A Cdx4-Sall4 Regulatory Module Controls the Transition from Mesoderm Formation to Embryonic Hematopoiesis

Elizabeth J. Paik,1 Shaun Mahony,2 Richard M. White,3 Emily N. Price,3 Anthony DiBiase,1 Bilguujin Dorjsuren,1 Christian Mosimann,4 Alan J. Davidson,5 David Gifford,6 and Leonard I. Zon1,*

1Stem Cell Program and Division of Hematology/Oncology, Boston Children’s Hospital, Harvard Stem Cell Institute, Harvard Medical School and Howard Hughes Medical Institute, Boston, MA 02115, USA
2Department of Biochemistry and Molecular Biology, Center for Eukaryotic Gene Regulation, Pennsylvania State University, University Park, PA 16802, USA
3Department of Cancer Biology and Genetics, Memorial Sloan-Kettering Cancer Center, New York, NY 10065, USA
4Institute of Molecular Life Sciences, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland
5Department of Molecular Medicine and Pathology, School of Medical Sciences, The University of Auckland, Auckland 1142, New Zealand
6Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
*Correspondence: zon@enders.tch.harvard.edu

This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial-No Derivative Works License, which permits non-commercial use, distribution, and reproduction in any medium, provided the original author and source are credited.

SUMMARY

Deletion of caudal/cdx genes alters hox gene expression and causes defects in posterior tissues and hematopoiesis. Yet, the defects in hox gene expression only partially explain these phenotypes. To gain deeper insight into Cdx4 function, we performed chromatin immunoprecipitation sequencing (ChiP-seq) combined with gene-expression profiling in zebrafish, and identified the transcription factor spekk-like 4 (sall4) as a Cdx4 target. ChiP-seq revealed that Sall4 bound to its own gene locus and the cdx4 locus. Expression profiling showed that Cdx4 and Sall4 coregulate genes that initiate hematopoiesis, such as hox, scl, and lmo2. Combined cdx4/sall4 gene knockdown impaired erythropoiesis, and overexpression of the Cdx4 and Sall4 target genes scl and lmo2 together rescued the erythroid program. These findings suggest that auto- and cross-regulation of Cdx4 and Sall4 establish a stable molecular circuit in the mesoderm that facilitates the activation of the blood-specific program as development proceeds.

INTRODUCTION

Primitive hematopoietic progenitors arise from the yolk sac in mammals and generate red blood cells (RBCs), thereby providing oxygen to the rapidly developing embryos (Orkin and Zon, 2008). In zebrafish, equivalent cells arise from the lateral plate mesoderm (LPM), where anteriorly located cells give rise to myeloid cells and posteriorly located cells produce mostly RBCs. The first hematopoietic progenitors appear bilaterally from the 2- to 3-somite stage and express the transcription factor (TF) genes fli1a, scl, and lmo2 (Liao et al., 1998; Thompson et al., 1998). By the 5-somite stage, posterior LPM cells are specified to the RBC lineage expressing gata1 (Davidson et al., 2003; Detrich et al., 1995). As embryos develop, these bilateral stripes merge, creating the intermediate cell mass (ICM) region (Detrich et al., 1995). By 24 hr postfertilization (hpf), the embryonic RBCs start circulating.

Cdx4, a member of the caudal family, has been linked to embryonic hematopoiesis and leukemogenesis (Bansal et al., 2006; Davidson et al., 2003; Wang et al., 2005). The Cdx genes encode homeodomain-containing TFs that are known as master regulators of the Hox genes, and help establish the anterior-posterior (A-P) axis (Pownall et al., 1996; Subramanian et al., 1995). Mammals have three paralogs of the Cdx family (Cdx1, Cdx2, and Cdx4) that are expressed in the posterior tissues of the embryo (Young and Deschamps, 2009). Targeted knockout of Cdx genes demonstrated their roles in paraxial mesoderm, neuroectoderm, and endoderm formation in mice (Chawengsaksophak et al., 2004; Gao et al., 2009; van den Akker et al., 2002; van Nes et al., 2006; Young et al., 2009). For example, Cdx2/4 compound knockout mice show a truncated axial skeleton, decreased presomitic mesoderm, and defective caudal hindgut endoderm and placenta, indicating that Cdx genes function redundantly in mesendodermal tissue formation (van Nes et al., 2006; Young et al., 2009). Zebrafish also have three cdx paralogs: cdx1a, cdx1b, and cdx4. cdx4−/− embryos display shortened tail and neurectoderm defects, which are enhanced when cdx1a is also knocked down (Davidson et al., 2003; Davidson and Zon, 2006; Shimizu et al., 2006).

