Stem Cell Reports
Resource
OPEN ACCESS

A Quantitative Proteomic Analysis of Hemogenic Endothelium Reveals Differential Regulation of Hematopoiesis by SOX17

Raedun L. Clarke,[1,4] Aaron M. Robitaille,[2,4] Randall T. Moon,[2,3,*] and Gordon Keller[1,*]
[1]McEwen Centre for Regenerative Medicine, University Health Network, Toronto, ON M5G 1L7, Canada
[2]Institute for Stem Cell and Regenerative Medicine and Department of Pharmacology, University of Washington School of Medicine, Seattle, WA 98109, USA
[3]Howard Hughes Medical Institute, Chevy Chase, MD 20815, USA
[4]Co-first author
*Correspondence: rtmoon@uw.edu (R.T.M.), gkeller@uhnresearch.ca (G.K.)
http://dx.doi.org/10.1016/j.stemcr.2015.07.008

SUMMARY

The in vitro derivation of hematopoietic stem cells (HSCs) from pluripotent stem cells (PSCs) is complicated by the existence of multiple overlapping embryonic blood cell programs called primitive, erythromyeloid progenitor (EMP), and definitive. As HSCs are only generated during the definitive stage of hematopoiesis, deciphering the regulatory pathways that control the emergence of this program and identifying markers that distinguish it from the other programs are essential. To identify definitive specific pathways and marker sets, we used label-free proteomics to determine the proteome of embryo-derived and mouse embryonic stem cell-derived VE-CADHERIN^{+}CD45^{-} definitive hematopoietic progenitors. With this approach, we identified *Stat1* as a marker that distinguishes the definitive erythroid lineage from the primitive- and EMP-derived lineages. Additionally, we provide evidence that the generation of the *Stat1*$^{+}$ definitive lineage is dependent on *Sox17*. These findings establish an approach for monitoring the emergence of definitive hematopoiesis in the PSC differentiation cultures.

INTRODUCTION

The generation of hematopoietic stem cells (HSCs) from pluripotent stem cells (PSCs) is dependent on our ability to accurately recapitulate the intricate steps of early embryonic hematopoietic development in the differentiation culture. The mammalian embryonic hematopoietic system consists of at least three distinct programs—primitive, erythromyeloid progenitor (EMP)/yolk sac definitive, and intraembryonic definitive—that differ in their potential as well as spatial and temporal patterns of development (McGrath et al., 2011; Medvinsky and Dzierzak, 1996; Palis et al., 1999). Primitive hematopoiesis is the first to develop, initiates in the yolk sac at embryonic day 7.0 (E7.0), and produces a restricted repertoire of hematopoietic cells consisting of primitive erythrocytes that express embryonic forms of globin and macrophages. Shortly after the onset of primitive hematopoiesis, the yolk sac generates a cohort of erythroid, myeloid, and megakaryocytic progenitors that constitute a distinct phase of embryonic hematopoiesis known as the erythromyeloid progenitor program, also referred to as the yolk sac definitive program (McGrath et al., 2011). This program appears to be positioned developmentally between primitive and definitive hematopoiesis. In the mouse, it generates erythroid cells that express adult and embryonic globins but does not give rise to lymphoid lineage cells or hematopoietic stem cells (McGrath et al., 2011; Palis et al., 1999). HSCs are generated during the intraembryonic-definitive stage of hematopoiesis that is specified following the emergence of the two yolk sac programs at different sites within the embryonic arterial vasculature, the most well characterized being the developing aorta, gonad, and mesonephros (AGM) (Medvinsky and Dzierzak, 1996). In addition to HSC and lymphoid potential, murine definitive hematopoiesis is distinguished from the two earlier programs by the fact that the erythroid lineage expresses exclusively adult forms of globin. Although both the EMP and intraembryonic-definitive hematopoietic programs are considered to be "definitive" due to their ability to generate erythroid precursors, to distinguish between them in this study, we use the terms EMP for the yolk-sac-derived progenitors and definitive for those derived from intraembryonic sites.

There is now compelling evidence indicating that both the EMP and definitive programs derive from progenitors known as hemogenic endothelial cells (HECs) that display endothelial markers and are associated with the arterial vasculature (Bertrand et al., 2005; Boisset et al., 2010; Clarke et al., 2013; Kissa and Herbomel, 2010). Live imaging and histological analyses have demonstrated that hematopoietic progenitors develop from the HECs through a process known as the endothelial to hematopoietic transition (EHT) (Boisset et al., 2010; Kissa and Herbomel, 2010). Currently, there are no markers to distinguish EMP-HECs from definitive HECs. As the EMP program develops earlier than the definitive program, the EMP-derived lineages can predominate the PSC differentiation cultures and are often misinterpreted as representing definitive

hematopoiesis. Developing strategies to distinguish definitive- from EMP-derived hematopoiesis in the PSC differentiation cultures is essential for identifying the signaling pathways that regulate the development of definitive HECs and their specification to HSCs.

We have previously shown that SOX17 is expressed in definitive HECs and that it is required for the establishment of the definitive, but not the primitive, hematopoietic program (Clarke et al., 2013). Our analysis of differentiated hematopoietic progeny from HECs generated from SOX17-null mouse embryonic stem cells (mESCs) revealed the presence of primitive erythroid progenitors, myeloid progenitors, and erythroid progenitors that gave rise to large colonies distinct from the primitive erythroid colonies. The presence of this population of erythroid progenitors suggests that the EMP program may also be *Sox17* independent. Our goals in the present study were to define the role of *Sox17* in EMP hematopoiesis, understand how *Sox17* regulates the generation of definitive hemogenic endothelium, and identify additional markers to distinguish these programs. To achieve these goals, we deciphered the proteome of AGM-derived HECs and of HECs generated from *Sox17*-wild-type and *Sox17*-knockout mESCs using label-free quantitative proteomics. The findings from these studies demonstrate the overall utility of analyzing the proteome of small numbers of cells and show that expression of the transcription factor *Stat1* and dependency on *Sox17* distinguish the definitive erythroid lineage from the primitive and EMP-derived erythroid lineages.

RESULTS

Proteomic Analyses of Wild-Type mESC- and Embryo-Derived HEC-Containing Populations

To identify markers that distinguish the different HEC populations and regulators that control their development, we quantitatively measured the global protein expression profile of mESC- and embryo-derived progenitors using label-free proteomics. In the first set of experiments, we isolated and analyzed the VE-CADHERIN$^+$ (VEC) CD45$^-$ population from the AGM region of E10.5 and E11.5 mouse embryos. This population contains non-hemogenic endothelial cells (ECs), HECs, and type I pre-HSCs at E10.5 and E11.5 (Rybtsov et al., 2011). The corresponding VEC$^-$CD45$^-$ fractions from the same embryos were analyzed as representing the non-EC/HEC populations. For the second set of samples, we focused on mESC-derived populations. We previously described a chemically defined serum-free and feeder-free differentiation protocol for mESCs in which the primitive hematopoietic program is identified as an early arising FLK-1$^+$ population (day 3.25 [D3.25]) and the definitive program as a FLK-1$^+$ population that emerges at D5.25 (Clarke et al., 2013; Irion et al., 2010). When isolated and cultured for an additional 2 days (total day 7 [D7]) these definitive D5.25 FLK-1$^+$ cells give rise to an endothelial VEC$^+$ population that contains HECs that undergo the EHT to generate CD45$^+$ hematopoietic cells at D9. The D5.25 FLK-1$^+$ progenitors and derivative D7 VEC$^+$CD45$^-$ cells were isolated from two separate mESC lines (mESC-1 and mESC-2) and subjected to proteomic analyses (Figure 1A).

For all analyses, protein digests were measured by nano-liquid chromatography-tandem mass spectrometry (nano-LC-MS/MS) on a Q Exactive mass spectrometer and quantified using label-free methods (Luber et al., 2010). In total, 36,678 peptides corresponding to 4,044 proteins were quantified in both the in vivo and in vitro studies (Table S1). Protein quantification was reproducible across in vitro and in vivo measurements, with a Pearson's correlation between replicates of 0.89 and 0.90, respectively (Figures 1B and 1C).

To validate that the proteomics analysis could accurately distinguish between the EC/HEC and non-EC/HEC fractions, we analyzed the E11.5 VEC$^+$ and VEC$^-$ populations for proteins known to be expressed in HECs, including ALDH1A2 (RALDH2), SCL/TAL1, CBFβ, and ESAM. All were enriched in the VEC$^+$ fraction, confirming the utility of this proteomic approach to accurately identify the HECs within this heterogeneous cell population (Chanda et al., 2013; Van Handel et al., 2012; Yokota et al., 2009) (Figure S1A; Tables S2 and S3). Next, we asked if our proteomics analysis could identify differences between the E10.5 VEC$^+$CD45$^-$ and E11.5 VEC$^+$CD45$^-$ populations. The E11.5 VEC$^+$CD45$^-$ population contains more potent or higher numbers of pre-HSCs than the E10.5 population (Rybtsov et al., 2011; Yokomizo et al., 2012). We identified 85 proteins that were differentially expressed between these populations (Figures S1B and S1C; Table S4). Among the proteins expressed at higher levels in the E11.5 population than in the E10.5 population, we identified STAM2 and SLFN1. STAM2 is a critical adaptor protein downstream of growth factor and inflammatory cytokine signaling pathways (Pandey et al., 2000). The identification of STAM2 in the E11.5 population is of interest, as recent studies indicate that inflammatory signaling plays a crucial role in the emergence of HSCs (Espín-Palazón et al., 2014; Li et al., 2014; Sawamiphak et al., 2014). SLFN1 is a member of the Schlafen family of proteins that are expressed in endothelial and hematopoietic cells and are emerging as important regulators of the immune response (Kuang et al., 2014; Schwarz et al., 1998). By contrast, ESAM (endothelial cell selective adhesion molecule), which is a marker of emerging multipotent stem and progenitor cells, and members of the RAS signaling pathway (RAB17 and SIPA1) were found at higher levels in the E10.5 population.

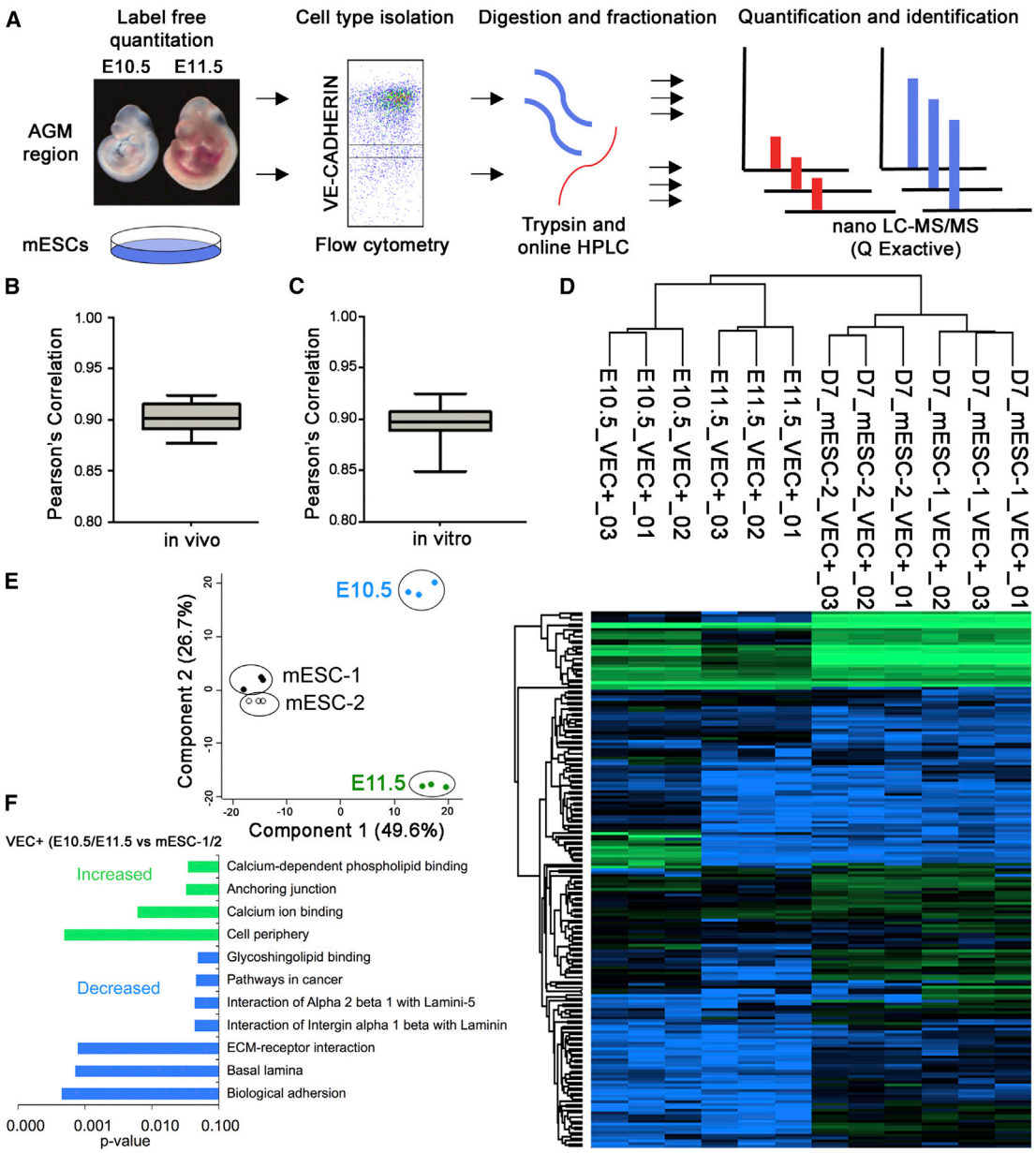

Figure 1. Quantitative Proteomic Profiling of Hemogenic Endothelium

(A) Proteomic workflow used to identify differentially regulated protein expression in vitro and in vivo.

