Identification of novel regulators of developmental hematopoiesis using endoglin regulatory elements as molecular probes

Citation for published version:

Digital Object Identifier (DOI):
10.1182/blood-2016-02-697870

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
Blood

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
Identification of Novel Regulators of Developmental Hematopoiesis Using *Endoglin*

Regulatory Elements as Molecular Probes

Running title: *Endoglin* GREs Target Distinct Hemogenic Precursors

Rabab Nasrallah¹, Eva Fast², Parham Solaimani³*, Kathy Knezevic¹, Alexia Eliades⁴, Rahima Patel⁴, Roshana Thambyrajah⁴, Ashwin Unnikrishnan¹, Julie Thoms¹, Dominik Beck¹, Chris S. Vink³, Aileen Smith⁶**, Jason Wong¹, Mairi Shepherd⁶, David Kent⁶, Rahul Roychoudhuri⁷, Fabian Paul⁸, Julia Klippert⁸, Annette Hammes⁸, Thomas Willnow⁸, Bertie Göttsgens⁶, Elaine Dzierzak³,⁵, Leonard I. Zon², George Lacaud⁴#, Valerie Kouskoff⁴#, John E. Pimanda¹,⁹#

¹Lowy Cancer Research Centre and the Prince of Wales Clinical School, UNSW Australia, Sydney, NSW 2052, Australia. ²Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge, Massachusetts, USA. ³Erasmus MC Stem Cell Institute, Department of Cell Biology, Erasmus University Medical Center, 3000 CA Rotterdam, Netherlands, ⁴Cancer Research UK Manchester Institute, The University of Manchester, UK. ⁵MRC Centre for Inflammation Research, University of Edinburgh, UK. ⁶Department of Haematology, Wellcome Trust and MRC Cambridge Stem Cell Institute and Cambridge Institute for Medical Research, Cambridge University, Cambridge CB2 0XY, UK. ⁷The Babraham Institute, Babraham Research Campus, Cambridge CB22 3AT. ⁸Max-Delbrueck-Center for Molecular Medicine, 13125 Berlin, Germany. ⁹Department of Haematology, The Prince of Wales Hospital, Sydney, NSW 2032, Australia.

*Current address- BeDataDriven B.V. Binckhorstlaan 36 M4.12, 2615 BE, The Hague, Netherlands

**Current address- GSK Pharmaceutical Company, London, UK

#Correspondence:
Dr John E Pimanda: j.pimanda@unsw.edu.au, +61293851003 (Tel), +61293851510 (Fax). Dr Valerie Kouskoff: Valerie.Kouskoff@cruk.manchester.ac.uk, +44(0)1614468381 (Tel), +44(0)1614463109 (Fax). Dr Georges Lacaud: Georges.Lacaud@cruk.manchester.ac.uk, +44(0)161 446 8380 (Tel, +44(0)1614463109 (Fax)

Abstract- 200, Text- 5032, Figures-7, References- 55

Supplemental data: Supplemental Figures-10, Supplemental Tables-2, Supplemental methods

Scientific category: Hematopoiesis and Stem Cells
KEY POINTS:

• Endoglin regulatory elements target hemogenic mesoderm and hemogenic endothelium
• Hemogenic progenitors can be enriched using these elements as molecular probes to discover novel regulators of hematopoiesis
ABSTRACT

Enhancers are the primary determinants of cell identity and specific promoter/enhancer combinations of Endoglin (ENG) have been shown to target blood and endothelium in the embryo. Here, we generated a series of embryonic stem cell lines, each targeted with reporter constructs driven by specific promoter/enhancer combinations of ENG, to evaluate their discriminative potential and value as molecular probes of the corresponding transcriptome. The *Eng* promoter (P) in combination with the -8/+7/+9kb enhancers, targeted cells in FLK1 mesoderm that were enriched for blast colony forming potential, whereas the P/-8kb enhancer targeted TIE2+/c-KIT+/CD41- endothelial cells that were enriched for hematopoietic potential. These fractions were isolated using reporter expression and their transcriptomes profiled by RNA-seq. There was high concordance between our signatures and those from embryos with defects at corresponding stages of hematopoiesis. Of the six genes that were up-regulated in both hemogenic mesoderm and hemogenic endothelial fractions targeted by the reporters, LRP2, a multiligand receptor, was the only gene that had not previously been associated with hematopoiesis. We show that LRP2 is indeed involved in definitive hematopoiesis and by doing so validate the use of reporter gene coupled enhancers as probes to gain insights into transcriptional changes that facilitate cell fate transitions.
INTRODUCTION

With advances in microscopy and histology, different cell types can now readily be distinguished from one another. However, the molecular characteristics that make each cell type unique and help distinguish stem cells from their more differentiated progeny in a tissue, are still obscure. Harvesting pure populations of stem cells is a pre-requisite to probing their molecular identity. Over the years, protocols combining flow cytometry with single cell serial transplantation assays have been progressively refined to purify mouse and human adult hematopoietic stem cells (HSCs)\textsuperscript{1,2}.

