Cell Stem Cell
Resource

Efficient Endoderm Induction from Human Pluripotent Stem Cells by Logically Directing Signals Controlling Lineage Bifurcations

Kyle M. Loh,[1,2,11,*] Lay Teng Ang,[1,11,*] Jingyao Zhang,[1,12] Vibhor Kumar,[1,12] Jasmin Ang,[1] Jun Qiang Auyeong,[1] Kian Leong Lee,[3] Siew Hua Choo,[1] Christina Y.Y. Lim,[1] Massimo Nichane,[1] Junru Tan,[1] Monireh Soroush Noghabi,[1] Lisa Azzola,[4] Elizabeth S. Ng,[4,5] Jens Durruthy-Durruthy,[2] Vittorio Sebastiano,[2] Lorenz Poellinger,[3] Andrew G. Elefanty,[4,5] Edouard G. Stanley,[4,5] Qingfeng Chen,[6,7,8] Shyam Prabhakar,[1] Irving L. Weissman,[2] and Bing Lim[1,9,10,*]

[1]Stem Cell and Developmental Biology Group, Genome Institute of Singapore, Agency for Science, Technology and Research (A*STAR), Singapore 138672, Singapore

[2]Department of Developmental Biology, Stanford Institute for Stem Cell Biology and Regenerative Medicine, Stanford University School of Medicine, Stanford, CA 94305, USA

[3]Cancer Science Institute of Singapore, Centre for Translational Medicine, National University of Singapore, Singapore 117599, Singapore

[4]Department of Anatomy and Developmental Biology, Monash University, Clayton, VIC 3800, Australia

[5]Murdoch Childrens Research Institute, The Royal Children's Hospital, Parkville, VIC 3052, Australia

[6]Humanized Mouse Unit, Institute of Molecular and Cell Biology, A*STAR, Singapore 138673, Singapore

[7]Interdisciplinary Research Group in Infectious Diseases, Singapore-MIT Alliance for Research and Technology, Singapore 138602, Singapore

[8]Department of Microbiology, Yong Yoo Lin School of Medicine, National University of Singapore, Singapore 119228, Singapore

[9]Department of Medicine, Harvard Medical School, Boston, MA 02115, USA

[10]Division of Hematology/Oncology, Beth Israel Deaconess Medical Center, Boston, MA 02115, USA

[11]These authors contributed equally to this work

[12]These authors contributed equally to this work

*Correspondence: kyleloh@stanford.edu (K.M.L.), anglt1@gis.a-star.edu.sg (L.T.A.), limb1@gis.a-star.edu.sg (B.L.)

http://dx.doi.org/10.1016/j.stem.2013.12.007

SUMMARY

Human pluripotent stem cell (hPSC) differentiation typically yields heterogeneous populations. Knowledge of signals controlling embryonic lineage bifurcations could efficiently yield desired cell types through exclusion of alternate fates. Therefore, we revisited signals driving induction and anterior-posterior patterning of definitive endoderm to generate a coherent roadmap for endoderm differentiation. With striking temporal dynamics, BMP and Wnt initially specified anterior primitive streak (progenitor to endoderm), yet, 24 hr later, suppressed endoderm and induced mesoderm. At lineage bifurcations, cross-repressive signals separated mutually exclusive fates; TGF-β and BMP/MAPK respectively induced pancreas versus liver from endoderm by suppressing the alternate lineage. We systematically blockaded alternate fates throughout multiple consecutive bifurcations, thereby efficiently differentiating multiple hPSC lines exclusively into endoderm and its derivatives. Comprehensive transcriptional and chromatin mapping of highly pure endodermal populations revealed that endodermal enhancers existed in a surprising diversity of "pre-enhancer" states before activation, reflecting the establishment of a permissive chromatin landscape as a prelude to differentiation.

INTRODUCTION

At developmental junctures, multipotent progenitors choose between multiple fates (Graf and Enver, 2009; Loh and Lim, 2011). Extrinsic signals often instruct a particular fate while repressing alternate lineages. It is critical to decipher the extrinsic signals that direct such lineage segregations in order to efficiently differentiate human pluripotent stem cells (hPSCs) into pure populations of desired cell types in the absence of mutually exclusive, unwanted lineages. However, the precise lineage outcomes specified by these signals at particular bifurcations remain to be fully clarified, despite informative insights from in vivo genetic perturbations (Tam and Loebel, 2007; Zorn and Wells, 2009) and explant approaches (Bernardo et al., 2011; Deutsch et al., 2001). Pertinent issues include how alternate lineages are segregated at each branchpoint as well as the exact order and kinetics of dynamic signaling switches that drive successive cell fate transitions (Wandzioch and Zaret, 2009).

The present work revisits signaling dynamics that drive induction and anterior-posterior patterning of the definitive endoderm (DE) germ layer and subsequent organ formation. DE is the embryonic precursor to organs including the thyroid, lungs, pancreas, liver, and intestines (Švajger and Levak-Svajger, 1974). The pluripotent epiblast (E5.5 in mouse embryogenesis) differentiates into the anterior primitive streak (E6.5), which generates DE (E7.0–E7.5) (Lawson et al., 1991; Tam and Beddington, 1987). DE is then patterned along the anterior-posterior axis into distinct foregut, midgut, and hindgut territories (E8.5), and endoderm organ primordia arise

from specific anteroposterior domains (E9.5) (Zorn and Wells, 2009).

Various methods to differentiate hPSCs toward DE employ animal serum, feeder coculture, or defined conditions (Cheng et al., 2012; D'Amour et al., 2005; Touboul et al., 2010), but they typically yield a mixture of DE and other contaminating lineages, with induction efficiencies fluctuating between hPSC lines (Cohen and Melton, 2011; McKnight et al., 2010). Viewed from the perspective of lineage bifurcations, these mixed lineage outcomes might stem from incomplete exclusion of alternate fates at such junctures. Heterogeneous early DE populations harboring contaminating lineages complicate the subsequent generation of endodermal organ derivatives (McKnight et al., 2010).

In vertebrate embryos and during PSC differentiation, TGF-β/nodal/activin signaling is imperative for DE specification, whereas BMP broadly induces mesodermal subtypes (e.g., Bernardo et al., 2011; D'Amour et al., 2005; Dunn et al., 2004). Yet, TGF-β signaling (even with additional factors) is insufficient to specify homogeneous DE (quantified by Chetty et al., 2013). BMP, fibroblast growth factor (FGF), VEGF, and Wnt have also been employed together with TGF-β signals to generate DE (Cheng et al., 2012; Green et al., 2011; Kroon et al., 2008; Nostro et al., 2011; Touboul et al., 2010). However, these factors have also been implicated in mesoderm formation (Davis et al., 2008), and their precise involvement in DE induction remains to be clarified.

We have systematically elucidated how mutually exclusive lineages are separated at four consecutive steps of endoderm development: PS induction, segregation of endoderm versus mesoderm germ layers, DE anterior-posterior patterning, and bifurcation of liver and pancreas. Accurately defining which signals instructed or repressed specific fates at each endodermal bifurcation enabled homogeneous hPSC differentiation down one path or the other. Knowledge of precise temporal signaling dynamics, combined with efficient differentiation throughout successive developmental steps, culminated in a single strategy to universally differentiate diverse hPSC lines into pure populations of endodermal lineages by excluding alternate lineages at each branchpoint. Altogether, this provides a coherent view of signaling logic underlying multiple steps of endoderm induction and patterning. This also furnishes the means to molecularly profile highly homogeneous endoderm populations, allowing us to comprehensively capture transcriptional and chromatin dynamics underlying endoderm specification.

RESULTS

BMP, FGF, TGF-β, and Wnt Initially Establish the Primitive Streak and Anteroposteriorly Pattern It

This work was preceded by findings that activin, in conjunction with FGF, BMP, and a phosphatidinositol 3-kinase (PI3K) inhibitor ("AFBLy") (Touboul et al., 2010) or together with animal serum (D'Amour et al., 2005), induced hESCs toward DE. However, we and others (Chetty et al., 2013) observed that these methods still yielded mixed lineage outcomes, which was evident during the differentiation of five hESC lines (Figures 1A and 2B and 2C; Figures S1–S3 available online). For example, AFBLy (Touboul et al., 2010) concurrently generated mesoderm, upregulating skeletal, vascular, and cardiac genes ($p < 10^{-8}$; Figure 1A; Figures S1A–S1D), whereas activin and serum treatment (D'Amour et al., 2005) yielded a proportion of undifferentiated cells (Figures 2B, 2C, and 2F). Creation of impure early DE populations might explain the emergence of nonendoderm lineages after downstream differentiation (Kroon et al., 2008; Rezania et al., 2012).

Guided by prior in vivo and in vitro findings, we selectively perturbed developmental signals (>3,200 signaling conditions) at specific embryonic stages of hPSC differentiation in serum-free conditions and assessed resultant lineage outcomes by qPCR (yielding >16,000 data points, Figure S1–S4). These signaling perturbations revealed elements of the signaling logic underlying DE induction (Figures 1, 2, 3, and 4).

In vivo, DE arises from the primitive streak (PS, E6.5) (Lewak-Svajger and Švajger, 1974). The anteriormost PS (APS) generates DE (E7.0–E7.5), whereas posterior PS (PPS) forms mesoderm (Lawson et al., 1991; Tam and Beddington, 1987). Determinants of anterior versus posterior PS from hPSCs remain to be elucidated.

We found both APS and PPS were combinatorially induced by BMP, FGF, and Wnt on day 1 of hESC differentiation. These signals have been individually implicated in PS induction (Bernardo et al., 2011; Blauwkamp et al., 2012; Gadue et al., 2006), but their roles in PS patterning have not been dissected in detail. If BMP, FGF, or Wnt were inhibited, both APS and PPS formation failed (Figure 1B), corroborating the lack of PS in BMP and Wnt pathway knockout mice (Beppu et al., 2000; Liu et al., 1999; Mishina et al., 1995). FGF signaling was equally permissive for both APS and PPS emergence, and endogenous FGF was sufficient to drive either outcome (Figure 1Bi, Figures S2A–S2C). However, exogenous Wnt (either Wnt3a or GSK3 inhibition [CHIR]) was necessary to maximize PS induction, and Wnt broadly promoted both APS and PPS (Figure 1Bii and 1Biii). Limited PS formation could occur without exogenous Wnt, but was dependent on endogenous Wnt (Figure 1Bii). BMP levels arbitrated between APS and PPS; lower (endogenous) BMP levels elicited APS, whereas higher BMP yielded PPS (Figure 1Biv; Figure S2B). Nonetheless, the absolute necessity of BMP for *MIXL1-GFP*$^{+}$ APS induction (Figure 1Di, $p < 0.025$) was unexpected, because BMP was typically thought to be posteriorizing (Bernardo et al., 2011). Therefore, FGF, Wnt, and low BMP were essential for APS specification.

