seq data of GATA6-KO and GATA6 chromatin immunoprecipitation sequencing (ChIP-seq) data (Chia et al., 2019; Fisher et al., 2017; Li et al., 2019; Tsankov et al., 2015) (Figure S3H), and performed GSEA, which indicated that GATA6 target genes (Table S3) were significantly enriched in those genes affected by *GATA6-AS1* knockdown in differentiated endoderm (Figure 3F). These data together supported that *GATA6-AS1* likely regulated GATA6 and shared similar downstream targets in endoderm differentiation.

To confirm GATA6 was the functional target of *GATA6-AS1*, we put GATA6 coding region back to *GATA6-AS1*-depletion cells and checked whether the phenotype of endoderm differentiation defect could be rescued. Overexpression of GATA6 did not disrupt the pluripotency in ESC (Figures S3I–S3K). However, GATA6 overexpression indeed rescued the gene expression of endoderm markers (Figure 3H) and recovered endoderm differentiation in *GATA6-AS1* knockdown cells (Figure 3G) as well. These data demonstrated that GATA6 was the functional target of *GATA6-AS1* in human endoderm differentiation.

GATA6-AS1 Interacts with SMAD2/3 to Promote Definitive Endoderm Differentiation

Previous research reported that SMAD2/3 transcriptionally activated endoderm-specific transcription factors, including *GATA6*, to promote endoderm differentiation (Chia et al., 2019; Kim et al., 2011; Li et al., 2019). Therefore, we were wondering whether *GATA6-AS1* was involved in the SMAD2/3-mediated regulation of GATA6. To this end, we performed SMAD2/3 ChIP assay on *GATA6-AS1* knockdown and wild-type endoderm cells. The ChIP-qPCR results showed that SMAD2/3 indeed bound to the promotor region of GATA6 (Figure 4A); moreover, *GATA6-AS1* deletion reduced SMAD2/3 occupation on the GATA6 locus (Figure 4A). We further performed an SMAD2/3

Figure 3. GATA6 Is the Functional Target of *GATA6-AS1* in Endoderm Differentiation

(A) Subcellular localization of *GATA6-AS1* in human endoderm cells. *GAPDH*, *ACTB*, and *SOX17* transcripts served as cytoplasm-located control, and *MALAT1* and *NEAT1* served as nuclei-located controls (n = 6). Red represents nuclei fraction and blue represents cytoplasmic fraction.

(B) Time course (day 0 to 5) expression pattern of *GATA6* RNA during endoderm differentiation (n = 3). All data were normalized to day 0.

(C) The correlation analysis of *GATA6* and *GATA6-AS1* transcripts in different samples. The data presented in Figure 3C are shown in Table S4.

(D) The expression level of *GATA6* RNA upon *GATA6-AS1* knockdown in 293T, differentiated endoderm cells (HUES8) (n = 3).

(E) Venn diagram indicates the significant overlap of differentially expressed genes due to GATA6-KO or *GATA6-AS1* knockdown in endoderm cells. The gene list is shown in Table S3.

(F) GSEA profile of GATA6 target genes in differentiated control endoderm cells and *GATA6-AS1* knockdown endoderm cells. The list of GATA6 target genes is shown in Table S3.

(G) Flow cytometric analysis showed the percentage of CXCR4-positive cells in differentiated HUES8 (WT), *GATA6-AS1* knockdown (AS1-KD), and GATA6-overexpressing cells with *GATA6-AS1* knockdown (AS1-KD; GATA6-O/E) (n = 3).

(H) mRNA levels of endoderm marker genes, including *FOXA2*, *SOX17*, and *CXCR4* detected by qRT-PCR for the samples shown in G (n = 3).

Please cite this article in press as: Yang et al., *GATA6-AS1* Regulates GATA6 Expression to Modulate Human Endoderm Differentiation, Stem Cell Reports (2020), https://doi.org/10.1016/j.stemcr.2020.07.014

