Genome-wide analysis shows that *Ldb1* controls essential hematopoietic genes/pathways in mouse early development and reveals novel players in hematopoiesis

Running title: The role of *Ldb1* in hemangioblast development

Athina Mylona[1,2*], Charlotte Andrieu-Soler[1,3,*], Supat Thongjuea[4*], Andrea Martella[1], Eric Soler[1,5,6], Ruud Jorna[1], Jun Hou[1,7], Christel Kockx[8], Wilfred van Ijcken[8], Boris Lenhard[4,9,10] and Frank Grosveld[1,6,7,11]

[1]Department of Cell Biology, Erasmus MC, Dr. Molenwaterplein 50, 3065GE Rotterdam, the Netherlands; [2]Currently Department of Hematology, Erasmus MC, Dr. Molenwaterplein 50, 3065GE Rotterdam, the Netherlands; [3]INSERM U872, Université René Descartes Sorbonne Paris Cité, Team 17, Centre de Recherche des Cordeliers, Paris, France; [4]Computational Biology Unit - Bergen Center for Computational Science and Sars Centre for Marine Molecular Biology, University of Bergen, Thormøhlensgate 55, N-5008 Bergen, Norway; [5]Laboratory of Hematopoiesis and Leukemic Stem Cells (LSHL), CEA / INSERM U967, Fontenay-aux-Roses, France; [6]Cancer Genomics Centre (www.cancergenomics.nl); [7]Netherlands Consortium for Systems Biology (www.genomics.nl/GenomicsCentres/Enabling%20technologies/NCSB.aspx); [8]Centre for Biomics Erasmus MC, Rotterdam, the Netherlands; [9]Department of Molecular Sciences, Institute of Clinical Sciences, Faculty of Medicine, Imperial College London; [10]MRC Clinical Sciences Centre, Hammersmith Hospital Campus, Du Cane Road, London, W12 0NN, United Kingdom; [11]Centre for Biomedical Genetics (www.biomedicalgenetics.nl).

[*]These authors contributed equally to this work.

Corresponding author: Prof. Frank Grosveld f.grosveld@erasmusmc.nl

Tel.: +31-10-7043593/Fax: +31-10-7044743

Corresponding author for computational biology: Dr. Boris Lenhard b.lenhard@imperial.ac.uk

Key points

- Lack of yolk sac hematopoiesis in the Ldb1$^{-/-}$ mouse results from a decreased number of hemangioblasts and a differentiation block.
- Identification of genes and pathways regulated by Ldb1 in the hemangioblast reveals potential targets for early developmental manipulation.

Abstract

The first site exhibiting hematopoietic activity in mammalian development is the yolk sac blood island, which originates from the hemangioblast. Here we performed differentiation assays, as well as genome-wide molecular and functional studies in BL-CFCs to gain insight into the function of the essential Ldb1 factor in early primitive hematopoietic development. We show that the previously reported lack of yolk sac hematopoiesis and vascular development in *Ldb1*$^{-/-}$ mouse result from a decreased number of hemangioblasts and a block in their ability to differentiate into erythroid and endothelial progenitor cells. Transcriptome analysis and correlation with the genome-wide binding pattern of Ldb1 in hemangioblasts revealed a number of direct target genes and pathways misregulated in the absence of *Ldb1*. The regulation of essential developmental factors by Ldb1 defines it as an upstream transcriptional regulator of hematopoietic/endothelial development. We show the complex interplay that exists between transcription factors and signalling pathways during the very early stages of hematopoietic/endothelial development and the specific signalling occurring in hemangioblasts in contrast to more advanced hematopoietic developmental stages. Finally, by revealing novel genes and pathways, not previously associated with early development, our study provides novel candidate targets to manipulate the differentiation of hematopoietic and/or endothelial cells.

Key words

Hematopoietic development; Transcriptional regulation; Genome-wide DNA binding analysis

Introduction

Primitive erythroid and endothelial progenitors emerge almost simultaneously in mesodermal blood islands from a common hemangioblast progenitor[1] that is characterized by the expression of Flk1; the receptor for VEGF (Vascular Endothelial Growth Factor) (Fig. S1). The other mesodermal-derived lineage expressing Flk1 is the cardiovascular progenitor cell, which is distinguishable from the hemangioblast by PDGF (platelet-derived growth factor) receptor expression. The hemangioblast has been studied in mouse, chick and zebrafish embryos[2] and its existence is supported by the observation that hematopoietic and endothelial progenitors share the expression of common genes.[3-7]

Deletion of *Flk1* leads to embryonic lethality between 8.5 and 9.5 dpc possibly due to the absence of yolk sac blood islands.[8, 9] The hemangioblast stage can be more specifically studied *in vitro* through the differentiation of ES cells into embryoid bodies (EBs), widely considered as a relevant model of embryonic development. Within EBs blast colony forming cells (BL-CFCs) can generate blast colonies, able to differentiate into hematopoietic and endothelial progenitors. They are therefore considered as the *in vitro* equivalent of the hemangioblast.[10]

Recently Li *et al.* demonstrated that *Ldb1* is also required during primitive hematopoiesis in the mouse, for the development of megakaryocytes and the maintenance of long-term hematopoietic stem cells.[11, 12] *Ldb1* deletion is embryonic lethal after 9.5 dpc, with severe developmental defects including small size, anterior

truncation, absence of heart/foregut structures, impairment of yolk sac hematopoiesis and vascular development.[13] However, no functional studies or molecular mechanisms are known describing *Ldb1* function in early hemangioblasts. Ldb1 molecular function has been well-studied in definitive erythropoiesis, where it plays an essential role by forming large transcriptional complexes with Gata1, Lmo2, Scl, E12/E47, HEB, Lyl1, E2-2, Cbfa2t3 (Eto2), Mtgr1, Cdk9, Lmo4,[14] binding E-box/GATA motifs genome-wide.[15]

In this study, we generated *Ldb1*$^{-/-}$ ES cells and performed differentiation assays and genome-wide molecular and functional studies in BL-CFCs to unravel the function of Ldb1 in early primitive hematopoietic development. We showed that *Ldb1* deletion actually results in two separate defects; a decreased number of hemangioblasts and their inability to differentiate further down the hematopoietic and endothelial lineages. To gain insight into the mechanism underlying these defects, we determined the direct and indirect targets of Ldb1 in these progenitor cells (in contrast to more advanced hematopoietic developmental stages). The genome-wide binding pattern of Ldb1 revealed that it is directly bound to a number of gene loci involved in mammalian hematopoietic development. Their expression is misregulated in the absence of *Ldb1* placing it as one of the most upstream factors controlling hematopoietic development. Furthermore our study reveals novel genes and pathways that are potentially important for early hematopoietic/endothelial development and shows the complex interplay between transcription factors and signalling pathways in very early primitive development stage, providing novel candidate targets that can be used to manipulate the differentiation of hematopoietic and/or endothelial cells.

Methods

Ldb1$^{-/-}$ ES cell line generation/genotyping

Both *Ldb1* alleles were inactivated using the same construct (Suppl. Methods/Fig. S2A). After targeting the first allele a single ES cell clone was transiently transfected with a Cre recombinase expressing vector to delete exons 1/2 and the PMC-neo cassette, followed by targeting of the second allele. Southern blotting confirmed homologous recombination and deletion. Ldb1 protein levels were determined in nuclear extracts from *Ldb1*$^{+/+}$ and *Ldb1*$^{-/-}$ ES cells with anti-Ldb1 N-18 (Santa Cruz®, Cat. SC-11198).

ES cell differentiation

Ldb1$^{+/+}$ and *Ldb1*$^{-/-}$ ES cells were grown in suspension at 10.000 cells/ml on non-adherent dishes in IMDM medium (without Leukemia Inhibitory Factor (LIF)) with 15% FCS, 1% P/S, 1% L-glutamine (Gibco, Cat.25030-08), 0.05µg/ml transferin (Roche, Cat.652-202), 0.05µg/ml ascorbic acid (Sigma, Cat.A-4544), 3µl/ml monothioglycerol (Sigma, Cat.M-6145) for day 4 EBs and 1.8µl/ml for day 6/day 8 EBs. 5% protein free hybridoma medium II (Gibco, Cat.12040-077) was added for day 6/day 8 EBs.

Blast **C**olony **F**orming **C**ell (BL-CFC) assay

Day 4 EBs were disrupted with trypsin-EDTA and cells were transferred in methycellulose-based media with 10% FCS, 1% P/S, 1% L-glutamine, 0.25µg/ml transferin, 0.25µg/ml ascorbic acid, 2µl/ml monothioglycerol, 0.01µg/ml mIL6 (R&D Systems, Cat.406-ML), 0.005µg/ml hVEGF (R&D Systems, Cat.293-VE) for 3 days. Colonies were scored according to morphology and number under an inverted microscope.

Colony Forming Cell (CFC) assay

Ldb1$^{+/+}$ and *Ldb1*$^{-/-}$ day 6 EBs were disrupted with 2.5% collagenase. Cells were transferred in methycellulose-based media with 10% FCS, 1% L-glutamine, 0.25µg/ml transferin, 0.25µg/ml ascorbic acid, 2µl/ml monothioglycerol, 5% protein free hybridoma medium II, 0.01µg/ml mIL6, 0.001µg/ml IL3 (R&D Systems, Cat.403-ML), 0.005µg/ml hIL11 (R&D Systems, Cat.418-ML), 0.003µg/ml GM-CSF (R&D Systems, Cat.415-ML), 4U/ml EPO (R&D Systems, Cat.959-ME), 0.005µg/ml TPO (R&D Systems, Cat.488-TO), 0.1 µg/ml SCF (R&D Systems, Cat.455-MC). Red primitive erythroid colonies and white macrophage colonies (composed of round cells growing in clumps) were identified microscopically after 6 days according to morphology and colour.

Flow Cytometry

Day 4 EB single cell suspensions were labelled with Flk1-PE clone Avas 12α1 (BD Pharmigen™, Cat.555308) in 1% BSA/PBS. Dead cells were excluded by Hoechst 33258 (Molecular Probes). Flk1^{+} cells were sorted from day 4 EBs on FACSAria using the Diva 5.1 software (BD). Day 6/8 EB single cell suspensions were labelled in 1% BSA/PBS with CD41-PE clone MWReg30 (Santa Cruz®, Cat.sc-19963) or CD31-FITC clone MEC 13.3 (BD Pharmigen™, Cat.553372). Dead cells were excluded with 7-AAD (Invitrogen™). FACS was performed on FACScan (Becton Dickinson) and analysed with Cell Quest.