A Dynamic Switch in BMP and Wnt Signaling Induces Primitive Streak but Subsequently Suppresses DE Emergence

To further differentiate APS toward DE, prior studies used similar factors to induce both lineages over 3–5 days (Nostro et al., 2011; Touboul et al., 2010). Instead, we found that APS and DE were sequentially driven by diametrically opposite signals within 24 hr of differentiation. BMP and Wnt initially specified APS from hESCs on day 1, but, 24 hr later, BMP and Wnt induced mesoderm and reciprocally repressed DE formation from PS on days 2–3 (Figures 1Ci and 1Cii). Interestingly, not only removing exogenous BMP but neutralizing endogenous BMP (using noggin or DM3189/LDN-193189) was critical to eliminate mesoderm and to reciprocally divert PS differentiation toward DE (Figure 1Ci). This was evinced by 3,000-fold downregulation of *MESP1* and concurrent upregulation of *SOX17*, *HHEX*, *FOXA1*, and *FOXA2* in two hESC lines (Figures S1C–S1E). Given that

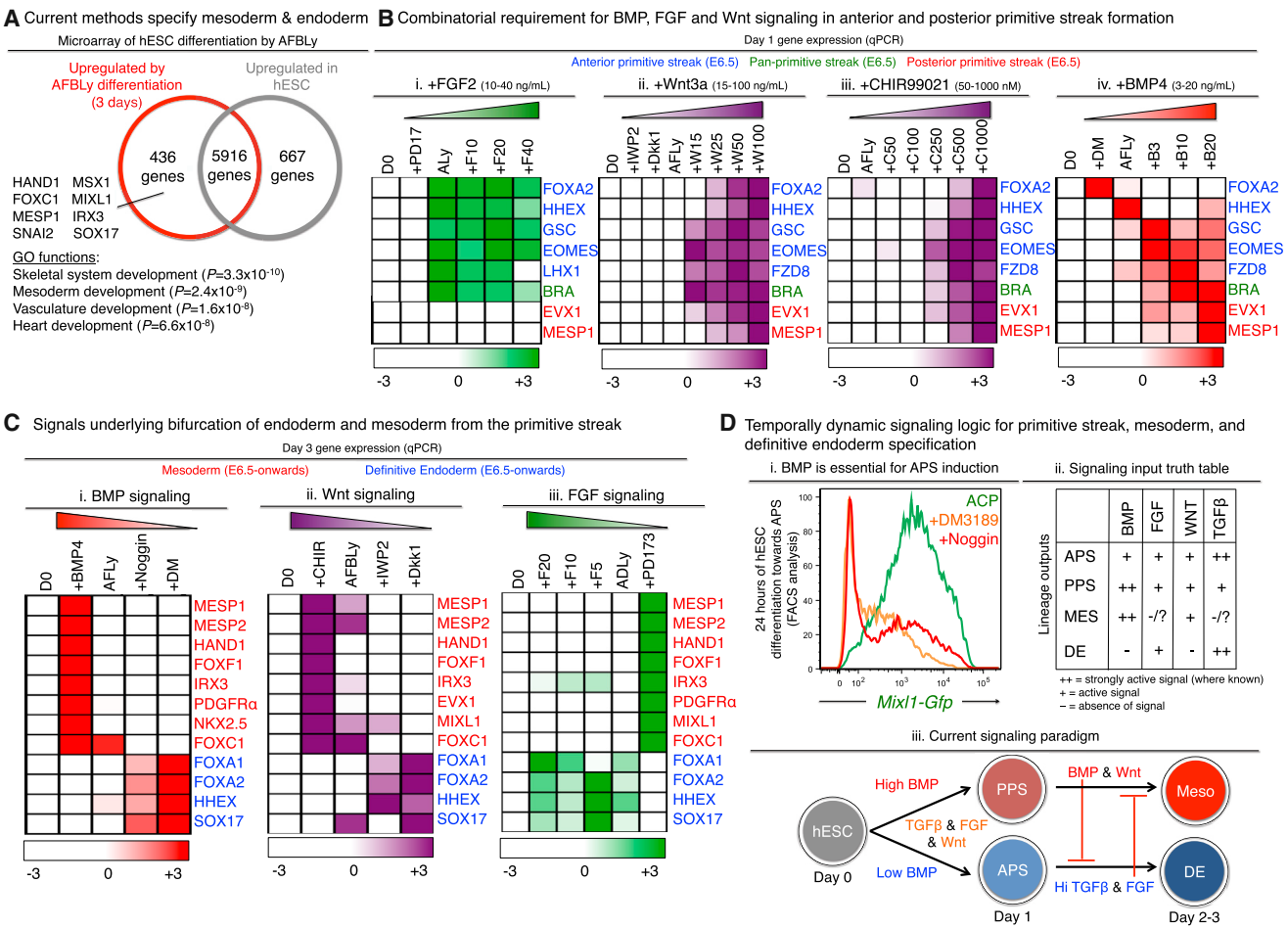

Figure 1. Dynamic Signaling Switch for Primitive Streak and Endoderm Formation

(A) Microarray analysis of genes upregulated >2-fold during AFBLy treatment of H9 hESC (Touboul et al., 2010) and GO analysis.

(B) To test effects of increasing FGF2 (10–40 ng/ml), Wnt3a (15–100 ng/ml), CHIR99021 (50–1,000 nM), or BMP4 (3–20 ng/ml) (panels Bi, Bii, Biii, and Biv, respectively) and respective inhibitors (100 nM PD173074, 2 μM IWP2, 150 ng/ml Dkk1 and 250 nM DM3189) on PS formation, H1 hESCs were differentiated toward PS for 24 hr with indicated base combinations of *a*ctivin (100 ng/ml), *F*GF2 (20 ng/ml) and 10 μM *LY*294002 ("AFLy" or "ALy") in conjunction with the indicated signaling perturbations, and qPCR was performed (day 1).

(C) To test effects of increasing BMP, FGF, or Wnt signaling (10 ng/ml BMP4, 3 μM CHIR, and 5–20 ng/ml FGF2; panels Ci, Cii, and Ciii, respectively) on DE versus mesoderm emergence from PS, H1 hESCs were initially differentiated with AFBLy toward PS for 24 hr and then subsequently differentiated with AFLy, AFBLy, or *ALy* + 250 nM *D*M3189 ("ADLy") for 48 subsequent hours with indicated signaling perturbations, and qPCR was performed (day 3).

(D) HES3 *MIXL1-GFP*$^{+}$ PS (Davis et al., 2008) induced by 100 ng/ml *a*ctivin, 2 μM *C*HIR, and 50 nM *P*I-103 (ACP) by day 1 of differentiation was blocked by concomitant addition of (Di) BMP inhibitors (300 ng/ml noggin or 250 nM DM3189), (Dii) truth table, (Diii) and schematic of dynamic signaling during differentiation. See also Figure S1.

prolonged BMP and Wnt were known to induce mesoderm (Bernardo et al., 2011; Gadue et al., 2006), our results altogether argue against prior sustained BMP treatment to induce DE (Cheng et al., 2012; Nostro et al., 2011; Touboul et al., 2010), which we show abrogated DE and, instead specified mesoderm. Timed BMP inhibition also improved DE induction from mouse ESCs (mESCs), although which developmental step(s) it benefited remain unclear (Sherwood et al., 2011). In summary, understanding the precise kinetics of BMP signaling was essential to thwart extraneous mesoderm production.

Similarly, endogenous Wnt/β-catenin signals directed PS toward mesoderm, such that inhibiting endogenous Wnt (using IWP2, Dkk1, or XAV939) on days 2 and 3 blocked mesoderm formation from two hESC lines (Figure 1Cii; Figures S1G and S1H). However, individually inhibiting either BMP or Wnt was sufficient to abolish mesoderm, indicating that inhibiting both was redundant (Figure S1H). Thus, we subsequently only inhibited BMP to derive DE from PS. Finally, our results contrast with prolonged Wnt treatment to induce DE (Sumi et al., 2008), which we show specified mesoderm from PS and blocked DE instead. In summary, BMP and Wnt induced mesoderm from PS and suppressed endoderm; therefore, their inhibition ablated mesoderm and diverted differentiation toward DE.

Whereas BMP and Wnt specified mesoderm (Gertow et al., 2013), we found DE formation from PS was jointly driven by FGF (Figure 1Ciii) and TGF-β signaling (Bernardo et al., 2011; D'Amour et al., 2005). If FGF was inhibited, mesoderm formation was re-enabled, even in the absence of promesodermal BMP

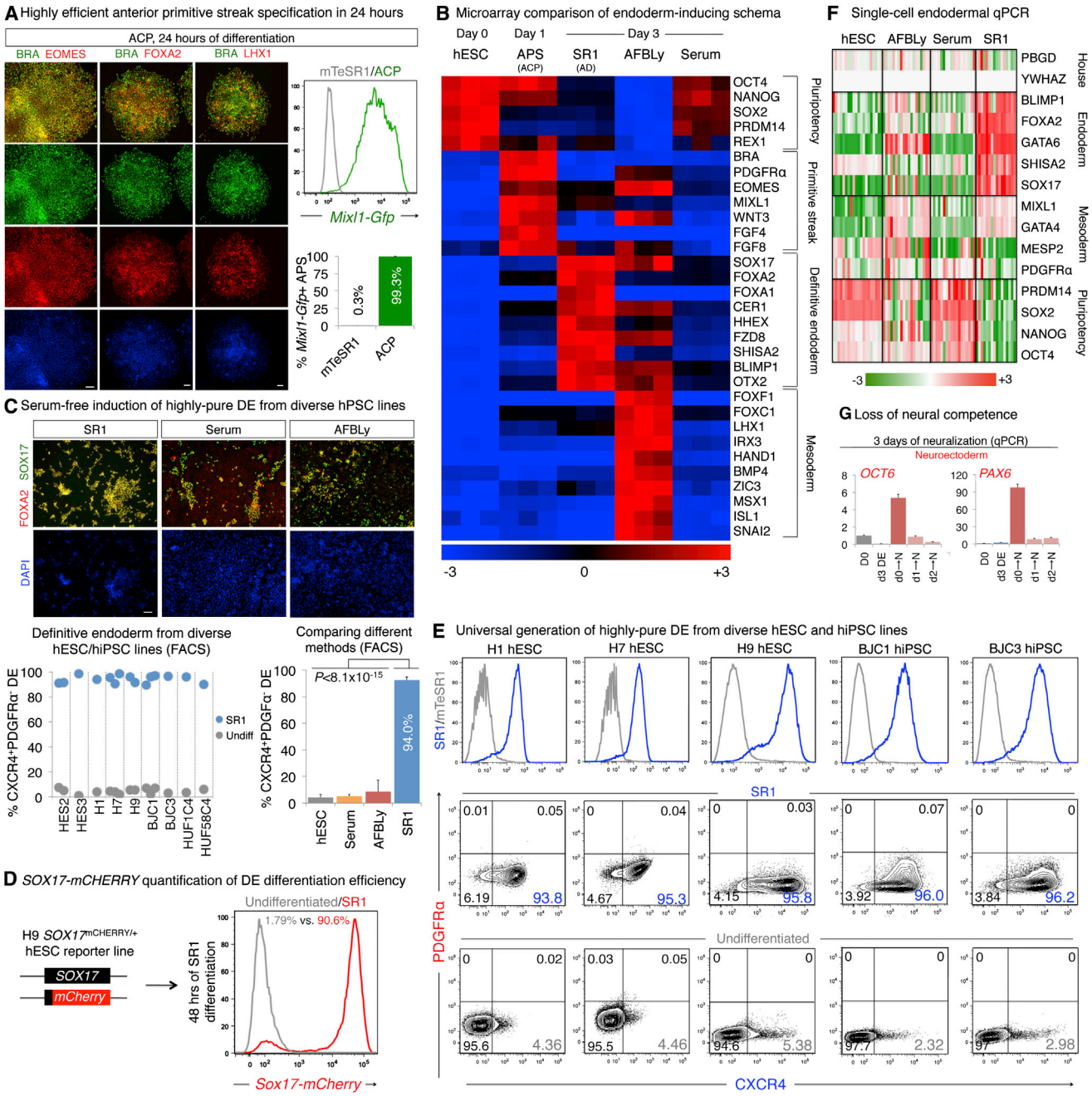

Figure 2. Efficient DE Induction in Defined Conditions by SR1

(A) H1 hESCs differentiated by ACP for 24 hr stained for BRACHYURY, FOXA2, EOMES, and LHX1 (nuclear staining by DAPI); scale bar, 100 μm for all subsequent figures (left); FACS shows >99% of HES3 hESCs are *MIXL1-GFP*$^+$ (Davis et al., 2008) after 24 hr of ACP treatment (right).

(B) Microarray heatmap of independent triplicates; undifferentiated HES3 hESCs (day 0), ACP-induced APS (day 1), SR1-induced DE (day 3), or hESC differentiated by AFBLy (Touboul et al., 2010) or serum (D'Amour et al., 2005) for 3 days.

(C) FOXA2 and SOX17 staining of SR1-, serum-, or AFBLy-treated H1 hESCs after 3 days of differentiation (top); summary of CXCR4$^+$PDGFRα$^-$ DE percentages in hPSCs (gray) or after SR1 differentiation (blue) from seven hPSC lines, dots depict experimental replicates (bottom left); histogram summarizing CXCR4$^+$PDGFRα$^-$ DE percentages after various differentiation protocols, error bars depict standard deviation (bottom right).

(D) FACS analysis of H9 *SOX17-mCHERRY* hESCs; reporter expression before or after 2 days of SR1 differentiation.