(B) Label-free quantification of protein expression in vitro is reproducible. Tryptic digestions of cell lysates were analyzed by nano-LC MS/MS. n = 3 individual measurements; Pearson correlation of 0.89.

(C) Label-free quantification of protein expression in vivo is reproducible. Tryptic digestions of mouse tissue were analyzed as in (B). n = 3 individual measurements; Pearson correlation of 0.90.

(D) Protein expression was analyzed by unsupervised hierarchal cluster analysis (HC). Green denotes high, relative protein expression and blue low, relative protein expression.

(E) Protein expression was analyzed by principle-component analysis.

(F) GO terms linked to proteins enriched in VE-CADHERIN+ populations were analyzed in G-profiler using the cocoa algorithm.

See also Figure S1

Lastly we compared our D7 mESC-derived ECs/HECs to the E10.5 and E11.5 VEC$^+$CD45$^-$ ECs/HECs (Figure 1D). Principle-component analysis revealed that the VEC$^+$ populations from the two mESC lines were similar to one another. Component 2 reflects the variance between E10.5 AGM and E11.5 AGM. Based on this analysis, the mESC-derived populations appear to represent a stage of development between them. Component 1 is an indication of the variance between in vivo and in vitro culture conditions. As expected, cells grown in vitro do not perfectly recapitulate in vivo development (Figure 1E). Gene Ontology (GO) analysis showed that proteins linked to anchoring junctions, including calcium ion binding and calcium-dependent phospholipid binding, were more highly expressed in the AGM VEC$^+$CD45$^-$ and VEC$^-$CD45$^-$ cells than the D7 mESC-derived cells (Figure 1F). Collectively, these findings highlight the power of the label-free proteomics approach for assessing protein expression in a small number of cells, and in doing so, they identify candidate signaling pathways that may play a role in the generation of HSCs from HECs in vitro.

Determination of the SOX17-Dependent Proteome in mESC-Derived HEC-Containing Populations

We previously reported that SOX17 is a critical regulator of hemogenic endothelium development in vivo and during the directed differentiation of mESCs in vitro (Clarke et al., 2013). However, the direct targets of SOX17 required for the generation of HECs are unknown. In our previous studies, we demonstrated that the doxycycline-inducible overexpression of SOX17 in FLK-1$^+$ mESC-derived endothelial progenitors increased the formation of HECs while repressing the endothelial to hematopoietic transition. We therefore hypothesized that the quantitative measurement of changes in protein expression following the induced expression of SOX17 would identify both direct and indirect target proteins of SOX17 during the generation of HECs. D5.25 FLK-1$^+$ endothelial progenitors were isolated and treated with doxycycline for 1–3 days to induce the expression of SOX17. Samples were isolated from D6 to D9 and analyzed for total protein expression using label-free proteomics (Figure S2A). In total, SOX17 directly or indirectly regulated the expression of 110 proteins (Figure S2B; Table S5) that we classified into early, mid, or late effectors (Figure S2D). SOX17 has been implicated as both a transcriptional activator and repressor, suggesting that proteins that exhibit a significant increase or decrease immediately following the induction of SOX17 may be direct transcriptional targets of SOX17 in HECs. We not only quantified changes in SOX17 expression (Figures S2B and S2C) but also identified several proteins that have been suggested to be targets of SOX17 in other cell systems, including UCHL1 (extraembryonic endoderm development) (Sinner et al., 2004) and CDKN1C (p57/KIP2, respiratory epithelial cells) (Lange et al., 2009) (Table S5). Consistent with our previous observation that the enforced expression SOX17 promotes the maintenance and expansion of HECs, thus postponing the EHT (Clarke et al., 2013), we identified that SOX17 overexpression decreased the expression of proteins linked to the erythrocyte function during the time course of directed differentiation (Figures S2E and S2F). Taken together, we expect that this proteomic dataset will be a useful resource for researchers seeking to identify regulators of the expansion and maintenance of HECs.

As SOX17 is required for the establishment of definitive, but not primitive, hematopoiesis, we hypothesized that analysis of the SOX17-dependent proteome would reveal previously unidentified insights into the global pathways involved in the regulation and identification of these disparate hematopoietic programs. To investigate this, we next analyzed the proteome of D7 VEC$^+$CD45$^-$ cells generated from control and *Sox17*$^{-/-}$ mESCs (Figure 2A). We found that 344 proteins were differentially expressed between these populations (Figure 2B; Table S6). Among the differences, we noted that the SOX17-null VEC$^+$ cells expressed higher levels of erythrocyte-linked proteins (Figure 2C), specifically those associated with the primitive and EMP hematopoietic programs including the embryonic forms of β-globin, HBB-εγ and HBB-βH1 (Table S6). This is consistent with our previous observation that the D7 *Sox17*$^{-/-}$ VEC$^+$ population is heterogeneous and contains committed erythroid and myeloid progenitors as well as HECs (Clarke et al., 2013). Interestingly, GO analysis identified that *Stat* transcription factors had decreased expression in the *Sox17*$^{-/-}$ population compared to the control cells (STAT1, STAT3, and STAT5B) (Figure 2C; Table S6). Notably, STAT1 expression was significantly reduced in the D7 *Sox17*$^{-/-}$ VEC$^+$ cells, suggesting that it may be a previously unidentified transcriptional target of SOX17. Consistent with the reduction in STAT1, we also observed a significant reduction in its direct transcriptional target ISG15 (Malakhov et al., 2003) (Figures 2D and S3A).

To determine whether SOX17 can directly activate *Stat1* transcription, HEK293T cells were transfected with a *Sox17* expression plasmid along with a construct containing a 2.5-kb region of the *Stat1* promoter that includes four potential SOX17-binding sites upstream of luciferase (JASPER database; data not shown). Transfection of *Sox17* did not lead to induction of luciferase activity (Figure S3B). Consistent with this observation, we did not detect binding of SOX17 to the *Stat1* promoter by chromatin immunoprecipitation (ChIP) analysis (data not shown). Finally, we were unable to demonstrate an upregulation of *Stat1* following the enforced expression of SOX17 in the D5.25 FLK-1$^+$ cells using a doxycycline-inducible mESC line

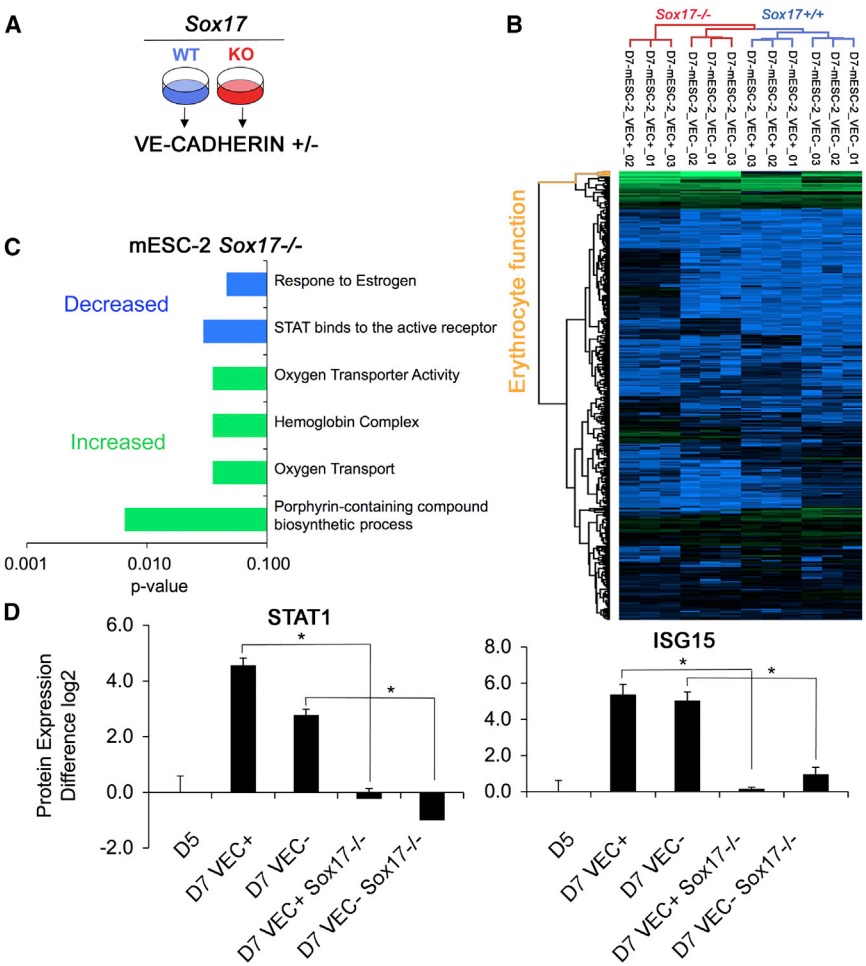

Figure 2. *Sox17*-Regulated Protein Expression

(A) Schematic outlining *Sox17* proteomic comparison.

(B) Protein expression was analyzed by unsupervised hierarchal cluster analysis (HC). Green denotes high, relative protein expression and blue low, relative protein expression.

(C) GO terms linked to proteins regulated by *Sox17* knockout were analyzed in G-profiler using the cocoa algorithm.

(D) STAT1 and ISG15 protein expression was quantified by nano-LC-MS/MS on a Q Exactive. n = 3 individual measurements. *p < 0.01.

See also Figure S2.

(Figure S3C). Collectively, these findings provide strong evidence that *Stat1* is not a direct target of SOX17 in the VEC$^+$ cells and in doing so support an alternative interpretation in which *Stat1* is expressed in a definitive hematopoietic progenitor that is regulated by SOX17.

Expression of *Stat1* in Embryo-Derived Erythroid Populations

The observed differences in globin gene expression between the wild-type and SOX17-null VEC$^+$CD45$^-$ populations indicate that they are heterogeneous and additional markers are needed to identify the type of hematopoietic progenitors present within them. As previous studies have shown that CD41 is expressed on emerging primitive and definitive hematopoietic progenitors, including those of the erythroid lineage (Ferkowicz et al., 2003; Mitjavila-Garcia et al., 2002), we next analyzed D7 wild-type and *Sox17*$^{-/-}$ VEC$^+$CD45$^-$ populations for the presence of this marker. Approximately half the wild-type VEC$^+$CD45$^-$ population expressed CD41, whereas more than 80% of the SOX17-null cells were positive (Figure S3D). These findings show that the VEC$^+$CD45$^-$ populations used for the proteomic analyses contained CD41$^+$ hematopoietic progenitors and raise the interesting possibility that the differences in *Stat1* and globin expression between the two reflect differences in their erythroid progenitor content. Specifically, they support the hypothesis that *Stat1* is expressed in an erythroid progenitor that is regulated by SOX17. Differential expression of the *Stat* regulators in different erythroid populations is in line with a recent computational study that provided evidence that STAT3 signaling is required for primitive erythropoiesis, whereas *Stat1* is preferentially expressed during the definitive stage (Greenfest-Allen et al., 2013). EMP-derived erythroid progenitors were not examined in these experiments.