Zebrafish cdx4−/− embryos display a severe anemia, with decreased levels of scl, lmo2, and gata1 expression in the posterior LPM (Davidson et al., 2003). This function of Cdx4 in hematopoiesis is conserved in murine embryonic stem cells (ESCs), where Cdx4 overexpression increases erythroid, megakaryocyte, granulocyte, and macrophage lineage formation (Wang et al., 2005). Compound Cdx knockout in murine ESCs also leads to failures in hematopoietic differentiation (Wang et al., 2008). In addition to embryonic hematopoiesis, Cdx members are implicated in leukemogenesis, as shown by CDX2 translocations in acute myeloid leukemia (AML) patients, and leukemia...
seen in mice with Cdx4 overexpression in bone marrow (Bansal et al., 2006; Scholl et al., 2007).

Many of the functions of Cdx have been linked to its ability to regulate Hox gene transcription (Young and Deschamps, 2009). However, an impact on Hox gene regulation does not explain all the defects in Cdx mutants. For example, conditional Cdx2 knockout mice show defects in endoderm and presomatic mesoderm formation independently of Hox genes (Gao et al., 2009; Savory et al., 2009a). In addition, overexpression of anterior hox genes does not rescue the hindbrain defects seen in cdx4+/−/cdx1a−/− zebrafish (Skromne et al., 2007). These results suggest that Cdx function cannot be explained solely by Hox genes, and there must be other critical downstream genes.

Advances in chromatin immunoprecipitation sequencing (ChIP-seq) technology have helped investigators decipher complex transcriptional networks in many biological systems in a global manner. In ESCs, ChIP-seq experiments revealed that Oct4, Nanog, Sox2, and Sall4 share cis-regulatory modules and regulate each other to maintain pluripotency (Lim et al., 2008; Loh et al., 2006). Similar global studies were conducted in developing Drosophila and zebrafish embryos (Morley et al., 2009; Sandmann et al., 2006, 2007). Here, we used a genome-wide approach to identify direct Cdx4 target genes in zebrafish embryos by implementing ChIP-seq and microarray expression profiling. We show that the zinc-finger TF gene sall4 is a downstream target of Cdx4. ChIP-seq analysis of Sall4 similarly demonstrates that it binds to its own promoter and cdx4 regulatory elements. Gene-expression studies demonstrate that Cdx4 and Sall4 directly activate hox genes and hematopoietic genes. Cdx4 and Sall4 genetically interact during ventral mesoderm development, consistent with a role for these TFs in the early expression of hematopoietic TFs. Together, these results suggest that auto- and cross-regulatory interactions between cdx4 and sall4, as well as coactivation of common gene targets, establish a key regulatory module that directs the transition of the mesoderm into blood.

RESULTS

Identification of Direct Cdx4 Downstream Target Genes
To identify direct Cdx4 downstream target genes, we conducted ChIP-seq using zebrafish embryos in the bud stage, the stage just prior to the emergence of fli1a+, scl+, and lmo2+ hematopoietic progenitors. Since ChIP-quality zebrafish Cdx4 antibodies were not available, a full-length, C-terminally myc-tagged zebrafish Cdx4 was overexpressed and myc antibodies were used for immunoprecipitation. This technique has been used successfully for Nanog ChIP-seq analysis in zebrafish (Xu et al., 2012). To reduce nonspecific Cdx4 binding, cdx4-myc mRNA was injected at a dose that rescued cdx4+/− animals but did not cause morphological changes in cdx4−/− animals (Figure S1A available online).

Cdx4 ChIP-seq yielded 5,166 binding events, which corresponded to 1,343 proximal genes (distance to the transcriptional start site: ±10 kb). De novo motif analysis using the top 500 bound sequences gave an ATAAA motif, which was reported as the Cdx2 consensus DNA-binding motif (Figure 1A; Nishiyama et al., 2009; Verzi et al., 2010). The Cdx family members share a highly similar homeodomain and are functionally redundant (Savory et al., 2009b). The most well-known targets of the Cdx family members are Hox genes (Charité et al., 1998; Knittel et al., 1995; Subramanian et al., 1995), and strong Cdx4 binding was found across all seven clusters of hox genes in the zebrafish genome (Figures 1B and S1B). Cdx4 bound to its own locus, suggesting possible autoregulation (Figure 1C).