(E) FACS analysis of CXCR4 and PDGFRα expression before or after SR1 differentiation from indicated hPSC lines.

(F) Single-cell qPCR heatmap of 80 single cells (H7 hESCs, or those differentiated by SR1, AFBLy or serum for 2 days).

(G) To test their neural competence, H1 hESCs, after 0–2 days of SR1 induction, were transferred ("→") into neuralizing media ("N," 3 days), and neural gene expression was compared to SR1-induced DE ("day 3 DE"); see Supplemental Experimental Procedures.

See also Figures S2 and S3.

(Figure 1Ciii), showing that FGF prevented illegitimate conversion of prospective DE to mesoderm. FGF is also essential for DE formation from mESCs, yet, paradoxically, it was previously found that exogenous FGF was detrimental to DE induction (Hansson et al., 2009), which we did not observe (Figure 1Ciii).

In conclusion, these data uncovered a signaling cross antagonism in which BMP and Wnt versus FGF and TGF-β respectively induced mesoderm versus endoderm from the PS and did so by cross-repressing the alternate fate (Figures 1Dii–1Diii). Furthermore, BMP and Wnt yielded dichotomous lineage outcomes, depending on the developmental time of exposure; their effects became reversed within 24 hr (Figure 1D; Figure S1F).

Universal Generation of Highly Purified DE from Diverse hPSC Lines through Sequential APS Formation and Mesoderm Suppression

The above findings that APS and DE were sequentially specified by opposing signals, together with the necessity of BMP inhibition to eliminate mesoderm from the PS, motivated a serum-free monolayer approach ("SR1") for DE induction. We first differentiated hPSC to APS in 24 hr (Figure 2A) while excluding ectoderm by combining high activin/TGF-β with CHIR (emulating Wnt/β-catenin signaling) and PI3K/mTORC inhibition (Figures S2C–S2E), abbreviated "ACP." This yielded a 99.3% ± 0.1% *MIXL1-GFP*$^+$ PS population (Davis et al., 2008) in which pan-PS TF BRACHYURY was coexpressed with APS-specific TFs EOMES, FOXA2 and LHX1 (Figure 2A; Figure S2H). Twenty-four hours later, CHIR was withdrawn and APS was subsequently differentiated into DE for forty-eight hours by high activin concomitant with BMP blockade (DM3189) to exclude mesoderm. Exogenous FGF was superfluous as endogenous FGF sufficed (Figure 1Ciii; Figure S2A).

Sequential APS formation followed by DE induction universally yielded a 94.0% ± 3.1% CXCR4$^+$PDGFRα$^-$ DE population from 9 diverse hESC (H1, H7, H9, HES2, and HES3) and hiPSC (BJC1, BJC3, HUF1C4, and HUF58C4) lines by day 3 of differentiation (Figures 2B–2E; Figure S2I), overcoming line-to-line induction variability. SR1 abundantly elicited SOX17$^+$FOXA2$^+$ DE (Figure 2C; Figure S3C) and downregulated hPSC marker CD90 (Figure S2J). hESC (94.0% ± 3.1%) and hiPSC (94.0% ± 3.4%) did not markedly differ in DE induction efficiencies ($p > 0.97$, Figure S3D). We further exploited a *SOX17-mCHERRY* knockin hESC reporter line to quantify differentiation efficiencies and found SR1 induced >90% SOX17-mCHERRY$^+$ DE (Figure 2D). SR1 generated definitive instead of extraembryonic endoderm (ExEn) as evinced by lack of PDGFRα and *SOX7* (Figure 2E; Figure S3A).

We directly compared DE induction by SR1 against two prevailing protocols, AFBLy (Touboul et al., 2010) or activin and serum treatment (D'Amour et al., 2005), across five hESC lines (Figure S3A). SR1 differentiation exclusively yielded DE (*SOX17*, *FOXA1*, *FOXA2*, *CER1*, and *FZD8*) from all five hESC lines with minimal mesoderm, neuroectoderm, or ExEn (Figure 2B; Figure S3A). In contrast, the other DE protocols generated modest amounts of SOX17$^+$FOXA2$^+$ DE (Figure 2B and 2C; Figure S3C) and produced mixed lineage outcomes; AFBLy upregulated mesoderm TFs (*FOXF1*, *HAND1*, *MSX1*, and *ISL1*), whereas pluripotency TF expression (*OCT4*, *SOX2*, and *NANOG*) persisted after serum induction across all five lines (Figure 2B; Figure S3A). At a clonal level, both FACS quantification (Figure 2C; Figure S3B) and single-cell qPCR (Figure 2F) confirmed SR1 yielded purer DE than either AFBLy or serum treatment; 20/20 of SR1-differentiated cells were *FOXA2*$^+$, whereas few cells after AFBLy (1/20 cells) or serum induction (2/20 cells) highly expressed *FOXA2* (Figure 2F). Thus, even though all three differentiation protocols utilized high activin, clearly, activin alone was insufficient to generate pure DE.

Finally, neural competence was relinquished within 24 hr of SR1 induction (Figure 2G), showing that mutually exclusive ectoderm potential was lost upon APS commitment.

Anteroposterior Patterning of hESC-Derived DE into Mutually Exclusive AFG, PFG, and MHG domains by BMP, FGF, RA, TGF-β, and Wnt Signaling

After its initial specification in vivo, DE is patterned along the anteroposterior axis into distinct domains, which are the regional antecedents to endodermal organs (Zorn and Wells, 2009). The anterior foregut (AFG) gives rise to lungs and thyroid, the posterior foregut (PFG) to pancreas and liver, and the midgut/hindgut (MHG) to small and large intestines (Figures 3A and 3B). Therefore, having induced mostly homogeneous DE from hPSCs by day 3, we next attempted to anteroposteriorly pattern it into distinct AFG, PFG, or MHG populations by 4 subsequent days of differentiation (Figure 3A) based on increasing knowledge of signals controlling DE patterning in vivo (Zorn and Wells, 2009) and in vitro (e.g., Green et al., 2011; Sherwood et al., 2011; Spence et al., 2011).

In vertebrate embryos, tailbud mesoderm expresses *BMP4*, *FGF4/8*, and *WNT3A* and is juxtaposed with posterior endoderm, suggesting these signals might posteriorly pattern the nearby MHG. In vitro, we found BMP markedly posteriorized DE (Figure 3Ci) by inducing MHG TFs (e.g., *CDX2*, *EVX1*, and 5′ *HOX* genes), congruent with zebrafish data (Tiso et al., 2002). Wnt (emulated by CHIR) was similarly posteriorizing (Figure 3Cii), and FGF could also partially posteriorize PFG into MHG (Figure S4A), confirming prior work (Sherwood et al., 2011; Spence et al., 2011). BMP, FGF, and Wnt all reciprocally suppressed anterior endoderm TF *SOX2* (Figure 3C; Figure S4A). Hence, we used a combination of BMP, CHIR, and FGF to pattern day 3 DE into >99% CDX2$^+$ MHG (Figure 3D) while suppressing foregut (Figure 3E) in serum-free conditions.

Conversely, inhibiting posteriorizing BMP signals broadly yielded anterior endoderm (foregut). Combining BMP inhibition with TGF-β inhibition (Green et al., 2011) yielded >98% OTX2$^+$ AFG pharyngeal endoderm (Figure 3D) by day 7 of differentiation. Separately, BMP inhibition in conjunction with RA signaling generated PFG (Figures 3E and 3F), consistent with how RA regionalizes the PFG in vivo (Stafford and Prince, 2002). AFG and PFG were functionally distinct, because only PFG was competent to subsequently form liver and pancreas (Figure 3G).

Invoking the above signaling logic, we generated separate AFG, PFG, and MHG populations from DE in a mutually exclusive manner. Global microarray profiling of distinct patterned populations revealed that anteroposterior marker expression was clearly developmentally demarcated (Figure 3E and 3F, reproduced in two hESC lines). Graded, spatially collinear *HOX* gene expression (Zorn and Wells, 2009) was observed after in vitro patterning, whereby PFG expressed 3′ anterior *HOX* genes

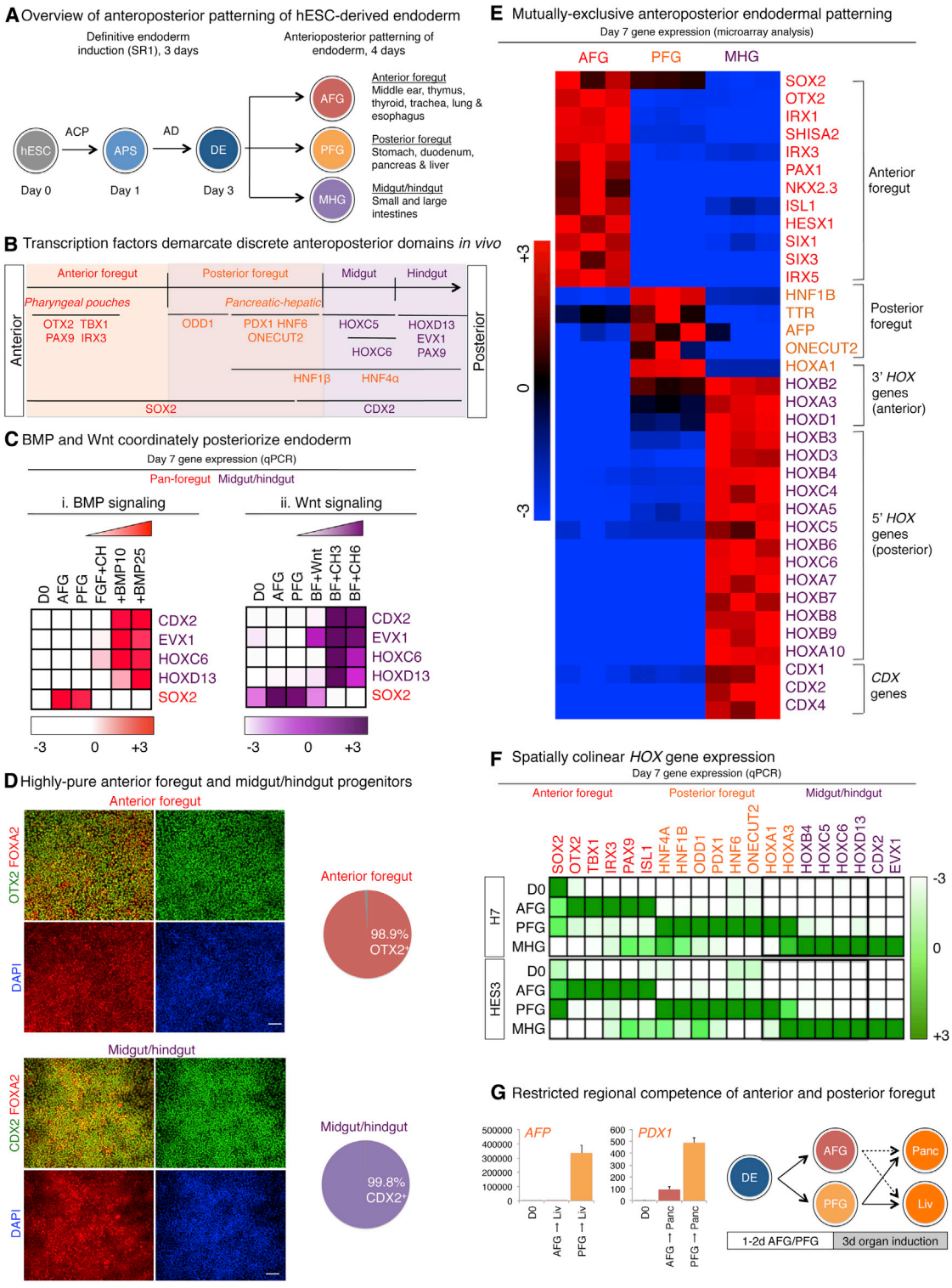

Figure 3. Anteroposterior Patterning of hESC-Derived DE

(A) Overview of anteroposterior patterning strategy.

(B) TF expression in anteroposteriorly patterned endoderm in vivo, see Table S1.