To determine if expression of *Stat1* distinguishes the erythroid lineages generated from the different embryonic hematopoietic programs, we monitored its expression as well as that of other *Stat* genes in the primitive, EMP-derived, and definitive erythroid cells (Figure 3A) obtained from wild-type and SOX17-null embryos and mESCs

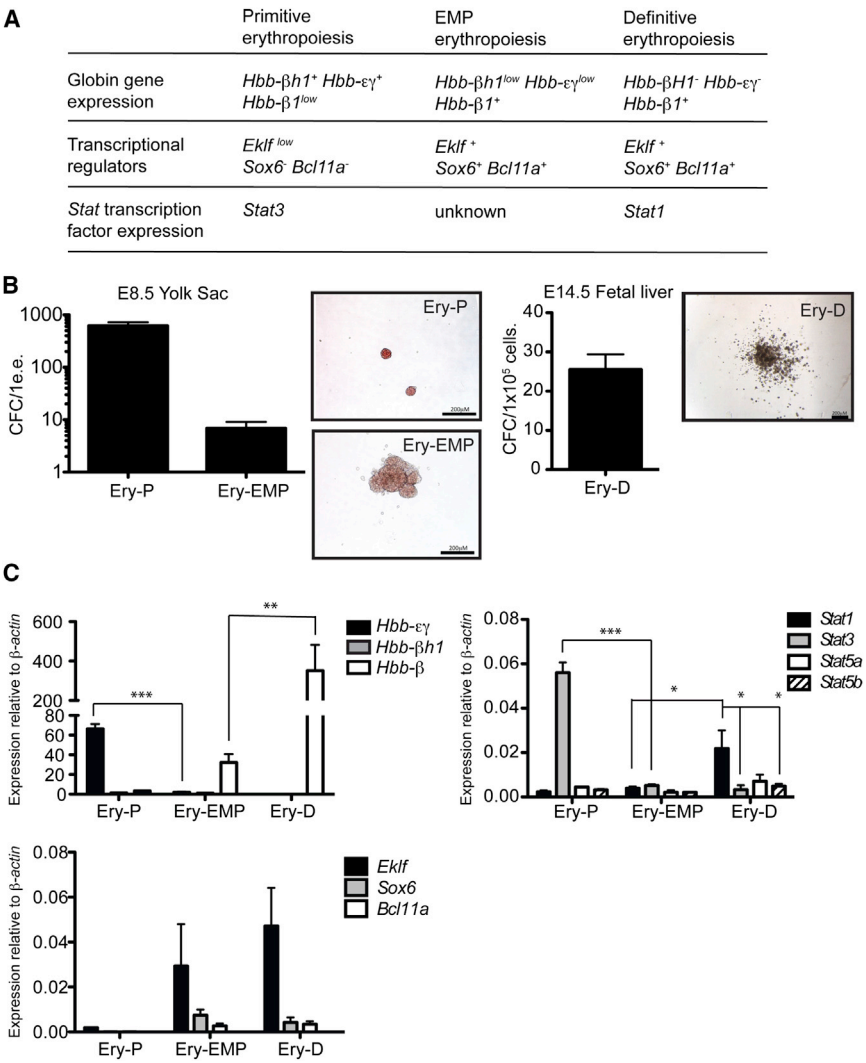

Figure 3. Primitive, EMP, and Definitive Erythroid Cells Have Distinct Gene Expression Signatures In Vivo

(A) Previously published expression patterns of globin genes and transcriptional regulators in mouse erythroid cells

(B) Erythroid potential of E8.5 yolk sac (YS) and E14.5 fetal livers. Bars represent SEM number of colonies per yolk sac from 7 E8.5 embryos and per 1 × 10^5 cells from the livers of 5 E14.5 embryos. The tissues were obtained from two dams.

(C) Real-time qPCR analysis of the indicated genes in Ery-P, Ery-EMP, and Ery-D colonies from E8.5 YS and E14.5 fetal livers. Values represent SEM; n = 5 individual experiments. ***p < 0.005, **p < 0.01, *p < 0.05. Ery-P colonies were picked following 5 days of culture, and Ery-EMP and Ery-D colonies were picked following 8 days of culture. Colonies were pooled for RNA extraction. See also Figure S3.

differentiated in culture. In the mouse, the primitive erythroid (Ery-P), EMP-derived (Ery-EMP), and definitive erythroid (Ery-D) cells can be distinguished based on distinct β-globin gene expression patterns. Ery-P cells express high levels of the embryonic β-globin HBB-βH1 and HBB-εγ and low levels of the adult forms β-globin Hbb-β1/β2. Ery-EMP cells express low levels of HBB-βH1 and HBB-εγ, and an increased amount of HBB-β1/β2 (McGrath et al., 2011). Ery-D cells express exclusively the adult β-globins HBB-β1/β2. In addition to patterns of globin expression, the development of these erythroid populations can be monitored by the expression of specific transcription factors. EKLF/KLF1 regulates the expression of several erythroid-specific genes, including both embryonic and adult globin, and is present in all three erythroid populations (Basu et al., 2007; Hodge et al., 2006). *Sox6* and *Bcl11A* are expressed in the EMP and definitive erythroid cells, where they function to suppress the embryonic globin genes (Sankaran et al., 2009; Yi et al., 2006). Neither is expressed in the primitive erythroid lineage.

E8.5 yolk sacs (YS) and E14.5 fetal livers (FL) were isolated and the cells from each cultured in methylcellulose for the generation of primitive, EMP (E8.5 YS), and definitive (E14.5 FL) erythroid colonies, respectively (Figure 3B). The subsets of colonies were identified based on phenotype, then picked, pooled, and analyzed by real-time qPCR for gene expression profiles (Figure 3A). As expected, primitive erythroid colonies expressed high levels of *Hbb-εγ* and low levels of *Hbb-β1*. Additionally, they expressed *Stat3*, as previously reported (Greenfest-Allen et al., 2013). The primitive erythroid colonies expressed only low levels of the other *Stat* genes and *Eklf*. They had no detectable levels of *Sox6* or *Bcl11a*. In contrast, the fetal-liver-derived definitive colonies expressed high levels of adult β-globin (*Hbb-β1*) but no embryonic (*Hbb-εγ*, *Hbb-βh1*) globin. These colonies also expressed *Stat1*, *Eklf*, *Sox6*, and

Bcl11a (Figure 3C). Only low levels of *Stat3*, *Stat5a*, and *Stat5b* were detected in the definitive cells. As a means to categorize the real-time qPCR values, we will consider a deltaCT (ΔCT) level of >50 as "high" for globin genes and >0.02 as high for transcription factors. The EMP-derived erythroid colonies expressed a mixture of low levels of embryonic and adult β-globin, as well as *Eklf*, *Sox6*, and *Bcl11a*. Interestingly, these erythroid colonies showed low expression levels of all STAT proteins. Together, these data show that embryo-derived primitive, EMP, and definitive erythroid cells have distinct gene expression signatures and that the expression of *Stat1* distinguishes between EMP and definitive erythroid cells.

Expression of *Stat1* in mESC-Derived Wild-Type and *Sox17*$^{-/-}$ Erythroid Populations

We next examined if this gene expression signature can be used to distinguish the ESC-derived erythroid lineages generated from both wild-type and *Sox17*$^{-/-}$ ESCs. The *Sox17*$^{-/-}$ populations were included, as our previous study indicated that development of primitive and EMP programs were SOX17 independent, whereas specification of the definitive lineage required SOX17. In the initial set of experiments, we analyzed the primitive and EMP lineages generated from D3.25 FLK-1^{+} mesoderm, a population that we have previously shown gives rise to the equivalent of yolk sac hematopoiesis (Clarke et al., 2013; Irion et al., 2010). For these studies, the wild-type (control) and *Sox17*$^{-/-}$ FLK-1^{+} cells were isolated by fluorescence-activated cell sorting (FACS) and aggregated overnight. The aggregates were cultured for 48 hr (Figure 4A). As an independent method to distinguish the different erythroid progenitors, we analyzed their dependency on NOTCH signaling, as previous studies have shown that the primitive and EMP hematopoietic programs are NOTCH independent whereas specification of the definitive program is dependent on this pathway (Hadland et al., 2004). To measure NOTCH dependency, we included the NOTCH pathway antagonist gamma secretase inhibitor (γSI) in the aggregate cultures. The resulting populations were assayed in methylcellulose. Both the wild-type and *Sox17*$^{-/-}$ D3.25 FLK-1+ populations generated typical small primitive erythroid (Ery-P) colonies as well as larger erythroid colonies that resembled those that develop from either EMP or definitive progenitors (Ery-EMP/D) (Figure 4B). Although the erythroid potential was intact in the *Sox17*$^{-/-}$ mesoderm, the cellularity as well as the frequency of Ery-P and Ery-EMP progenitors was significantly lower in the *Sox17*$^{-/-}$ cultures than in the wild-type cultures (Figure 4C). The reduced cellularity in *Sox17*$^{-/-}$ populations is consistent with observations in our previous study (Clarke et al., 2013). The addition of γSI did not significantly affect cell numbers or the generation of Ery-P progenitors from either mESC line (Figure 4D). It did, however, lead to a modest but significant reduction in the number of larger wild-type colonies (Ery-EMP/D) indicating that a subset of them originate from a NOTCH-dependent definitive hematopoietic progenitor or that NOTCH is involved in the proliferation, but not emergence, of EMP hematopoietic progenitors. (Figure 4D, black bars). The number of *Sox17*$^{-/-}$ Ery-EMP/D progenitors was not affected by the addition of γSI (Figure 4D, gray bars).

The wild-type and *Sox17*$^{-/-}$ Ery-P colonies had a gene expression signature similar to that of the E8.5 YS-derived primitive erythroid cells, including high levels of *Hbb-εγ* and *Stat3* (Figure 4E). They also expressed low levels of *Eklf* and the other *Stat* genes. *Sox6* and *Bcl11a* were not expressed in these colonies. The expression profile of the wild-type Ery-EMP colonies was also in line with that of the YS-derived colonies and showed significantly lower levels of embryonic globin *Hbb-εγ* and *Stat3* than the Ery-P colonies. As observed with the yolk sac colonies, the mESC-derived Ery-EMP colonies also had low levels of all the other *Stat* genes tested. These colonies did express *Eklf*, *Sox6*, and *Bcl11a* (Figure 4F, black bars). The *Sox17*$^{-/-}$-derived colonies had a similar "EMP" profile, including low levels of *Hbb-εγ* and the *Stat* genes. These colonies did, however, express lower levels of *Hbb-β1* and *Eklf* than the wild-type colonies. The reason for these differences is currently not clear.