Binding to specific signaling pathways and gene programs does not necessarily mean that Cdx4 regulates their transcription at the time of the analysis. To demonstrate that Cdx4 binding to these loci leads to transcriptional changes during the LPM-to-hematopoietic transition, we evaluated gene expression by microarrays. Given the redundancy between Cdx4 and Cdx1a, we used 2-somite stage embryos coinjected with cdx4 and cdx1a morpholinos (cdx4−/−/cdx1a−/−) because these embryos display the most severe loss-of-Cdx-function phenotype (Davidson and Zon, 2006). To systematically compare the ChIP-seq target list with the microarray data, we ranked differentially regulated genes in the cdx4−/−/cdx1a−/− microarray data using the Gene Set Enrichment Analysis (GSEA) and compared them with Cdx4 binding (Subramanian et al., 2005). This analysis showed that the Cdx4-bound genes were highly downregulated in the cdx4−/−/cdx1a−/− microarray (normalized enrichment score [NES] = 1.76, false discovery rate [FDR] = 0.1; Figure S1C), suggesting that Cdx4 works as a transcriptional activator, consistent with data previously reported with Cdx2 (Nishiyama et al., 2009; Verzi et al., 2011). For example, 26 of the 44 Cdx4-bound hox genes and the Cdx4 bound-Wnt pathway genes, including axin1, axin2, and wnt5b, were all downregulated, confirming the Cdx4 ChIP-seq data (Figure S1D). These results suggest that a subset of the Cdx4-binding sites detected by ChIP-seq is indeed transcriptionally regulated by Cdx4.

sall4 Is a Direct Downstream Target of Cdx4
spalt-like 4 (sall4) was strongly downregulated in cdx4−/−/cdx1a−/− and bound by Cdx4 (Figures 1D and S1E). The spalt gene was first identified in Drosophila, where its mutation caused homeotic changes in the posterior head and...
anterior tail region (Frei et al., 1988; Jürgens, 1988). Its homologs are highly expressed in the developing tail tip region in both chickens and frogs (Barembaum and Bronner-Fraser, 2004; Neff et al., 2005). In addition, SALL4 transgenic mice develop AML, reminiscent of mice that were transplanted with mCdx4-transduced bone marrow (Bansal et al., 2006; Ma et al., 2006). Because the cdx homolog, caudal, is involved in homeotic switches in Drosophila and is highly enriched in the posterior region, the sall4 gene made an excellent candidate to act downstream of cdx4.

To validate the microarray data indicating that a loss of Cdx function leads to decreased sall4 transcription, we conducted whole-mount in situ hybridization (WISH) was conducted in 10-somite stage cdx4mo/cdx1amo embryos. sall4 expression was downregulated in these embryos, consistent with the Cdx factors directly regulating sall4 transcription (Figure 1E). We also found that cdx4 and sall4 are spatially and temporally coexpressed (Figure S1F). Taken together, these results strongly suggest that Cdx4 regulates sall4 transcription during zebrafish development.

Sall4 Binds to the cdx4 Locus

To examine Sall4 downstream target genes, we conducted ChIP-seq on bud-stage zebrafish embryos injected with 100 pg of sall4-myc miRNA, which does not cause morphological defects (Figure S1G). Sall4 ChIP-seq resulted in 9,747 binding events that corresponded to 1,832 proximal genes. De novo motif analysis gave the ATTTGCAT motif, which is known as the Oct4 motif (Figure 1F; Loh et al., 2006). In addition, Sall4-myc bound to the oct4/pou5f1 locus, consistent with reports showing that Sall4 is a part of the core ESC complex with Oct4 (Figure 1G; Rao et al., 2010; van den Berg et al., 2010; Wu et al., 2006). Sall4 binds to its own locus, suggesting autoregulation, and to the cdx4 locus, indicating cross-regulatory interactions between Sall4 and Cdx4 (Figures 1H and S1H).

Cdx4 and Sall4 Coregulate Genes Responsible for the Mesoderm-to-Blood Transition

The Sall4 ChIP-seq analysis showed a substantial overlap in binding sites with Cdx4 across the genome (Figure 2A). When we compared the Cdx4-Sall4 cobound sites with the bud-stage zebrafish histone ChIP-seq data, we found that 44% of these sites overlapped with H3K27ac-enriched domains, whereas only 1.4% overlapped with H3K4me3-enriched domains. This indicates that Cdx4 and Sall4 cobinding corresponds to active enhancers (Figure 2B; Rada-Iglesias et al., 2011). A Gene Ontology (GO) analysis of the Cdx4-Sall4 cobound gene list revealed that the cobound group is enriched with genes associated with pattern specification and embryonic morphogenesis (Figure 2C).