Flk1^{+} Blast Colony Forming Cell (BL-CFC) isolation

Day 4 EBs were disrupted with dissociation buffer (StemPro Accutase-Invitrogen™). Dead cells were removed by Lymphoprep separation. Magnetic-assisted cell sorting (MACS) was carried out with autoMACS® Pro Separator (Miltenyi Biotec) using

Anti-PE MicroBeads (Miltenyi Biotec) and the Flk1-PE clone Avas 12α1 (BD Pharmigen™) antibody. Purity was assessed by flow cytometry.

Chromatin Immunoprecipitation (ChIP) sequencing

ChIP-sequencing of endogenous Ldb1 (using anti-Ldb1 antibody N-18, Santa Cruz®) on $Flk1^+$ BL-CFCs isolated from day 4 EBs was performed as previously described.[15, 16]

Gene expression profiling

RNA was isolated from *Ldb1*$^{+/+}$ and *Ldb1*$^{-/-}$ $Flk1^+$ cells with the QIAGEN RNeasy Mini Kit and integrity was checked on the Agilent 2100 Bioanalyzer.

Microarray

RNA was converted to biotin-labelled cRNA, hybridised on the Mouse Genome 430 2.0 Array and analysed with the Affymetrix GeneChip® Scanner 3000 according to the manufacturer protocol.

RNA sequencing

RNA sequencing was performed on Illumina HiSeq 2000 platform according to the manufacturer instructions.

* For ChIP/RNA sequencing and microarray bioinformatic analysis see supplemental methods.

ChIP/RNA sequencing and microarray data sets described in this study have been deposited in the NCBI Gene Expression Omnibus (Accession Number: GSE43044).

Results

No primitive erythropoiesis in the *Ldb1*$^{-/-}$ mouse

Generation of the *Ldb1*$^{-/-}$ mouse is presented in Fig. S2A-D. The first striking observation in the embryos is the complete absence of any blood or vasculature from the extra-embryonic yolk sac, which also fails to completely surround the whole embryo, showing that erythropoiesis and vascular development are severely impaired in the absence of *Ldb1* (Fig.1A, in agreement with Li et al. 2011).[11] *Ldb1*$^{+/-}$ mice were viable, able to breed and did not exhibit any of the phenotypic characteristics of the *Ldb1*$^{-/-}$ mice (not shown, [11]).

Ldb1$^{-/-}$ EBs contain no erythroid and less endothelial progenitor cells

We additionally generated an *Ldb1*$^{-/-}$ ES cell line to carry out differentiation, molecular and functional studies in hemangioblast-equivalent cells. The second *Ldb1* locus allele in *Ldb1*$^{+/-}$ ES cells was deleted (Fig. S2E) and the absence of Ldb1 protein confirmed by western blotting and immunofluoresence (Fig. S2F-G). The morphology and growth of *Ldb1*$^{-/-}$ ES cells appeared similar to normal ES cells (not shown and Li et al., 2011).[11, 17] We then tested their differentiation potential into EBs. Since *Ldb1*$^{+/-}$ mice did not exhibit any phenotypic differences as compared to *Ldb1*$^{+/+}$ mice they were not used for further investigations. *Ldb1*$^{+/+}$ and *Ldb1*$^{-/-}$ EBs did not show any phenotypic differences at days 3 and 4 appearing similar in size and structure (Fig. 1B). However after 8 days we observed a striking absence of erythroid clusters within the *Ldb1*$^{-/-}$ EBs (Fig. 1B), correlating with the hematopoietic defect observed in the knock-out mouse. Of note, *Ldb1*$^{-/-}$ EBs survived at least until day 10 without apparent apoptosis, but were more adherent than the wild type ones. After 10 days the majority was found attached to the bottom of the culture dish (not shown). Colony Forming Cell (CFC) assays were performed to test the potential of *Ldb1*$^{-/-}$ EBs

to give rise to primitive erythroid and macrophage colonies. Both bright red primitive erythroid colonies and white macrophage colonies, present in *Ldb1*$^{+/+}$ day 6 EBs, were absent in the *Ldb1*$^{-/-}$ cultures. This defect was fully rescued by exogenous re-expression of Ldb1 (Fig. 1C,D), indicating that those effects are specifically due to the Ldb1 absence. In addition, the presence of primitive erythroid progenitors was examined by flow cytometry using CD41 (integrin alpha-IIb) as marker.[18] *Ldb1*$^{+/+}$ EBs contained 2.6% and 12% of CD41^{+} cells at day 6 and day 8 respectively, while *Ldb1*$^{-/-}$ EBs did not contain any CD41^{+} erythroid progenitor cells in accordance with the absence of primitive erythroid colonies in *Ldb1*$^{-/-}$ CFC cultures (Fig.1E). The presence of endothelial progenitors in day 6 and day 8 EBs was also examined using CD31 (PECAM-1) as a marker, since *Ldb1*$^{-/-}$ embryos suffer from a severe defect in vascular development.[19] Day 6 and 8 *Ldb1*$^{+/+}$ EBs contained 21% and 16% CD31^{+} cells respectively. These cells were also found to be present in *Ldb1*$^{-/-}$ EBs at the same time points but reduced to less than half when compared to wild type (Fig. 1E).

***Ldb1*$^{-/-}$ EBs do not give rise to blast colonies**

We then investigated whether BL-CFCs, the *in vitro* hemangioblast equivalent, were present in the *Ldb1*$^{-/-}$ EBs. Day 4 BL-CFCs were isolated using the VEGF receptor Flk1 as a marker (Fig. S1). Flk1^{+} BL-CFCs were present in both *Ldb1*$^{+/+}$ and *Ldb1*$^{-/-}$ EBs, however a striking difference was observed as *Ldb1*$^{-/-}$ EBs repeatedly contained half the number of Flk1^{+} cells (Fig. 1F). In three independent experiments, blast colony forming assays also showed that blast colonies did form in the wild type but not in the knockout cultures (Fig. 1G). In conclusion, BL-CFCs are present in *Ldb1*$^{-/-}$ EBs but fail to generate fully grown blast colonies. *Ldb1*$^{-/-}$ Flk1^{+} cells eventually die in culture (not shown).

Transcriptome analysis identifies the genes and associated signalling pathways active in BL-CFCs

To gain insight into the gene expression network and signalling pathways involved in the maintenance of the hemangioblast *in vitro* equivalent cells, transcriptome analysis of $Flk1^+$ BL-CFCs by RNA-sequencing showed 12,468 expressed genes. Pathway enrichment analysis of the expressed genes revealed the signalling pathways potentially active in these progenitor cells (Table 1, Column A).

Ldb1 plays a specific and major role in the BL-CFCs gene regulatory network

To identify the direct genomic targets of Ldb1 in $Flk1^+$ BL-CFCs, ChIP (chromatin immunoprecipitation)-sequencing in sorted $Flk1^+$ cells showed 4,252 significant Ldb1 binding peaks. Of all genes expressed in $Flk1^+$ cells over 1/4 or 3,264 genes contain Ldb1 peak(s) in their vicinity. The group of Ldb1-bound genes were found to participate in approximately half of the signalling pathways active in BL-CFCs (Table 1, Columns A versus B).

We compared the Ldb1 ChIP-sequencing dataset of primitive $Flk1^+$ BL-CFCs to the Li *et al*. dataset in more committed adult hematopoietic Lin^- bone marrow progenitors.[11] This revealed that the majority of the 4,252 Ldb1 binding events in $Flk1^+$ cells is specific to these early progenitor cells (79.6% or 3,384 binding events) (Fig. 2A). Of note, among those specific binding events, a few (respectively 35 and 168 events) have also been detected in previously published Ldb1 ChIP-sequencing datasets from fetal liver or definitive erythroleukemic cells.[15] We therefore excluded those 203 binding events from the 3,384 $Flk1^+$-specific events for further investigation. The resulting 3,181 $Flk1^+$-specific binding events are referred to as "Set1" and will be compared with "Set2", the common binding events between $Flk1^+$ cells and Lin^- bone marrow cells (20.4% or 868 Ldb1 binding events in $Flk1^+$ cells).

We then performed a functional study to identify genes and associated signalling pathways regulated by Ldb1 in BL-CFCs. Day 4 *Ldb1*$^{-/-}$ Flk1^{+} cells were used for microarray and RNA-sequencing. The corresponding gene expression profiles were compared to the one of wild-type Flk1^{+} cells. Of the 12,468 genes found to be expressed in Flk1^{+} cells, 2,675 genes were identified as differentially expressed between the *Ldb1*$^{+/+}$ and the *Ldb1*$^{-/-}$ Flk1^{+} cells (Table S1). Of those, 1,424 genes were down-regulated (53.23%) and 1,251 were up-regulated (46.77%). The differentially expressed genes were classified into three independent sets, Set1*, Set2* and Set3 respectively depending on whether they were included in the Ldb1-bound genes of Set1, Ldb1-bound genes of Set2 or were unbound in Flk1^{+} cells (Fig. 2B). Interestingly, the majority of Ldb1 direct targets (Set1* and Set2*) were observed down-regulated genes; 65% (483 genes) were down-regulated while 35% (259 genes) were up-regulated in Set1* and 69.3% (149 genes) were down-regulated with 30.7% (66 genes) up-regulated in Set2*. This trend is absent in indirect target genes (Set3), which present 46.1% (792 genes) and 53.9% (926 genes) of down- and up-regulated genes respectively. The fact that the majority of direct target genes is down-regulated as a result of *Ldb1* deletion shows that Ldb1 primarily acts as a positive transcriptional regulator in BL-CFCs. However the 1/3 of up-regulated genes in this group suggests an additional role for Ldb1 complexes as repressors of transcription. A search for enriched DNA binding motifs carried-out on sequences targeted by Ldb1 in Set1 revealed the known E-box(Tal1):Gata motif (motif 5) in only 3,6% of peaks (114 out of 3,181) (Fig. 3A, C). The motifs found with the highest frequency were those associated with Klf transcription factors (motif 1: 513 peaks out of 3,181 or 16%), SP1 (motif 2: 569 peaks or 18%) and PU.1 (motif 3: 333 peaks or 10%). In addition, 3% of Ldb1 peaks from Set1 are associated with a CTCF motif (motif 8: 98 out of

3,181 peaks) (Fig. 3A, C). Conversely, in Set2 peaks (common with adult bone marrow progenitor peaks), a prominent E-box:Gata motif was found (motif 1: 419 peaks out of 868 or 48%) as described,[11, 15] and no CTCF motif was detected (Fig. 3B, C). This suggests that Ldb1 forms different complexes in primitive $Flk1^+$ cells targeting different regulatory sites. Alternatively, Ldb1 complex activity may be regulated by different co-factors in these cells as suggested by the differences in recognition motifs, potentially explaining the different repressing and activating transcriptional functions of Ldb1 in BL-CFCs.