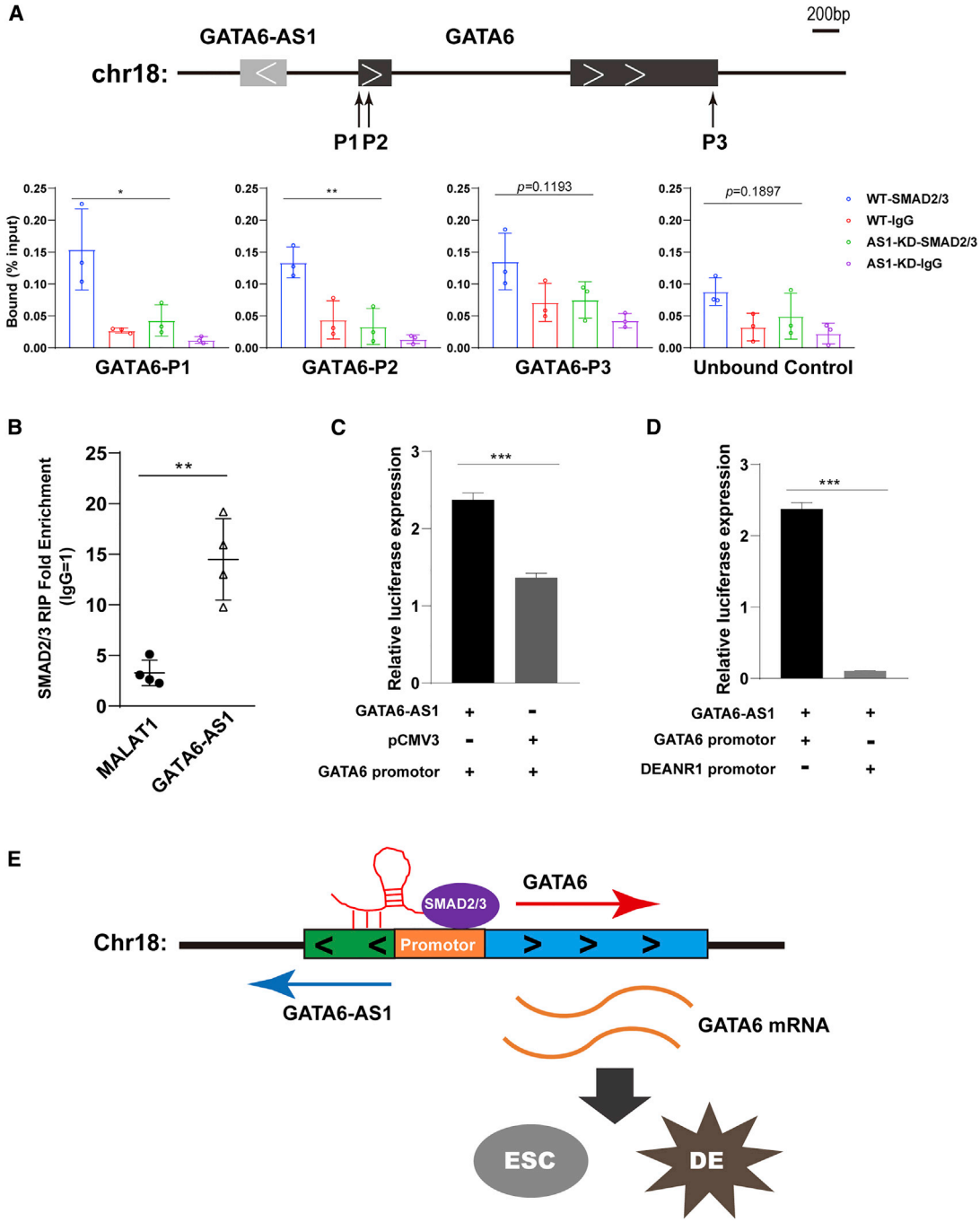


Figure 4. *GATA6-AS1* Interacts with SMAD2/3 and Regulates GATA6 Transcription Activity

(A) ChIP-qPCR analysis of SMAD2/3 bound to GATA6 locus upon *GATA6-AS1* depletion (n = 3). Arrows represent the genomic position of PCR primers.

(B) Immunoprecipitation demonstrated that SMAD2/3 interacted with *GATA6-AS1*. *MALAT1* served as lncRNA control (n = 4).

(C) *GATA6-AS1* transcript increased the transcription activity of GATA6 promoter in 293T cells (n = 3).

(D) *GATA6-AS1* transcript showed specific activation for *GATA6* promoter rather than *DEANR1* promoters (n = 3).