To generate ESC-derived definitive progenitors, D5.25 FLK-1^{+} cells were isolated, aggregated, and cultured overnight (Figure 5A). The aggregates were then plated as a monolayer in the presence of cytokines known to induce the EHT and the generation of CD45^{+} hematopoietic progenitors. γSI was included in some of the cultures during the aggregation and Matrigel stages. The wild-type and *Sox17*$^{-/-}$ D9 populations gave rise to erythroid colonies that phenotypically resembled those derived from Ery-EMP or Ery-D progenitors (Figure 5B). Primitive erythroid colonies were not detected in these cultures. The addition of γSI did not impact overall cellularity, but it did lead to a reduction in the proportion of CD41^{+} cells and in the number of erythroid colonies generated in the wild-type population, consistent with the interpretation that they represent the definitive lineage (Figures 5C–5E, black bars). In contrast, neither the size of the *Sox17*$^{-/-}$ CD41^{+} population nor the number of *Sox17*$^{-/-}$ erythroid colonies was affected by the addition of γSI, indicating that these cells represent the EMP program. Gene expression analysis confirmed these differences as the control erythroid colonies revealed an expression pattern comparable to the FL-derived definitive lineage, including *Hbb-β1*^{hi}, *Hbb-εγ*^{neg}, *Stat1*^{hi}, *Eklf*^{hi}, *Sox6*$^{+}$, and *Bcl11a* (Figure 5F, black bars). The *Sox17*$^{-/-}$ colonies, by contrast, displayed an EMP profile that included significantly lower levels of

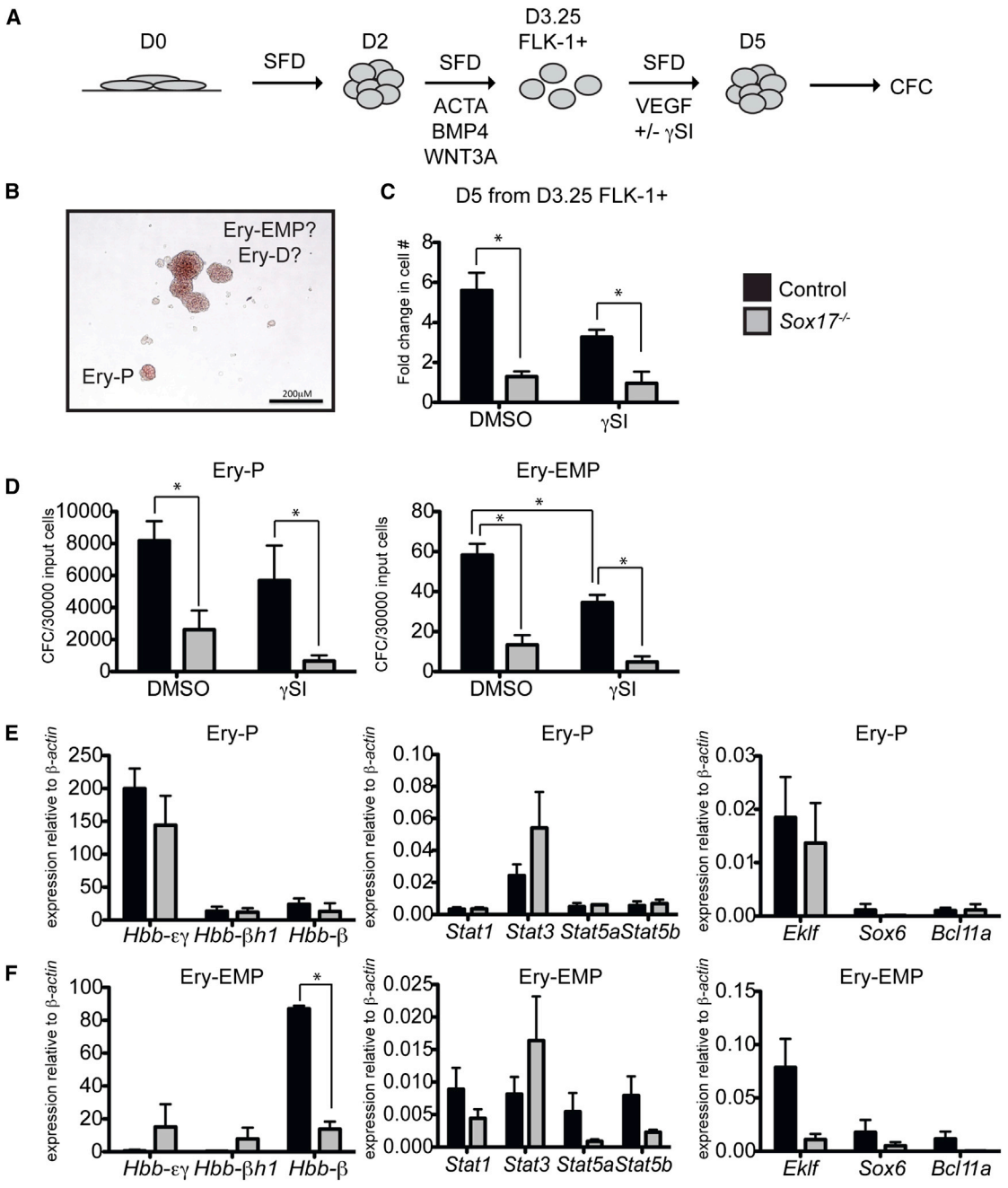

Figure 4. Day 3.25 FLK-1$^+$ Wild-Type and *Sox17*$^{-/-}$ mESC-Derived Populations Contain Primitive and EMP Progenitors

(A) Schematic representation of the differentiation protocol for generating different hematopoietic populations from mESCs.

(B) Representative image of erythroid colonies from a control D5 population generated from D3.25 FLK-1$^+$ cells after 5 days in culture.

(C) Fold change in cell numbers following 48 hr of culture of control and *Sox17*$^{-/-}$ D3.25 FLK-1$^+$ cells in the presence and absence of γ-secretase inhibitor (γSI, L685458, Tocris, 10 μM). Cells cultured in the vehicle DMSO represent the control. DMSO p = 0.00958, γSI p = 0.0285. n = 3 individual experiments (SEM).

(D) Erythroid potential of the D5 population generated from D3.25 FLK-1$^+$ cells. n = 3 individual experiments (SEM). Control EryP DMSO:*Sox17*$^{-/-}$ DMSO p = 0.031, control EryP γSI:*Sox17*$^{-/-}$ γSI p = 0.012, control Ery-EMP DMSO:control Ery-EMP γSI p = 0.024, control Ery-EMP DMSO:*Sox17*$^{-/-}$ Ery-EMP DMSO p = 0.0039, and control Ery-EMP γSI:*Sox17*$^{-/-}$ Ery-EMP γSI p = 0.0032.

(legend continued on next page)

Stat1 and *Hbb-β1* expression than observed in the wild-type definitive colonies. Taken together, these findings show that combination of *Stat1* and globin gene expression can be used to distinguish the different mESC-derived erythroid programs. Additionally, they confirm our earlier interpretation that the primitive and EMP programs are not dependent on SOX17, whereas generation of the definitive lineage is SOX17 dependent. Finally, they show that the EMP-derived erythroid lineage persists to the definitive stage in the *Sox17*$^{-/-}$ cultures.

SOX17 Is Not Required for Development of the Primitive and EMP-Derived Erythroid Lineage In Vivo

To evaluate the SOX17 dependency of primitive and EMP erythroid lineages in vivo, we next analyzed colonies generated from the yolk sacs of E8.5 *Sox17*$^{+/+}$, *Sox17*$^{+/-}$, and *Sox17*$^{-/-}$ embryos for the above expression signature. As observed with the mESC-derived populations, the SOX17-deficient (*Sox17*$^{+/-}$ and *Sox17*$^{-/-}$) embryos contained a lower frequency of primitive erythroid progenitors than the wild-type littermates, possibly due to defects in the yolk sac vasculature resulting from SOX17 deficiency (Kim et al., 2007). In spite of the reduced numbers, the colonies from the three different sets of embryos had very similar expression profiles that included high levels of *Hbb-εγ* and *Stat3* and low levels of the other *Stat* genes (Figures 6A and 6B). Both the *Sox17*$^{+/-}$ and *Sox17*$^{-/-}$ embryos contained progenitors that generated EMP-type colonies. The expression pattern of these colonies was consistent with that of the EMP, with the exception that those derived from the *Sox17*$^{-/-}$ progenitors expressed lower levels *Hbb-β1* than the colonies from the *Sox17*$^{+/-}$ and *Sox17*$^{+/+}$ progenitors, similar to the pattern seen in the *Sox17*$^{-/-}$ mESC-derived EMPs (Figure 5F). Collectively, these data show that SOX17 is not required for the emergence of primitive and EMP-derived hematopoietic program in the embryo.

DISCUSSION

Label-free proteomics represents an emerging technology for identifying and quantifying patterns of protein expression. While differences in patterns of protein expression have previously been shown to correlate with differentiated cell types, it has been much more challenging to develop technologies that can detect regulatory proteins such as transcription factors that specify specific cellular fates. Emerging embryonic HSCs and their progeny have been comprehensively examined via global transcriptional profiling, which revealed key insights into hematopoietic specification and leukemogenesis (Boisset et al., 2010; McKinney-Freeman et al., 2012). The direct precursor to the HSC, the HEC, has not been extensively examined, as only an abbreviated transcriptional profile of HECs within the murine embryo was recently elucidated using a Fluidigm-based approach (Swiers et al., 2013). Here, we have utilized label-free proteomics to decipher key regulators of embryonic hematopoiesis both in vivo and during the directed differentiation of ESCs in vitro. This approach to cellular profiling has the distinct advantage of identifying only translated functional proteins present in the cell, thus eliminating the unknown effects of post-transcriptional modification that may affect mRNA translation.

Our first analysis focused on the proteomic landscape of E10.5 and E11.5 VEC^{+} populations that contain ECs, HECs, and pre-HSCs (Rybtsov et al., 2011). With this approach, we were able to accurately identify and quantify known regulators of the early hematopoietic programs such as SCL/TAL1 and ALDH1A2, as well as a small cohort of proteins that exhibited dramatic expression changes between E10.5 and E11.5, providing a focused and manageable list of candidates to interrogate as putative regulators of the generation of type I pre-HSCs. To date, the generation of PSC-derived HSCs has been difficult, likely due to the undefined nature of PSC differentiation culture conditions and the lack of accurate comparison of the in vitro populations to potential equivalent in vivo embryonic stages. Comparison of D7 mESC-derived HEC populations to the E10.5 and E11.5 HEC populations revealed that our D7 HEC population represents a developmental intermediate. Therefore, interrogation of this dataset will help identify the absence of key regulatory proteins, thus providing a potential explanation for the inability of these cells to faithfully generate HSCs in vitro.

Understanding the regulatory pathways that control the establishment of the different hematopoietic programs and developing methods to identify and isolate progenitors for each are essential to accurately model hematopoiesis in ESC and induced PSC differentiation cultures and ultimately to generate HSCs in vitro. Building on our previous observations that SOX17 is a critical regulator of definitive hematopoiesis (Clarke et al., 2013), in this study, we proceeded to analyze the SOX17-dependent proteome in

(E and F) Real-time qPCR analysis of the erythroid colonies. n = 3 individual experiments (SEM). Control Ery-EMP:*Sox17*$^{-/-}$ Ery-EMP *Hbb-β* p = 0.00012. In addition, but not labeled in the figure, the control EryP and Ery-EMP colonies were significantly different from each other for the following genes; *Hbb-εγ* (p = 0.002), *Hbb-β* (p = 0.0023), and *Stat3* (p = 0.09). The *Sox17*$^{-/-}$ EryP and Ery-EMP colonies were significantly different from each other for the following genes; *Hbb-εγ* (p = 0.05).

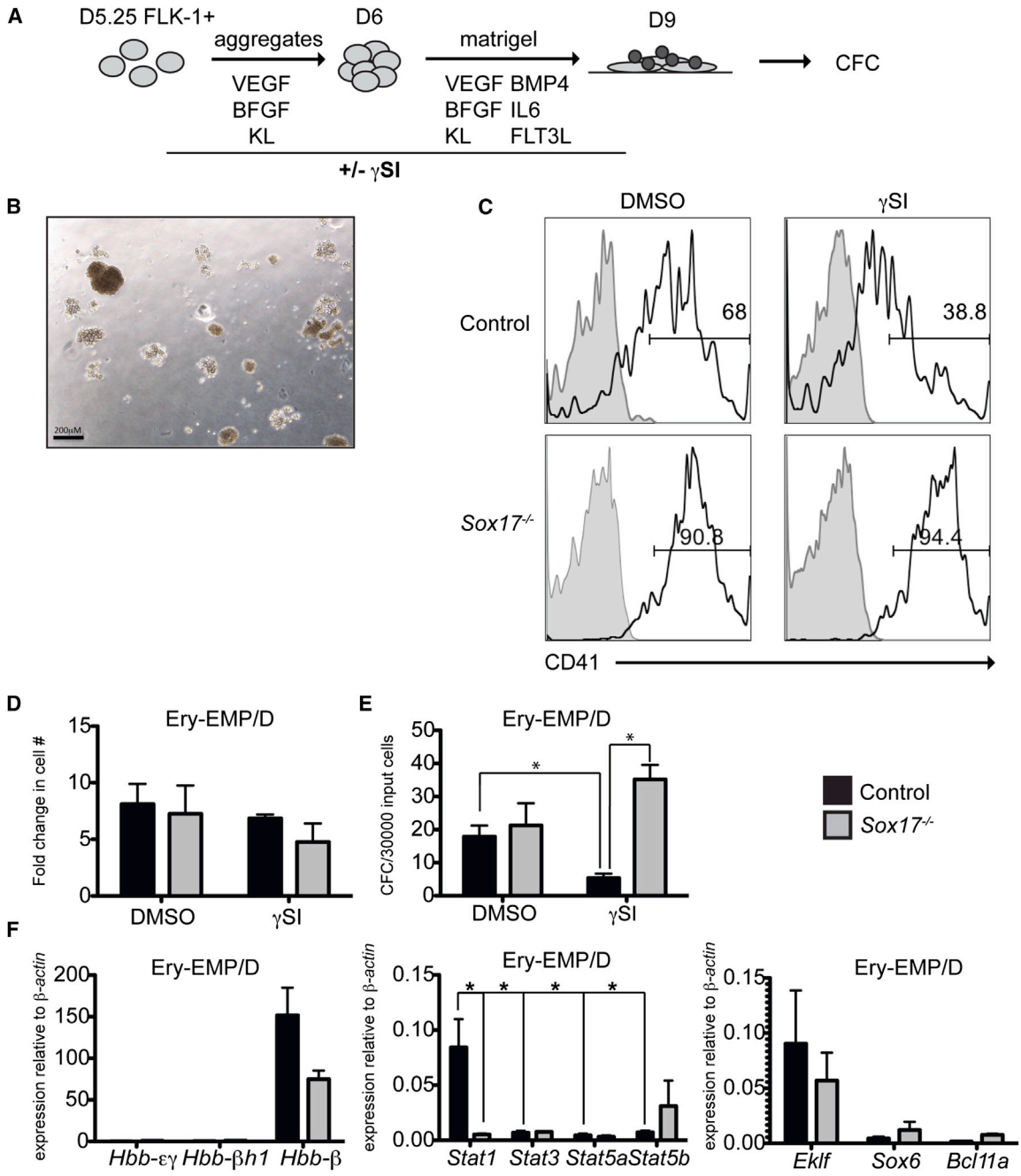

Figure 5. Day 9 mESCs Generate *Stat1*$^{+}$, NOTCH-Dependent, and SOX17-Dependent Definitive Erythroid Cells