Ldb1 preferentially controls developmental genes in BL-CFCs and regulates a wide range of important signalling pathways

Among the misregulated genes in *Ldb1*$^{-/-}$ $Flk1^+$ BL-CFCs directly bound by Ldb1 (Set1*), we found genes previously shown to participate in the regulation of mouse embryonic development and hematopoiesis such as *Gata2*,[20, 21] *Gata1*,[22] *Runx1*,[23] *Fli1*,[24] *Lyl1*,[25] *Id2*, *Id4*,[26] *Klf4*,[27] *Gfi1b*,[28] *Cbfa2t3*/*Eto2*,[29] *Myb*,[30] *Dkk1*, *Sfrp1*,[31, 32] *Meis1*,[33] *Raldh2*,[34] *Fgfr2*.[35] The changes observed in some of those essential genes were confirmed by real-time PCR (Fig. S3). Of interest is the down-regulation of *Stat5a* and *Sox7, Sox17* and *Sox18* all having been reported as hematopoietic regulators.[36-38] Fig. 4 shows 4 examples of hematopoietic-essential Set1* genes. As BL-CFCs can differentiate into both the hematopoietic and endothelial lineages, we further analysed the hematopoietic and endothelial specific genes.[39] Fig. 5 shows that the majority of these is down-regulated in the absence of Ldb1, illustrating the role of Ldb1 as an activator of transcription and to a lesser extent as a repressor in hematopoiesis and endothelial development (like the larger group of Set1* containing those genes, see Fig. 2B).

To decipher the essential functions of Ldb1 in the maintenance/differentiation of BL-CFCs, we performed a pathway enrichment analysis on the total number of genes misregulated in Flk1^{+} cells in the absence of *Ldb1* (Table 1, Column C). Approximately 1/3 (37 of 127) of the pathways enriched for genes expressed in Flk1^{+} cells are misregulated in the *Ldb1*$^{-/-}$ Flk1^{+} cells. To distinguish between direct and indirect functions of Ldb1, we carried out in parallel pathway enrichment analyses on the Ldb1 direct target genes (Set1* Table 1, Column D) and the indirect Ldb1-regulated genes (Set3 Table 1, Column E). Approximately half of the misregulated pathways in the *Ldb1*$^{-/-}$ Flk1^{+} cells was directly targeted by Ldb1 (Table1 column C versus D). Among these, Acute Myeloid Leukemia, Basal cell carcinoma, Interleukin signalling pathway, Melanoma and T cell receptor signalling pathway were found. Four examples of perturbed signalling pathways are provided illustrating how the genes are affected by the absence of Ldb1 (Fig. S4-7). Signalling pathways such as WNT (Wingless), MAPK (Mitogen-Activated Protein Kinase), Hedgehog, ErbB, Focal adhesion, Integrin, Insulin, and Angiogenesis were commonly enriched for both direct and indirect genes of Set1* and Set3. Not unexpectedly we also found myeloid and lymphoid specific pathways in that context. In contrast, the VEGF-, EGF-(Epidermal Growth Factor), PDGF- (Platelet-Derived Growth Factor), and Endocytosis signalling pathways were enriched only in the indirect Set3 genes as compared to Set1*.

In conclusion, the large number of pathways that are (potentially) affected in the *Ldb1*$^{-/-}$ Flk1^{+} BL-CFCs easily explains the severe phenotype of the *Ldb1*$^{-/-}$ embryos. Importantly, the fact that some Flk1^{+} cells still emerge during *Ldb1*$^{-/-}$ ES cell differentiation suggests that most of these pathways may play some role in

hemangioblast formation but become essential at subsequent developmental stages before the hematopoietic and endothelial lineages diverge.

Discussion

Primitive hematopoiesis begins in the yolk sac after the migration of brachyury expressing mesodermal cells through the primitive streak and their differentiation toward hematopoietic or endothelial precursors that collectively form blood islands. Hematopoietic precursors give rise to primitive erythroblasts and endothelial precursors to the vasculature. Both are absent in *Ldb1*$^{-/-}$ embryo yolk sacs, questioning the role of *Ldb1* in this process.

Many of the misregulated factors identified in the gene expression profiling of Flk1^{+} BL-CFCs from day 4 *Ldb1*$^{+/+}$ and *Ldb1*$^{-/-}$ EBs act in the nuclear "response" of well known signalling pathways. WNT signalling has a positive role in regulating primitive hematopoiesis.[32, 40] Addition of the inhibitor Dkk1 (down-regulated in the *Ldb1*$^{-/-}$ Flk1^{+} cells) blocks the pathway reducing the potential of BL-CFCs to generate colonies.[31] Both WNT and Notch signalling pathways are involved in the differentiation of hemangioblast cells towards the primitive erythroid lineage. WNT signalling is active in Flk1^{+} BL-CFCs during the first hours of differentiation towards blast colonies, while Notch signalling remains inactive through *Numb* inhibition. Between 12 and 24 hours of differentiation the inhibitory effect exerted on Notch signalling is reduced and its activation leads to the expression of WNT inhibitors blocking the WNT pathway.[31] Both pathways were found to be active in Flk1^{+} cells, however only WNT signalling was perturbed in *Ldb1*$^{-/-}$ Flk1^{+} cells. We propose that WNT misregulation can be attributed to changes in expression of genes bound by Ldb1. In *Drosophila* the importance of interplay between Notch, FGF (Fibroblast

Growth Factor), MAPK and EGFR signalling pathways has been documented as a requirement for blood specification, [41] with the last two pathways affected by the deletion of *Ldb1* in Flk1^{+} cells.

Changes in gene expression of the WNT signalling associated transcription factors *Lef1*, *Tcf7* and *Tcf7l1* in *Ldb1*$^{-/-}$ Flk1^{+} cells confirms the importance of this pathway in early hematopoiesis. The down-regulation of *Etv2/Er71* in knockout cells leads to the same conclusion as it acts downstream of not only the WNT but also the BMP (Bone Morphogenic Protein) and Notch pathways.[42, 43] Also expression of the TGF-β signalling receptors, endoglin and Alk-1 (Acvrlk1), was reduced in *Ldb1*$^{-/-}$ Flk1^{+} cells. ES cell knockout and rescue experiments showed that these receptors are important for hemangioblast development and primitive hematopoiesis.[44] Recently Ldb1 has been implicated in cell migration and focal adhesion through an interaction with the Ste20-like kinase (SLK), a microtubule associated protein necessary for migration.[45] Our data show that expression of a number of focal adhesion proteins is affected, although SLK did not appear to change. Among the misregulated genes in *Ldb1*$^{-/-}$ Flk1^{+} BL-CFCs we found *Stat5a*, *Sox7*, *Sox17* and *Sox18*. Interestingly, *Stat5a* over-expression in differentiating ES cells leads to an increase in hematopoietic progenitors.[36] *Sox7*, *Sox18* and *Sox17* participate in different stages of hematopoietic development in the mouse embryo with *Sox7* and *Sox18* being involved at the earliest stages of yolk sac hematopoiesis and *Sox17* participating at later stages with the emergence of long-term repopulating hematopoietic stem cells.[37, 38]

Ldb1 acts in a core complex with the essential hematopoietic regulators Scl, Lmo2, Gata1/Gata2, and associates with other essential factors such as Runx1.[14, 15] When *Scl* is absent primitive hematopoiesis is impaired and hematopoietic colonies are absent in culture similar to *Ldb1*$^{-/-}$ embryo phenotype.[20, 46, 47] *In vitro* generation of *Scl*$^{-/-}$ EBs

further showed that this factor is dispensable for hemangioblast generation, but essential for its commitment towards hematopoietic and endothelial lineages.[48] *Gata1* deficient embryos fail to generate erythroid cells but *Gata1* null ES cells can form other hematopoietic lineages.[49, 50] *Gata2* deficiency shows reduced yolk sac hematopoiesis and reduced number of ES cell derived colonies. Finally *Lmo2* absence results in a phenotype similar to that of *Scl* and *Gata2*, although macrophage colonies were still present in colony assays.[47] Our data show that all of these factors are regulated by Ldb1, placing it as a key regulator of the transcription factor network at the origin of the hematopoietic system. Accordingly, Ldb1 deletion in the mouse results in a more severe hematopoietic phenotype than observed with the factors mentioned above. The additional defect observed in endothelial development can also be explained by the EB results. There are still some $CD31^+$ endothelial cells in the *Ldb1*$^{-/-}$ EBs, but their number is severely reduced. We conclude that the defect in primitive hematopoiesis in the absence of *Ldb1* occurs at a stage in development before the hematopoietic and endothelial lineages diverge, i.e. *Ldb1* is involved in the proliferation and differentiation of $Flk1^+$ BL-CFCs resulting in half the number of BL-CFCs that do not develop into blast colonies when compared to normal.

It is yet not clear how the hemangioblast develops from the embryonic mesoderm and how differentiation into hematopoietic and endothelial cells is regulated. Our data show that this must be a process that is orchestrated by a number of signal transduction pathways and transcription factors, many of which are directly or indirectly affected by the loss of Ldb1. One of the best studied transcription factors in hematopoietic development *Runx1* is affected by the absence of *Ldb1*. It is expressed at low levels after the first 24 hours of blast colony formation and increases as the first hematopoietic progenitors of the erythroid and macrophage lineages emerge.[31]

Although *Runx1*$^{-/-}$ blast colonies and primitive hematopoietic colonies are still able to grow, [23] its absence results in the emergence of fewer colonies from EBs.[31] Thus down-regulation of *Runx1* in the *Ldb1*$^{-/-}$ EBs would be yet another cause of the decrease seen in the number of BL-CFCs. A large number of key hematopoietic transcription factors is misregulated in the absence of Ldb1, showing that this factor is critically required to properly orchestrate the gene expression network of hemangioblasts and their progeny (Fig. 6).

A number of factors are also up-regulated and it is interesting to note that expression of the mesodermal marker *brachyury* in *Ldb1*$^{-/-}$ BL-CFCs and *Ldb1*$^{-/-}$ EBs is up-regulated (data not shown). *Brachyury* expression is lost within the first 24 hours of blast colony formation [31] and our data show that it is suppressed by Ldb1. The *brachyury* up-regulation also suggests that *Ldb1*$^{-/-}$ BL-CFCs have not yet lost their mesodermal identity/properties and are therefore unable to differentiate into blast colonies.