One of the utilitarian benefits of determining the molecular fingerprint of a HSC is that it could serve as a measurable goal when developing protocols aimed at generating HSCs from differentiated cells\textsuperscript{3}. The failure of current protocols to generate long-term repopulating HSCs from ES/iPS cells is attributed in part to our incomplete understanding of the developmental journey that mesodermal progenitors traverse in the embryo when generating the complement of HSCs that are resident in the bone marrow of a newborn\textsuperscript{4}. Determining the molecular identities of embryonic HSC precursors is complicated by the lack of consensus regarding the precise HSC intermediates in the embryo, functional assays that are less than ideal for assessment of these intermediates and knowledge that these intermediates are transitory cell populations that are present in very small numbers\textsuperscript{5}. FLK1 expressing mesodermal cells in the posterior primitive streak when isolated from the embryo and cultured \textit{in vitro} generate blast colonies that have blood, endothelial and vascular smooth muscle potential\textsuperscript{6}. Blast colony forming cell (BL-CFC) potential in FLK1+ mesoderm has been estimated to be \textasciitilde1:300\textsuperscript{7}. Hemogenic potential in TIE2+c-KIT+ hemogenic endothelial or VE-CAD+CD45-CD41- pre-HSC cells in the dorsal aorta that transit to hematopoietic cells range from 1:100-300\textsuperscript{8-10}. These functional estimates are too low to probe the molecular
identities of either the early hemangioblast or hemogenic endothelial cell populations in the
developing embryo using currently available protocols.

Cell identity is encoded within the sequences of tissue specific gene regulatory
elements that direct and coordinate gene expression in a cell. A number of regulatory
elements of hematopoietic transcription factors have previously been shown to direct reporter
expression to developing blood cells in the mouse embryo and include enhancers of \( Scl \),
\( Runx1, Gata2, Erg, Fli1, Lmo2 \) and \( Lyl1 \), which also form a recursive circuit in human adult
HSCs. The \( Runx1 +23 \) enhancer marks a population of early hemogenic endothelial cells
that transit to HSCs and has been used to isolate cells from different embryonic stages for
transcriptomic analysis. \( Ly6a/Sca1 \) and Endoglin (\( Eng; CD105 \)) serve as useful cell surface
markers for isolation of murine HSC fractions. The promoter of \( Ly6a \) and
promoter/enhancer combinations of \( Eng \) also target embryonic hematopoiesis and in the case
of the former have been used in conjunction with a reporter to isolate hemogenic endothelial
cells and HSCs from early embryos.

ENG is an accessory receptor and modulator of TGF-\( \beta \) superfamily signaling. ENG
is expressed on FLK1+ mesoderm and is required for normal BL-CFC development and its
expression facilitates the hematopoietic program in these cells. ENG null mice die at
E9.5 with vascular defects due to abnormal endothelial and pericyte development. It is also
a marker of adult murine HSCs that was identified using a \( Scl +19 \) driven fluorescent reporter
coupled with transcriptomic and proteomic assessment of purified cells. An emerging
concept of developmental hematopoiesis posits that HSC development from the dorsal aorta
at E10 reflects maturation of cells that were fated earlier during embryogenesis towards the
hematopoietic lineage. As such we rationalised that transcriptional regulation of ENG,
which is functionally important for the development of hematopoietic intermediates could be instructive in helping elucidate the transcriptional environment of these cells. We have previously shown that sequence information within the promoter and hemato-endothelial enhancers of Eng determine how reporter genes are targeted to either endothelial or blood and endothelial tissues in the embryo\textsuperscript{17,22}. Given the spectrum of cell types that are involved in the developmental journey of embryonic HSCs and the deterministic role that ENG plays in their development, we hypothesised that distinct combinations of promoter/enhancers of this gene are used by different hematopoietic intermediates to regulate ENG expression. We rationalised that if distinct promoter/enhancer constructs indeed targeted functionally distinct hematopoietic intermediates, they could be used as molecular probes to profile the transcriptional environment of these cells.