(E) Model of *GATA6-AS1*-regulated GATA6 via interacting with SMAD2/3 and promoting the transcription activity. *GATA6-AS1* might form an RNA-DNA triple helix with the GATA6 promotor and interacted with SMAD2/3 to activate GATA6 transcription, thus further promoting human ESCs to differentiate into definitive endoderm.

immunoprecipitation assay in endoderm cells and found that SMAD2/3 interacted with *GATA6-AS1*, rather than unrelated lncRNA, such as *MALAT1* (Figure 4B). Luciferase assay showed that the *GATA6-AS1* transcript could activate the transcription of the *GATA6* promoter (Figure 4C), and that this activation was specific to the *GATA6* promoter rather than the *DEANR1* promoter (Figure 4D). These results demonstrated that *GATA6-AS1* was required for SMAD2/3-mediated GATA6 transcriptional activation to modulate human endoderm differentiation (Figure 4E).

DISCUSSION

LncRNAs, especially *cis*-acting lncRNAs, have attracted more attention in studies of development and differentiation; however, the function of the majority of lncRNAs is still unclear due to numerous transcripts and diverse cell types (Gil and Ulitsky, 2019). Here, we revealed that *GATA6-AS1* was functional in regulating GATA6 expression by interacting with and mediating SMAD2/3 to bind to the *GATA6* promoter region. Biologically, depletion of *GATA6-AS1* severely blocked human endoderm differentiation, and GATA6 overexpression rescued the endoderm differentiation deficiency. Our study identified a novel lncRNA in human endoderm differentiation and replenished a regulation model of key factor GATA6.

Recently *GATA6-AS1* was studied in two different biological systems. Neumann et al., (2018) showed that *GATA6-AS1* was upregulated in endothelial cells during hypoxia and that *GATA6-AS1* knockdown disrupted endothelial-mesenchymal transition and promoted formation of blood vessels. They further found that *GATA6-AS1* negatively regulated LOXL2, which led to the deamination of active marker H3K4me3, and *GATA6-AS1* knockdown led to reduced H3K4me3 and repressed the expression of angiogenesis-related genes in endothelial cells (Neumann et al., 2018). Zhu et al. (2018) studied *GATA6-AS1* (named *lncGata6*) in the mouse intestine system and reported that *lncGata6* was highly expressed in mice intestine stem cells and contributed to regeneration and colorectal tumorigenesis. In this context, lncGata6 recruited the NURF complex onto the Ehf promoter and promoted Lgr4/5 expression by Ehf. In our endoderm differentiation context, we also checked whether those two mechanisms existed. First, we found that expression of both LOXL2 and EHF was very low during endoderm differentiation (Figure S3L), and that *GATA6-AS1* knockdown had no significant effect on LOXL2 expression (Figure S3M). Second, our data clearly suggested that *GATA6-AS1* regulated GATA6 expression demonstrated by both RNA and protein analysis (Figures 2G and 3D) and, most importantly, that overexpression of GATA6 was able to rescue the phenotype of defected endoderm differentiation due to *GATA6-AS1* depletion (Figures 3G and 3H). Supporting our conclusion, several studies showed the expression levels of *GATA6-AS1* was tightly correlated with GATA6 in diverse cell types, including stem cells and cancer cells (Kuo et al., 2019; Liu et al., 2018; Luo et al., 2016). These data suggested that the relationship between *GATA6-AS1* and GATA6 might be more universal. Recently, Kuo et al. (2019) reported that *GATA6-AS1* could form RNA-DNA triple helices in the GATA6 promoter through a computational method called the Triplex Domain Finder. This raised a hypothesis that the *GATA6-AS1* transcript might form RNA-DNA triple helices on its own gene body region, recruiting and assisting SMAD2/3 bound at the promotor of GATA6 to activate GATA6 transcription, which further contributed to various biological contexts, including endoderm differentiation.

Definitive endoderm differentiation is mainly induced by the transforming growth factor β/SMAD signaling pathway. SMAD regulates target gene expression via binding to SMAD-binding elements in promotor regions. Recent studies showed that lncRNA was also necessary in the SMAD regulation network. LncRNA *GAS5* competitively bound SMAD3 protein via multiple RNA SBEs to inhibit SMAD3 binding to the SMC gene promoter to block smooth muscle cell differentiation (Tang et al., 2017). *Lnc-TSI* interacted with MH2 domain of SMAD3, blocking interaction of SMAD3 with SMAD7 in renal fibrogenesis (Wang et al., 2018). In addition, LncRNA ELIT-1 bound to SMAD3 and functioned as a cofactor to recruit SMAD3 to the promoters of *Snail* and other SMAD target genes to boost epithelial-mesenchymal transition progression (Sakai et al., 2019). In our study, we revealed that *GATA6-AS1* was able to function as a cofactor of SMAD2/3 to positively regulate GATA6 for endoderm differentiation. This finding together with the *DEANR1* report (Jiang et al., 2015) suggested that lncRNAs may play a special role for SMAD functions in pluripotency exit and differentiation initiation in human pluripotent stem cells, which provides a meaningful insight in understanding SMAD signaling and the switch between pluripotency and differentiation.