(C) To test effects of (Ci) increasing BMP4 (10–25 ng/ml) or (Cii) increasing CHIR (3–6 μM) on MHG induction, day 3 DE was differentiated for 4 subsequent days with indicated base conditions together with designated signaling perturbations until day 7, with AFG and PFG controls indicated (subsumed by Figure S4A); (Ci) FGF+CHIR, 100 ng/ml FGF2 + 3 μM CHIR; (Cii) BF, 10 ng/ml BMP4 + 100 ng/ml FGF2.

(D) OTX2, FOXA2, and CDX2 staining of H1-derived day 7 AFG and MHG, respectively, with quantification.

(legend continued on next page)

(e.g., *HOXA1*), but MHG exclusively expressed 5′ posterior *HOX* genes and *CDX* genes (Figures 3E and 3F).

TGF-β Competes with BMP/MAPK Signaling to Specify Mutually Exclusive Bifurcation of Pancreatic and Hepatic Fates

In vivo, liver and pancreas develop from a common PFG precursor (Chung et al., 2008; Deutsch et al., 2001). During PSC differentiation, BMP and FGF are typically used to induce liver, whereas Hedgehog inhibition and FGF are applied to generate pancreas (e.g., Cho et al., 2012; Kroon et al., 2008). We executed a signaling analysis encompassing >500 conditions (Figure 4A; Figure S4B) to clarify how pancreas versus liver might be segregated in a mutually exclusive way (Figure 4B).

We found TGF-β signaling promoted *PDX1*$^+$ pancreas formation, whereas BMP and FGF/MAPK signaling specified *AFP*$^+$ liver (Figure 4A). Importantly, we clarified that each of these signals reciprocally repressed formation of the alternate lineage (Figure 4A), explaining why the PFG lineage decision is bistable (Chung et al., 2008). Due to such cross repression, eliminating propancreatic TGF-β reciprocally expanded liver (Figures 4Ai and 4Aii), whereas inhibition of prohepatic FGF/MAPK (Deutsch et al., 2001) diverted differentiation toward pancreas (Figure 4Aiv). Our findings differ from prior work and may explain previous inefficiencies in liver or pancreas induction. Prior use of FGF for pancreatic induction (Cho et al., 2012; Kroon et al., 2008; Nostro et al., 2011) may, in fact, block pancreas and instead specify liver (Figure 4Aiv), as suggested by embryonic studies (Deutsch et al., 2001). On the other hand, provision of TGF-β for hepatic induction (Cho et al., 2012) may abrogate liver and, instead drive pancreas (Figures 4Ai and 4Aii).

In summary, a dichotomy in TGF-β versus BMP in respectively specifying pancreas versus liver (Figure 4B) has not been, to our knowledge, previously elucidated and is reminiscent of how these signaling pathways often cross repress each other's transduction (Candia et al., 1997). We further identified combinatorial interactions between these morphogens. For example, TGF-β signaling *and* MAPK inhibition was essential for pancreas formation, because MAPK inhibition was ineffective if TGF-β was inhibited in parallel (Figure S4Bi). Conversely, hepatic induction cooperatively required TGF-β inhibition *and* MAPK signaling (Figure 4Aiv; Figure S4Bi), because TGF-β inhibition failed to efficiently create liver if MAPK was simultaneously inhibited.

hESC-Derived Hepatic Progeny Engraft Long Term into Unconditioned Mouse Liver

To differentiate DE toward liver while explicitly inhibiting pancreas, we induced DE toward PFG for 1 day (Figure 4Bi; Figure S4Biv) and then employed BMP and other factors together with inhibition of propancreatic TGF-β signaling to direct PFG toward liver over 3 subsequent days with minimal pancreatic contamination (Figure S4C). We generated 72.3% ± 6.3% AFP$^+$ early hepatic progenitors (Figure 4C) from four hESC lines within 7 days of differentiation, which is twice as rapid as prior methods. Moreover liver markers were induced 60–210 times higher than what was obtainable with earlier protocols (Figure S4D).

To validate the hepatic potential of early AFP$^+$ liver progenitors, they were empirically matured in vitro with oncostatin M and dexamethasone (Kamiya et al., 1999) into a mixed albumin (hALB)$^+$ hepatoblast population (Figure S5A), which exhibited some CYP3A4 metabolic activity (Figure 4Di), expressed low density lipoprotein receptor (LDLR) and could uptake cholesterol (Figure 4Dii). When transplanted into neonatal mouse livers, early AFP$^+$ hepatic progenitors failed to engraft (Figure S5B). However when their differentiated hALB$^+$ progeny were transplanted, human albumin (mean 15.1 ng/ml) was detected in the blood of 47% of recipients 2–3 months posttransplantation, indicating long-term engraftment (Figure 4E). Indeed, foci of hALB$^+$ hESC-derived hepatic cells (marked with constitutively expressed GFP prior to transplantation) were present in all lobes of the adult liver (Figure 4E; Figure S5B). This suggested hALB$^+$ hepatic cells had integrated and/or migrated throughout the liver and they were not simply locally persisting at the site of transplantation. Finally, hALB$^+$ cells coexpressed human hepatic marker HepPar1 (Figure S5C), but did not detectably express fetal marker AFP (Figure S5D), suggesting they had progressed past the fetal stage. To our knowledge, this is one of the first demonstrations that hESC-derived hepatic cells could engraft long term into normal mouse livers that were not compromised by extensive pharmacologic or genetic damage (cf. Yusa et al., 2011).

Comprehensive Transcriptional and Chromatin State Mapping of Endoderm Induction and Anteroposterior Patterning

Capitalizing on our ability to obtain rather homogenous populations of hESC-derived endodermal lineages, we captured genome-wide transcriptional and chromatin dynamics during endoderm development by profiling a hierarchy of six pure progenitor populations (hESC, APS, DE, AFG, PFG, and MHG) using RNA sequencing (RNA-seq) and chromatin immunoprecipitation sequencing (ChIP-seq) for four histone H3 modifications (K4me3, K27me3, K27ac, and K4me2; Figures 5, 6, and 7; Figures S5–S7). This yielded 30 transcriptional and chromatin state maps spanning four embryonic stages (epiblast, PS, DE, and anteroposterior patterning), totaling >1.3 billion aligned reads (Figure S5E) and providing a global view of molecular events driving endoderm development.

Our analyses captured acute developmental transitions. RNA-seq revealed dramatic transcriptional changes within 24 hr during synchronous transit from pluripotency to APS in vitro (Figure 5A), mirroring how epiblast (E5.5) and PS (E6.5) arise within 1 day in the mouse. The *BRACHYURY* and *NODAL* promoters were bivalently marked by activation-associated K4me3 and repression-associated K27me3 in hESCs. Yet, within 24 hr of APS induction, they were unilaterally resolved, losing repressive K27me3 and gaining active marks K27ac and K4me3

(E) Microarray heatmap of HES3-derived AFG, PFG, and MHG populations on day 7 in independent triplicate.
(F) qPCR of day 7 AFG, PFG, and MHG populations from H7 and HES3 hESC lines; *HOX* genes boxed.
(G) To test their pancreatic or hepatic competence, day 3 DE was patterned into AFG or PFG for 1–2 days, and each was then subsequently differentiated toward pancreas or liver for 3 further days; see Supplemental Experimental Procedures.
See also Figure S4.

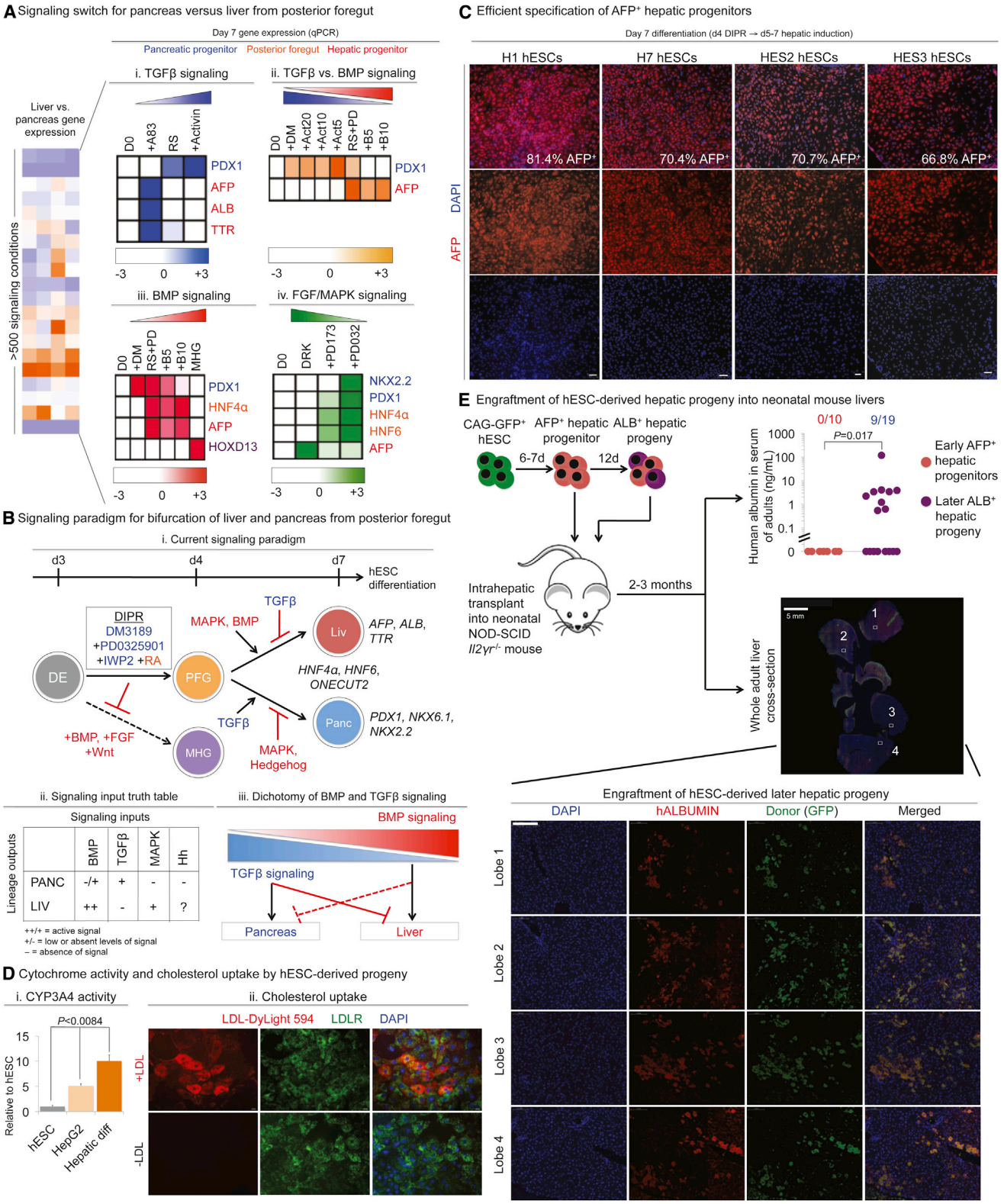

Figure 4. Bifurcation of Liver versus Pancreas from Posterior Foregut

(A) To test effects of increasing amounts of (Ai–Aiii) BMP/TGF-β signaling or (Aiv) FGF/MAPK signaling on pancreas versus liver induction, day 3 DE was differentiated with indicated conditions with (Ai and Aii) 5–20 ng/ml activin or (Aii and Aiii) 5–10 ng/ml BMP4 and respective inhibitors (1 μM A8301, 250 nM DM3189, 100 nM PD173074, 500 nM PD0325901) where indicated. Abbreviations for base conditions: (Ai) RS, 2 μM RA + SANT1; (Aii and Aiii) RS+PD, RS + PD0325901; (Aiv) DRK, DM3189 + RA + KAAD-cyclopamine.

(legend continued on next page)

concomitant with rapid *BRACHYURY* and *NODAL* upregulation in APS (Figure 5B).

Endoderm Enhancer Activation Is Associated with EOMES, SMAD2/3/4, and FOXH1 Co-occupancy

To map K27ac-marked active enhancers (Rada-Iglesias et al., 2011) throughout all six profiled lineages, we employed DFilter (Kumar et al., 2013) to identify distal elements with significant K27ac enrichment. Distinct batteries of active enhancers were invoked during each endodermal lineage transition (Figure 5C). APS enhancers (e.g., *BRACHYURY* and *NODAL*) were rapidly activated within 24 hr (Figure 5B). During DE patterning, distinct cohorts of enhancers were commissioned in each anteroposterior domain in AFG (*SIX1* and *TBX1*; Figure S7A), PFG (*HOXA1*; Figure S7B), and MHG (*CDX2* and *PAX9*; Figure 5D; Figure S5G).