Here we show using ES cells with single copy reporter coupled transgenes targeted to the constitutively active HPRT locus that distinct promoter/enhancer combinations of ENG are used by FLK1+ mesoderm and hemogenic endothelium that are enriched for BL-CFC and hematopoietic potential respectively. Using these reporter coupled transgenes as probes to harvest cell populations from ESC differentiation assays, we performed RNA-seq to identify gene sets that were associated with functional enrichment of hematopoietic potential and show their complementarity with primary mouse tissues at matching stages of development. Of the six genes that were up-regulated in both hemogenic mesoderm and hemogenic endothelial fractions targeted by the reporters, LRP2, a multiligand receptor, was the only gene that had not previously been associated with hematopoiesis. Here, we show that LRP2 is indeed involved in definitive hematopoiesis and by doing so validate the use of reporter gene coupled enhancers as a discovery tool.
MATERIALS AND METHODS

Murine ES cell culture
The Bry/GFP\textsuperscript{7} and HM1 ES cells\textsuperscript{23} were cultured as previously described. See Supplemental data.

Generating $Hprt$ targeting constructs
To generate the $Hprt$ targeting constructs DNA fragments corresponding to the $Eng$ -8, P, +7, +17,\textsuperscript{22} mutant +7 (+7\Delta, mutant Gata), and mutant +9 (+9\Delta, mutant Ets) regions (Fig. S1B) along with the $LacZ$ reporter gene were ligated in multiple combinations into the pMP8 targeting vector. See Supplemental data.

$Hprt$ targeting of HM1 ES cells
To prepare the cells for targeting, HM1 ES cells were cultured in DMEM-ES medium containing 1× 6-Thioguanine (6-TG, Sigma Aldrich) for a week before electroporation. See Supplemental data. Genomic DNA was extracted and analysed for successful recombination by PCR and southern blots.

Isolation of Genomic DNA
Genomic DNA isolated from recombinant ES cell clones was analysed for successful integration of the DNA fragments by qPCR using primers listed in Supplemental data.

Southern blotting
10μg of genomic DNA was digested with HpaI and SacI (NEB) at 37°C O/N and blotted as detailed in Supplemental data.
ES cell differentiation into EBs

To generate EBs, ES cells were collected and cultured as detailed in Supplemental data. Cells were then seeded on ultra-low attachment 60mm plates (Sterilin) and placed in an incubator at 37°C and 5%CO₂. These culture conditions were optimal for growth of day 1-4 EBs.

LacZ staining

Fluorescein di-β-D-galactopyranoside (FDG, Sigma Aldrich) was added to the cells at a 1:1 ratio and incubated at 37C for 1 min. The cell suspension was then added to IMDM/20%FBS and placed on ice for 1 hour. Further staining with primary and secondary antibodies was performed after the LacZ staining procedure. See Supplemental data.

Flow cytometry and cell sorting

Cells were collected from EBs and liquid blast cultures and dissociated into a single cell suspension. Details of procedure and antibodies are listed in Supplemental data.

Methylcellulose blast colony forming (BL-CFC) assay

2x10⁴ FLK1+/LacZ- and FLK1+/LacZ+ cells sorted from day 3 EBs were seeded in methylcellulose mix. The blast potential in Flk1+ cells was assayed in triplicate dishes of 1ml. The plates were then placed in a humidified incubator 37°C and 5% CO₂. Colonies were scored following four days of culture. See Supplemental data.

Liquid BL-CFC assay

Day 3 EBs were dissociated into single cells using trypsin (Gibco), and stained with a Flk1-PE or Flk1-bio antibody (eBiosciences). Sorted FLK1+ cells were seeded on gelatin coated plates at a density of 7.5-8.5x10⁴ cells/9.6cm² (in a humidified incubator 37°C, 5% CO₂). The
culture medium consisted of IMDM supplemented with 10% FBS pre-tested for
differentiation, 1% L-Glutamine, 0.6% of 30mg/ml transferrin, 0.3% of 0.15M MTG, 0.5% of
5mg/ml ascorbic acid, 15% D4T conditioned medium, 0.1% of 5μg/ml VEGF, and 0.1% of
10μg/ml IL-6 (R&D systems).

Hematopoietic methylcellulose colony-forming assay
Cells isolated from day 2 or 4 of liquid blast cultures and seeded for CFU-C assays as
detailed in Supplemental data. Primitive colonies were scored after 4-5 days of culture,
whereas most definitive colonies were scored after 7-10 days. All cytokines used in this assay
were purchased from R&D systems.

RT-PCR
List of primers and methods are in Supplemental data.

Chromatin immunoprecipitation (ChIP) assay
ChIP assays were performed as detailed previously. See Supplemental data for a list of
primers and experimental details.

Mouse embryo immunostaining and imaging
10um thick cryosections of E10.5 fixed mouse embryos embedded in a gelation/sucrose
solution were thawed and rehydrated in PBS prior to immunostaining. Details of procedure
and antibodies are listed in Supplemental data.