Recently, GATA6 has been well studied in human pancreatic development by different groups, and the dosage-sensitive requirement of GATA6 in the differentiation of both definitive endoderm and pancreas has been identified (Chia et al., 2019; Shi et al., 2017; Tiyaboonchai et al., 2017). However, there were very few reports showing how GATA6 was regulated in endoderm and pancreas differentiation as well as in disease pathogenesis. We noticed that many SNPs located in the *GATA6/GATA6-AS1* locus were highly associated with the clinical traits, including pancreatic agenesis, congenital heart disease, or acute myocardial infarction (Allen et al., 2011; Sun et al., 2019). The variants within coding region and changing the

protein sequence very likely act by affecting the function of the GATA6 protein (Allen et al., 2011), but how those SNPs located in noncoding regions contribute to disease are poorly studied. Very recently, Gadue's group reported that a noncoding SNP rs12953985 located at the downstream target of the GATA6 locus contributed to the regulation of GATA6 and thus to pancreas differentiation and pancreatic agenesis (Kishore et al., 2020). Here, we reported that *GATA6-AS1* positively regulated GATA6 transcription; moreover, *GATA6-AS1* knockdown cells failed to differentiate into endodermal pancreatic lineage (Figure S3F). The fact that several noncoding SNPs locate in the gene body of *GATA6-AS1* and our functional dissection of *GATA6-AS1* indicate a possibility that *GATA6-AS1* might be a target for these diseases, and further studies on the relationship of these noncoding variants and *GATA6-AS1* in different contexts would facilitate our understanding on the role of noncoding variants in disease.

In summary, we dissected the biological function of *GATA6-AS1* in human endoderm differentiation. Moreover, we found that *GATA6-AS1* regulated GATA6 expression by interacting with SMAD2/3 and mediating the transcriptional activation of *GATA6*. Our results not only uncovered a new lncRNA contributing to early lineage specification, but also provided evidence to support the regulation ship between GATA6 and the divergent transcript *GATA6-AS1*.

EXPERIMENTAL PROCEDURES

Cell Culture and Differentiation

Two different human ESC lines, HUES8 and H9, were used in this study. They were both cultured with mTeSR1 (STEMCELL Technologies, cat. no. AB217641) on Matrigel-coated plates. Human embryonic kidney 293T cells were cultured with DMEM containing 10% FBS and 1% penicillin-streptomycin. The protocol to differentiate human ESCs into definitive endoderm was based on a previous report (Jiang et al., 2013a) with minor modifications. Generally, DMEM-F12 (Gibco, cat. no. C11330500BT) or RPMI 1640 (Gibco, cat. no. C22400500BT) was used as basal medium, supplemented with 0.2% BSA (YEASEN, cat. no. B57370), 1% penicillin-streptomycin (Gibco, cat. no. 15140163). Activin A (100 ng/mL, PeproTech, cat. no. 120-14P) and 2.5 μM CHIR99021 (Selleck, cat. no. S2924) were added at day 1, followed by 100 ng/mL Activin A alone for the next 2 or 4 days. For following pancreatic lineage differentiation, the differentiated definitive endoderm cells were cultured in MCDB131 (Gibco, cat. no. 10372019) supplemented with 1.5 g/L sodium bicarbonate (Gibco, cat. no. 25080094), 10 mM glucose (Invitrogen, cat. no. A2494001), 0.5% BSA, 0.25 mM ascorbic acid (Sigma, cat. no. A5960), 1× ITS-X (BasalMedia, cat. no. S452J7), 1× GlutaMAX (Gibco, cat. no. 35050061), 1% penicillin-streptomycin, 50 ng/mL KGF (PeproTech, cat. no. 100-19), and 2 μM XAV-939 (Selleck, cat. no. S1180) for 2 days. Next, 0.5 μM SANT1 (Selleck, cat. no. S7092), 100 nM TTNPB (Selleck, cat. no. S4627), and 500 nM PDBU (Sigma, cat. no. P1269) were additionally added to the above medium and cultured for 5 days.