Finally our data also show that a number of genes involved in cardiovascular development are affected suggesting that Ldb1 is also important in cardiovascular development presumably in complex with the cardiac Gata factors.

In summary *Ldb1* was identified as an essential regulator of mouse hemangioblast proliferation and differentiation. The severity of the hematopoietic phenotype is the result of a decreased number of hemangioblast cells and their inability to differentiate further down the hematopoietic and endothelial lineages. The expression of essential transcription factors is directly affected by *Ldb1*, revealing its crucial role in the early hematopoietic/endothelial regulatory networks (Fig.6). In addition the identification of Ldb1 target genes in Flk1^{+} cells reveals a number of novel pathways contributing

to hematopoietic/endothelial development and provides better understanding of the early developmental steps of these lineages.

Acknowledgements:

We thank the members of the Grosveld laboratory for technical help and discussion. Cell sorting was done by Reinier van der Linden. This work was supported by NWO (NL), the Dutch Academy of Sciences (KNAW), the EU (the integrated projects EUTRACC and SyBoss), and the European People Marie Curie actions Program (Marie Curie European Reintegration Grants ERG, Call: FP7-PEOPLE-2010-RG).

Author Contributions:

A.M., C.A.S. and F.G. designed the research. A.M., C.A.S., A.M., E.S. and R.J. carried out the research. C.K. carried out Illumina sequencing, S.T., and J.H. performed bioinformatics analyses. W.V.I., and B.L. supervised Illumina data processing and bioinformatics analyses. A.M. C.A.S., E.S., S.T. and F.G. wrote the manuscript.

The authors declare no conflict of interest.

References

1. Baron MH. Early patterning of the mouse embryo: implications for hematopoietic commitment and differentiation. *Exp Hematol.* Sep 2005;33(9):1015-1020.
2. Xiong JW. Molecular and developmental biology of the hemangioblast. *Dev Dyn.* May 2008;237(5):1218-1231.
3. Huber TL, Kouskoff V, Fehling HJ, Palis J, Keller G. Haemangioblast commitment is initiated in the primitive streak of the mouse embryo. *Nature.* Dec 2 2004;432(7017):625-630.
4. Kallianpur AR, Jordan JE, Brandt SJ. The SCL/TAL-1 gene is expressed in progenitors of both the hematopoietic and vascular systems during embryogenesis. *Blood.* Mar 1 1994;83(5):1200-1208.

5. Nishikawa SI, Nishikawa S, Hirashima M, Matsuyoshi N, Kodama H. Progressive lineage analysis by cell sorting and culture identifies FLK1+VE-cadherin+ cells at a diverging point of endothelial and hemopoietic lineages. *Development.* May 1998;125(9):1747-1757.
6. Costa G, Kouskoff V, Lacaud G. Origin of blood cells and HSC production in the embryo. *Trends Immunol.* May 2012;33(5):215-223.
7. Lacaud G, Keller G, Kouskoff V. Tracking mesoderm formation and specification to the hemangioblast in vitro. *Trends in cardiovascular medicine.* Nov 2004;14(8):314-317.
8. Shalaby F, Rossant J, Yamaguchi TP, et al. Failure of blood-island formation and vasculogenesis in Flk-1-deficient mice. *Nature.* Jul 6 1995;376(6535):62-66.
9. Choi K. The hemangioblast: a common progenitor of hematopoietic and endothelial cells. *Journal of hematotherapy & stem cell research.* Feb 2002;11(1):91-101.
10. Keller GM. In vitro differentiation of embryonic stem cells. *Curr Opin Cell Biol.* Dec 1995;7(6):862-869.
11. Li L, Jothi R, Cui K, et al. Nuclear adaptor Ldb1 regulates a transcriptional program essential for the maintenance of hematopoietic stem cells. *Nat Immunol.* Feb 2011;12(2):129-136.
12. Li L, Lee JY, Gross J, Song SH, Dean A, Love PE. A requirement for Lim domain binding protein 1 in erythropoiesis. *J Exp Med.* Nov 22 2010;207(12):2543-2550.
13. Mukhopadhyay M, Teufel A, Yamashita T, et al. Functional ablation of the mouse Ldb1 gene results in severe patterning defects during gastrulation. *Development.* Feb 2003;130(3):495-505.
14. Meier N, Krpic S, Rodriguez P, et al. Novel binding partners of Ldb1 are required for haematopoietic development. *Development.* Dec 2006;133(24):4913-4923.
15. Soler E, Andrieu-Soler C, de Boer E, et al. The genome-wide dynamics of the binding of Ldb1 complexes during erythroid differentiation. *Genes & development.* Feb 1 2010;24(3):277-289.
16. Soler E, Andrieu-Soler C, Boer E, et al. A systems approach to analyze transcription factors in mammalian cells. *Methods.* Feb 2011;53(2):151-162.
17. Song SH, Kim A, Ragoczy T, Bender MA, Groudine M, Dean A. Multiple functions of Ldb1 required for beta-globin activation during erythroid differentiation. *Blood.* Sep 30 2010;116(13):2356-2364.
18. Ferkowicz MJ, Starr M, Xie X, et al. CD41 expression defines the onset of primitive and definitive hematopoiesis in the murine embryo. *Development.* Sep 2003;130(18):4393-4403.
19. Baldwin HS, Shen HM, Yan HC, et al. Platelet endothelial cell adhesion molecule-1 (PECAM-1/CD31): alternatively spliced, functionally distinct isoforms expressed during mammalian cardiovascular development. *Development.* Sep 1994;120(9):2539-2553.
20. Tsai FY, Keller G, Kuo FC, et al. An early haematopoietic defect in mice lacking the transcription factor GATA-2. *Nature.* Sep 15 1994;371(6494):221-226.
21. Tsai FY, Orkin SH. Transcription factor GATA-2 is required for proliferation/survival of early hematopoietic cells and mast cell formation, but

not for erythroid and myeloid terminal differentiation. *Blood.* May 15 1997;89(10):3636-3643.
22. Weiss MJ, Keller G, Orkin SH. Novel insights into erythroid development revealed through in vitro differentiation of GATA-1 embryonic stem cells. *Genes Dev.* May 15 1994;8(10):1184-1197.
23. Lacaud G, Gore L, Kennedy M, et al. Runx1 is essential for hematopoietic commitment at the hemangioblast stage of development in vitro. *Blood.* Jul 15 2002;100(2):458-466.
24. Liu F, Walmsley M, Rodaway A, Patient R. Fli1 acts at the top of the transcriptional network driving blood and endothelial development. *Curr Biol.* Aug 26 2008;18(16):1234-1240.
25. Capron C, Lecluse Y, Kaushik AL, et al. The SCL relative LYL-1 is required for fetal and adult hematopoietic stem cell function and B-cell differentiation. *Blood.* Jun 15 2006;107(12):4678-4686.
26. Nogueira MM, Mitjavila-Garcia MT, Le Pesteur F, et al. Regulation of Id gene expression during embryonic stem cell-derived hematopoietic differentiation. *Biochem Biophys Res Commun.* Sep 24 2000;276(2):803-812.
27. Gardiner MR, Gongora MM, Grimmond SM, Perkins AC. A global role for zebrafish klf4 in embryonic erythropoiesis. *Mech Dev.* Sep-Oct 2007;124(9-10):762-774.
28. Saleque S, Cameron S, Orkin SH. The zinc-finger proto-oncogene Gfi-1b is essential for development of the erythroid and megakaryocytic lineages. *Genes Dev.* Feb 1 2002;16(3):301-306.
29. Goardon N, Lambert JA, Rodriguez P, et al. ETO2 coordinates cellular proliferation and differentiation during erythropoiesis. *EMBO J.* Jan 25 2006;25(2):357-366.
30. Mucenski ML, McLain K, Kier AB, et al. A functional c-myb gene is required for normal murine fetal hepatic hematopoiesis. *Cell.* May 17 1991;65(4):677-689.
31. Cheng X, Huber TL, Chen VC, Gadue P, Keller GM. Numb mediates the interaction between Wnt and Notch to modulate primitive erythropoietic specification from the hemangioblast. *Development.* Oct 2008;135(20):3447-3458.
32. Nostro MC, Cheng X, Keller GM, Gadue P. Wnt, activin, and BMP signaling regulate distinct stages in the developmental pathway from embryonic stem cells to blood. *Cell Stem Cell.* Jan 10 2008;2(1):60-71.
33. Pillay LM, Forrester AM, Erickson T, Berman JN, Waskiewicz AJ. The Hox cofactors Meis1 and Pbx act upstream of gata1 to regulate primitive hematopoiesis. *Dev Biol.* Apr 15 2010;340(2):306-317.
34. Goldie LC, Lucitti JL, Dickinson ME, Hirschi KK. Cell signaling directing the formation and function of hemogenic endothelium during murine embryogenesis. *Blood.* Oct 15 2008;112(8):3194-3204.
35. Nakazawa F, Nagai H, Shin M, Sheng G. Negative regulation of primitive hematopoiesis by the FGF signaling pathway. *Blood.* Nov 15 2006;108(10):3335-3343.
36. Schuringa JJ, Wu K, Morrone G, Moore MA. Enforced activation of STAT5A facilitates the generation of embryonic stem-derived hematopoietic stem cells that contribute to hematopoiesis in vivo. *Stem Cells.* 2004;22(7):1191-1204.
37. Gandillet A, Serrano AG, Pearson S, Lie ALM, Lacaud G, Kouskoff V. Sox7-sustained expression alters the balance between proliferation and