Upon DE specification, 10,543 enhancers were activated (Table S5), gaining K27ac despite being largely inactive in hESCs. Active DE enhancers flanked archetypic DE regulators, e.g., *SOX17* (Figure 6G) and *CXCR4* (Figure S5F). Gene ontology analyses (McLean et al., 2010) associated these enhancers most significantly with endoderm development ($p < 3.84 \times 10^{-26}$) and gastrulation ($p < 7.92 \times 10^{-26}$; Figure 6A), affirming the purity of differentiated DE populations. Genes adjacent to active DE enhancers were upregulated in gastrula-stage endoderm in vivo ($p < 1.38 \times 10^{-39}$, Figure 6A) and upon DE differentiation in vitro (Figure 6B). Active DE enhancers coincided with euchromatic mark K4me2 (Figure S6A), were devoid of repression-associated K27me3 (Figure S6A), were evolutionarily conserved (Figure 6C), and were broadly inactive in other lineages (Figure S6B).

DE enhancers previously remained elusive, because most prior work only assessed promoter marks (Kim et al., 2011; Xie et al., 2013). However, enhancer profiling of hESC-derived DE was recently reported (Gifford et al., 2013), and, therefore, we compared our two DE data sets using identical analytic methods (Table S6). Unexpectedly, DE enhancers from the former data set (Gifford et al., 2013) were highly enriched for neural functions ($p < 3.93 \times 10^{-28}$; Figure 6D), because enhancers for neural TFs *BRN2* and *PAX3* were activated, but *SOX17* enhancers were virtually dormant (Figure S6C). Association of DE enhancers with neural genes led to the prior conclusion that endoderm and ectoderm development are related (Gifford et al., 2013), which contradicts the in vivo order of germ layer segregations (cf. Tzouanacou et al., 2009). By contrast, neural terms were largely absent in SR1-derived DE (Figure 6A), and, ultimately, only 4.8% of DE enhancers were shared between our and their data sets. Thus, molecular profiling of mixed DE populations (potentially enriched for ectoderm) may have precluded accurate molecular description of endoderm development.

How DE enhancers are inaugurated during differentiation remains obscure. Motifs for multiple TFs, including DE specifiers EOMES and FOXA2 as well as TGF-β signaling effectors SMAD2/3 and FOXH1 ($p = 10^{-59}$–10^{-197}), were enriched in DE enhancers (Figure 6E), which is consistent with how these TFs specify DE in vivo (e.g., Dunn et al., 2004; Teo et al., 2011). Interestingly, we found EOMES, SMAD2/3, SMAD4, and FOXH1 (Kim et al., 2011; Teo et al., 2011) co-occupied an extensive series of DE enhancers (Figure 6F), including a *SOX17* enhancer (Figure 6G). Although EOMES individually engaged some elements, colocalization of EOMES with TGF-β signaling effectors SMAD2/3/4 and FOXH1 correlated with maximal enhancer acetylation (Figure 6F, $p < 10^{-300}$). Thus, convergence of both lineage-specifying and signaling-effector TFs may propel full-fledged enhancer activation upon differentiation (Calo and Wysocka, 2013).

Endoderm Enhancers Reside in a Diversity of "Pre-enhancer" States in Uncommitted Cells prior to Activation

It remains unclear how DE enhancers are swiftly engaged upon hESC differentiation. SMAD2/3/4 and FOXH1 occupy DE enhancers upon differentiation, but infrequently do so in the uncommitted state (Figure S6A). Perhaps these enhancers are instead primed for activation at the level of chromatin. Premarking of developmental enhancers by euchromatic K4me1 in ESCs signifies a "window of opportunity" for subsequent enhancer activation (Calo and Wysocka, 2013; Rada-Iglesias et al., 2011). We looked back in developmental progression, assessing occupancy of DE enhancers by >24 histone modifications and chromatin regulators (Ernst et al., 2011) in hESCs prior to enhancer activation (Figure 7A). Unexpectedly, K4me1 labeled less than one-third of future DE enhancers in hESCs, implying that "poising" by K4me1 in hESCs is not always essential for immediate enhancer activation (Figures 7A and 7B). Thus, we sought to systematically discover all possible "pre-enhancer" chromatin states of DE enhancers in hESCs.

Unsupervised clustering revealed 25% of DE enhancers existed in a pre-enhancer state (cluster 1) in hESCs largely defined by histone variant H2AZ and no other known chromatin marks (Figure 7A; Figure S6D). Despite virtual absence of K4me1, H2AZ-marked pre-enhancers became rapidly activated within 3 days of DE induction (Figure 7A). DE enhancers less frequently resided in a repressed state designated by heterochromatic mark K9me3 (cluster 2) (Zhu et al., 2012) or a "latent" pre-enhancer state largely lacking known histone modifications (cluster 5, Figure 7A) (Ostuni et al., 2013). Only 10% of DE pre-enhancers were marked by K27me3 in hESCs (Figure 7B), suggesting Polycomb (Rada-Iglesias et al., 2011) was not always necessary to repress developmental enhancers in hESCs. Instead, perhaps the absence of K27ac/histone acetyltransferases (HATs) was sufficient to confer inactivity. Only a minority

(B) Depictions of (Bi) dynamic signaling inputs, (Bii) truth table and (Biii) dichotomy of BMP and TGF-β signaling for liver versus pancreas induction.
(C) AFP immunostaining of day 7 early liver progenitors and quantification.
(D) Substrate luciferase assay for CYP3A4 metabolic activity (Di) and staining for LDLR expression and LDL-DyLight 594 uptake (Dii) in hESC-derived late hepatic progeny.
(E) CAG-GFP$^+$ hESC differentiated into early hepatic progenitors or late hepatic progeny were transplanted (Chen et al., 2013) (top left); human albumin levels in mouse sera, each dot is an individual mouse (fractions of successfully engrafted mice indicated; top right); recipient whole-liver cross-section with different lobes and subfields indicated; scale bar, 5 mm (middle right); costaining for human albumin and GFP in four distinct hepatic lobes, fields numbered above (bottom). See also Figures S4 and S5.

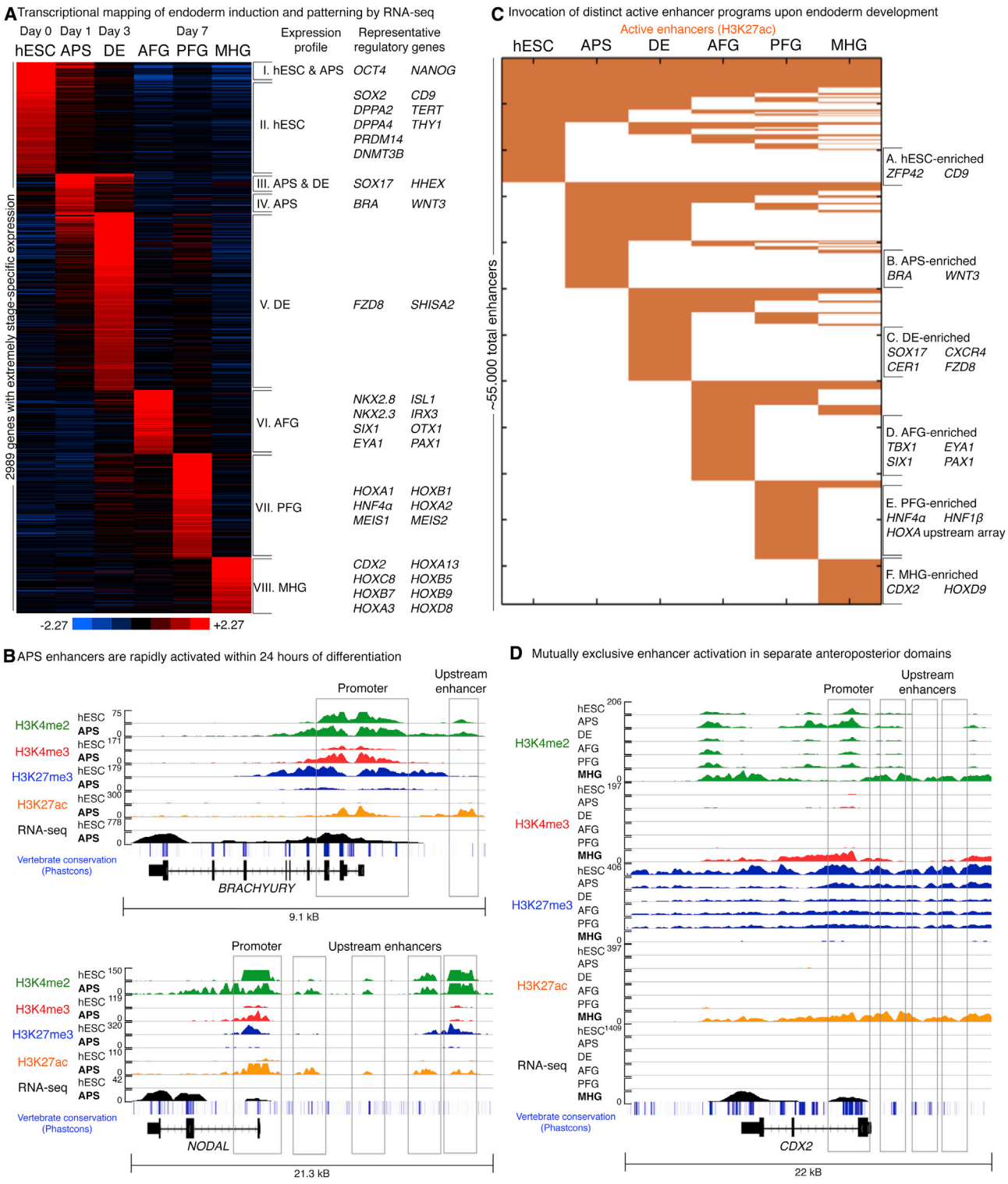

Figure 5. Comprehensively Mapping Transcriptional and Epigenetic Dynamics during Endodermal Development

(A) RNA-seq heatmap of stage-specific genes upregulated at indicated lineage transitions (Supplemental Experimental Procedures).

(B and D) Compiled ChIP-seq (histone modifications), RNA-seq (gene expression), vertebrate conservation (Phastcons), and coding gene structure at selected genomic loci with cell types and genomic distance indicated. Numbers indicate fold enrichment over input (ChIP-seq) and FPKM values (RNA-seq).

(C) Binary heatmap of H3K27ac-marked active enhancers activated at respective differentiation phases (Supplemental Experimental Procedures); each row is an individual enhancer.

See also Figures S5–S7.