Construction of Knockdown

We utilized shRNA for knockdown. A lentiviral vector system, containing the murine U6-shRNA cassette and a puromycin selection marker, was used in this study. After infection, we used 2 μg/mL puromycin for selection to establish stable knockdown human ESC lines. We constructed two efficient shRNA vectors against *GATA6-AS1*. The sequences of oligoes were listed as below:

shRNA no. 1: GAAACCGTTCTCATCCAAT
shRNA no. 2: GAAAGGATTTCTTCCGACA
shControl: GAAGTATTCCGCGTACGTT.

RNA-Binding Protein Immunoprecipitation and ChIP

RNA-binding protein immunoprecipitation (RIP) assay was performed using Magna RIP Kit (Millipore) according to the manufacturer's instruction. For each sample, cell lysates from 2×10^7 cells were prepared and 5 μg control IgG (R&D, cat. no. AB-108-C) or antibody against SMAD2/3 (R&D, cat. no. AF3797) was used.

ChIP was performed using the protocol reported previously (Chia et al., 2019). For each sample, cell lysates from 1×10^7 definitive endoderm cells were prepared. For each IP, 10 μg SMAD2/3 (R&D, cat. no. AF3797) or IgG was used. Primers used to detect chromatin enrichment are listed in Table S5.

RNA-Seq and Data Analysis

We performed RNA-seq for undifferentiated ESCs, differentiated definitive endoderm samples from *GATA6-AS1* knockdown and wild-type genotypes on Illumina HiSeq X Ten PE150. To identify differentially expressed genes among ESCs, definitive endoderm, and *GATA6-AS1* knockdown cells, we used Bioconductor package DESeq2 to determine gene expression differences (Love et al., 2014). GSEA was performed on GSEA software to identify specifically enriched signal pathways (Mootha et al., 2003; Subramanian et al., 2005).

To identify GATA6 target genes, we downloaded the following datasets: RNA-seq data of GATA6-KO (E-MTAB-5958 [Chia et al., 2019]), ChIP-seq data of GATA6 (GSE77360 [Fisher et al., 2017], GSE109524 [Li et al., 2019], and GSE61475 [Tsankov et al., 2015]). The genes with both GATA6 binding and changed expression upon GATA6 knockout were identified as GATA6 target genes, which were used in Figure 3F.

Statistical Methods

All experiments were performed as at least three independent experiments. Data were analyzed in GraphPad Prism 8. Data are shown as means ± SD. Comparisons were conducted via Student's t test (two-tailed, equal variance) and p values are shown with $^*p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Data and Code Availability

The Gene Expression Omnibus accession numbers for the RNA-seq raw data reported in this work are GSE137208 (undifferentiated HUES8) and GSE143499 (endoderm samples).

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.stemcr.2020.07.014.

AUTHOR CONTRIBUTIONS

W.J. conceived the project, and designed the experiment together with J.Y. and P.L. J.Y. performed most of the bench experiments with help from P.L. and M.L. P.L. performed the initial screening. J.Y. analyzed the next-generation sequencing data with help from C.Y. and T.Z. J.Y. drafted the manuscript, and W.J. and J.Y. finalized the manuscript. All authors contributed to and approved the final manuscript.

ACKNOWLEDGMENTS

We thank Yinglei Li, Jing Lv, Ran Liu, and other laboratory members for technical help and helpful discussion. This work was supported by the National Key Research and Development Program of China (No. 2016YFA0503100), the National Natural Science Foundation of China (No. 91740102 and 31970608), the Fundamental Research Funds for the Central Universities, and the Medical Science Advancement Program of Wuhan University (No. TFZZ2018053 and TFJC2018005).

Received: February 2, 2020
Revised: July 14, 2020
Accepted: July 14, 2020
Published: August 13, 2020

REFERENCES

Allen, H.L., Flanagan, S.E., Shaw-Smith, C., De Franco, E., Akerman, I., Caswell, R., International Pancreatic Agenesis, C., Ferrer, J., Hattersley, A.T., and Ellard, S. (2011). GATA6 haploinsufficiency causes pancreatic agenesis in humans. Nat. Genet. *44*, 20–22.

Chen, L.L. (2016). Linking long noncoding RNA localization and function. Trends Biochem. Sci. *41*, 761–772.