differentiation of hematopoietic progenitors at the onset of blood specification. *Blood.* Nov 26 2009;114(23):4813-4822.
38. Serrano AG, Gandillet A, Pearson S, Lacaud G, Kouskoff V. Contrasting effects of Sox17- and Sox18-sustained expression at the onset of blood specification. *Blood.* May 13 2010;115(19):3895-3898.
39. Takase H, Matsumoto K, Yamadera R, et al. Genome-wide identification of endothelial cell-enriched genes in the mouse embryo. *Blood.* Jul 26 2012;120(4):914-923.
40. Lengerke C, Schmitt S, Bowman TV, et al. BMP and Wnt specify hematopoietic fate by activation of the Cdx-Hox pathway. *Cell Stem Cell.* Jan 10 2008;2(1):72-82.
41. Grigorian M, Mandal L, Hakimi M, Ortiz I, Hartenstein V. The convergence of Notch and MAPK signaling specifies the blood progenitor fate in the Drosophila mesoderm. *Developmental biology.* May 1 2011;353(1):105-118.
42. Kataoka H, Hayashi M, Nakagawa R, et al. Etv2/ER71 induces vascular mesoderm from Flk1+PDGFRalpha+ primitive mesoderm. *Blood.* Dec 22 2011;118(26):6975-6986.
43. Lee D, Park C, Lee H, et al. ER71 acts downstream of BMP, Notch, and Wnt signaling in blood and vessel progenitor specification. *Cell Stem Cell.* May 8 2008;2(5):497-507.
44. Zhang L, Magli A, Catanese J, Xu Z, Kyba M, Perlingeiro RC. Modulation of TGF-{beta} signaling by endoglin in murine hemangioblast development and primitive hematopoiesis. *Blood.* Jul 7 2011;118(1):88-97.
45. Storbeck CJ, Wagner S, O'Reilly P, et al. The Ldb1 and Ldb2 transcriptional cofactors interact with the Ste20-like kinase SLK and regulate cell migration. *Mol Biol Cell.* Oct 2009;20(19):4174-4182.
46. Robb L, Lyons I, Li R, et al. Absence of yolk sac hematopoiesis from mice with a targeted disruption of the scl gene. *Proceedings of the National Academy of Sciences of the United States of America.* Jul 18 1995;92(15):7075-7079.
47. Warren AJ, Colledge WH, Carlton MB, Evans MJ, Smith AJ, Rabbitts TH. The oncogenic cysteine-rich LIM domain protein rbtn2 is essential for erythroid development. *Cell.* Jul 15 1994;78(1):45-57.
48. D'Souza SL, Elefanty AG, Keller G. SCL/Tal-1 is essential for hematopoietic commitment of the hemangioblast but not for its development. *Blood.* May 15 2005;105(10):3862-3870.
49. Pevny L, Lin CS, D'Agati V, Simon MC, Orkin SH, Costantini F. Development of hematopoietic cells lacking transcription factor GATA-1. *Development (Cambridge, England).* Jan 1995;121(1):163-172.
50. Pevny L, Simon MC, Robertson E, et al. Erythroid differentiation in chimaeric mice blocked by a targeted mutation in the gene for transcription factor GATA-1. *Nature.* Jan 17 1991;349(6306):257-260.

Legends

Table 1: Signalling pathways enriched in *Ldb1*$^{+/+}$ and *Ldb1*$^{-/-}$ derived Flk1^{+} cells.

Column A: Signalling pathways enriched for genes expressed in Flk1^{+} cells. **Column B**: Signalling pathways enriched for genes expressed in Flk1^{+} cells and bound by Ldb1. **Column C**: Signalling pathways perturbed in *Ldb1*$^{-/-}$ Flk1^{+} cells. **Column D**: Signalling pathways enriched with genes differentially expressed in *Ldb1*$^{-/-}$ Flk1^{+} cells and bound by Ldb1 (Set1*). **Column E**: Signalling pathways enriched with genes differentially expressed in *Ldb1*$^{-/-}$ Flk1^{+} cells and not bound by Ldb1 (Set3). The presented signalling pathways were selected based on significance (p-value ≤ 0.05) unless stated otherwise. Stars (*) indicate a p-value between 0.07-0.09.

Figure 1: *Ldb1* deletion leads to defects in primitive hematopoiesis and hemangioblast development.

(A) At 9.5 dpc the blood and a vascular network in the *Ldb1*$^{-/-}$ embryo yolk sac is completely absent. **(B)** *Ldb1*$^{+/+}$ and *Ldb1*$^{-/-}$ ES cells were differentiated into EBs. After 3 and 4 days *Ldb1*$^{+/+}$ and *Ldb1*$^{-/-}$ EBs look similar. After 8 days *Ldb1*$^{-/-}$ EBs lack erythroid clusters (grey arrows). **(C)** CFC assay with cell suspensions from *Ldb1*$^{+/+}$ and *Ldb1*$^{-/-}$ day 6 EBs. Primitive erythroid and macrophage colonies were able to grow from *Ldb1*$^{+/+}$ but not *Ldb1*$^{-/-}$ cultures. **(D)** CFC assay with cell suspensions from *Ldb1*$^{+/+}$ and *Ldb1*$^{-/-}$ day 6 EBs as well as *Ldb1*$^{-/-}$ EBs at day 6 expressing exogenous Ldb1. The colonies that formed were counted as primitive erythroid or macrophage colonies according to morphology and colour. The figure shows that while Ldb1 absence impairs the formation of primitive erythroid and macrophage colonies, this phenotype can be rescued with Ldb1 expression. **(E)** FACS analysis of day 6 and day 8 *Ldb1*$^{+/+}$ and *Ldb1*$^{-/-}$ EBs. CD41^{+} primitive erythroid progenitors are not present in the *Ldb1*$^{-/-}$ EBs that contain less than half the number of CD31^{+} endothelial cells in

comparison with the *Ldb1*$^{+/+}$ EBs. Cell suspensions were stained with either CD41-PE or CD31-FITC. CD41-PE^{+} cells are detected through the Fl2 channel and CD31-FITC^{+} cells were detected through the Fl1 channel. **(F)** FACS analysis of day 4 EBs. *Ldb1*$^{-/-}$ EBs contain approximately 50% less Flk1^{+} BL-CFCs than *Ldb1*$^{+/+}$ EBs. (**G)** *Ldb1*$^{+/+}$ Flk1^{+} BL-CFCs give fully grown blast colonies but *Ldb1*$^{-/-}$ Flk1^{+} BL-CFCs do not.

Figure 2: Genome-wide Ldb1 binding in BL-CFCs versus adult hematopoietic progenitor cells and transcriptional function of Ldb1 on BL-CFC gene expression.

(A) Venn diagram comparing Ldb1 binding events in Flk1^{+} BL-CFCs versus Lin^{-} bone marrow hematopoietic progenitor cells.[11] **(B)** Correlative crossing between Ldb1 ChIP-sequencing dataset in BL-CFCs and gene expression dataset of *Ldb1*$^{-/-}$ versus *Ldb1*$^{+/+}$ BL-CFCs.

Figure 3: Enrichment in DNA binding motifs within BL-CFC Ldb1 binding peaks.

(A) Most prevalent motifs enriched within Ldb1 binding peaks of Set1. **(B)** Most prevalent motifs enriched within Ldb1 binding peaks of Set2. Frequency of each motif within the total number of studied binding peaks is given. **(C)** Each identified motif has been associated with one or more transcription factors found in the JASPAR database and the significance of this association is determined by the p-value. Please note that each factor represents a family of transcription factors binding the same motif.

Figure 4: Bubble plot representation of Ldb1 binding peaks around 4 selected hematopoiesis-specific genes down-regulated in the *Ldb1*$^{-/-}$ BL-CFCs.

Each of the figures A-D are divided into three panels. The upper panel shows the studied gene and the neighbouring genomic genes. The middle panel shows the differential expression (presented as log2fold change on the y-axis) of the studied gene in *Ldb1*$^{-/-}$ Flk1^{+} cells in correlation with the Ldb1 binding peaks identified within or around the gene body. The light grey colour given to the circles (allocated to binding peaks) represents down-regulation of gene expression. On the bottom panel ChIP binding peak heights are represented on the y-axis and genome location on the x-axis. The dark grey colour represents Ldb1 ChIP peaks and the black colour represents IgG ChIP peaks (used as control).

Figure 5: Bubble plot representation of Ldb1 binding peaks around differentially regulated developmental hematopoietic/endothelial genes in Ldb1$^{-/-}$ BL-CFCs.

Each bubble in the plot represents one Ldb1 binding peak allocated to a hematopoietic specific gene (purple bubble), an endothelial specific gene (yellow bubble) or a gene involved both in hematopoietic/endothelial development (red bubble). The distance of each binding peak relative to a transcription start site (TSS) of each gene is represented on the x-axis, while changes in gene expression in the *Ldb1*$^{-/-}$ versus *Ldb1*$^{+/+}$ Flk1^{+} cells are represented on the y-axis as the log2fold change. The height of the binding signal/peak is represented by the diameter of the bubble.

Hematopoietic and endothelial specific genes were identified through the use of IPA (Ingenuity® Systems, www.ingenuity.com) and showed a 13.67% (29 out of 212 genes) overlap with the subset of endothelial genes identified through the study of Takase *et al*. 2012.[39] The lists of hematopoietic and endothelial specific genes as well

as the common hematopoietic/endothelial genes used to generate the plot are provided in Table S2.

Figure 6: Speculative model of Ldb1 function in hemangioblast cells.

Schematic representation of Ldb1 functions in the hemangioblast based on whole genome and transcriptome analyses, revealing *Ldb1* involvement in the regulation of key hematopoietic gene expression and essential developmental signalling pathways.

Table 1

Pathways	Column A	Column B	Column C	Column D	Column E
	Signalling pathways enriched with genes expressed in Flk1^{+} cells	**Signalling pathways enriched with genes expressed in Flk1^{+} cells (and bound by Ldb1)**	**Signalling pathways perturbed in *Ldb1*$^{-/-}$ Flk1^{+} cells**	**Signalling pathways perturbed in *Ldb1*$^{-/-}$ Flk1^{+} cells (and bound by Ldb1) (Set1*)**	**Signalling pathways perturbed in *Ldb1*$^{-/-}$ Flk1^{+} cells (and not bound by Ldb1) (Set3)**
Focal adhesion	√	√	√	√	√
Adherens junction	√	√	√	√	√
Angiogenesis	√	√	√	√	√
Colorectal cancer	√	√	√	√	√
Hedgehog signalling pathway	√	√	√	√	√
ErbB signalling pathway	√	√	√	√	√
Insulin signalling pathway	√	√	√	√	√
Integrin Signalling Pathway	√	√	√	√	√
Leukocyte transendothelial migration	√	√	√	√	√
Pathways in cancer	√	√	√	√	√
MAPK signalling pathway	√	√	√	√	√
Regulation of actin cytoskeleton	√	√	√	√	√
Wnt signalling pathway	√	√	√	√	√
Acute myeloid leukemia	√	√	√	√	
Basal cell carcinoma		√	√	√	
Interleukin signalling pathway	√*	√*	√*	√	
Melanoma	√	√	√	√	