Generating zebrafish morpholinos and analysis
Details are listed in Supplemental data.
Statistical analysis

RT-PCR data, BL-CFCs and hematopoietic colony counts were statistically analysed using Student’s T-test or Paired Student’s T-test. Significant differences are marked * for p<0.05, ** for p<0.01, *** for p<0.001 and not significant differences are marked NS.

RNA sequencing and analysis

RNA was isolated from cells as per standard method (For details about samples refer to supplementary table S3), and amplified using the Ovation RNA amplification system V2 (Nugen). Single-stranded TruSeq cDNA libraries were generated and sequenced using the Illumina HiSeq2000 analyzer (BGI, Hong Kong). The data has been deposited in GEO under the accession number GSE77390. See Supplemental data for details.

Pathway and Gene-set enrichment analysis

The Ingenuity IPA Core Analysis Tool (version 17199142) was used to establish ‘Molecular and Cellular Functions’ that correlate with the lists of differentially expressed genes. Gene-set enrichment analysis was performed using the GSEA Java Desktop tool (v 2.0.13). See Supplemental data for details.
RESULTS

Mesoderm to hemangioblast transition is accompanied by increased Eng expression and chromatin accessibility at hemato-endothelial regulatory elements. The promoter of ENG when coupled with -8kb, +7kb and +9kb enhancers have previously been shown to direct reporter expression to either endothelial or blood and endothelial tissues in the embryo (Figure 1A; 17,22). The Bry-GFP ESC line has been used extensively to investigate the developmental progression of pre-mesoderm (GFP-/FLK1-) to pre-hemangioblast mesoderm (GFP+/FLK1-; G+/F-) to the hemangioblast (GFP+/FLK1+; G+F+) (Figure 1B). We used this cell line to first evaluate expression of Eng and chromatin accessibility at hemato-endothelial regulatory elements of Eng17,22 as cells progressed from pre-hemangioblast mesoderm to hemangioblast mesoderm. Eng expression increased by ~ 3-fold (Figure 1C) and enrichments of H3K9 acetylation (an active chromatin mark) increased ~ 10-20 fold at the Eng promoter and -8kb, +7kb and +9kb Eng enhancers (Figure 1D). There was no change in H3K9Ac at -4kb, a region that is highly conserved across species but shows no enhancer activity22.

The Eng promoter when combined with the -8, +7 and +9 hemato-endothelial enhancers, targets FLK1+ mesodermal cells enriched for BL-CFC potential. HM1 ESCs have a disrupted Hprt locus that can be reconstituted by homologous recombination of a targeting vector23. They serve as a useful tool to evaluate reporter activity of single copies of gene regulatory elements at a constitutively active locus at different stages of ESC differentiation. We took advantage of this system to introduce combinations of Eng regulatory elements with blood and endothelial activity in in vivo transgenic assays (Figure S1). Successful recombination and generation of ESC lines with -8/P/lacZ, -8/P/lacZ/+7, -8/P/lacZ/+9, -8/P/lacZ/+7/+9, -8/P/lacZ/+7Δ (GATA)/+9 and -8/P/lacZ/+7Δ (GATA)/+9 Δ
(ETS), was confirmed by RT-PCR and southern blotting (Figure S2). We used these ESC lines as a tool-kit with which to track, evaluate and compare the activity of each of these gene regulatory elements (GREs) during different stages of hematopoietic development and to fractionate cells for functional validation and transcriptomic analysis.

To identify which, if any of the Eng GREs, targeted FLK1+ mesoderm enriched for hemangioblast potential, we generated embryoid bodies (EBs) from each ESC line and fractionated FLK1+lacZ- (F+L-) and FLK1+lacZ+ (F+L+) cells and performed blast colony forming cell (BL-CFC) assays (Figure 2A). The Eng -8P/lacZ/+7/+9 construct, which showed robust blood and endothelial staining in vivo targeted a fraction of the FLK1+ mesoderm that showed increased (~ 4 fold) BL-CFC potential (Figure 2B and Figure S3A). We have previously shown that mutating the GATA binding motifs in +7 and ETS binding motifs in +9 diminished endothelial activity and extinguished hematopoietic activity of the Eng -8P/lacZ/+7/+9 construct in transgenic assays. There was a corresponding reduction or failure of the mutant constructs to preferentially target cells with BL-CFC potential (Figure 2C (i)-(ii) and Figure S3B). The -8P/lacZ construct showed strong endothelial but no hematopoietic activity in transgenic assays. FLK1+ cells targeted by this construct (F+L+) showed significantly lower BL-CFC potential than F+L- cells (Figure 2D (i) and Figure S3A). For the -8P/lacZ/+7 and -8P/lacZ/+9 constructs, which showed strong endothelial and low- to moderate hematopoietic activity in in vivo transgenic assays, FLK1+ mesoderm (F+L+) had either lower or equivalent BL-CFC activity than FLK1+ (F+L-) cells (Figure 2D (ii)-(iii) and Figure S3A). It is important to note that the total number of BL-CFCs generated by FLK1+ mesoderm will vary from clone to clone but that comparisons are of BL-CFC potential of L+ and L- sorted FLK1+ cells from each clone.
Taken together, these data show that the GREs of ENG that showed increased chromatin accessibility as pre-hemangioblast mesoderm progressed to hemangioblast mesoderm (i.e. -8/P/+7/+9; Figure 1D) act collectively to target reporter gene expression to BL-CFCs in FLK1+ mesoderm. It is noteworthy that in in vivo transgenic assays, it was also this construct (-8/P/lacZ/+7/+9) that had the strongest and most specific activity in blood and endothelium in the developing embryo17.