To determine whether Cdx4 and Sall4 cobinding leads to a cooperative transcriptional output, we obtained gene-expression data from cdx4mo, sall4mo, and cdx4mo/sall4mo
embryos. To examine whether knockdown of both genes leads to a greater transcriptional change compared with single knockdowns, we compared the expression level in the double morphants with that in the single morphants (see Experimental Procedures for details). In both the 3- and 10-somite stage embryos, the double morphants exhibited enhanced down- or upregulation of mRNA expression compared with the single morphants (Figures 3D and S2A). At the 10-somite stage, the top ten most downregulated genes from double morphants included hox genes, indicating that in addition to being cobound to the hox gene loci, Cdx4 and Sall4 cooperate to regulate the hox gene transcriptional program (Figure S2B). In addition, these double morphants exhibit downregulation of non-hox gene members such as gfi1.1 and morc3b (Figure S2B).
We noted that a number of genes affected in the double morphants were those that are typically associated with primitive hematopoiesis (e.g., gata1, gfi1.1, and znfl2) (Detrich et al., 1995; Galloway et al., 2008; Wei et al., 2008). To globally address whether double morphants show defects in hematopoiesis, we used GSEA to compare the genes affected in cdx4mo/sall4mo with the gene signatures of hematopoietic cells subjected to fluorescence-activated cell sorting (FACS) or morphants deficient in blood. The gene signature of the fli1a:GFP+ cells (the sorted population from the 10-somite stage Tg(fli1a:GFP) that express GFP in both blood and endothelial cells; Lawson and Weinstein, 2002) was highly enriched in the cdx4mo/sall4mo gene-expression set (NES = 2.50, FDR = 0.0; Figure 3D). Genes that were downregulated in the sclmo (which lacks a hemangioblast population; Dooley et al., 2005; Patterson et al., 2005) were also strongly enriched in the cdx4mo/sall4mo (NES = 1.87, FDR = 0.0; Figure 3E). To determine whether the cdx4mo/sall4mo enrichment was specific to blood, was conducted additional GSEAs using endothelial (Wong et al., 2009), renal (O’Brien et al., 2011; Wingert and Davidson, 2011; Wingert et al., 2007), muscle (de la Serna et al., 2005), and neuron (Lein et al., 2007) gene sets (Figures S2C–S2F). Among these, only the endothelial gene set showed significant enrichment (NES = 2.47, FDR = 0.0; Figure S2C). This result is not surprising, because blood and endothelial populations share common TFs, especially during primitive hematopoiesis (Davidson and Zon, 2004). In contrast, the enrichment scores for other gene sets were not significant, indicating that cdx4mo/sall4mo embryos show more specific defects in the hemangioblast population (Figures S2D–S2F). Taken together, the ChIP-seq and microarray data indicate that Cdx4 and Sall4 coregulate a set of genes involved in both embryonic pattern formation and primitive hematopoiesis.

**Figure 3. sall4 Cooperates with cdx4 in Zebrafish Hematopoiesis**

(A) gata1 WISH of 18-somite-stage wild-type or cdx4−/− embryos that were uninjected or sall4mo injected. sall4 knockdown in the cdx4−/− background leads to a decrease in gata1+ cells (note the high-powered view).

(B) o-Dianisidine staining of 3 dpf wild-type or cdx4−/− embryos that were uninjected or sall4mo injected. cdx4−/−/sall4mo embryos have fewer hemoglobinized RBCs. Arrows point to the stained RBCs in the embryos. Scale bar, 200 μm.

See also Figure S3.

We noted that a number of genes affected in the double morphants were those that are typically associated with primitive hematopoiesis (e.g., gata1, gfi1.1, and znfl2) (Detrich et al., 1995; Galloway et al., 2008; Wei et al., 2008). To globally address whether double morphants show defects in hematopoiesis, we used GSEA to compare the genes affected in cdx4mo/sall4mo with the gene signatures of hematopoietic cells subjected to fluorescence-activated cell sorting (FACS) or morphants deficient in blood. The gene signature of the fli1a:GFP+ cells (the sorted population from the 10-somite stage Tg(fli1a:GFP) that express GFP in both blood and endothelial cells; Lawson and Weinstein, 2002) was highly enriched in the cdx4mo/sall4mo gene-expression set (NES = 2.50, FDR = 0.0; Figure 3D). Genes that were downregulated in the sclmo (which lacks a hemangioblast population; Dooley et al., 2005; Patterson et al., 2005) were also strongly enriched in the cdx4mo/sall4mo (NES = 1.87, FDR = 0.0; Figure 3E). To determine whether the cdx4mo/sall4mo enrichment was specific to blood, was conducted additional GSEAs using endothelial (Wong et al., 2009), renal (O’Brien et al., 2011; Wingert and Davidson, 2011; Wingert et al., 2007), muscle (de la Serna et al., 2005), and neuron (Lein et al., 2007) gene sets (Figures S2C–S2F). Among these, only the endothelial gene set showed significant enrichment (NES = 2.47, FDR = 0.0; Figure S2C). This result is not surprising, because blood and endothelial populations share common TFs, especially during primitive hematopoiesis (Davidson and Zon, 2004). In contrast, the enrichment scores for other gene sets were not significant, indicating that cdx4mo/sall4mo embryos show more specific defects in the hemangioblast population (Figures S2D–S2F). Taken together, the ChIP-seq and microarray data indicate that Cdx4 and Sall4 coregulate a set of genes involved in both embryonic pattern formation and primitive hematopoiesis.