(A) Schematic representation of the differentiation protocol for generating definitive hematopoietic progenitors from mESCs.

(B) Representative image of control erythroid colonies from a D9 population generated from D5.25 FLK-1^{+} cells after 8 days in culture.

(C) Representative flow cytometric profiles of cells in D9 populations stained with CD41.

(D) Fold change in cell numbers of D9 populations generated in the presence and absence of γSI.

(E) Erythroid potential of the D9 populations; n = 3 individual experiments (SEM). Control DMSO: control γSI p = 0.024. Control γSI:*Sox17*$^{-/-}$ γSI p = 0.0028.

(F) Real-time qPCR analysis of the erythroid colonies. n = 3 individual experiments (SEM). Control:*Sox17*$^{-/-}$ *Stat1* p = 0.0032, control *Stat1:Stat3* p = 0.03, control *Stat1:Stat5a* p = 0.03, and control *Stat1:Stat5b* p = 0.03.

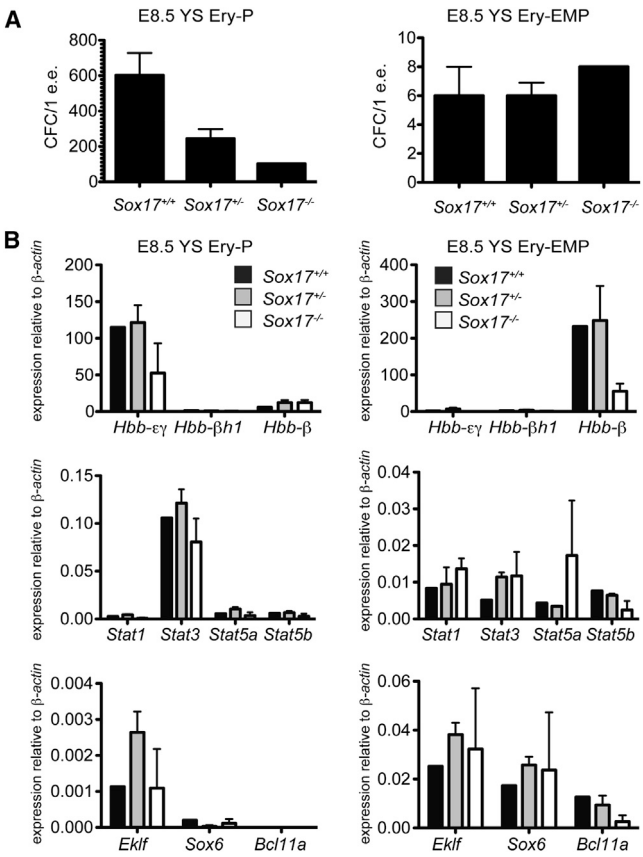

Figure 6. *Sox17* Is Not Required for the Formation of Ery-EMP In Vivo

(A) Erythroid potential of E8.5 YS from *Sox17*$^{+/+}$ (n = 3), *Sox17*$^{+/-}$ (N = 6) and *Sox17*$^{-/-}$ (n = 3) embryos from two dams.

(B) Real-time qPCR analysis of pooled Ery-P and Ery-EMP erythroid colonies. n = 3 individual experiments (SEM).

HECs and hematopoietic progenitors, and through this approach, we were able to identify potential regulatory pathways of the definitive hematopoietic program and markers to distinguish it from the earlier primitive and EMP programs. Specifically, we identified *Stat1* as a marker whose expression distinguishes the definitive erythroid lineage from the primitive and EMP-derived erythroid lineages. The differences observed between the EMP and definitive programs are particularly important, as they are difficult to distinguish by other methods. The observation that the EMP program persists through the later definitive stage of differentiation in the *Sox17*$^{-/-}$ cultures demonstrates that EMP progeny can persist in the PSC-derived differentiation cultures for extended periods of time. Based on these differential patterns of *Stat1* expression, we were able to demonstrate that SOX17 dependence is a distinguishing factor that separates the definitive hematopoietic program from the EMP and primitive hematopoietic programs.

The expression analyses of the mESC-derived populations showing that the D9 progenitors express *Stat1* in a SOX17-dependent fashion provide additional evidence that we were able to specify the definitive program in the differentiation cultures. The analyses of erythroid *Stat1* expression will complement the use of T cell potential as a method to distinguish the PSC-derived hematopoietic programs and monitor the emergence of the definitive lineage in vitro. In addition to demonstrating the emergence of the definitive lineage, our studies show that the early specification of the primitive and EMP lineages can be monitored by the co-expression of *Stat* and globin genes. The observation that the D3.5 FLK-1^{+} progenitors give rise to both the primitive and EMP erythroid lineages that are found in the yolk sac provides further support that the in vitro PSC system accurately recapitulates in vivo embryonic hematopoiesis.

Our STAT protein expression data support the recently published computational prediction of a differential role for STAT signaling in primitive and definitive erythropoiesis (Greenfest-Allen et al., 2013). STAT signaling is central to hematopoietic cell biology, as this pathway is activated in response to numerous cytokines, including erythropoietin (EPO) and stem cell factor (SCF/KL). The observations that EPO activates *Stat1*, *Stat3*, *Stat5a*, and *Stat5b* in mouse and human cell lines in culture, and that STAT1- and STAT5-deficient fetuses display a wide array of erythroid defects, provide clear evidence that STAT proteins are critical regulators of developmental erythropoiesis and erythroblast maturation (Cui et al., 2004; Halupa et al., 2005; Kirito and Komatsu, 2002; Kirito et al., 1997; Richmond et al., 2005; Socolovsky et al., 1999, 2001). Analysis of the role of STAT3 during erythropoiesis is hampered by the fact that *Stat3*$^{-/-}$ mice are embryonic lethal at E7.5 due to a lack of mesoderm formation. However, small-molecule inhibition of STAT3 activation resulted in decreased numbers of primitive, but not adult definitive, erythroblasts in culture, suggesting a differential role of STAT3 during distinct stages of hematopoiesis (Greenfest-Allen et al., 2013). A mechanism by which STAT1 may preferentially regulate the maturation of definitive erythroid cells is through its transcriptional target ISG15. ISG15 is strongly upregulated during erythroid differentiation in the bone marrow (Maragno et al., 2011) and was identified in our proteomics screen as a potential SOX17 target exhibiting a parallel expression pattern to STAT1 (Figure 2D). ISG15 can be covalently linked to lysine residues of target proteins as a method of post-translational modification (Zhang and Zhang, 2011). STAT5 and globin proteins have been shown to be ISGylated in erythroblasts, and ISG15$^{-/-}$ erythroblasts are unable to terminally differentiate, thus providing a testable hypothesis by which STAT1 regulates definitive erythroid differentiation. Collectively, the observation

that distinct developmental subsets of erythroid cells express different STAT proteins presents a previously unidentified mode of regulation of embryonic erythropoiesis.

In summary, the data presented in this study demonstrate that quantitative proteomics is an informative and accurate method to identify global patterns of protein expression in small numbers of cells. By applying this approach to early hematopoietic progenitor populations, we have gained insights into potential pathways that regulate embryonic hematopoietic development and uncovered strategies for identifying the definitive hematopoietic program in PSC differentiation cultures. Further comparison of our proteomics analysis to previously published transcriptome datasets will provide a systems biology view of the early stages of embryonic hematopoiesis and allow us to obtain a better understanding of the signaling and transcription factor networks that underlie the specification of HSCs.

EXPERIMENTAL PROCEDURES

Sample Preparation for Mass Spectrometry

50,000 cells underwent FACS. The cells were then washed in 1× PBS and flash frozen. Cell pellets were lysed in 1 M urea and 50 mM ammonium bicarbonate (pH 7.8) and heated to 50°C for 20 min. Cell debris was removed by centrifugation (1,000 × *g*, 2 min). Following a bicinchoninic acid assay, normalized quantities of protein were reduced with 2 mM DTT, alkylated with 15 mM iodoacetamide, and digested overnight with a 1:50 ratio of trypsin to total protein. The resulting peptides were desalted on Waters Sep-Pak C18 cartridges.

Nano-LC-MS/MS Measurements

Peptides were measured by nano-LC-MS/MS on a Q Exactive (Thermo Scientific). Peptides were separated online by reverse-phase chromatography using a heated 50°C 30-cm C18 columns (75-mm ID packed with Magic C18 AQ 5-μM/100-A beads) in a 120-min gradient (1%–45% acetonitrile with 0.1% formic acid) separated at 250 nl/min. The Q Exactive was operated in the data-dependent mode with the following settings: 70,000 resolution, 300–2,000 m/z full scan, Top 10, and an 1.8-m/z isolation window.

Data Processing and Analysis

Identification and label-free quantification of peptides was done with MaxQuant 1.3.0.5 using a 1% false discovery rate (FDR) against the mouse Swiss-Prot/TrEMB database downloaded from UniProt on October 11, 2013 (Cox and Mann, 2008). The databases contained forward and reverse mouse sequences as well as common contaminants. We analyzed three replicates per condition. Peptides were searched for variable modification of n-term protein acetylation, oxidation (M), deamidation (NQ), with a 6-ppm mass error and a match between run window of 2 min. Proteins that were significantly regulated between conditions were identified using a permutation-based t test (S1, FDR 5%) in Perseus 1.4.1.3. GO terms were analyzed in G-profiler (Reimand et al., 2011).

Real-Time qPCR

Total RNA was prepared with the RNAqueous-Micro Kit (Ambion) and treated with RNase-free DNase (Ambion). 500 ng to 1 μg RNA was reverse transcribed into cDNA using random hexamers and Oligo (dT) with Superscript III Reverse Transcriptase (Invitrogen). qPCR was performed on a MasterCycler EP RealPlex (Eppendorf) using QuantiFast SYBR Green PCR Kit (QIAGEN). Expression levels were normalized to the housekeeping gene *β-actin* (*Actβ*). The oligonucleotide sequences are listed in Table S7.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, three figures, and seven tables and can be found with this article online at http://dx.doi.org/10.1016/j.stemcr.2015.07.008.

AUTHOR CONTRIBUTIONS

R.L.C., A.M.R., G.K., and R.T.M. designed experiments. R.L.C. and A.M.R. performed experiments and analysis. R.L.C. and A.M.R. wrote the manuscript. G.K. and R.T.M. edited the manuscript.

ACKNOWLEDGMENTS

This work is supported in part by the University of Washington's Proteomics Resource (UWPR95794) and by the NIH (U01HL100395). R.T.M. is an Investigator of the HHMI.