Global transcriptomic analysis of FLK1 mesoderm targeted by Eng -8/P/LacZ/+7/+9 identifies genes associated with hemangioblast activity. To discover genes associated with the activation of these GREs and increased activity of the reporters, we performed RNA-sequencing on sorted lacZ+ and lacZ- cell fractions from three independent experiments. As expected Kdr (Flk1) expression was comparable in both fractions and Eng transcripts were increased in the lacZ+ fraction consistent with Eng GRE driven reporter activity (Figure 3A). There was also a shared set of genes that was consistently differentially expressed between the lacZ+ and lacZ- cell fractions (Figure 3B (i)) and included 107 up-regulated and 101 down-regulated genes. These included cell surface receptors and transcription factors known to be associated with blood and endothelium (fold change ≥ 2 and p-value <0.5; Table S1 and Figure S4A) and genes that have been associated with hemangioblast development in the LifeMap Sciences embryonic development compendium (Figure 3B (ii)). Individually the expression of many genes known to be associated with early mesoderm (e.g. Bmy/T and Bmp4), blood (e.g. Gata1 and Tal1) and endothelial (e.g. Foxc2 and Etv2) development did not vary significantly between these cell fractions (Figure 3B (ii)). Indeed, as hemangioblasts are a sub-population of FLK1+ mesoderm with multi-lineage differentiation potential, it would have been unusual to see significant differences in expression of individual genes that are strongly associated with commitment to a specific lineage. However, ingenuity pathway
analysis (IPA) revealed that when differentially expressed genes were considered as a collective, there were strong associations with blood and blood vessel development for genes in the FLK1+/lacZ+ set (Figure S4B). Consistent with these biological functions, this gene set also showed significant associations with signaling pathways that govern endothelial development and eNOS signaling (Figure S4C-F).

To investigate the \textit{in vivo} relevance of our gene set, we used GSEA analysis to compare expression overlaps with gene expression data from FLK1+ cells in ETV2\textsuperscript{25,26} and LDB1\textsuperscript{27} KO embryos, both of which are defective in hemangioblasts (Figure 3C). There were strong overlaps between genes expressed in FLK1+ mesoderm targeted by \textit{ENG} -8/P/lacZ/+7/+9 and genes expressed in hemangioblast competent WT ETV2 and LDB1 embryos compared to ETV2+/- or LDB1-/- embryos respectively. Therefore the molecular signature of BL-CFC enriched FLK1+ mesoderm that was identified using differential reporter activity of \textit{ENG} GREs is consistent with \textit{in vivo} functional capacity.

\textbf{The Eng promoter in combination with the -8 endothelial enhancer targets hemogenic endothelial cells enriched for hematopoietic potential.} Definitive hematopoiesis in the embryo progresses through a TIE2+/c-KIT+/CD41- hemogenic endothelial (HE) intermediate\textsuperscript{9}. We used a cell culture system that mirrors this \textit{in vivo} transition to investigate whether any of the reporter ESC lines preferentially targeted HE cells and whether they could be used to isolate cell fractions that were enriched for hematopoietic potential\textsuperscript{9}. To this end, FLK1+ cells were sorted from day 3EBs and seeded into liquid blast culture media (Figure 4A). At day 2 of culture, lacZ+ and lacZ- HE cells were isolated by FACS and re-seeded into liquid blast media (LBM) for two further days followed by flowcytometry and CFU-C assays to evaluate the hematopoietic potential of each fraction. Of the reporter ESC lines, \textit{Eng} -
8/P/lacZ was unique in that it was active in a fraction of HE cells that generated more TIE2-/CD41+ and CD45+ cells after 48 hours in culture (Figure 4B (i)- (ii)) and contained almost all CFU-C potential (Figure 4B (iii)). Whereas Eng -8/P/lacZ/+7/+9 targeted FLK1+ mesoderm with increased BL-CFC potential (Figure 2), it did not target HE cells with increased hematopoietic potential (Figure 4C (i) - (iii)). Indeed this and each of the other constructs, targeted HE cells that had lower hematopoietic potential (Figure 4D-E; Figure S5). Taken together, these data showed that not only was there a specific combination of ENG GREs that targeted HE cells but that the combination was distinct from that which targeted BL-CFCs in FLK1+ mesoderm.