**sall4 and cdx4 Functionally Interact during the Formation of RBCs**

To demonstrate that sall4 functionally interacts with cdx4 during embryonic hematopoiesis, sall4 was knocked down in cdx4+/+ embryos. cdx4+/−/sall4−/− embryos were crossed and sall4mo was injected into one-cell-stage embryos (Davidson et al., 2003; Harvey and Logan, 2006). Knockdown of sall4 in either a cdx4+/+ or cdx4−/− background resulted in a slight shortening of the A-P axis by 28 hpf, but had no effect on erythropoiesis. Because cdx4−/− have fewer gata1+ cells in the ICM region and fewer RBCs later, we examined these phenotypes in sall4 knockdowns in the background of cdx4+/+, cdx4−/−/sall4−/−, and cdx4−/− embryos. At the 18-somite stage, the cdx4/sall4mo embryos had fewer gata1+ cells compared with the uninjected cdx4−/− embryos (figure 3A). The cdx4+/−/sall4mo or cdx4−/−/sall4mo embryos did not show a decrease in gata1. Although the hematopoietic defect in the cdx4−/−/sall4mo embryos was not as severe as that shown in the cdx4mo/cdx1amo embryos that completely lack gata1 (Davidson and Zon, 2006), these studies indicate that Sall4 and Cdx4 functionally interact during erythroid progenitor formation (figure 3A).

The injected embryos were grown to 3 days postfertilization (dpf), and o-dianisidine staining was conducted to
examine hemoglobinized RBCs (Ransom et al., 1996). As previously reported (Davidson et al., 2003), $cdx4^{-/-}$ embryos showed a decreased number of hemoglobinized RBCs compared with wild-type or $cdx4^{+/+}$ embryos. When $sall4^{mo}$ was injected, these $cdx4^{-/-}$ embryos showed even fewer hemoglobinized RBCs, consistent with $sall4$ cooperating with $cdx4$ during erythropoiesis (Figure 3B, arrow). The posterior mesoderm was also critically affected in $cdx4^{-/-}/sall4^{mo}$ embryos, as evidenced by the size of the tail (Figure S3).

**Loss of scl and lmo2 Is Responsible for the Lack of RBCs in $cdx4^{-/-}/sall4^{mo}$ Embryos**

The loss of RBCs seen in the $cdx4^{-/-}/sall4^{mo}$ embryos could be caused by an earlier defect in LPM differentiation. The TF Scl and its cofactor Lmo2 are key determinants of hematopoietic precursors that are expressed as bilateral stripes in the LPM at the 3-somite stage (Liao et al., 1998; Thompson et al., 1998). Loss of $scl$ or $lmo2$ in the zebrafish leads to a lack of RBCs (Dooley et al., 2005; Patterson et al., 2005, 2007). At the 10-somite stage, $cdx4^{-/-}$ embryos displayed only a few posterior $scl^+$ and $lmo2^+$ cells, whereas $cdx4^{-/-}/sall4^{mo}$ embryos showed no detectable $scl$ or $lmo2$ expression (Figure 4A). This suggests that early hematopoietic progenitors do not form properly in $cdx4^{-/-}/sall4^{mo}$ embryos. In contrast, $cdx4^{mo}/sall4^{mo}$ embryos have $fli1a^+$ and $draculin^+$ cells in the LPM, indicating that the general lateral mesoderm is present in these embryos. $myoD$ (somite), $pax2a$ (kidney), and $tbx16$ (presomitic mesoderm) expression is minimally affected, showing that $cdx4$ and