Chia, C.Y., Madrigal, P., Denil, S., Martinez, I., Garcia-Bernardo, J., El-Khairi, R., Chhatriwala, M., Shepherd, M.H., Hattersley, A.T., Dunn, N.R., et al. (2019). GATA6 cooperates with EOMES/SMAD2/3 to deploy the gene regulatory network governing human definitive endoderm and pancreas formation. Stem Cell Reports *12*, 57–70.

Chu, L.F., Leng, N., Zhang, J., Hou, Z., Mamott, D., Vereide, D.T., Choi, J., Kendziorski, C., Stewart, R., and Thomson, J.A. (2016). Single-cell RNA-seq reveals novel regulators of human embryonic stem cell differentiation to definitive endoderm. Genome Biol. *17*, 173.

D'Amour, K.A., Agulnick, A.D., Eliazer, S., Kelly, O.G., Kroon, E., and Baetge, E.E. (2005). Efficient differentiation of human embryonic stem cells to definitive endoderm. Nat. Biotechnol. *23*, 1534–1541.

Daneshvar, K., Pondick, J.V., Kim, B.M., Zhou, C., York, S.R., Macklin, J.A., Abualteen, A., Tan, B., Sigova, A.A., Marcho, C., et al. (2016). DIGIT is a conserved long noncoding RNA that regulates GSC expression to control definitive endoderm differentiation of embryonic stem cells. Cell Rep. *17*, 353–365.

Delas, M.J., and Hannon, G.J. (2017). lncRNAs in development and disease: from functions to mechanisms. Open Biol. *7*, 170121.

Drukker, M., Tang, C., Ardehali, R., Rinkevich, Y., Seita, J., Lee, A.S., Mosley, A.R., Weissman, I.L., and Soen, Y. (2012). Isolation of primitive endoderm, mesoderm, vascular endothelial and trophoblast progenitors from human pluripotent stem cells. Nat. Biotechnol. *30*, 531–542.

Fagerberg, L., Hallstrom, B.M., Oksvold, P., Kampf, C., Djureinovic, D., Odeberg, J., Habuka, M., Tahmasebpoor, S., Danielsson, A., Edlund, K., et al. (2014). Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. Mol. Cell Proteomics *13*, 397–406.

Fatica, A., and Bozzoni, I. (2014). Long non-coding RNAs: new players in cell differentiation and development. Nat. Rev. Genet. *15*, 7–21.

Fisher, J.B., Pulakanti, K., Rao, S., and Duncan, S.A. (2017). GATA6 is essential for endoderm formation from human pluripotent stem cells. Biol. Open *6*, 1084–1095.

Frank, S., Ahuja, G., Bartsch, D., Russ, N., Yao, W., Kuo, J.C., Derks, J.P., Akhade, V.S., Kargapolova, Y., Georgomanolis, T., et al. (2019). yylncT defines a class of divergently transcribed lncRNAs and safeguards the T-mediated mesodermal commitment of human PSCs. Cell Stem Cell *24*, 318–327.e8.

Gil, N., and Ulitsky, I. (2019). Regulation of gene expression by *cis*-acting long non-coding RNAs. Nat. Rev. Genet. *21*, 102–117.

Jiang, S., Cheng, S.J., Ren, L.C., Wang, Q., Kang, Y.J., Ding, Y., Hou, M., Yang, X.X., Lin, Y., Liang, N., et al. (2019). An expanded landscape of human long noncoding RNA. Nucleic Acids Res. *47*, 7842–7856.

Jiang, W., Liu, Y., Liu, R., Zhang, K., and Zhang, Y. (2015). The lncRNA DEANR1 facilitates human endoderm differentiation by activating FOXA2 expression. Cell Rep. *11*, 137–148.

Jiang, W., Wang, J., and Zhang, Y. (2013a). Histone H3K27me3 demethylases KDM6A and KDM6B modulate definitive endoderm differentiation from human ESCs by regulating WNT signaling pathway. Cell Res. *23*, 122–130.

Jiang, W., Zhang, D., Bursac, N., and Zhang, Y. (2013b). WNT3 is a biomarker capable of predicting the definitive endoderm differentiation potential of hESCs. Stem Cell Reports *1*, 46–52.

Kang, Y.J., Yang, D.C., Kong, L., Hou, M., Meng, Y.Q., Wei, L., and Gao, G. (2017). CPC2: a fast and accurate coding potential calculator based on sequence intrinsic features. Nucleic Acids Res. *45*, W12–W16.