Received: May 4, 2015
Revised: July 14, 2015
Accepted: July 22, 2015
Published: August 11, 2015

REFERENCES

Basu, P., Lung, T.K., Lemsaddek, W., Sargent, T.G., Williams, D.C., Jr., Basu, M., Redmond, L.C., Lingrel, J.B., Haar, J.L., and Lloyd, J.A. (2007). EKLF and KLF2 have compensatory roles in embryonic beta-globin gene expression and primitive erythropoiesis. Blood *110*, 3417–3425.

Bertrand, J.Y., Giroux, S., Golub, R., Klaine, M., Jalil, A., Boucontet, L., Godin, I., and Cumano, A. (2005). Characterization of purified intraembryonic hematopoietic stem cells as a tool to define their site of origin. Proc. Natl. Acad. Sci. USA *102*, 134–139.

Boisset, J.C., van Cappellen, W., Andrieu-Soler, C., Galjart, N., Dzierzak, E., and Robin, C. (2010). In vivo imaging of haematopoietic cells emerging from the mouse aortic endothelium. Nature *464*, 116–120.

Chanda, B., Ditadi, A., Iscove, N.N., and Keller, G. (2013). Retinoic acid signaling is essential for embryonic hematopoietic stem cell development. Cell *155*, 215–227.

Clarke, R.L., Yzaguirre, A.D., Yashiro-Ohtani, Y., Bondue, A., Blanpain, C., Pear, W.S., Speck, N.A., and Keller, G. (2013). The

expression of Sox17 identifies and regulates haemogenic endothelium. Nat. Cell Biol. *15*, 502–510.

Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat. Biotechnol. *26*, 1367–1372.

Cui, Y., Riedlinger, G., Miyoshi, K., Tang, W., Li, C., Deng, C.X., Robinson, G.W., and Hennighausen, L. (2004). Inactivation of Stat5 in mouse mammary epithelium during pregnancy reveals distinct functions in cell proliferation, survival, and differentiation. Mol. Cell. Biol. *24*, 8037–8047.

Espín-Palazón, R., Stachura, D.L., Campbell, C.A., García-Moreno, D., Del Cid, N., Kim, A.D., Candel, S., Meseguer, J., Mulero, V., and Traver, D. (2014). Proinflammatory signaling regulates hematopoietic stem cell emergence. Cell *159*, 1070–1085.

Ferkowicz, M.J., Starr, M., Xie, X., Li, W., Johnson, S.A., Shelley, W.C., Morrison, P.R., and Yoder, M.C. (2003). CD41 expression defines the onset of primitive and definitive hematopoiesis in the murine embryo. Development *130*, 4393–4403.

Greenfest-Allen, E., Malik, J., Palis, J., and Stoeckert, C.J., Jr. (2013). Stat and interferon genes identified by network analysis differentially regulate primitive and definitive erythropoiesis. BMC Syst. Biol. *7*, 38.

Hadland, B.K., Huppert, S.S., Kanungo, J., Xue, Y., Jiang, R., Gridley, T., Conlon, R.A., Cheng, A.M., Kopan, R., and Longmore, G.D. (2004). A requirement for Notch1 distinguishes 2 phases of definitive hematopoiesis during development. Blood *104*, 3097–3105.

Halupa, A., Bailey, M.L., Huang, K., Iscove, N.N., Levy, D.E., and Barber, D.L. (2005). A novel role for STAT1 in regulating murine erythropoiesis: deletion of STAT1 results in overall reduction of erythroid progenitors and alters their distribution. Blood *105*, 552–561.

Hodge, D., Coghill, E., Keys, J., Maguire, T., Hartmann, B., McDowall, A., Weiss, M., Grimmond, S., and Perkins, A. (2006). A global role for EKLF in definitive and primitive erythropoiesis. Blood *107*, 3359–3370.

Irion, S., Clarke, R.L., Luche, H., Kim, I., Morrison, S.J., Fehling, H.J., and Keller, G.M. (2010). Temporal specification of blood progenitors from mouse embryonic stem cells and induced pluripotent stem cells. Development *137*, 2829–2839.

Kim, I., Saunders, T.L., and Morrison, S.J. (2007). Sox17 dependence distinguishes the transcriptional regulation of fetal from adult hematopoietic stem cells. Cell *130*, 470–483.

Kirito, K., and Komatsu, N. (2002). [Signal transduction in hematopoiesis: a functional role of Stat 1 and Stat 3 in erythropoiesis and megakaryopoiesis]. Rinsho Ketsueki *43*, 278–281.

Kirito, K., Uchida, M., Yamada, M., Miura, Y., and Komatsu, N. (1997). A distinct function of STAT proteins in erythropoietin signal transduction. J. Biol. Chem. *272*, 16507–16513.

Kissa, K., and Herbomel, P. (2010). Blood stem cells emerge from aortic endothelium by a novel type of cell transition. Nature *464*, 112–115.

Kuang, C.Y., Yang, T.H., Zhang, Y., Zhang, L., and Wu, Q. (2014). Schlafen 1 inhibits the proliferation and tube formation of endothelial progenitor cells. PLoS ONE *9*, e109711.

Lange, A.W., Keiser, A.R., Wells, J.M., Zorn, A.M., and Whitsett, J.A. (2009). Sox17 promotes cell cycle progression and inhibits TGF-beta/Smad3 signaling to initiate progenitor cell behavior in the respiratory epithelium. PLoS ONE *4*, e5711.

Li, Y., Esain, V., Teng, L., Xu, J., Kwan, W., Frost, I.M., Yzaguirre, A.D., Cai, X., Cortes, M., Maijenburg, M.W., et al. (2014). Inflammatory signaling regulates embryonic hematopoietic stem and progenitor cell production. Genes Dev. *28*, 2597–2612.

Luber, C.A., Cox, J., Lauterbach, H., Fancke, B., Selbach, M., Tschopp, J., Akira, S., Wiegand, M., Hochrein, H., O'Keeffe, M., and Mann, M. (2010). Quantitative proteomics reveals subset-specific viral recognition in dendritic cells. Immunity *32*, 279–289.

Malakhov, M.P., Kim, K.I., Malakhova, O.A., Jacobs, B.S., Borden, E.C., and Zhang, D.E. (2003). High-throughput immunoblotting. Ubiquitiin-like protein ISG15 modifies key regulators of signal transduction. J. Biol. Chem. *278*, 16608–16613.

Maragno, A.L., Pironin, M., Alcalde, H., Cong, X., Knobeloch, K.P., Tangy, F., Zhang, D.E., Ghysdael, J., and Quang, C.T. (2011). ISG15 modulates development of the erythroid lineage. PLoS ONE *6*, e26068.

McGrath, K.E., Frame, J.M., Fromm, G.J., Koniski, A.D., Kingsley, P.D., Little, J., Bulger, M., and Palis, J. (2011). A transient definitive erythroid lineage with unique regulation of the β-globin locus in the mammalian embryo. Blood *117*, 4600–4608.

McKinney-Freeman, S., Cahan, P., Li, H., Lacadie, S.A., Huang, H.T., Curran, M., Loewer, S., Naveiras, O., Kathrein, K.L., Konantz, M., et al. (2012). The transcriptional landscape of hematopoietic stem cell ontogeny. Cell Stem Cell *11*, 701–714.

Medvinsky, A., and Dzierzak, E. (1996). Definitive hematopoiesis is autonomously initiated by the AGM region. Cell *86*, 897–906.

Mitjavila-Garcia, M.T., Cailleret, M., Godin, I., Nogueira, M.M., Cohen-Solal, K., Schiavon, V., Lecluse, Y., Le Pesteur, F., Lagrue, A.H., and Vainchenker, W. (2002). Expression of CD41 on hematopoietic progenitors derived from embryonic hematopoietic cells. Development *129*, 2003–2013.

Palis, J., Robertson, S., Kennedy, M., Wall, C., and Keller, G. (1999). Development of erythroid and myeloid progenitors in the yolk sac and embryo proper of the mouse. Development *126*, 5073–5084.

Pandey, A., Fernandez, M.M., Steen, H., Blagoev, B., Nielsen, M.M., Roche, S., Mann, M., and Lodish, H.F. (2000). Identification of a novel immunoreceptor tyrosine-based activation motif-containing molecule, STAM2, by mass spectrometry and its involvement in growth factor and cytokine receptor signaling pathways. J. Biol. Chem. *275*, 38633–38639.

Reimand, J., Arak, T., and Vilo, J. (2011). g:Profiler–a web server for functional interpretation of gene lists (2011 update). Nucleic Acids Res. *39*, W307–W315.

Richmond, T.D., Chohan, M., and Barber, D.L. (2005). Turning cells red: signal transduction mediated by erythropoietin. Trends Cell Biol. *15*, 146–155.

Rybtsov, S., Sobiesiak, M., Taoudi, S., Souilhol, C., Senserrich, J., Liakhovitskaia, A., Ivanovs, A., Frampton, J., Zhao, S., and

Medvinsky, A. (2011). Hierarchical organization and early hematopoietic specification of the developing HSC lineage in the AGM region. J. Exp. Med. *208*, 1305–1315.

Sankaran, V.G., Xu, J., Ragoczy, T., Ippolito, G.C., Walkley, C.R., Maika, S.D., Fujiwara, Y., Ito, M., Groudine, M., Bender, M.A., et al. (2009). Developmental and species-divergent globin switching are driven by BCL11A. Nature *460*, 1093–1097.

Sawamiphak, S., Kontarakis, Z., and Stainier, D.Y. (2014). Interferon gamma signaling positively regulates hematopoietic stem cell emergence. Dev. Cell *31*, 640–653.

Schwarz, D.A., Katayama, C.D., and Hedrick, S.M. (1998). Schlafen, a new family of growth regulatory genes that affect thymocyte development. Immunity *9*, 657–668.

Sinner, D., Rankin, S., Lee, M., and Zorn, A.M. (2004). Sox17 and beta-catenin cooperate to regulate the transcription of endodermal genes. Development *131*, 3069–3080.

Socolovsky, M., Fallon, A.E., Wang, S., Brugnara, C., and Lodish, H.F. (1999). Fetal anemia and apoptosis of red cell progenitors in Stat5a-/-5b-/- mice: a direct role for Stat5 in Bcl-X(L) induction. Cell *98*, 181–191.

Socolovsky, M., Nam, H., Fleming, M.D., Haase, V.H., Brugnara, C., and Lodish, H.F. (2001). Ineffective erythropoiesis in Stat5a(-/-) 5b(-/-) mice due to decreased survival of early erythroblasts. Blood *98*, 3261–3273.

Swiers, G., Baumann, C., O'Rourke, J., Giannoulatou, E., Taylor, S., Joshi, A., Moignard, V., Pina, C., Bee, T., Kokkaliaris, K.D., et al. (2013). Early dynamic fate changes in haemogenic endothelium characterized at the single-cell level. Nat. Commun. *4*, 2924.

Van Handel, B., Montel-Hagen, A., Sasidharan, R., Nakano, H., Ferrari, R., Boogerd, C.J., Schredelseker, J., Wang, Y., Hunter, S., Org, T., et al. (2012). Scl represses cardiomyogenesis in prospective hemogenic endothelium and endocardium. Cell *150*, 590–605.

Yi, Z., Cohen-Barak, O., Hagiwara, N., Kingsley, P.D., Fuchs, D.A., Erickson, D.T., Epner, E.M., Palis, J., and Brilliant, M.H. (2006). Sox6 directly silences epsilon globin expression in definitive erythropoiesis. PLoS Genet. *2*, e14.

Yokomizo, T., Yamada-Inagawa, T., Yzaguirre, A.D., Chen, M.J., Speck, N.A., and Dzierzak, E. (2012). Whole-mount three-dimensional imaging of internally localized immunostained cells within mouse embryos. Nat. Protoc. *7*, 421–431.

Yokota, T., Oritani, K., Butz, S., Kokame, K., Kincade, P.W., Miyata, T., Vestweber, D., and Kanakura, Y. (2009). The endothelial antigen ESAM marks primitive hematopoietic progenitors throughout life in mice. Blood *113*, 2914–2923.

Zhang, D., and Zhang, D.E. (2011). Interferon-stimulated gene 15 and the protein ISGylation system. J. Interferon Cytokine Res. *31*, 119–130.