---

**Figure 4. scl and lmo2 Are Responsible for the Loss of RBCs Seen in $cdx4^{-/-}/sall4^{mo}$ Embryos.**

(A and B) $scl$ (A) and $lmo2$ (B) WISH at the 10-somite stage wild-type or $cdx4^{-/-}$ embryos that are either uninjected or $sall4^{mo}$ injected. Embryos were flat-mounted, with the anterior end pointing to the left and the posterior end pointing to the right. (C) $fli1a$, $draculin$, $pax2a$, $myoD$, and $tbx16$ WISH in 10-somite stage wild-type, $cdx4^{mo}$, $sall4^{mo}$, $cdx4^{mo}/sall4^{mo}$ embryos. All scale bars, 200 μm. (D and E) Gene track of the $scl$ (D) and $lmo2$ (E) loci, showing Cdx4 and Sall4 binding. See also Figure S4.
sall4 specifically affect early hematopoietic progenitors (Figure 4B; Kimmel et al., 1989; Majumdar et al., 2000; Weinberg et al., 1996). In accordance with a putative regulatory role in scl and lmo2 transcription, both the Cdx4 and Sall4 ChIP-seq profiles showed both factors binding near the scl and lmo2 loci, which was confirmed by ChIP-PCR analysis (Figures 5C, 5D, and S4).

To determine whether this early block in scl and lmo2 expression is responsible for the RBC loss in cdx4mo/sall4mo embryos, we overexpressed scl and lmo2 in transgenic zebrafish carrying the globin lcr:GFP reporter (Ganis et al., 2012; Figure 5A). The scl and lmo2 genes, together with a control plasmid, were driven under the control of the ubi promoter (Mosimann et al., 2011) and injected into cdx4mo/sall4mo/Tg(lcr:GFP) embryos. The ubi:mCherry plasmid was used as an additional control to assess injection mosaicism. As expected, injection of these plasmids led to ubiquitous expression of transgenes (Figure S5). In line with our previous results, cdx4mo/sall4mo/Tg(lcr:GFP) embryos lacked lcr:GFP expression, which could not be rescued with control plasmid injection or ubi:scl injection (Figure 5B, top and middle rows). In contrast, ubi:scl; ubi:lmo2 double injection in cdx4mo/sall4mo embryos led to lcr:GFP expression in the ICM (n = 9/21; Figure 5B, bottom row), as well as gata1 expression (n = 23/49; Figure 5C), indicating that scl and lmo2 can rescue the erythropoietic defects caused by Cdx4 and Sall4 deficiency. Taken together, these results reveal that cdx4 and sall4 genetically interact in a regulatory circuit involving cross- and autoregulation to control the expression of TFs involved in mesodermal and hematopoietic lineage specification.

**DISCUSSION**

Our findings illustrate a transcriptional circuit involving Cdx4 and Sall4 that is required for the developmental transition from mesoderm to blood formation in zebrafish. By integrating ChIP-seq and expression profiling data from zebrafish embryos, we show that Sall4 shares multiple common target genes with Cdx4, including the TFs themselves, hox genes, scl, and lmo2. Consistent with this finding, the cdx4mo/sall4mo embryos exhibit enhanced defects in embryonic hematopoiesis compared with cdx4mo embryos, which can be overcome by scl and lmo2 overexpression.

sall4 deficiency alone does not affect embryonic erythropoiesis, but it cooperates with the cdx4 mutant phenotype as shown by microarray, loss of RBCs, and severe tail truncation (Figures 2, 3, S2A, and S3). Similarly, the cdx1amo has a mild phenotype, but cdx4mo/cdx1amo embryos show a more severe axial truncation and a complete loss of erythropoiesis compared with cdx4amo (Davidson and Zon, 2006). This finding is attributed to redundancy between...
Cdx1a and Cdx4, given that Cdx paralogs compensate for each other (Savory et al., 2009b). Sall4 shares activities with the Cdx factors, and Sall4 binding likely enhances the transcriptional activity of Cdx-containing complexes on cobound target genes (Figure 2). Sall4 activates target genes synergistically with other TFs, such as Tcf4 and beta-catenin in gastrula-stage mouse embryos, and with genes synergistically with other TFs, such as Tcf4 and

![Diagram of Mesoderm and Blood Specification](image)

**Figure 6. Model**

Cdx4 and Sall4 autoregulate and cross-regulate each other (Sall4-to-Cdx4 regulation is dotted because it is unclear whether Sall4 regulates Cdx4 transcription). They regulate the LPM-to-blood transition in two ways: (1) they regulate hox genes for mesoderm specification, and (2) they regulate scl and lmo2 for blood specification.