T cell receptor signalling pathway	√	√	√	√	
VEGF signalling pathway	√	√	√		√
Apoptosis	√	√	√		√
Axon guidance	√	√	√		√
Chronic myeloid leukemia	√	√	√		√
PDGF Signalling Pathway	√	√	√		√
Small cell lung cancer	√	√	√		√
EGF Signalling Pathway	√*	√	√		√
Endocytosis	√	√	√		√
Glioma	√	√	√		√
Neurotrophin signalling pathway	√	√	√		√
TGF-beta signalling pathway	√	√	√		√
Prostate cancer	√	√	√		
Pancreatic cancer	√	√	√		
uCalpain and friends in Cell spread		√	√		
Bladder cancer	√		√		√
Dopamine receptor mediated signalling pathway	√*		√		√
Tight junction	√		√		√
Oxidative stress response	√		√		√
Melanogenesis	√		√		
B cell activation	√		√		
Telomeres, Telomerase, Cellular Aging, and Immortality	√		√		
Cell cycle	√	√		√	
mTOR signalling pathway	√	√		√	
Ras Pathway	√	√			√
Glutathione metabolism	√				√

Control of Gene Expression by Vitamin D Receptor	√*	√			
Aminoacyl-tRNA biosynthesis	√	√			
Amyotrophic lateral sclerosis (ALS)	√	√			
B cell receptor signalling pathway	√	√			
Basal transcription factors	√	√			
Endometrial cancer	√	√			
Glyoxylate and dicarboxylate metabolism	√	√			
Heme biosynthesis	√	√			
Huntington disease	√	√			
Interferon-gamma signalling pathway	√	√			
JAK/STAT signalling pathway	√	√			
Lysosome	√	√			
Transcription regulation by bZIP transcription factor	√	√			
Adipocytokine signalling pathway	√	√			
Alzheimer's disease	√	√			
Ubiquitin mediated proteolysis	√	√			
Non-small cell lung cancer	√	√			
Nucleotide excision repair	√	√			
Oxidative phosphorylation	√	√			
p38 MAPK Signalling Pathway	√	√			
p53 signalling pathway	√	√			
Parkinson's disease	√	√			
Proteasome	√	√			
Renal cell carcinoma	√	√			
Ribosome	√	√			
RNA degradation	√	√			

Spliceosome	√	√			
Apoptotic Signalling in Response to DNA Damage	√				
Arginine and proline metabolism	√				
Base excision repair	√				
Biosynthesis of unsaturated fatty acids	√				
Chemokine signalling pathway	√				
Chondroitin sulfate biosynthesis	√				
Circadian clock system	√				
Citrate cycle (TCA cycle)	√				
CXCR4 Signalling Pathway	√				
Cysteine and methionine metabolism	√				
DNA replication	√				
Dorso-ventral axis formation	√				
Endothelin signalling pathway	√				
Fatty acid metabolism	√				
Fc gamma R-mediated phagocytosis	√				
fMLP induced chemokine gene expression in HMC-1 cells	√				
Fructose and mannose metabolism	√				
Gap junction	√				
General transcription regulation	√				
Glycerophospholipid metabolism	√				
Glycosylphosphatidylinositol (GPI)-anchor biosynthesis	√				
GnRH signalling pathway	√				
HIV-I Nef: negative effector of Fas and TNF	√				

Homologous recombination	√				
Hypoxia response via HIF activation	√				
Influence of Ras and Rho proteins on G1 to S Transition	√				
Inositol phosphate metabolism	√				
Keratan sulfate biosynthesis	√				
Limonene and pinene degradation	√				
Links between Pyk2 and Map Kinases	√				
Long-term depression	√				
Long-term potentiation	√				
Lysine degradation	√				
Mismatch repair	√				
N-Glycan biosynthesis	√				
Notch signalling pathway	√				
One carbon pool by folate	√				
Oocyte meiosis	√				
Other glycan degradation	√				
Phosphatidylinositol signalling system	√				
Phospholipids as signalling intermediaries	√				
Progesterone-mediated oocyte maturation	√				
Propanoate metabolism	√				
Purine metabolism	√				
Pyrimidine metabolism	√				
Pyruvate metabolism	√				
Ras Signalling Pathway	√				
Regulation of eIF4e and p7 S6 Kinase	√				
RNA polymerase	√				

Role of BRCA1, BRCA2 and ATR in Cancer Susceptibility	√				
Role of Mitochondria in Apoptotic Signalling	√				
Selenoamino acid metabolism	√				
Signalling of Hepatocyte Growth Factor Receptor	√				
Skeletal muscle hypertrophy is regulated via AKT/mTOR pathway	√				
SNARE interactions in vesicular transport	√				
Steroid biosynthesis	√				
Amino sugar and nucleotide sugar metabolism	√				
Valine, leucine and isoleucine degradation	√				
Thyroid cancer	√				

Figure 1

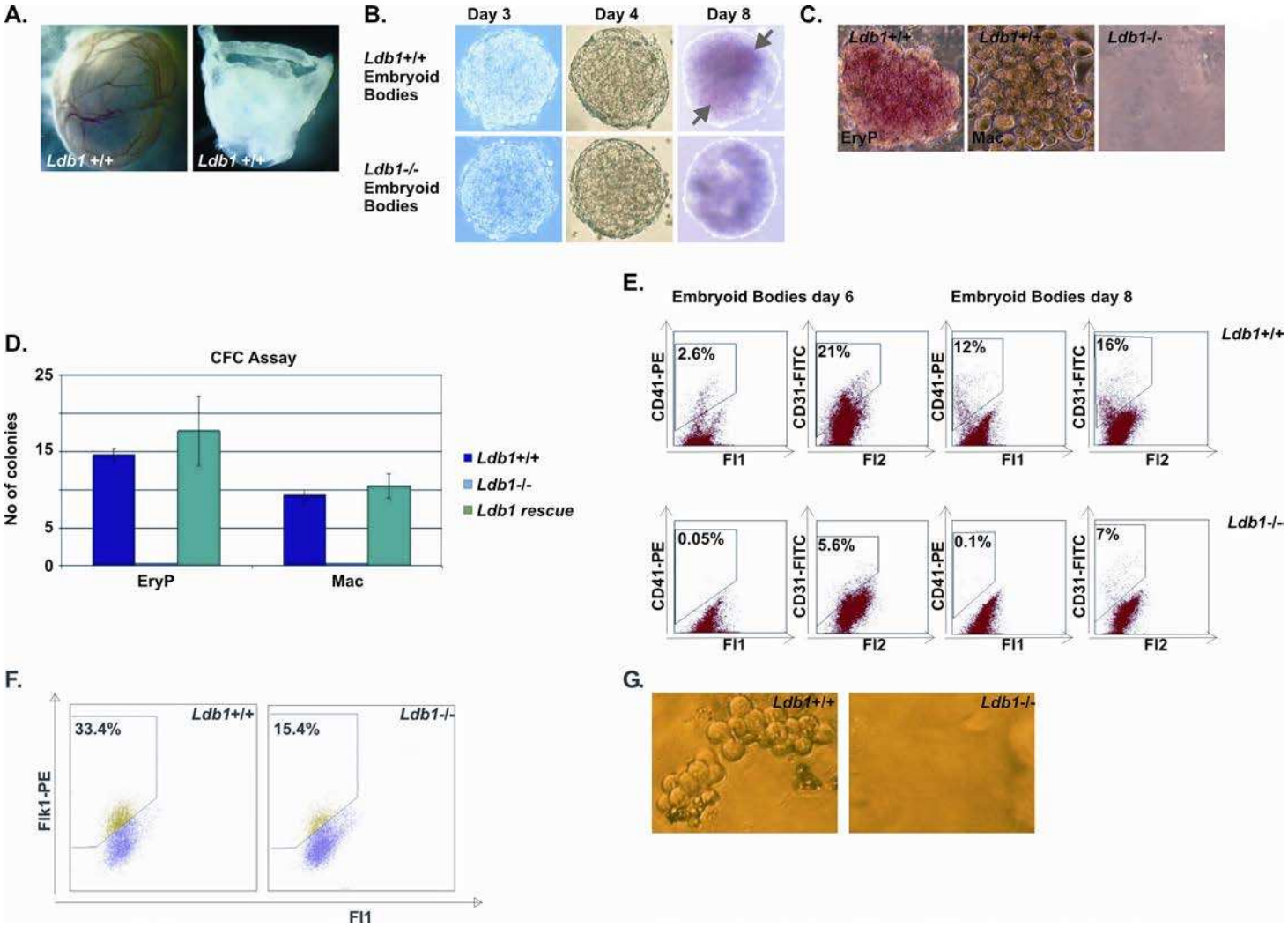

A.
Ldb1 +/+
Ldb1 +/+
B.
Day 3
Day 4
Day 8
Ldb1+/+ Embryoid Bodies
Ldb1-/- Embryoid Bodies
C.
Ldb1+/+
Ldb1+/+
Ldb1-/-
EryP
Mac
D.
CFC Assay
No of colonies
25
15
5
0
EryP
Mac
Ldb1+/+
Ldb1-/-
Ldb1 rescue
E.
Embryoid Bodies day 6
Embryoid Bodies day 8
Ldb1+/+
Ldb1-/-
CD41-PE
CD31-FITC
FI1
FI2
2.6%
21%
12%
16%
0.05%
5.6%
0.1%
7%
F.
Ldb1+/+
Ldb1-/-
33.4%
15.4%
Flk1-PE
FI1
G.
Ldb1+/+
Ldb1-/-