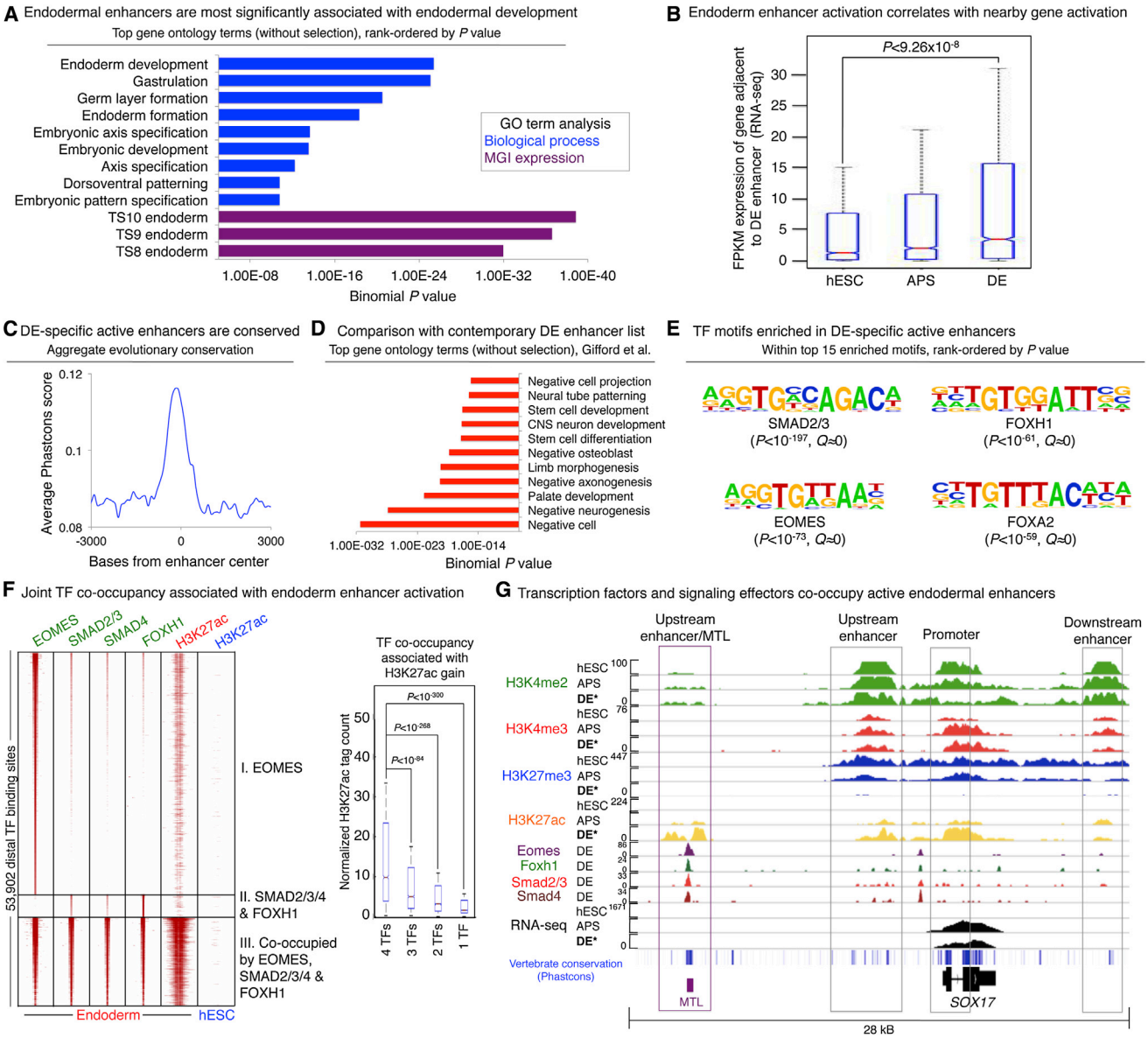

Figure 6. TGF-β Signaling Inaugurates Endodermal Active Enhancers

(A) Top-ranked GO terms associated with DE-specific active enhancers by GREAT (McLean et al., 2010) without preselection.

(B) Boxplot of RNA-seq FPKM expression values of genes adjacent to DE-specific enhancers at indicated in vitro differentiation stages.

(C) Phastcons score of DE-specific active DE enhancers.

(D) Top-ranked GO terms associated with DE enhancers identified from a previous data set (Gifford et al., 2013) using identical analytic methods (Table S6).

(E) TF motifs overrepresented in DE-specific active enhancers (Table S5B).

(F) Left: ChIP-seq signal heatmap based on all distal EOMES, SMAD2/3/4, and FOXH1 peaks in DE, showing TF overlap with one another and K27ac; each row is a single distal element (6 kB window size). Right: average H3K27ac tag count at DE distal elements bound by all 1, 2, 3, or 4 DE TFs (EOMES, SMAD2/3, SMAD4, and FOXH1).

(G) EOMES, FOXH1, and SMAD2/3/4 colocalize at conserved *SOX17* enhancer. MTL, multiple TF locus.

See also Figure S6.

of DE pre-enhancers (10%) were preloaded with HAT p300 (Rada-Iglesias et al., 2011) (Figure 7B), suggesting that rapid enhancer acetylation during differentiation may largely involve de novo HAT recruitment.

A pre-enhancer state solely delineated by H2AZ without other detectable distinguishing factors has not been previously described. H2AZ-laden nucleosomes are unstable and are readily displaced by TFs (Jin et al., 2009; Li et al., 2012). This may permit endoderm TFs to rapidly infiltrate DE enhancers upon differentiation, explaining rapid enhancer activation. Indeed, H2AZ-marked DE pre-enhancers in hESCs more readily attracted EOMES, SMAD2/3/4, and FOXH1 upon differentiation

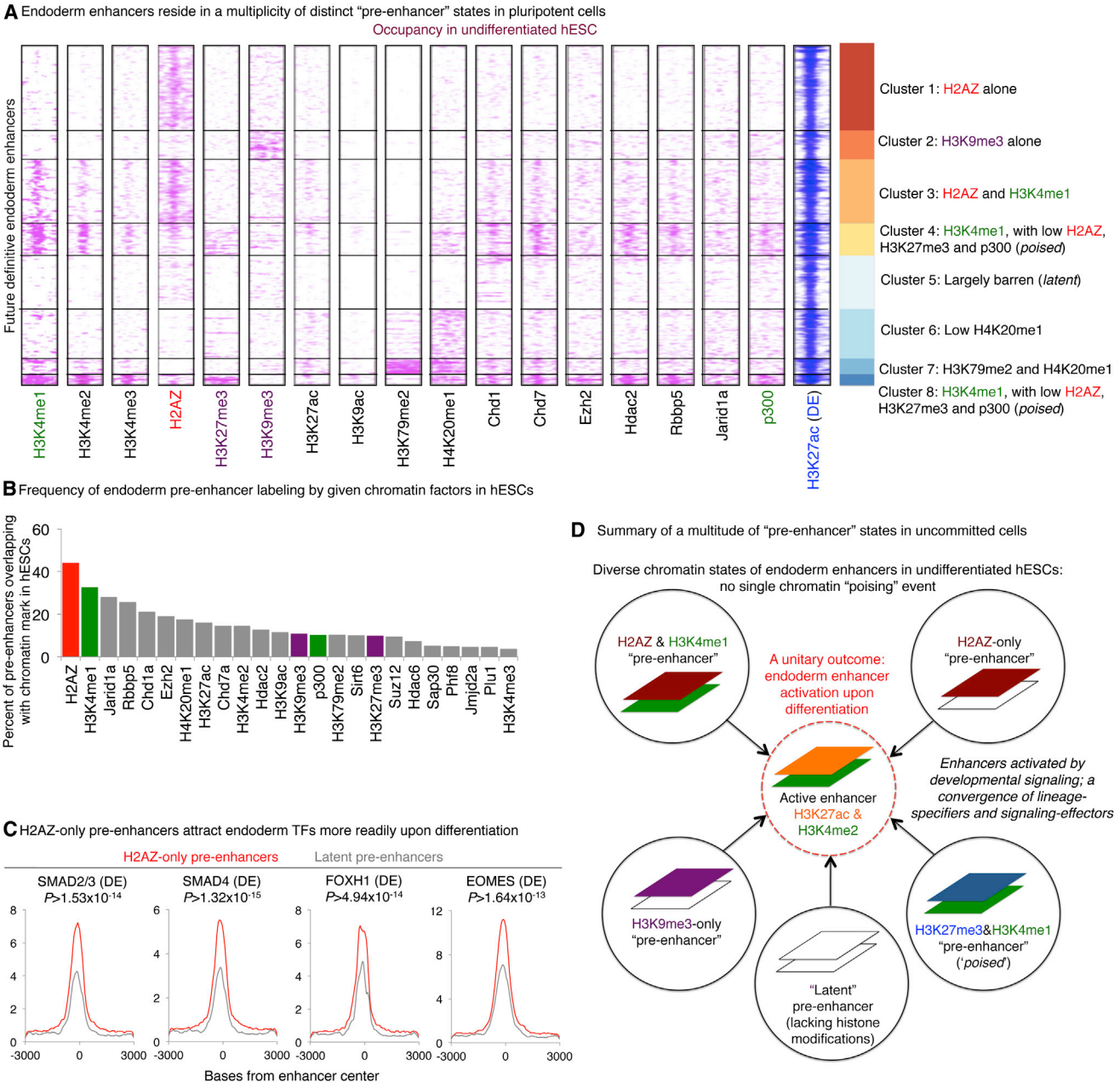

Figure 7. A Constellation of Diverse "Pre-enhancer" States

(A) ChIP-seq signal heatmap of indicated chromatin marks across future DE enhancer regions in hESCs, organized by unbiased clustering; each row is a single pre-enhancer.

(B) Frequency of DE pre-enhancers overlapping with a given chromatin mark in hESCs.

(C) Occupancy of DE enhancers by endoderm TFs in DE cells that were originally either H2AZ-only pre-enhancers (class 1) or latent pre-enhancers (class 5) in hESCs.

(D) Pre-enhancer state summary.

See also Figure S6.

(Figure 7C, p = 10^{-13}–10^{-15}) when compared to latent pre-enhancers.

In sum, initial K4me1 "poising" is not the only predictor of subsequent enhancer activation. We show that there is a diversity of pre-enhancer states characterized by different combinations of chromatin marks (Figure 7D).

DISCUSSION

PSC differentiation typically yields a range of developmental outcomes that vary between PSC lines. Contaminating lineages may generate undesired tissues upon transplantation and obscure molecular analyses of lineage commitment. To meet

this challenge, we delineated the signaling logic for induction and anteroposterior patterning of human endoderm from PSC and for subsequent bifurcation of pancreas versus liver, clarifying the separation of alternate lineages at each stage. Such knowledge permitted us to rationally exclude alternate fates at every step following the in vivo hierarchy of germ layer segregations (Tzouanacou et al., 2009). This approach yielded precise induction of a single lineage (endoderm) from diverse hESC and hiPSC lines without extraneous lineages typically induced by earlier protocols. This level of endodermal purity enabled accurate chromatin analysis of endoderm induction at a resolution previously unattainable due to contaminating lineages. Therefore the highly homogeneous DE populations described here constitute an ideal starting point to efficiently generate downstream endodermal derivatives (McKnight et al., 2010), a notion we validate by producing engraftable liver cells. In summary, this work expounds a coherent view of signaling logic and chromatin dynamics propelling endoderm specification and patterning, thereby availing both developmental biology and hPSC differentiation.

Developmental Segregation of Mutually Exclusive Endodermal Fates

Throughout four successive stages of endoderm development, we accurately defined the signals that instructed or repressed a given lineage, thus providing a clearer view of how endodermal lineage bifurcations are driven. In fact, this refined understanding suggested that previous protocols provided incorrect signals that repressed DE formation, thereby leading to inefficient differentiation. For example, BMP, FGF, TGF-β, and Wnt have been used to elicit both endoderm (Touboul et al., 2010) and mesoderm (Gadue et al., 2006), and, therefore, the exact lineages induced by these signals has remained ambiguous.

We generated DE in the virtual absence of mesoderm or ectoderm. We found that combined FGF, TGF-β, and Wnt together with low BMP signaling (Bernardo et al., 2011; Blauwkamp et al., 2012; Gadue et al., 2006) was necessary to specify APS (>99% MIXL1$^+$) and repress ectoderm, abolishing ectoderm competence within 24 hr of APS induction. After ectoderm exclusion, mesoderm was sequentially eliminated by BMP inhibition, which, when combined with TGF-β and endogenous FGF signaling (Bernardo et al., 2011; D'Amour et al., 2005), exclusively drove PS toward DE. It was crucial to suppress endogenous mesoderm-inducing BMP and Wnt signaling within PS to achieve pure DE populations. We also clarified nuances in the interpretation of combinations of signals, showing that reception of one signal altered the response to others. For example, although BMP inhibition typically eradicated mesoderm from the PS, if DE-inducing FGF was blocked in parallel, mesoderm formation was re-enabled. Thus, FGF was obligatory to consolidate DE commitment.

Following PFG formation, TGF-β and BMP signaling dueled to specify pancreas versus liver, and each bilaterally cross-repressed the alternate fate, reminiscent of in vivo findings (Chung et al., 2008; Deutsch et al., 2001). Therefore, efficient liver induction required TGF-β inhibition to eliminate pancreatic fates in conjunction with BMP and MAPK to positively drive liver and vice versa. In sum, we show that, in order to efficiently drive hPSC differentiation down a single developmental route, it is critical not only to provide the relevant positive inductive signals, but it is *equally important* to inhibit repressive signals that instead drive progression down alternate lineage pathways.