Scl and Lmo2 are key hematopoietic TFs whose deficiency cause severe anemia (Dooley et al., 2005; Patterson et al., 2005; Shivdasani et al., 1995; Warren et al., 1994). As scl overexpression could not rescue the gata1 loss seen in cdx4−/−, it has been hypothesized that Cdx4 acts to make the posterior mesoderm competent to respond to genes that specify the hematopoietic fate (Davidson et al., 2003). An alternative model would include the possibility that other Cdx target genes participate in hematopoietic induction. One intriguing possibility is that Scl is insufficient to rescue the cdx4−/− because it lacks the Scl cofactor, Lmo2. Our ChIP-seq and ChIP-PCR data demonstrate that scl and lmo2 genes are bound by Cdx4 and Sall4, and expression of scl and lmo2 is reduced in cdx4−/−/sall4−/− (Figures 4A, 4C, 4D, and S4). Our rescue data show that scl in combination with lmo2 can drive erythropoietic genes in the cdx4−/−/sall4−/−, indicating that lmo2 is another key factor that is responsible for the loss of RBCs (Figure 5). The lack of gata1 rescue to the wild-type level shows that there could be other additional Cdx4 and Sall4 targets. A recent ChIP study showed that Lmo2 transcription in hematopoietic cells is regulated by HoxA9 (Huang et al., 2012). Our results suggest that Cdx4 and Sall4 regulate hoxa9a transcription in zebrafish (Figure S2B). This suggests that lmo2 could be regulated in a tightly controlled manner during the differentiation of LPM to blood, directly by Cdx4 and Sall4, and secondarily by another Cdx4 and Sall4 target gene, Hoxa9a. We speculate that this type of transcriptional loop ensures a robust activation of the blood program.

Our ChIP-seq data indicate that Cdx4 and Sall4 engage in auto- and cross-regulatory loops. Such regulatory circuits are frequently seen during embryonic development. For example, during segmentation of rhombomere 4 in the mouse hindbrain, retinoic acid signaling first turns on HoxB1, followed by HoxB1, HoxB2, and HoxA2, maintaining their expression through tight auto- and cross-regulatory loops (Agarwal et al., 2011; Gavalas et al., 2003; Tümpel et al., 2007). The Cdx4-Sall4 circuit during zebrafish development may function in an analogous way with upstream regulators such as Wnt, BMP, and FGF, activating expression of Cdx4 and Sall4, followed by the stabilization of their expression by auto- and cross-regulatory loops. We hypothesize that a stable Cdx4-Sall4 circuit is required to make the mesoderm competent for the specification of posterior tissue lineages such as blood by turning on the hematopoietic master TFs Scl and Lmo2 (Figure 6).

Our current model also gives insights into leukemogenesis. Previously, Hox gene misregulation was the main focus of attempts to understand the mechanism behind Cdx genes causing leukemia (Bansal et al., 2006; Scholl et al., 2007). Our current model suggests another possibility; i.e., that Cdx genes directly regulate Scl and Lmo2. As mutation of both of these genes has been linked to various leukemias, such as T cell acute lymphoblastic leukemia (Bash et al., 1995; Boehm et al., 1991; Royer-Pokora et al.,
In conclusion, our data establish a Cdx4-Sall4 circuit that may be aberrantly regulated in leukemias that are caused by Cdx or Sall4 mutation. In the posterior mesoderm to facilitate blood cell formation and axial elongation. We propose that during the early gastrula stage, Cdx4 and Sall4 are activated by upstream regulators such as Wnt, and their auto- and cross-regulation act to stabilize the mesoderm state. At the end of gastrulation, Cdx4 and Sall4 bind to scl and lmo2 to induce the hematopoietic program in the mesoderm. As upstream signaling dissipates, the Cdx4-Sall4 regulatory loops are disrupted, with a concomitant reduction in cdx4 and sall4 expression, and further blood differentiation is driven by factors such as Scl and Lmo2 (Figure 6). This model is likely applicable to the formation of other tissues and provides a molecular framework to understand how Cdx and Sall factors induce leukemia.

EXPERIMENTAL PROCEDURES

Zebrafish
Zebrafish were maintained according to Institutional Animal Care and Use Committee guidelines. The Animal Care and Use Committee of Boston Children’s Hospital approved all of the animal protocols. Wild-type (Tu) and cdx4+f/+ incrossed embryos were collected at the one-cell stage and injected with mRNA or morpholino for further experiments.

ChIP
ChIP was performed as previously described (Lee et al., 2006; Lindeman et al., 2009). For detailed information, see the Supplemental Experimental Procedures. myc (abcam ab9132), H3K4me3 (Millipore 07-473), and H3K27ac (abcam ab4729) antibodies were used at 5 μg per ChIP sample.