Figure 2

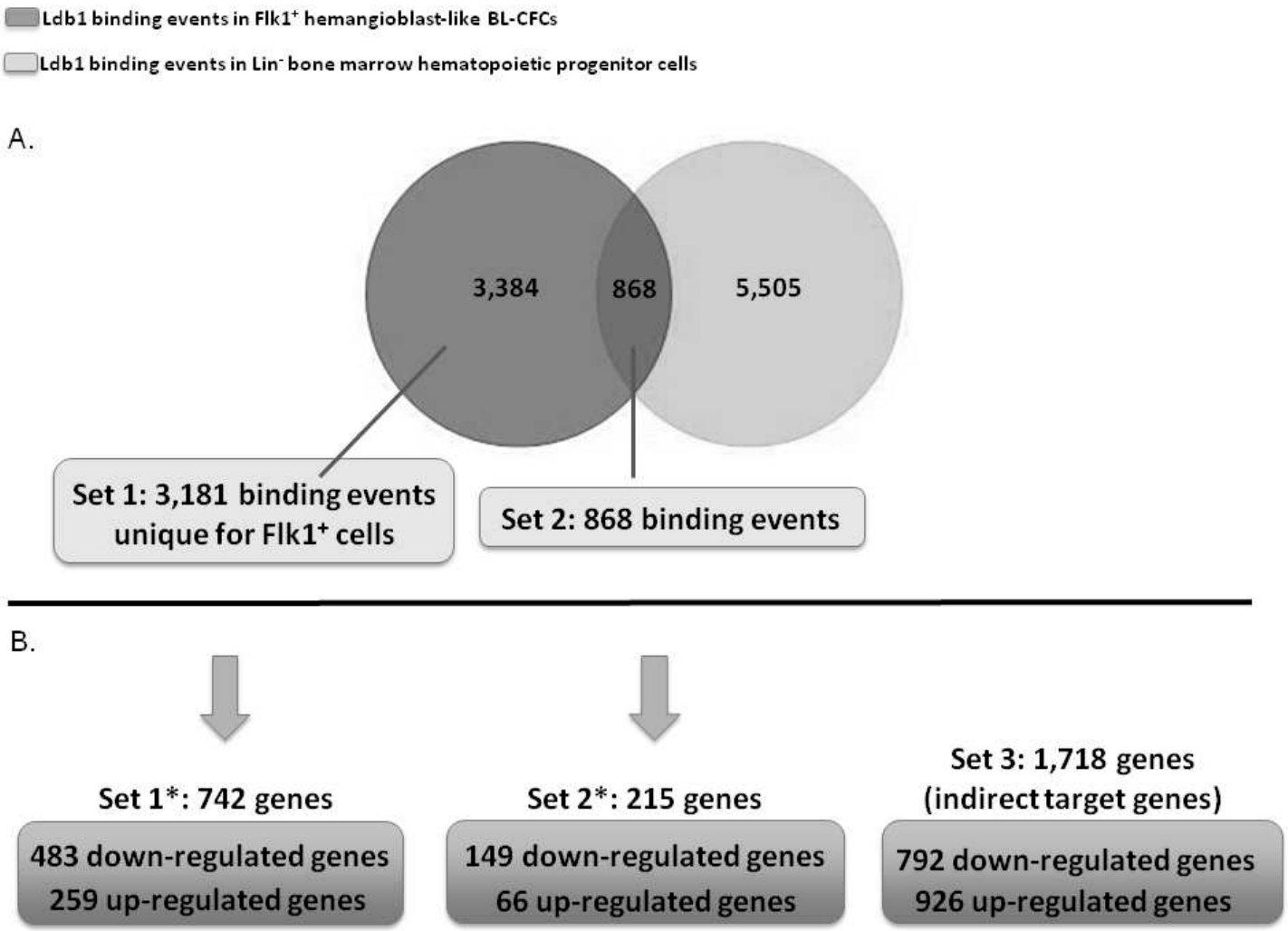
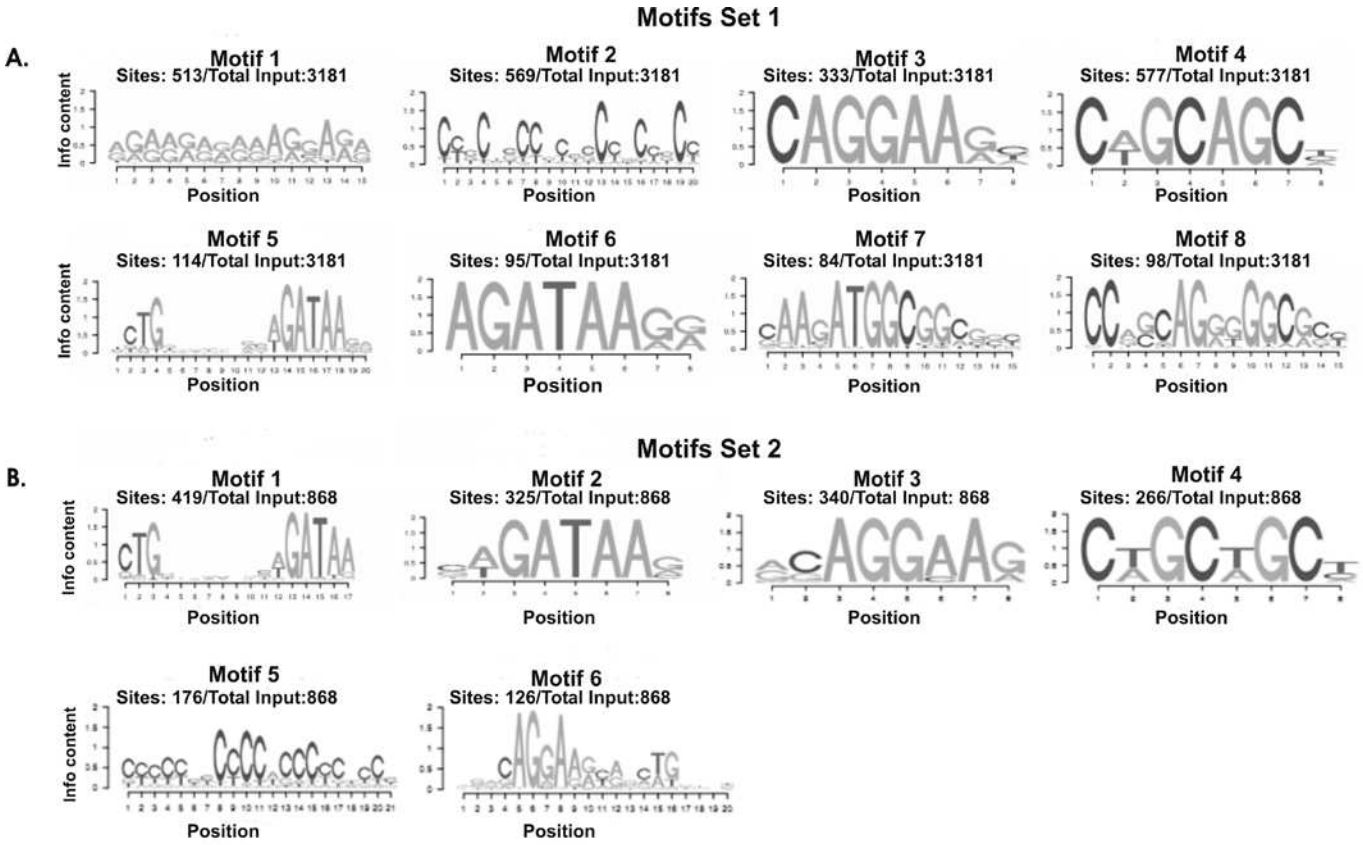

Ldb1 binding events in Flk1+ hemangioblast-like BL-CFCs
Ldb1 binding events in Lin- bone marrow hematopoietic progenitor cells
A.
3,384
868
5,505
Set 1: 3,181 binding events unique for Flk1+ cells
Set 2: 868 binding events
B.
Set 1*: 742 genes
483 down-regulated genes
259 up-regulated genes
Set 2*: 215 genes
149 down-regulated genes
66 up-regulated genes
Set 3: 1,718 genes (indirect target genes)
792 down-regulated genes
926 up-regulated genes

Figure 3



C.

Motifs Set 1	Predicted Transcription Factor	p-value
Motif 1	Klf4	3.52e-06
Motif 2	SP1	5.45e-07
Motif 3	PU.1	0.00018
Motif 4	Myf	9.46e-05
Motif 5	Tal:Gata	2.1e-18
Motif 6	Gata	1.59e-06
Motif 7	Smad3	0.00013
Motif 8	CTCF	1.03e-13

Motifs Set 2	Predicted Transcription Factor	p-value
Motif 1	Tal:Gata	5.51789e-17
Motif 2	Gata	1.31003e-11
Motif 3	PU.1	6.11e-05
Motif 4	Myf	9.99e-06
Motif 5	SP1	2.18e-08
Motif 6	PU.1/FEV/ELF5	0.00037/2.43e-05/5.56e-05