Hematopoietic potential is highest in Eng -8/P/lacZ targeted HE cells that do not as yet express surface ENG. Cell fate transitions are dynamic and our purpose was to use these reporter constructs to capture HE cells that were intrinsically fated towards the hematopoietic lineage at the earliest possible time point in culture. Based on the assumption that there would be a delay between transactivating the Eng GRE reporter and surface expression of ENG, we repeated the experiments described in Figure 4 using the Eng -8/P/lacZ ES cell line but here also incorporating surface ENG expression to isolate TIE2+/C-KIT+/CD41- HE fractions that were ENG+/lacZ-, ENG+/lacZ+ or ENG-/lacZ+ (Figure 5A). Interestingly, CFU-C potential within the lacZ+ fraction was highest in ENG-/lacZ+ HE cells (Figure 5B (i) – (ii)). ENG+/lacZ-, ENG+/lacZ+ and ENG-/lacZ+ HE cells were re-seeded in LBM and analysed by flow cytometry and CFU-C assays after two further days of culture. The proportions of TIE2-/CD41+ (Figure 5C (i)) and CD45+ (Figure 5C (ii)) cells and CFU-C potential (Figure 5C (iii)) were highest for cultured ENG-/lacZ+ HE cells. Taken together, these data show that the hematopoietic potential within HE cells can be targeted by Eng -
and that these ESCs could be used to interrogate the earliest transcriptional changes associated with this cell fate decision.

Transcriptomic analysis of HE fractions identifies genes associated with hemogenic endothelial to hematopoietic transition. To discover genes that act on Eng -8/P/lacZ and drive reporter gene activity and by extension are associated with hemogenic potential in TIE2+/C-KIT+/CD41- HE cells, we performed RNA-sequencing on sorted ENG+/lacZ-, ENG+/lacZ+ and ENG-/lacZ+ HE cell fractions from three independent experiments. As expected the fractions, which expressed surface ENG had abundant Eng transcripts, which were still comparatively low in ENG-/lacZ+ HE cells (Figure 6A). Consistent with its role as a major determinant of endothelial to hematopoietic transition (EHT)9,24,28, Runx1 transcripts were abundant in HE cells that were enriched with functional hemogenic cells (ENG-/lacZ+ HE) and relatively low in those (ENG+/lacZ- and ENG+/lacZ+ HE) that were not. In total, there were 707 up-regulated and 981 down-regulated genes in ENG-/lacZ+ HE cells compared with ENG+/lacZ- and ENG+/lacZ+ HE cells (Figure 6B (i); Table S2). It was interesting to note that only a subset of genes that have previously been attributed to mark HE cells based on cell surface protein expression were16 differentially expressed between these functionally distinct HE sub-populations (Figure 6B (ii)). This does not imply that these genes are not important but that their higher or lower expression is not associated with these early subtle transitions.

To investigate the in vivo relevance of our gene set, we used GSEA analysis to compare expression overlaps between ENG-/lacZ+ HE vs. ENG+/lacZ- HE and gene sets generated from primary embryonic endothelial cell (EC), hemogenic endothelial cells (HECs) and HSCs16. Consistent with our functional data, the gene sets associated with EC to HE
transition (Figure 6C (i) and HE to HSC transition (Figure 6C (ii)) showed strong overlaps with genes expressed in ENG-\(\text{lacZ}\) HE. Gene sets associated with HIF1a and DNA replication also showed strong overlaps with genes expressed in ENG-\(\text{lacZ}\) HE cells (Figure S6A). Genes that were UP in ENG-\(\text{lacZ}\) HE compared with ENG+\(\text{lacZ}\) HE cells feature prominently in IPA reconstructions of gene networks governing hematopoietic development (Figure S6B). Whereas genes that were UP in ENG-\(\text{lacZ}\) HE cells compared with either ENG+\(\text{lacZ}\) HE or ENG+/\(\text{lacZ}\) HE cells were associated with biological processes relating to blood development, genes that were DOWN in ENG-\(\text{lacZ}\) HE cells relative to the other two fractions were associated more with angiogenesis or vasculogenesis (Figure S6C-E). Interrogation of differentially expressed transcription factors (TFs) and cell surface receptors (CSRs) in the more functionally hemogenic ENG-\(\text{lacZ}\) HE fraction relative to the ENG+/\(\text{lacZ}\) HE fraction showed up-regulation of a number of TFs (e.g. \textit{Runx1} \textsuperscript{29}, \textit{Myb} \textsuperscript{30}, \textit{Gfi1b} \textsuperscript{31} etc.) and CSRs (\textit{Lgr5} \textsuperscript{32}) that are known to play a role in HSC development and down regulation of others (e.g. \textit{Sox17} \textsuperscript{33}), which are important for HE to HSC transition (Figure 6D).