ChIP-Seq Analysis
Sequence reads were aligned to the genome (danRer6) using Bowtie version 0.12.7 with options “-q–best–strata -m 1 -p 4–chunkmbs 1024’’ (Langmead et al., 2009). Only uniquely mapping reads were analyzed further. Binding events were detected using GPS as previously described (Guo et al., 2010). For detailed information on ChIP-seq analysis, see the Supplemental Experimental Procedures. The ChIP-seq data have been deposited in the GEO database under accession number GSE48254.

ChIP-PCR Analysis
Primers were designed for cdx4, sall4, scl, and lmo2 loci (listed in Table S1). Quantities are expressed as fold change compared with input controls.

Morpholinos
One-cell-stage embryos were injected with sall4⁻/⁻ (2 ng; Harvey and Logan, 2006), cdx4⁻/⁻ (1.5 ng), or cdx1a⁻/⁻ (0.5 ng).

mRNA
cdx4 and sall4 cDNA was described previously (Davidson et al., 2003; Harvey and Logan, 2006). mRNA was generated by the mMessage mMachine kit and injected into one-cell-stage embryos.

Microarray Analysis
Embryos were injected with either cdx4 or sall4, or both morpholinos together. Injected embryos, along with uninjected control embryos, were harvested at the 3- or 10-somite stage and RNA was prepared for microarray. For the arrays in these experiments, we used the prerelease Affymetrix Zebrafish 2.0 array, which was developed in the Zon laboratory in cooperation with Affymetrix. The cdx4⁻/⁻/cdx1a⁻/⁻ data set was generated on the Affymetrix Zebrafish 1.0 array. Detailed information on microarray analysis can be found in the Supplemental Experimental Procedures.

GSEA Analysis
Enrichment of Cdx4-Bound Genes in the cdx4/cdx1a Double-Morphant Gene-Expression Set
The fli1aGFP or sclmo gene lists were used as input to GSEA for queries into the entire cdx4⁻/⁻/sall4⁻/⁻ microarray data set. In this case, because the cdx4⁻/⁻/sall4⁻/⁻ arrays were performed using the Zebrafish Affymetrix 2.0 arrays, we used the Zv9 gene identifier as the input for GSEA. For endothelial, kidney, muscle, and neuron gene sets, the annotated gene list was probed. Enrichment of the double morphant compared with control embryos was calculated using the NES and FDR as described above. Detailed information regarding the gene list sources can be found in the Supplemental Experimental Procedures.

WISH and o-Dianisidine Staining
WISH and o-dianisidine staining were performed as previously described (Ransom et al., 1996; Thisse et al., 1993). The gata1, scl, lmo2, fli1a, myoD, anfrulin, pax2a, ttx16, and mon3b probes were described in previous publications (Detrich et al., 1995; Kimmel et al., 1989; Liao et al., 1998; Majumdar et al., 2000; Thompson et al., 1998; Weinberg et al., 1996).

DNA Injection
For DNA injection, 1 nl of a 16 ng/μl DNA mixture was injected into one-cell-stage embryos. For detailed information on the plasmids used, see the Supplemental Experimental Procedures and Table S2.

SUPPLEMENTAL INFORMATION
Supplemental Information includes Supplemental Experimental Procedures, five figures, and two tables and can be found with this article online at http://dx.doi.org/10.1016/j.stemcr.2013.10.001.

ACKNOWLEDGMENTS
We thank members of the Zon laboratory for helpful discussions, J. Ganis for providing Tg(lcrGFP), and M. Logan for providing the sall4 cDNA construct. We also thank the Whitehead Genome Technology Core for data production and analysis support. Microarray
studies were performed by the Molecular Genetics Core Facility at Children's Hospital Boston, supported by NIH grants P50-NS40828 and P30-HD18655. This work was supported by grants from the NHLBI (5R01HL048801-21 to L.I.Z. and 5P01HL32262-31), NIDDK (5P30 DK49126-19, DK53298-15, and R24 DK092760-02), and HHMI (to L.I.Z.). L.I.Z. is a founder and stockholder of Fate, Inc., and Scholar Rock, and a scientific advisor for Stemgent.

Received: February 26, 2013  
Revised: October 1, 2013  
Accepted: October 2, 2013  
Published: November 7, 2013

REFERENCES


Davidson, A.J., and Zon, L.I. (2006). The caudal-related homeobox genes cdx1a and cdx4 act redundantly to regulate hox gene expres-
Cooperative and antagonistic interactions between Sall4 and Tbx5 pattern the mouse limb and heart. Nat. Genet. 38, 175–183.