Figure 4

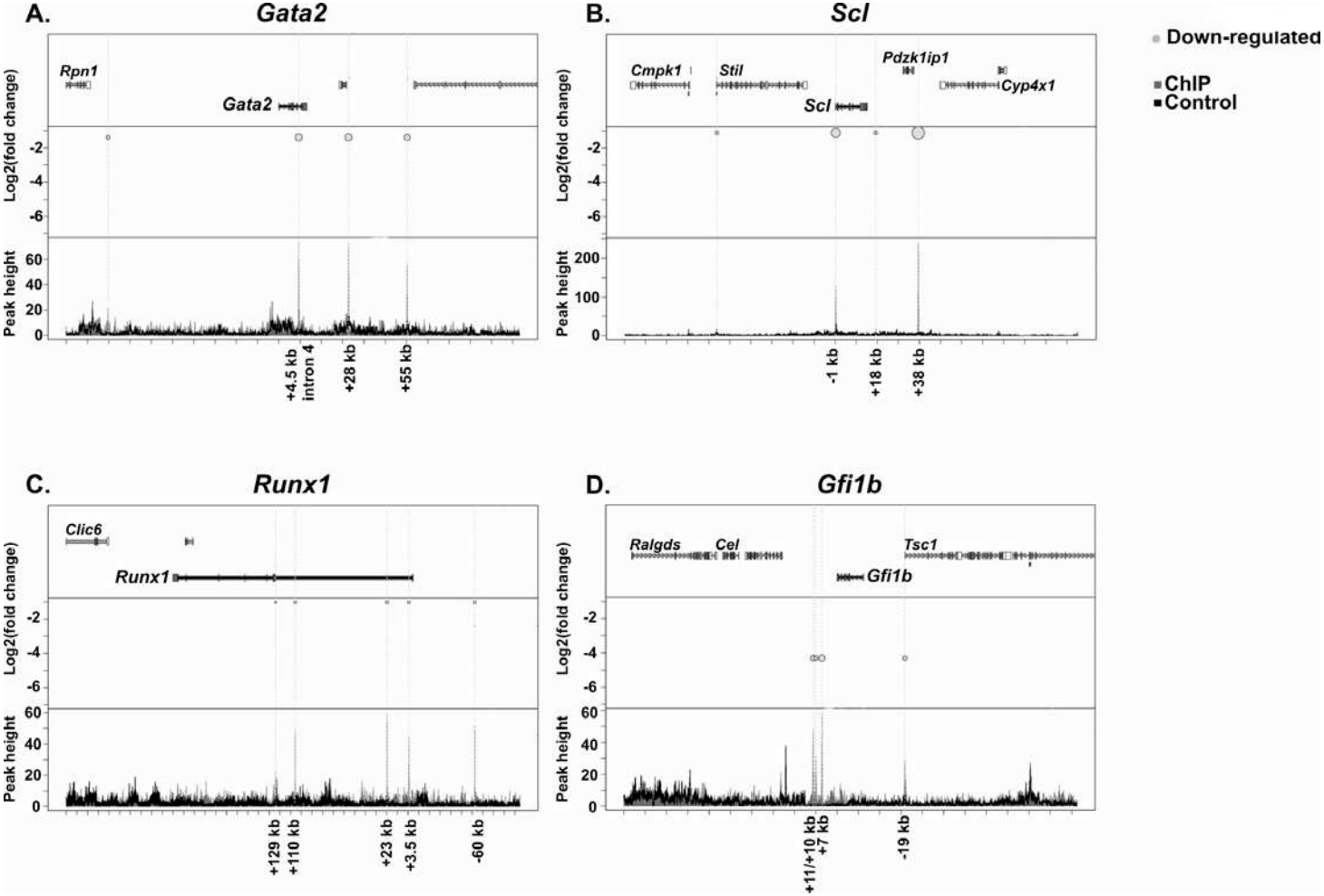

A.
Gata2
Rpn1
Gata2
Log2(fold change)
-2
-4
-6
Peak height
60
40
20
0
+4.5 kb intron 4
+28 kb
+55 kb
B.
Scl
Cmpk1
Stil
Pdzk1ip1
Cyp4x1
Scl
Log2(fold change)
-2
-4
-6
Peak height
200
100
0
-1 kb
+18 kb
+38 kb
Down-regulated
ChIP
Control
C.
Runx1
Clic6
Runx1
Log2(fold change)
-2
-4
-6
Peak height
60
40
20
0
+129 kb
+110 kb
+23 kb
+3.5 kb
-60 kb
D.
Gfi1b
Ralgds
Cel
Tsc1
Gfi1b
Log2(fold change)
-2
-4
-6
Peak height
60
40
20
0
+11/+10 kb
+7 kb
-19 kb

Figure 5

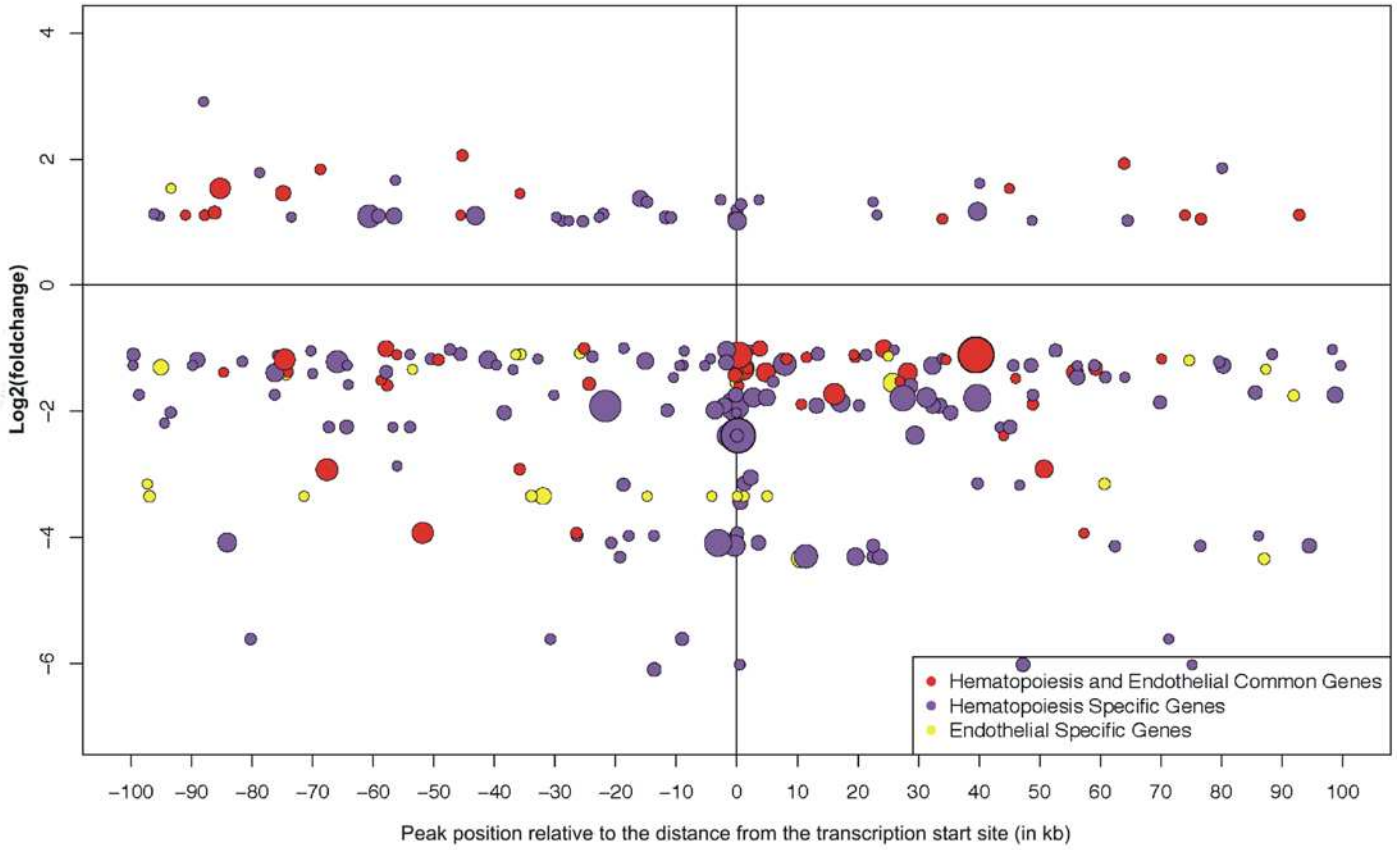
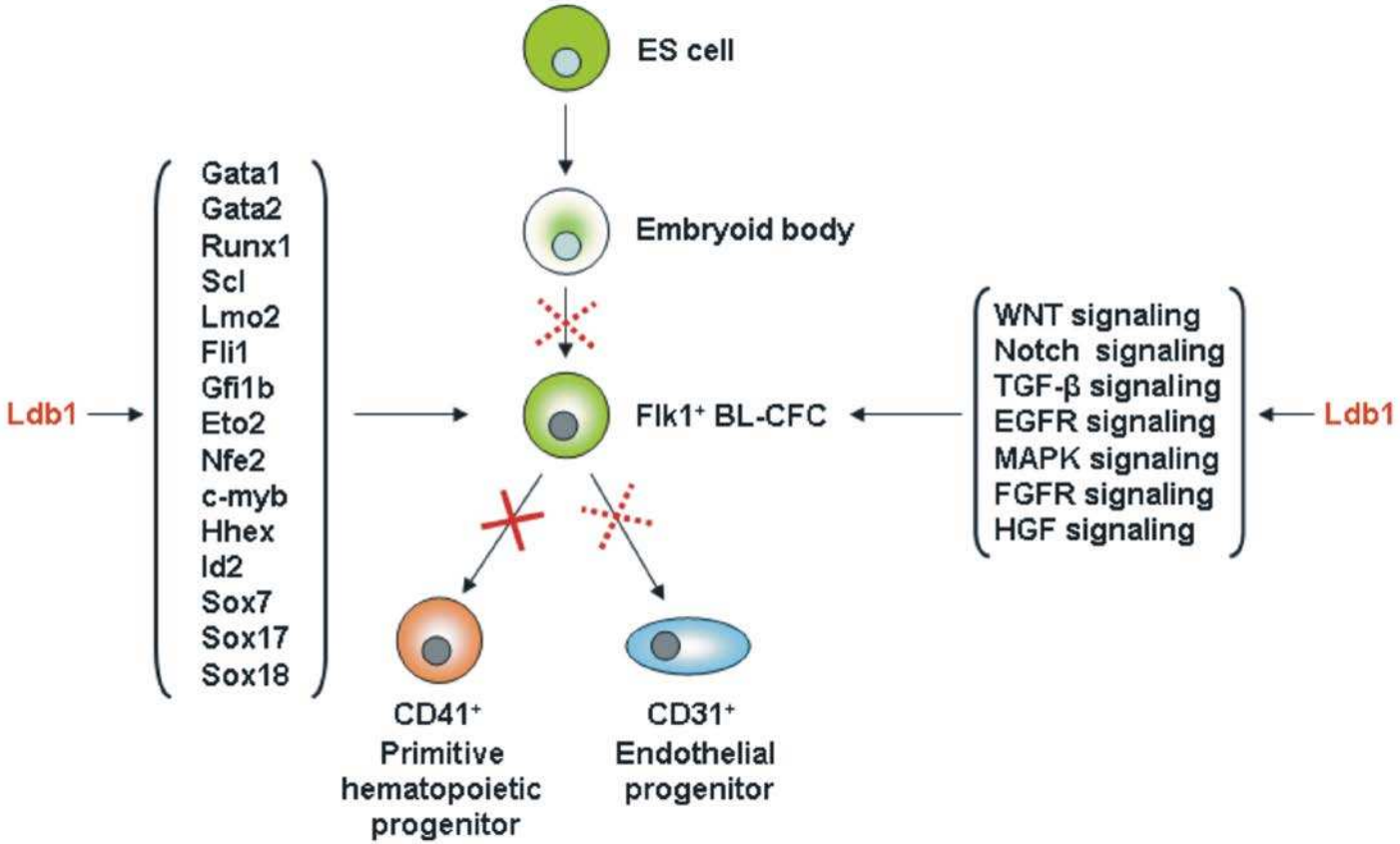

Log2(foldchange)
4
2
0
−2
−4
−6
−100 −90 −80 −70 −60 −50 −40 −30 −20 −10 0 10 20 30 40 50 60 70 80 90 100
Peak position relative to the distance from the transcription start site (in kb)
Hematopoiesis and Endothelial Common Genes
Hematopoiesis Specific Genes
Endothelial Specific Genes

Figure 6


ES cell
Embryoid body
Flk1+ BL-CFC
Ldb1
Gata1
Gata2
Runx1
Scl
Lmo2
Fli1
Gfi1b
Eto2
Nfe2
c-myb
Hhex
Id2
Sox7
Sox17
Sox18
WNT signaling
Notch signaling
TGF-β signaling
EGFR signaling
MAPK signaling
FGFR signaling
HGF signaling
Ldb1
CD41+
Primitive
hematopoietic
progenitor
CD31+
Endothelial
progenitor