\textbf{Lrp2 is required for normal blood emergence in the zebrafish aorta.} We then overlapped our gene sets to visualise associations between genes that were UP or DOWN in haemogenic mesoderm (HB) and/or hemogenic endothelium (HE) (Figure S7 and S8; Table S1-S2) to interrogate their function. Six genes were shared between the up-regulated groups (Figure 7A; Figure S7C) and eight genes between the down-regulated groups (Figure S8A). Genes that were DOWN in both HB and HE cells included several with no known association with haematopoiesis (Figure S8B). However, we focused on genes that were UP in both HB and HE cells (Figure 7A and S7C) for practical considerations given that their expression and functional role would be easier to validate. This group included hematopoietic transcription
factors (Gfi1\textsuperscript{31} and Lyl1\textsuperscript{34}), a platelet protein kinase C substrate (Plek\textsuperscript{35}) and granulocyte lysosomal and lysosomal membrane proteins (Mpo\textsuperscript{36} and Laptm5\textsuperscript{37}), all of which have known functions in the hematopoietic system. It also included a multifunctional ligand (Lrp2) with no previously described role in blood or blood development. Lrp2/Megalin is a member of an endocytic receptor complex that is involved in maternal-fetal transport of folate and other nutrients, lipids and morphogens such as sonic hedgehog (Shh) and retinoids\textsuperscript{38}. Given these associations we postulated that Lrp2 up-regulation in blood precursors was likely to be of functional significance.

The Ly6aGFP (Sca1) mouse model, in which all HSCs throughout development are GFP\textsuperscript{+}\textsuperscript{14,39} has facilitated the study of EHT. There mice were used to show in real-time, the transition of morphologically flat endothelial GFP+ cells in the E10.5 aorta to round GFP+ cells that co-express other HSC markers\textsuperscript{40}. Given that LRP2 was up-regulated in HE cells, we evaluated LRP2 expression in Ly6aGFP E10.5 AGM. LRP2 shows specific expression in endothelial cells with strong expression in Ly6aGFP+ endothelial cells and hematopoietic clusters (Figure 7B).

EHT is an evolutionarily conserved process in vertebrates and real-time imaging of transgenic zebrafish embryos has also shown the transition of aortic endothelial cells to hematopoietic cells\textsuperscript{41,42}. Lrp2 is highly conserved across different vertebrate species (Figure 7C; average sequence identity across all species shown = 70 %). The zebrafish genome has two closely related protein-coding genes, lrp2a on chr. 9 and lrp2b on chr. 12, both of which are expressed at 24-72hpf \textsuperscript{43}. To validate the involvement of LRP2 in HSC generation, we used a zebrafish morpholino oligo (MO) knockdown approach targeting both lrp2a and 2b together and each alone. At 36h post fertilization (hpf), morphants were assayed by ISH for
*cmyb* and *runx1*, markers for emerging blood progenitors in the aorta\(^4\). WT embryos showed robust *cmyb* and *runx1* expressing cells along the dorsal aorta in contrast to *lrp2a/b* morphants that showed severe reductions (Figure 7D (i); Figure S9A). There was partial rescue of AGM blood progenitors when *lrp2a/b* morphants were co-injected with *hLRP2* mRNA. The partial rescue was probably due to only partial homology of protein sequences between humans and fish (~65%) and quality of *in vitro* transcribed mRNA given the large size of *LRP2* cDNA (~14kb). To exclude non-specific toxicity related loss of *cmyb* and *runx1* expressing cells, we co-injected *lrp2a/b* MO with *tp53* MO and saw no restitution of *cmyb* expressing cells in the morphants (Figure S9B). Injection of *lrp2bMO* but not *lrp2aMO*, reduced the numbers of *cmyb* expressing AGM blood progenitors (*lrp2bMO*; Figure 7D (ii); *lrp2a*; data not shown). To establish that this defect in blood cell production was not a secondary to loss of vascular integrity, we injected *lrp2a/b* MO into flk:zsgreen transgenic embryos and saw no difference between morphants and controls at 32 hpf (Figure 7E). In addition to vascular integrity we also assessed blood flow in morphants. Both heart function as well as blood flow was indistinguishable from control embryos (data not shown). Taken together these data support a role for LRP2 during AGM hematopoiesis.
DISCUSSION