Kiecker, C., Bates, T., and Bell, E. (2016). Molecular specification of germ layers in vertebrate embryos. Cell. Mol. Life Sci. *73*, 923–947.

Kim, S.W., Yoon, S.J., Chuong, E., Oyolu, C., Wills, A.E., Gupta, R., and Baker, J. (2011). Chromatin and transcriptional signatures for Nodal signaling during endoderm formation in hESCs. Dev. Biol. *357*, 492–504.

Kishore, S., De Franco, E., Cardenas-Diaz, F.L., Letourneau-Freiberg, L.R., Sanyoura, M., Osorio-Quintero, C., French, D.L., Greeley, S.A.W., Hattersley, A.T., and Gadue, P. (2020). A non-coding disease modifier of pancreatic agenesis identified by genetic correction in a patient-derived iPSC line. Cell Stem Cell *27*, 137–146.e6.

Kulkarni, S., Lied, A., Kulkarni, V., Rucevic, M., Martin, M.P., Walker-Sperling, V., Anderson, S.K., Ewy, R., Singh, S., Nguyen, H., et al. (2019). CCR5AS lncRNA variation differentially regulates CCR5, influencing HIV disease outcome. Nat. Immunol. *20*, 824–834.

Kuo, C.C., Hanzelmann, S., Senturk Cetin, N., Frank, S., Zajzon, B., Derks, J.P., Akhade, V.S., Ahuja, G., Kanduri, C., Grummt, I., et al. (2019). Detection of RNA-DNA binding sites in long noncoding RNAs. Nucleic Acids Res. *47*, e32.

Li, J., Wu, X., Zhou, Y., Lee, M., Guo, L., Han, W., Mo, W., Cao, W.M., Sun, D., Xin, R., et al. (2018). Decoding the dynamic DNA methylation and hydroxymethylation landscapes in endodermal lineage intermediates during pancreatic differentiation of hESC. Nucleic Acids Res. *46*, 2883–2900.

Li, Q.V., Dixon, G., Verma, N., Rosen, B.P., Gordillo, M., Luo, R., Xu, C., Wang, Q., Soh, C.L., Yang, D., et al. (2019). Genome-scale screens identify JNK-JUN signaling as a barrier for pluripotency exit and endoderm differentiation. Nat. Genet. *51*, 999–1010.

Liu, Z., Dai, J., and Shen, H. (2018). Systematic analysis reveals long noncoding RNAs regulating neighboring transcription factors in human cancers. Biochim. Biophys. Acta Mol. Basis Dis. *1864*, 2785–2792.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550.

Luo, S., Lu, J.Y., Liu, L., Yin, Y., Chen, C., Han, X., Wu, B., Xu, R., Liu, W., Yan, P., et al. (2016). Divergent lncRNAs regulate gene expression and lineage differentiation in pluripotent cells. Cell Stem Cell *18*, 637–652.

Mootha, V.K., Lindgren, C.M., Eriksson, K.F., Subramanian, A., Sihag, S., Lehar, J., Puigserver, P., Carlsson, E., Ridderstrale, M., Laurila, E., et al. (2003). PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. Nat. Genet. *34*, 267–273.

Neumann, P., Jae, N., Knau, A., Glaser, S.F., Fouani, Y., Rossbach, O., Kruger, M., John, D., Bindereif, A., Grote, P., et al. (2018). The lncRNA GATA6-AS epigenetically regulates endothelial gene expression via interaction with LOXL2. Nat. Commun. *9*, 237.

Okabe, M. (2013). The cell biology of mammalian fertilization. Development *140*, 4471–4479.

Rinn, J.L., and Chang, H.Y. (2012). Genome regulation by long noncoding RNAs. Annu. Rev. Biochem. *81*, 145–166.

Sakai, S., Ohhata, T., Kitagawa, K., Uchida, C., Aoshima, T., Niida, H., Suzuki, T., Inoue, Y., Miyazawa, K., and Kitagawa, M. (2019). Long noncoding RNA ELIT-1 acts as a Smad3 cofactor to facilitate TGFbeta/smad signaling and promote epithelial-mesenchymal transition. Cancer Res. *79*, 2821–2838.

Schrode, N., Saiz, N., Di Talia, S., and Hadjantonakis, A.K. (2014). GATA6 levels modulate primitive endoderm cell fate choice and timing in the mouse blastocyst. Dev. Cell *29*, 454–467.