By inhibiting alternate fates at each juncture, we could universally differentiate nine diverse hESC/hiPSC lines into highly pure DE populations in defined conditions. This is contrary to the notion that different hPSC lines have distinct differentiation biases and that each might require customized signals to drive efficient commitment. Our observations are timely, because a prerequisite for cell replacement therapy is the consistent generation of homogeneous lineages from hPSCs under defined conditions (Cohen and Melton, 2011; McKnight et al., 2010). Recent strategies to generate "self-renewing" DE (Cheng et al., 2012) or liver buds (Takebe et al., 2013) from hPSCs are appealing, but require coculture with heterologous feeders and thus suit a different type of application.

Obligatory Endodermal Signaling Inputs Are Highly Temporally Dynamic

The precise sequence and kinetics of endoderm signaling transitions remain to be fully elucidated, despite their evident importance in vivo and in vitro (Green et al., 2011; Wandzioch and Zaret, 2009). For example, BMP and Wnt have been associated with mesoderm induction through studies of prolonged treatment over several days (Bernardo et al., 2011; Gadue et al., 2006). However we found that BMP and Wnt initially specified APS, but within 24 hr of differentiation, signaling requirements were reversed such that BMP and Wnt repressed DE from PS and instead induced mesoderm. Prior protocols reduced APS and DE induction into a single lengthy stage and persistently provided BMP for 3–5 days (Nostro et al., 2011; Touboul et al., 2010), likely generating contaminating mesoderm at later stages and inhibiting DE formation. The dynamism with which BMP and Wnt signals are interpreted during hPSC differentiation (within 24 hr) closely tracks how Wnt is initially inactive in E5.5 postimplantation epiblast, transiently elicited in E6.5 PS and then silenced once again in E7.5 DE in vivo (Maretto et al., 2003). Therefore, assigning BMP and Wnt as either proendoderm or antiendoderm is a misnomer because these signals can induce either outcome depending on timing within just 24 hr in vivo and in vitro.

Developmental Competence and a Diversity of Pre-enhancer States

To gain insight into endodermal lineage commitment mechanisms, we globally mapped transcriptional changes and regulatory element redeployment across multiple steps of endoderm induction and patterning. This resource could unveil novel drivers or markers of DE specification by identifying TFs upregulated at distinct stages. Here, we exploited the accompanying chromatin data to explore how endoderm competence is preconfigured in pluripotent cells.

Since Waddington's formalism of developmental competence (Waddington, 1940), its molecular basis has remained cryptic. Competence may be foreshadowed by permissive chromatin priming of developmental enhancers in progenitors (Calo and Wysocka, 2013). Various models proposed that such enhancers resided in "poised" or "latent" chromatin states prior

to activation (Ostuni et al., 2013; Rada-Iglesias et al., 2011). However, the prevalence of "poised" or "latent" pre-enhancer states (and whether they represented all pre-enhancer states) remained uncertain. With a priori knowledge of a catalog of DE enhancers, we systematically determined their antecedent "pre-enhancer" states in hESCs. Individual DE enhancers existed in a wide continuum of differentially marked pre-enhancer states prior to activation, extending beyond "poised" or "latent" states. Only a subset of DE enhancers were premarked by K4me1, p300 or other proposed "poising" factors in hESCs, showing there is no universal poising signature.

Strikingly, we found that many prospective DE enhancers were marked exclusively by H2AZ in the general absence of other chromatin marks. Thus, H2AZ is sometimes the earliest recognizable enhancer mark in lieu of K4me1. H2AZ prepositioning at DE enhancers enhanced future infiltration by EOMES, SMAD2/3/4, and FOXH1 upon differentiation and combinatorial occupancy by all of these TFs correlated with maximal enhancer activation. Indeed, H2AZ is essential for DE induction from mESCs, and it was shown that its presence at promoters increased FOXA2 recruitment (Li et al., 2012). Our related findings with DE *enhancers* suggest the primordial chromatin state of a DE enhancer in hESCs can influence its future engagement upon differentiation. Because some mesoderm enhancers are likewise exclusively marked by H2AZ in hESCs (Figure S6F), H2AZ prepositioning on developmental enhancers may broadly signal future fates available to uncommitted precursors. How H2AZ is deployed to these silent enhancers in ESCs remains unclear. It may be targeted by pluripotency TFs (e.g., Oct4), which physically interact with H2AZ depositor p400 (van den Berg et al., 2010), and might guide it to lineage specification genes in uncommitted ESCs to functionally presage future differentiation potential (Loh and Lim, 2011; Teo et al., 2011). Yet, half of endoderm enhancers apparently lack H2AZ in hESCs; therefore, to understand developmental competence, we must decipher the whole range of alternative pre-enhancer states.

EXPERIMENTAL PROCEDURES

SR1 DE Induction and Patterning

mTeSR1-grown hPSCs (Figure S2f) were passaged ~1:3 as small clumps, using collagenase IV onto fibronectin- or Matrigel-coated plates. One to two days later, they were washed and differentiated with Activin A (100 ng/ml), CHIR99021 (2 μM), and PI-103 (50 nM) in serumless CDM2 basal medium for 24 hr to specify APS (day 1), followed by Activin A (100 ng/ml) and DM3189 (250 nM) for 48 subsequent hr to specify DE (day 3). DE was antero-posteriorly patterned into AFG (A-83-01, 1 μM and DM3189, 250 nM), PFG (RA, 2 μM and DM3189, 250 nM), or MHG (BMP4, 10 ng/ml; CHIR99021, 3 μM; and FGF2, 100 ng/ml) for 4 subsequent days until day 7. For detailed differentiation methods, see Supplemental Experimental Procedures.

Hepatic Induction and Empirical Maturation

Day 3 DE was differentiated for 24 hr into early PFG by DM3189 (250 nM), IWP2 (4 μM), PD0325901 (500 nM), and RA (2 μM) and further differentiated into hepatic progenitors by A-83-01 (1 μM), BMP4 (10 ng/ml), IWP2, and RA for 3 further days until day 7. They were then empirically matured in vitro with BMP4 (2 days), followed by dexamethasone (10 μM) and oncostatin M (10 μg/ml) for 10 days (Kamiya et al., 1999) and then intrahepatically transplanted into newborn NOD-SCID *Il2rγ*$^{-/-}$ mice (Chen et al., 2013). All animal experiments were performed as ordained by A*STAR Institutional Animal Care and Use Committee guidelines.

Chromatin State Analysis

For ChIP-seq, H7-derived endoderm lineages were formaldehyde fixed, lysed to extract nuclei, sonicated, and precleared (Supplemental Experimental Procedures). Chromatin was probed overnight using K4me2, K4me3, K27ac, and K27me3 antibodies (Table S7) conjugated to Protein G Dynabeads (Invitrogen). Subsequently, chromatin was precipitated, rigorously washed (eight times), and cross-linking undone by overnight 65°C heating before RNase/Proteinase K treatment and column purification. Ten nanograms of chromatin were used to generate libraries (TruSeq Kit, Illumina) for Hi-Seq 2000 sequencing (Illumina, 36bp single-end reads; Figure S5E). Reads were aligned to hg19 (Bowtie), extended, and input normalized (MACS). DE enhancers (Table S5A) were assigned by DFilter (Kumar et al., 2013) as peaks with ≥4-fold more K27ac tags in DE than hESCs and were associated with gene ontology (GO) terms via the Genomic Regions Enrichment of Annotations Tool (GREAT) (McLean et al., 2010).

ACCESSION NUMBERS

Transcriptional and ChIP-seq data are available under Gene Expression Omnibus accession number GSE52658.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, seven figures, and seven tables and can be found with this article online at http://dx.doi.org/10.1016/j.stem.2013.12.007.

ACKNOWLEDGMENTS

We are grateful to the developmental biologists who set the precedent for this work decades ago. Margaret Fuller, Phillip Beachy, and Roel Nusse (Stanford Department of Developmental Biology) championed this work in all respects. We thank B.M. Brady, E. Lujan, H.T. Lim, E.K.M. Chia, S.W. Teo, S. Van Der Van, J.R. Tan, H. Ijo, J.J. Lum, X.Y. Huang, M. Rahmani, and J. Poschmann for technical assistance; S. Oh and S.Y. Ng for hESC; and A.M. Newman, R. Lu, P.W. Koh, E. Rim, P. Robson, M.M.J. Fischer, T. Meyer, and members of the Weissman and Lim groups for comments. We also thank R. Ettikian, T.A. Storm, P.A. Lovelace, K.Z.J. Chee, the GIS Solexa Group, the Biopolis Shared Facilities FACS Core, the Stanford Stem Cell Institute FACS core, and the Stanford PAN Microarray Core for logistical/core facility support. K.M.L. is supported by the Fannie and John Hertz Foundation, the U.S. National Science Foundation, and the Davidson Institute for Talent Development. L.T.A. and B.L. are supported by the Singapore Agency for Science, Technology, and Research (A*STAR). I.L.W. is supported by the California Institute for Regenerative Medicine (RT2-02060). L.A., E.S.N., A.G.E., and E.G.S. are supported by Stem Cells Australia, the Qatar National Research Foundation, and the Australia National Health and Medical Research Council (NHMRC) and A.G.E. and E.G.S. as NHMRC Senior Research Fellows. This work is dedicated to Dale L. Woodbury and M. William Lensch for their excellent mentorship.

Received: June 13, 2013
Revised: September 25, 2013
Accepted: December 16, 2013
Published: January 9, 2014

REFERENCES

Beppu, H., Kawabata, M., Hamamoto, T., Chytil, A., Minowa, O., Noda, T., and Miyazono, K. (2000). BMP type II receptor is required for gastrulation and early development of mouse embryos. Dev. Biol. *221*, 249–258.

Bernardo, A.S., Faial, T., Gardner, L., Niakan, K.K., Ortmann, D., Senner, C.E., Callery, E.M., Trotter, M.W., Hemberger, M., Smith, J.C., et al. (2011). BRACHYURY and CDX2 mediate BMP-induced differentiation of human and mouse pluripotent stem cells into embryonic and extraembryonic lineages. Cell Stem Cell *9*, 144–155.

Blauwkamp, T.A., Nigam, S., Ardehali, R., Weissman, I.L., and Nusse, R. (2012). Endogenous Wnt signalling in human embryonic stem cells generates

an equilibrium of distinct lineage-specified progenitors. Nat Commun *3*, 1070.

Calo, E., and Wysocka, J. (2013). Modification of enhancer chromatin: what, how, and why? Mol. Cell *49*, 825–837.

Candia, A.F., Watabe, T., Hawley, S.H., Onichtchouk, D., Zhang, Y., Derynck, R., Niehrs, C., and Cho, K.W. (1997). Cellular interpretation of multiple TGF-β signals: intracellular antagonism between activin/BVg1 and BMP-2/4 signaling mediated by Smads. Development *124*, 4467–4480.

Chen, Q., Khoury, M., Limmon, G., Choolani, M., Chan, J.K., and Chen, J. (2013). Human fetal hepatic progenitor cells are distinct from, but closely related to, hematopoietic stem/progenitor cells. Stem Cells *31*, 1160–1169.

Cheng, X., Ying, L., Lu, L., Galvão, A.M., Mills, J.A., Lin, H.C., Kotton, D.N., Shen, S.S., Nostro, M.C., Choi, J.K., et al. (2012). Self-renewing endodermal progenitor lines generated from human pluripotent stem cells. Cell Stem Cell *10*, 371–384.

Chetty, S., Pagliuca, F.W., Honore, C., Kweudjeu, A., Rezania, A., and Melton, D.A. (2013). A simple tool to improve pluripotent stem cell differentiation. Nat. Methods *10*, 553–556.

Cho, C.H.-H., Hannan, N.R.-F., Docherty, F.M., Docherty, H.M., Joåo Lima, M., Trotter, M.W.B., Docherty, K., and Vallier, L. (2012). Inhibition of activin/nodal signalling is necessary for pancreatic differentiation of human pluripotent stem cells. Diabetologia *55*, 3284–3295.

Chung, W.-S., Shin, C.H., and Stainier, D.Y.R. (2008). Bmp2 signaling regulates the hepatic versus pancreatic fate decision. Dev. Cell *15*, 738–748.