Stem Cell Reports, Volume *5*

Supplemental Information

A Quantitative Proteomic Analysis of Hemogenic Endothelium Reveals Differential Regulation of Hematopoiesis by SOX17

Aaron M. Robitaille, Raedun L. Clarke, Randall T. Moon, and Gordon Keller

SUPPLEMENTAL FIGURE TITLES AND LEGENDS

Figure S1. Proteomic analysis of VE-CADHERIN regulated proteins. Related to Figure 1. A. Protein expression of VEC^+ vs VEC^- populations was analyzed by unsupervised hierarchal cluster analysis [HC]. Green denotes high, relative protein expression and blue low, relative protein expression. B. Protein expression of E10.5 VEC^+ vs E11.5 VEC^+ was analyzed by HC. C. GO-terms linked to proteins enriched in E10.5 VEC^+ populations were analyzed in G-profiler using the cocoa algorithm.

Figure S2. Regulation of SOX17-dependent protein expression during a time course of directed differentiation. Related to Figure 2. A. Schematic outlining inducible SOX17 time course comparison. B. Volcano plot of differentially expressed proteins in D7_dox+ vs D7_dox-. Significant hits are shown in red (FDR<0.05). C. SOX17 protein expression was quantified by nano-LC-MS/MS on a Q Exactive. D. SOX17-regulated changes in protein expression were measured by nano-LC-MS/MS during a time course of directed differentiation to hemogenic endothelium. E. Expression of proteins linked to erythrocyte function is displayed. F. GO-terms linked to proteins regulated by SOX17 overexpression were analyzed in G-profiler using the cocoa algorithm.

Figure S3. *Stat1* is not a direct target of SOX17 in HE. Related to Figure 3. A. STAT3, STAT5A, STAT5B protein expression was quantified by nano-LC-MS/MS on a Q Exactive. B. Luciferase assays using HEK293T cells co-transfected with pGL3 control vector or the *Stat1* promoter contructs and increasing concentrations of the pCAG

expression plasmids. Bars represent mean luciferase intensity relative to pGL3-empty +/- SEM, N=3 independent experiments. C. qRT-PCR analysis of D5.25 Flk-1^{+} cells cultured for 2 days with (+DOX) or without (-DOX) 1ug/mL doxycycline to induce *Sox17* cDNA expression P > 0.05 D. Representative flow cytometric analysis of VE-CADHERIN (VEC), CD45 and CD41 on D7 cells derived from wildtype and *Sox17*$^{-/-}$ mESCs.

SUPPLEMENTAL TABLE TITLES AND LEGENDS

Table S1. All quantified proteins. Related to Figure 1. Peptides corresponding to 4044 proteins were measured by nano-LC-MS/MS on a Q Exactive mass spectrometer and quantified using label free methods across both the *in vivo* and in vitro samples. Common contaminants and reverse hits were removed.

Table S2. Proteins enriched in VE-CADHERIN^{+} or VE-CADHERIN^{-} cell populations in vivo. Related to Figure 1. Proteins that were differentially expressed between conditions were identified using a permutation-based t-test (S1, FDR 5%) in Perseus 1.4.1.3. Comparison of E10.5/E11.5 VEC^{+} to E10.5/E11.5 VEC^{-} populations.

Table S3. Proteins enriched in VE-CADHERIN^{+} or VE-CADHERIN^{-} cell populations in vitro. Related to Figure 1. Proteins that were differentially expressed between conditions were identified as in Table S2. Analysis of proteins expressed in mESC VEC^{+} versus mESC VEC^{-} cell populations.

Table S4. Proteins enriched in VE-CADHERIN$^+$ cell populations in vivo. Related to Figure 1. Proteins that were differentially expressed between conditions were identified as in Table S2. Comparison of proteins measured from E10.5 VEC$^+$CD45$^-$ and E11.5 VEC$^+$CD45$^-$ populations.

Table S5. Proteins regulated by inducible expression of SOX17 in mESCs. Related to Figure S2. D5.25 Flk-1$^+$ endothelial progenitors were treated with doxycycline for 1-3 days to induce the expression of SOX17. Samples were isolated from D6-D9 and analyzed for total protein expression using label free proteomics as in Table S2.

Table S6. Proteins enriched in *Sox17*$^{-/-}$ mESCs. Related to Figure 2. Total protein expression of D7 VEC$^+$CD45$^-$ cell populations from WT or *Sox17*$^{-/-}$ mESCs were analyzed as in Table S2.

Table S7. Primer Sequences. A list of the sequences used in primers for QPCR.

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

mESC maintenance and differentiation

mESC lines were maintained in serum-free/feeder-free culture conditions (SF-ES) supplemented with 2i (Gadue et al., 2006; Ying et al., 2008). mESC-1 is a derivative of the E14 mESC line from the 129/Ola mouse strain, mESC-2 is derived from the C57BL/6J x 129S6 mouse strain. mESCs were differentiated in a serum-free media

protocol. mESCs were trypsinized using Trypsin-LE (Invitrogen) and cultured in serum-free differentiation (SF-D) media in the absence of growth factors at a concentration of $2x10^5$/ml for 48hrs to form EBs. At day 2 (D2) EBs were dissociated with Typsin-LE and reaggregated at a concentration of $2x10^5$/ml in SF-D media containing recombinant human (rh) VEGF (5ng/ml, R&D, cat # 293-VE), rhBMP4 (1ng/ml, R&D, cat # 314-BP), recombinant mouse (rm) Wnt3a (3ng/ml, R&D, cat # 1324-WN) and ActivinA (1ng/ml, R&D, cat# 338-AC) unless otherwise indicated for 30hrs. At D3+6hr EBs were dissociated and reaggregated at a concentration of $5x10^5$/ml in 6 well tissue culture pluronic treated plates in SF-D containing ActA (1ng/ml) for 24hrs. At day 4+6 a final volume rhVEGF (5ng/ml), rhBMP4 (10ng/ml) and SB431452 (6uM, Sigma, cat# S-4317) were added to the cultures. At day 5+6 EBs were dissociated using TrypLE and Flk-1^+ cells were isolated by FACS. Flk-1^+ cells were aggregated overnight in 96 well locluster plates at a concentration of $3x10^4$/30uL in SF-D containing KL (2%vol/vol)-conditioned medium, rhVEGF (5ng/ml) and rhbFGF (10ng/ml, R&D, cat# 233-FB). The next morning aggregates were spoted on Matrigel (VWR, cat# 354230)-coated 24 well plates for 12hrs to allow adherance and then the wells were filled with SF-D containing KL (2%), rhVEGF (5ng/ml) and rhbFGF (10ng/ml) to form HE. At day 7+6 monolayers were trypsinized using Trypsin-LE and VEC^+CD45^- HE cells were isolated by FACS or maintained as monolayers and induced to undergo the EHT by culturing in SF-D containing KL (2%), rhVEGF (10ng/ml), rhbFGF (10ng/ml), rmTPO (50ng/ml, R&D, cat3 488-TO), rmIL6 (10ng/ml, R&D, cat# 406-ML), hIL11 (10ng/ml, R&D, cat# 218-IL), rmFlt3L (10ng/ml, R&D, cat3 427-FL) and rhBMP4 (30ng/ml) to induce hematopoiesis.

Embryo generation

Animals were maintained and bred the Animal Resource Center in the MaRS complex according to institutional guidelines. Embryos were generated through timed matings with the day of vaginal plug designated embryonic day 0.5. Pregnant dams were sacrificed and the developmental stage of the embryos was determined by counting the number of somite pairs.

Embryo cell preparation

Embryonic cells for methylcellulose assays were prepared as follows. The yolk sac and fetal livers from embryos were dissected and dissociated in 0.2 Collagenase B for 20' at 37°C. Cells were washed twice in IMDM and plated in methylcellulose assays at the following concentrations; E8.5 yolk sacs 1 yolk sac was plated per 1ml of methylcellulose, E14 fetal livers $5x10^4$ total cells was plated per 1 ml of methylcellulose.

Hematopoietic progenitor colony assay

Cells were plated in 1% methylcellulose (wt/vol, Sigma) containing 10% (vol/vol) plasma-derived serum (PDS; Antech), 5% (vol/vol) protein-free hybridoma medium (PFHM-II; Invitrogen) and the following cytokines: KL (2% (vol/vol) conditioned medium), TPO (50ng/ml), erythropoietin (2 U/ml), IL11 (25ng/ml) IL3 (1% (vol/vol) conditioned medium), GM-CSF (1% (vol/vol) conditioned medium), Transferrin (0.5%vol/vol) and IL6 (5 ng/ml). Cultures were maintained at 37°C, 5% CO_2 for 5-12 days. Mean and standard errors of 3 independent experiments were calculated.

Flow cytometry and cell sorting

EBs generated during differentiations were dissociated using Trypsin-LE and then stained in FACS buffer (PBS+3%FCS+0.02%NaN3) with the following antibodies: anti-Flk-1 (hybridoma), anti-CD41 APC (cat# 17-0411, eBioscience), anti-CD144 PECy7 (cat# 25-1441, eBioscience), anti-CD45 eF450 (cat# 48-0451, eBioscience). Isotype control antibodies were used for the negative control (cat# 16-4301, eBioscience,). Cells were analyzed on an LSR II flow cytometer (BD Biosciences) and analyzed with FlowJo (Tree Star, Ashland OR). Cells were sorted on a FACSAria (BD Biosciences).

Luciferase assays

Approximately 2.5KB of the *Stat1* sequence upstream of the transcriptional start site was clone into the pGL3-basic vector (Kindly provided by Dr. John Crispino, Northwestern University, Chicago IL). HEK293 ($5x10^4$) were plated in triplicate and transfected with 100ng of the indicated promoter/reporter constructs, 10ng pRL-TK, 3-300ng pCag-Sox17 or vector control using X-tremeGENE 9 (Roche). 24hrs after stimulation firefly and *Renilla* substrate activity was measured using the Dual-Luciferase Reporter Assay System (Promega). Firefly values were normalized to *Renilla* and then all normalized values set relative to the pGL3 control vector.

SUPPLEMENTAL REFERENCES

Gadue, P., Huber, T.L., Paddison, P.J., and Keller, G.M. (2006). Wnt and TGF-beta signaling are required for the induction of an in vitro model of primitive streak formation

using embryonic stem cells. Proceedings of the National Academy of Sciences of the United States of America *103*, 16806-16811.
Ying, Q.L., Wray, J., Nichols, J., Batlle-Morera, L., Doble, B., Woodgett, J., Cohen, P., and Smith, A. (2008). The ground state of embryonic stem cell self-renewal. Nature *453*, 519-523.