Shi, Z.D., Lee, K., Yang, D., Amin, S., Verma, N., Li, Q.V., Zhu, Z., Soh, C.L., Kumar, R., Evans, T., et al. (2017). Genome editing in hPSCs reveals GATA6 haploinsufficiency and a genetic interaction with GATA4 in human pancreatic development. Cell Stem Cell *20*, 675–688.e6.

Sigova, A.A., Mullen, A.C., Molinie, B., Gupta, S., Orlando, D.A., Guenther, M.G., Almada, A.E., Lin, C., Sharp, P.A., Giallourakis, C.C., et al. (2013). Divergent transcription of long noncoding RNA/mRNA gene pairs in embryonic stem cells. Proc. Natl. Acad. Sci. U S A *110*, 2876–2881.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., et al. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. U S A *102*, 15545–15550.

Sun, Z., Pang, S., Cui, Y., and Yan, B. (2019). Genetic and functional variants analysis of the GATA6 gene promoter in acute myocardial infarction. Front. Genet. *10*, 1100.

Tan, M., Jiang, L., Li, Y., and Jiang, W. (2019). Dual inhibition of BMP and WNT signals promotes pancreatic differentiation from human pluripotent stem cells. Stem Cells Int. *2019*, 5026793.

Tang, R., Zhang, G., Wang, Y.C., Mei, X., and Chen, S.Y. (2017). The long non-coding RNA GAS5 regulates transforming growth factor beta (TGF-beta)-induced smooth muscle cell differentiation via RNA Smad-binding elements. J. Biol. Chem. *292*, 14270–14278.

Thomson, J.A., Itskovitz-Eldor, J., Shapiro, S.S., Waknitz, M.A., Swiergiel, J.J., Marshall, V.S., and Jones, J.M. (1998). Embryonic stem cell lines derived from human blastocysts. Science *282*, 1145–1147.

Tiyaboonchai, A., Cardenas-Diaz, F.L., Ying, L., Maguire, J.A., Sim, X., Jobaliya, C., Gagne, A.L., Kishore, S., Stanescu, D.E., Hughes, N., et al. (2017). GATA6 plays an important role in the induction of human definitive endoderm, development of the pancreas, and functionality of pancreatic beta cells. Stem Cell Reports *8*, 589–604.

Tsankov, A.M., Gu, H., Akopian, V., Ziller, M.J., Donaghey, J., Amit, I., Gnirke, A., and Meissner, A. (2015). Transcription factor binding dynamics during human ES cell differentiation. Nature *518*, 344–349.

Wamaitha, S.E., and Niakan, K.K. (2018). Human pre-gastrulation development. Curr. Top. Dev. Biol. *128*, 295–338.

Wang, L., Park, H.J., Dasari, S., Wang, S., Kocher, J.P., and Li, W. (2013). CPAT: coding-Potential Assessment Tool using an alignment-free logistic regression model. Nucleic Acids Res. *41*, e74.

Wang, P., Luo, M.L., Song, E., Zhou, Z., Ma, T., Wang, J., Jia, N., Wang, G., Nie, S., Liu, Y., et al. (2018). Long noncoding RNA lncTSI inhibits renal fibrogenesis by negatively regulating the TGF-beta/Smad3 pathway. Sci. Transl Med. *10*, eaat2039.

Xin, R., Everett, L.J., Lim, H.W., Patel, N.A., Schug, J., Kroon, E., Kelly, O.G., Wang, A., D'Amour, K.A., Robins, A.J., et al. (2013). Dynamic chromatin remodeling mediated by polycomb proteins orchestrates pancreatic differentiation of human embryonic stem cells. Cell Stem Cell *12*, 224–237.

Yiangou, L., Ross, A.D.B., Goh, K.J., and Vallier, L. (2018). Human pluripotent stem cell-derived endoderm for modeling development and clinical applications. Cell Stem Cell *22*, 485–499.

Zhu, P., Wu, J., Wang, Y., Zhu, X., Lu, T., Liu, B., He, L., Ye, B., Wang, S., Meng, S., et al. (2018). LncGata6 maintains stemness of intestinal stem cells and promotes intestinal tumorigenesis. Nat. Cell Biol. *20*, 1134–1144.

Zorn, A.M., and Wells, J.M. (2009). Vertebrate endoderm development and organ formation. Annu. Rev. Cell Dev. Biol. *25*, 221–251.