Cohen, D.E., and Melton, D. (2011). Turning straw into gold: directing cell fate for regenerative medicine. Nat. Rev. Genet. *12*, 243–252.

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.

Davis, R.P., Ng, E.S., Costa, M., Mossman, A.K., Sourris, K., Elefanty, A.G., and Stanley, E.G. (2008). Targeting a GFP reporter gene to the MIXL1 locus of human embryonic stem cells identifies human primitive streak-like cells and enables isolation of primitive hematopoietic precursors. Blood *111*, 1876–1884.

Deutsch, G., Jung, J., Zheng, M., Lóra, J., and Zaret, K.S. (2001). A bipotential precursor population for pancreas and liver within the embryonic endoderm. Development *128*, 871–881.

Dunn, N.R., Vincent, S.D., Oxburgh, L., Robertson, E.J., and Bikoff, E.K. (2004). Combinatorial activities of Smad2 and Smad3 regulate mesoderm formation and patterning in the mouse embryo. Development *131*, 1717–1728.

Ernst, J., Kheradpour, P., Mikkelsen, T.S., Shoresh, N., Ward, L.D., Epstein, C.B., Zhang, X., Wang, L., Issner, R., Coyne, M., et al. (2011). Mapping and analysis of chromatin state dynamics in nine human cell types. Nature *473*, 43–49.

Gadue, P., Huber, T.L., Paddison, P.J., and Keller, G.M. (2006). Wnt and TGF-β signaling are required for the induction of an in vitro model of primitive streak formation using embryonic stem cells. Proc. Natl. Acad. Sci. USA *103*, 16806–16811.

Gertow, K., Hirst, C.E., Yu, Q.C., Ng, E.S., Pereira, L.A., Davis, R.P., Stanley, E.G., and Elefanty, A.G. (2013). WNT3A promotes hematopoietic or mesenchymal differentiation from hESCs depending on the time of exposure. Stem Cell Rev. *1*, 53–65.

Gifford, C.A., Ziller, M.J., Gu, H., Trapnell, C., Donaghey, J., Tsankov, A., Shalek, A.K., Kelley, D.R., Shishkin, A.A., Issner, R., et al. (2013). Transcriptional and epigenetic dynamics during specification of human embryonic stem cells. Cell *153*, 1149–1163.

Graf, T., and Enver, T. (2009). Forcing cells to change lineages. Nature *462*, 587–594.

Green, M.D., Chen, A., Nostro, M.-C., d'Souza, S.L., Schaniel, C., Lemischka, I.R., Gouon-Evans, V., Keller, G., and Snoeck, H.-W. (2011). Generation of anterior foregut endoderm from human embryonic and induced pluripotent stem cells. Nat. Biotechnol. *29*, 267–272.

Hansson, M., Olesen, D.R., Peterslund, J.M.L., Engberg, N., Kahn, M., Winzi, M., Klein, T., Maddox-Hyttel, P., and Serup, P. (2009). A late requirement for Wnt and FGF signaling during activin-induced formation of foregut endoderm from mouse embryonic stem cells. Dev. Biol. *330*, 286–304.

Jin, C., Zang, C., Wei, G., Cui, K., Peng, W., Zhao, K., and Felsenfeld, G. (2009). H3.3/H2A.Z double variant-containing nucleosomes mark 'nucleosome-free regions' of active promoters and other regulatory regions. Nat. Genet. *41*, 941–945.

Kamiya, A., Kinoshita, T., Ito, Y., Matsui, T., Morikawa, Y., Senba, E., Nakashima, K., Taga, T., Yoshida, K., Kishimoto, T., and Miyajima, A. (1999). Fetal liver development requires a paracrine action of oncostatin M through the gp130 signal transducer. EMBO J. *18*, 2127–2136.

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.

Kroon, E., Martinson, L.A., Kadoya, K., Bang, A.G., Kelly, O.G., Eliazer, S., Young, H., Richardson, M., Smart, N.G., Cunningham, J., et al. (2008). Pancreatic endoderm derived from human embryonic stem cells generates glucose-responsive insulin-secreting cells in vivo. Nat. Biotechnol. *26*, 443–452.

Kumar, V., Muratani, M., Rayan, N.A., Kraus, P., Lufkin, T., Ng, H.H., and Prabhakar, S. (2013). Uniform, optimal signal processing of mapped deep-sequencing data. Nat. Biotechnol. *31*, 615–622.

Lawson, K.A., Meneses, J.J., and Pedersen, R.A. (1991). Clonal analysis of epiblast fate during germ layer formation in the mouse embryo. Development *113*, 891–911.

Levak-Svajger, B., and Švajger, A. (1974). Investigation on the origin of the definitive endoderm in the rat embryo. J. Embryol. Exp. Morphol. *32*, 445–459.

Li, Z., Gadue, P., Chen, K., Jiao, Y., Tuteja, G., Schug, J., Li, W., and Kaestner, K.H. (2012). Foxa2 and H2A.Z mediate nucleosome depletion during embryonic stem cell differentiation. Cell *151*, 1608–1616.

Liu, P., Wakamiya, M., Shea, M.J., Albrecht, U., Behringer, R.R., and Bradley, A. (1999). Requirement for Wnt3 in vertebrate axis formation. Nat. Genet. *22*, 361–365.

Loh, K.M., and Lim, B. (2011). A precarious balance: pluripotency factors as lineage specifiers. Cell Stem Cell *8*, 363–369.

Maretto, S., Cordenonsi, M., Dupont, S., Braghetta, P., Broccoli, V., Hassan, A.B., Volpin, D., Bressan, G.M., and Piccolo, S. (2003). Mapping Wnt/β-catenin signaling during mouse development and in colorectal tumors. Proc. Natl. Acad. Sci. USA *100*, 3299–3304.

McKnight, K.D., Wang, P., and Kim, S.K. (2010). Deconstructing pancreas development to reconstruct human islets from pluripotent stem cells. Cell Stem Cell *6*, 300–308.

McLean, C.Y., Bristor, D., Hiller, M., Clarke, S.L., Schaar, B.T., Lowe, C.B., Wenger, A.M., and Bejerano, G. (2010). GREAT improves functional interpretation of cis-regulatory regions. Nat. Biotechnol. *28*, 495–501.

Mishina, Y., Suzuki, A., Ueno, N., and Behringer, R.R. (1995). Bmpr encodes a type I bone morphogenetic protein receptor that is essential for gastrulation during mouse embryogenesis. Genes Dev. *9*, 3027–3037.

Nostro, M.C., Sarangi, F., Ogawa, S., Holtzinger, A., Corneo, B., Li, X., Micallef, S.J., Park, I.-H., Basford, C., Wheeler, M.B., et al. (2011). Stage-specific signaling through TGFβ family members and WNT regulates patterning and pancreatic specification of human pluripotent stem cells. Development *138*, 861–871.

Ostuni, R., Piccolo, V., Barozzi, I., Polletti, S., Termanini, A., Bonifacio, S., Curina, A., Prosperini, E., Ghisletti, S., and Natoli, G. (2013). Latent enhancers activated by stimulation in differentiated cells. Cell *152*, 157–171.

Rada-Iglesias, A., Bajpai, R., Swigut, T., Brugmann, S.A., Flynn, R.A., and Wysocka, J. (2011). A unique chromatin signature uncovers early developmental enhancers in humans. Nature *470*, 279–283.

Rezania, A., Bruin, J.E., Riedel, M.J., Mojibian, M., Asadi, A., Xu, J., Gauvin, R., Narayan, K., Karanu, F., O'Neil, J.J., et al. (2012). Maturation of human embryonic stem cell-derived pancreatic progenitors into functional islets capable of treating pre-existing diabetes in mice. Diabetes *61*, 2016–2029.

Sherwood, R.I., Maehr, R., Mazzoni, E.O., and Melton, D.A. (2011). Wnt signaling specifies and patterns intestinal endoderm. Mech. Dev. *128*, 387–400.

Spence, J.R., Mayhew, C.N., Rankin, S.A., Kuhar, M.F., Vallance, J.E., Tolle, K., Hoskins, E.E., Kalinichenko, V.V., Wells, S.I., Zorn, A.M., et al. (2011). Directed differentiation of human pluripotent stem cells into intestinal tissue in vitro. Nature *470*, 105–109.

Stafford, D., and Prince, V.E. (2002). Retinoic acid signaling is required for a critical early step in zebrafish pancreatic development. Curr. Biol. *12*, 1215–1220.

Sumi, T., Tsuneyoshi, N., Nakatsuji, N., and Suemori, H. (2008). Defining early lineage specification of human embryonic stem cells by the orchestrated balance of canonical Wnt/β-catenin, Activin/Nodal and BMP signaling. Development *135*, 2969–2979.

Švajger, A., and Levak-Svajger, B. (1974). Regional developmental capacities of the rat embryonic endoderm at the head-fold stage. J. Embryol. Exp. Morphol. *32*, 461–467.

Takebe, T., Sekine, K., Enomura, M., Koike, H., Kimura, M., Ogaeri, T., Zhang, R.-R., Ueno, Y., Zheng, Y.-W., Koike, N., et al. (2013). Vascularized and functional human liver from an iPSC-derived organ bud transplant. Nature *499*, 481–484.

Tam, P.P., and Beddington, R.S. (1987). The formation of mesodermal tissues in the mouse embryo during gastrulation and early organogenesis. Development *99*, 109–126.

Tam, P.P.L., and Loebel, D.A.F. (2007). Gene function in mouse embryogenesis: get set for gastrulation. Nat. Rev. Genet. *8*, 368–381.

Teo, A.K.K., Arnold, S.J., Trotter, M.W.B., Brown, S., Ang, L.T., Chng, Z., Robertson, E.J., Dunn, N.R., and Vallier, L. (2011). Pluripotency factors regulate definitive endoderm specification through eomesodermin. Genes Dev. *25*, 238–250.

Tiso, N., Filippi, A., Pauls, S., Bortolussi, M., and Argenton, F. (2002). BMP signalling regulates anteroposterior endoderm patterning in zebrafish. Mech. Dev. *118*, 29–37.

Touboul, T., Hannan, N.R.F., Corbineau, S., Martinez, A., Martinet, C., Branchereau, S., Mainot, S., Strick-Marchand, H., Pedersen, R., Di Santo, J., et al. (2010). Generation of functional hepatocytes from human embryonic stem cells under chemically defined conditions that recapitulate liver development. Hepatology *51*, 1754–1765.

Tzouanacou, E., Wegener, A., Wymeersch, F.J., Wilson, V., and Nicolas, J.-F. (2009). Redefining the progression of lineage segregations during mammalian embryogenesis by clonal analysis. Dev. Cell *17*, 365–376.

van den Berg, D.L.C., Snoek, T., Mullin, N.P., Yates, A., Bezstarosti, K., Demmers, J., Chambers, I., and Poot, R.A. (2010). An Oct4-centered protein interaction network in embryonic stem cells. Cell Stem Cell *6*, 369–381.

Waddington, C.H. (1940). Organisers and Genes. (Cambridge, UK: Cambridge University Press).

Wandzioch, E., and Zaret, K.S. (2009). Dynamic signaling network for the specification of embryonic pancreas and liver progenitors. Science *324*, 1707–1710.

Xie, 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.

Yusa, K., Rashid, S.T., Strick-Marchand, H., Varela, I., Liu, P.-Q., Paschon, D.E., Miranda, E., Ordóñez, A., Hannan, N.R.F., Rouhani, F.J., et al. (2011). Targeted gene correction of α1-antitrypsin deficiency in induced pluripotent stem cells. Nature *478*, 391–394.

Zhu, Y., van Essen, D., and Saccani, S. (2012). Cell-type-specific control of enhancer activity by H3K9 trimethylation. Mol. Cell *46*, 408–423.

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

Note Added in Proof

After this manuscript was accepted, we found that using the differentiation approach described here with the SOX17-mCHERRY knockin reporter line, which provides a sensitive means of tracking endoderm specification, we are able to generate 99.1% ± 0.4% SOX17+ DE from hESCs (n = 3 independent experiments). This finding further substantiates our conclusion that the differentiation protocol that we describe enables highly efficient endoderm induction.