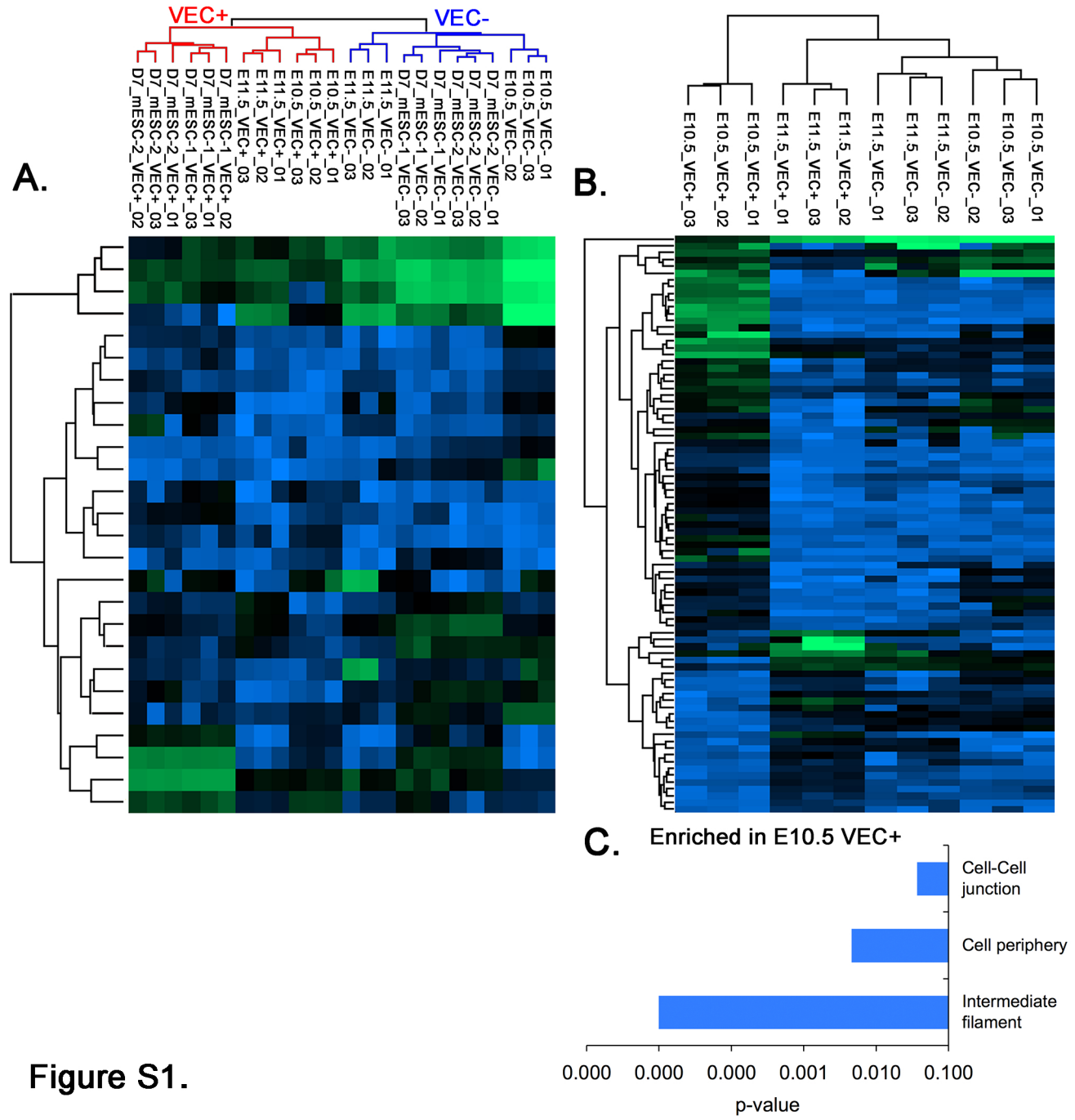
A.
VEC+
VEC-
D7_mESC-2_VEC+_02
D7_mESC-2_VEC+_03
D7_mESC-2_VEC+_01
D7_mESC-1_VEC+_03
D7_mESC-1_VEC+_01
D7_mESC-1_VEC+_02
E11.5_VEC+_03
E11.5_VEC+_02
E11.5_VEC+_01
E10.5_VEC+_03
E10.5_VEC+_02
E10.5_VEC+_01
E11.5_VEC-_03
E11.5_VEC-_02
E11.5_VEC-_01
D7_mESC-1_VEC-_03
D7_mESC-1_VEC-_02
D7_mESC-1_VEC-_01
D7_mESC-2_VEC-_03
D7_mESC-2_VEC-_02
D7_mESC-2_VEC-_01
E10.5_VEC-_03
E10.5_VEC-_02
E10.5_VEC-_01
B.
E10.5_VEC+_03
E10.5_VEC+_02
E10.5_VEC+_01
E11.5_VEC+_01
E11.5_VEC+_03
E11.5_VEC+_02
E11.5_VEC-_01
E11.5_VEC-_03
E11.5_VEC-_02
E10.5_VEC-_02
E10.5_VEC-_03
E10.5_VEC-_01
C.
Enriched in E10.5 VEC+
Cell-Cell junction
Cell periphery
Intermediate filament
0.000
0.000
0.000
0.001
0.010
0.100
p-value

Figure S1.

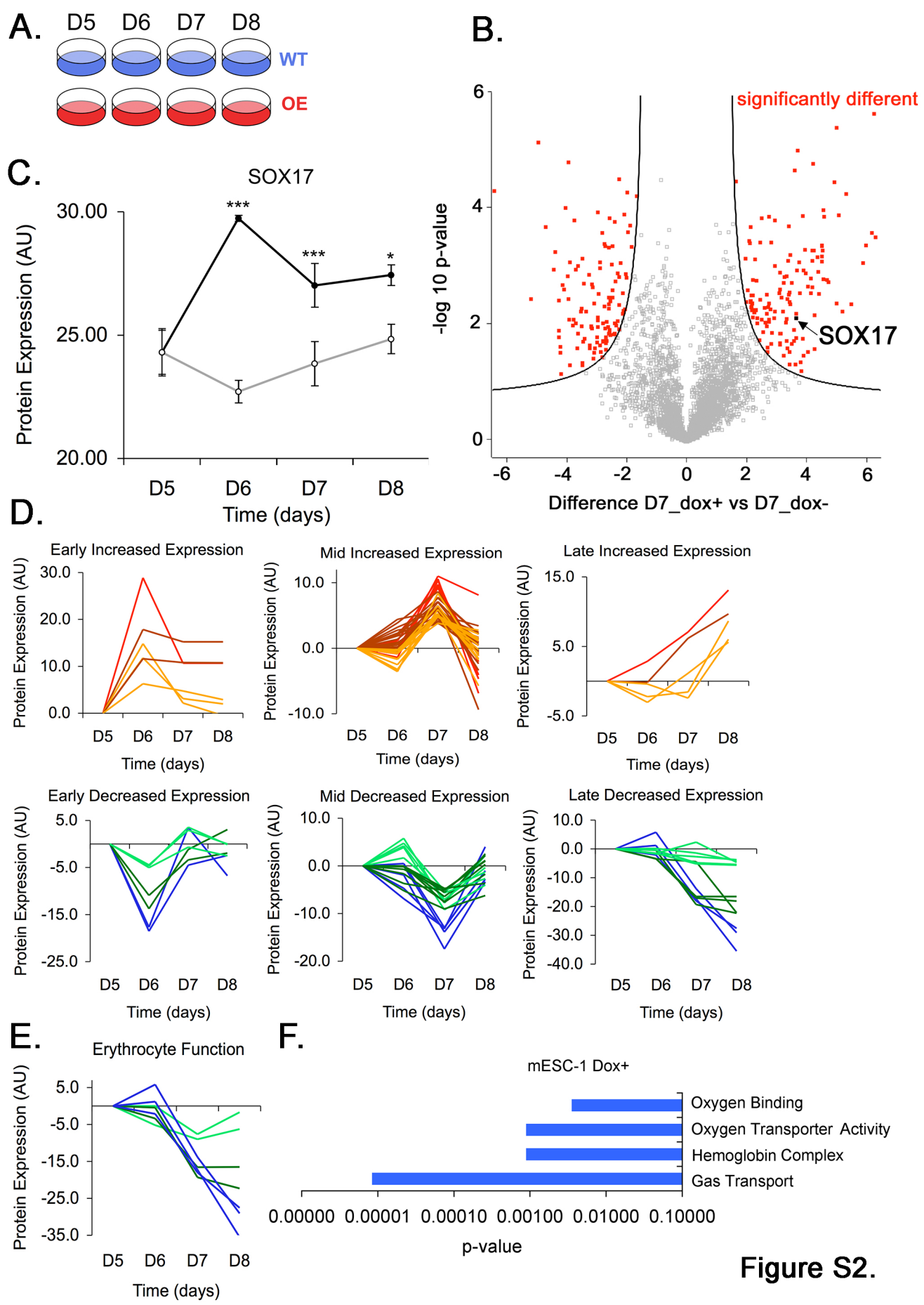
A.
D5
D6
D7
D8
WT
OE
B.
significantly different
SOX17
-log 10 p-value
Difference D7_dox+ vs D7_dox-
C.
SOX17
Protein Expression (AU)
Time (days)
D.
Early Increased Expression
Mid Increased Expression
Late Increased Expression
Early Decreased Expression
Mid Decreased Expression
Late Decreased Expression
E.
Erythrocyte Function
F.
mESC-1 Dox+
Oxygen Binding
Oxygen Transporter Activity
Hemoglobin Complex
Gas Transport
p-value

Figure S2.

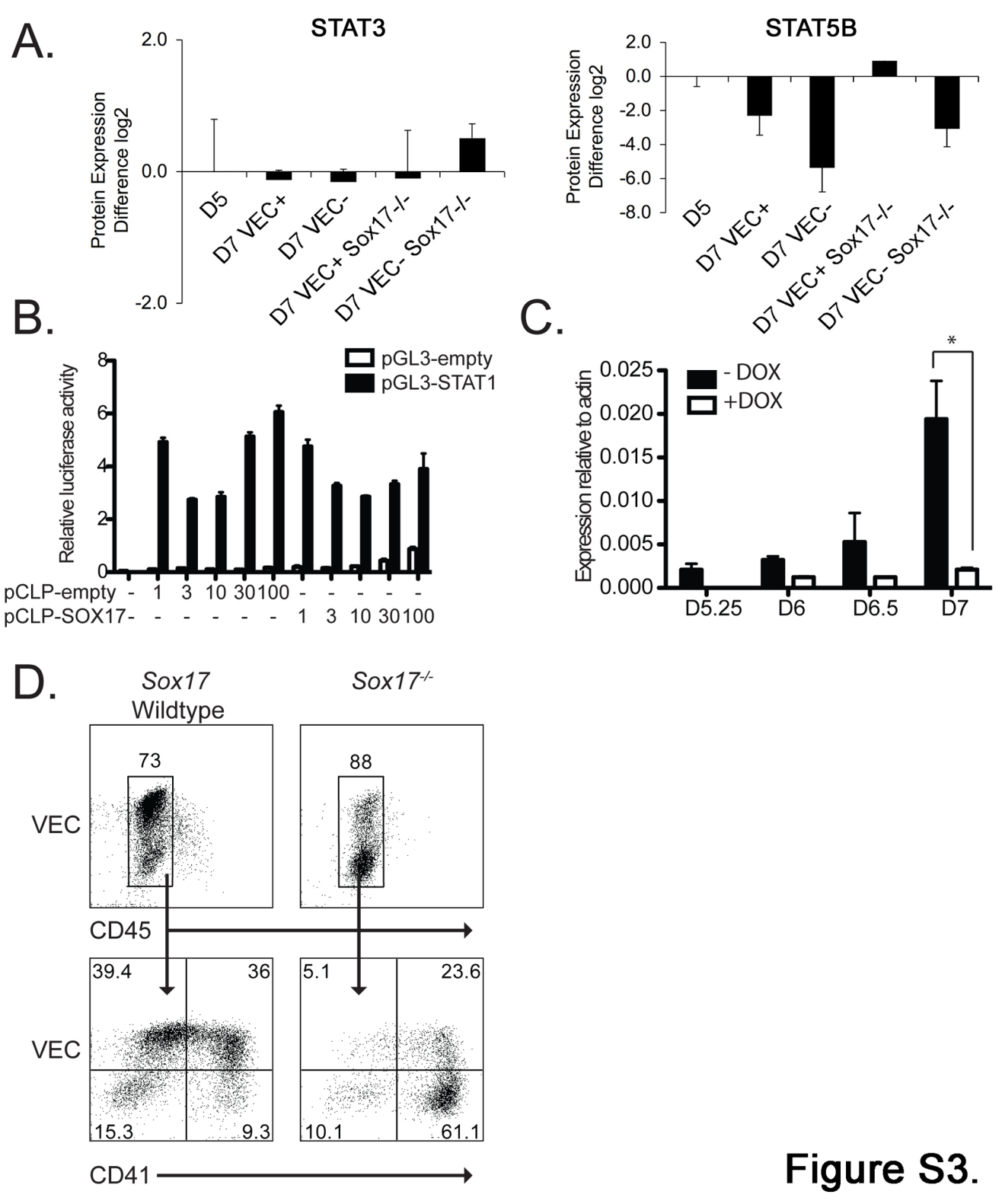
A.
STAT3
Protein Expression Difference log2
2.0
0.0
-2.0
D5
D7 VEC+
D7 VEC-
D7 VEC+ Sox17-/-
D7 VEC- Sox17-/-
STAT5B
Protein Expression Difference log2
2.0
0.0
-2.0
-4.0
-6.0
-8.0
D5
D7 VEC+
D7 VEC-
D7 VEC+ Sox17-/-
D7 VEC- Sox17-/-
B.
pGL3-empty
pGL3-STAT1
Relative luciferase activity
8
6
4
2
0
pCLP-empty - 1 3 10 30 100 - - - - -
pCLP-SOX17- - - - - - 1 3 10 30 100
C.
*
- DOX
+DOX
Expression relative to actin
0.025
0.020
0.015
0.010
0.005
0.000
D5.25
D6
D6.5
D7
D.
Sox17
Wildtype
Sox17-/-
73
88
VEC
CD45
39.4
36
15.3
9.3
5.1
23.6
10.1
61.1
VEC
CD41

Figure S3.