Regulatory elements of genes that demonstrate tissue specific expression have previously been used to target and characterise various cell populations in ESC systems\textsuperscript{45-48}. They have also been used in conjunction with cell surface markers to isolate cell fractions at specific stages during embryonic development for the \textit{de novo} discovery of regulators of hematopoiesis (e.g. Runx\textsuperscript{13} and Ly6a (Sca1)\textsuperscript{16}). These studies have in the main utilized reporter gene knock-in ESCs that disrupted a functional allele or transgenic mouse lines derived by pro-nuclear injections, which were subject to copy number and/or position effect variegation. Here we used a method to mitigate these caveats and leveraged prior knowledge that promoter/enhancer combinations of ENG specifically targeted endothelium or both blood and endothelium in the developing embryo\textsuperscript{17,22} to interrogate the transcriptomes of hemogenic progenitors.

When initiating these experiments, we did not envisage that distinct combinations of ENG promoter/enhancers would target haemangioblast potential in FLK1 mesoderm and hemogenic potential in TIE2+/C-KIT+/CD41- hemogenic endothelium. In retrospect, given the distinct transcriptomes and functional properties of haemangioblasts and hemogenic endothelium, this should not have come as a surprise. Nor did we predict that hemogenic potential would be enriched in ENG- HE1 cells targeted by the ENG\textsuperscript{-8PlacZ} transgene. Given that F+L+ cells expressed higher levels of ENG (Figure 3A), these data raise the question whether HE1 cells emerge from F+L- ENG low cells which are less able than their F+L+ counterparts to generate BL-CFCs or whether F+L+ ENG high cells subsequently shutdown ENG expression to facilitate their haemogenic potential in HE1 cells. As these populations were targeted by different transgenes (F+L+; Eng-8/P/lacZ/+7/+9 and ENG-L+ HE; Eng-8/P/lacZ), this could not be directly tested. However, ES cells targeted with dual
reporters each driven by either \textit{Eng-8/P/+7/+9} or \textit{Eng-8/P} may assist in addressing this specific question.

\textit{Lrp2}, a gene that encodes megalin, a multiligand uptake receptor that regulates circulating levels of diverse compounds\textsuperscript{49} emerged as a novel regulator of hematopoiesis. Mutations in LRP2 result in impaired neuro-epithelial development and are causative of Donnai-Barrow and facio-oculo-acoustico-renal syndromes\textsuperscript{50}. It has been implicated in balancing BMP4 and SHH signaling in neuro-epithelium by acting as a clearance receptor for BMP4 and by concentrating or depleting SHH by ligand recycling or clearance respectively in a cell type and context dependent manner\textsuperscript{51}. This is of mechanistic interest as the BMP4-SHH gradient between the neural tube and dorsal aorta has also been implicated in the induction of the HSC developmental program in the ventral wall of the dorsal aorta\textsuperscript{52}. On a C57Bl/6N background the LRP2 mutation causes lethality in mice around the time of birth and there are no mutant pups although embryo collections at all embryonic stages to E18.5 show expected Mendelian ratios. LRP2 mutations on 129 or CD1 backgrounds also do not yield survivors (Hammes et al, unpublished data). On a FVB/N background however, LRP2 null mice are viable with neural tube defects and this receptor has previously been implicated in folate endocytosis in the developing neural tube\textsuperscript{53}. However, peripheral blood and bone marrow hematopoietic stem and progenitor cell numbers were comparable in FVB/N wild type and mutant adult mice at 6-9 months of age (Figure S10). A more detailed analysis of embryonic hematopoiesis in mutant mice on both C57Bl/6N and FVB/N backgrounds will be required to establish whether the numbers of emergent HSCs differ at various time points and the identity of any modifier genes in FVB/N that compensate for the loss of LRP2 and these investigations are ongoing. However, taken together with the zebrafish data, which shows reduction rather than loss of HSCs, LRP2 is likely to facilitate (or provide a fitness
advantage) rather than be absolutely required for EHT. Indeed, it is important to keep in mind that \textit{Lrp2} transcripts were higher in HB cells with greater BL-CFC potential and HE cells with greater CFU-C potential but cells with lower numbers of transcripts were also able to generate BL-CFCs and CFU-Cs.

Deficiency of dietary folate also results in impaired neural tube development and megaloblastic anemia\textsuperscript{54}. Targeted inactivation of the reduced folate carrier (RFC1), which facilitates folate delivery into cells results in embryonic lethality at E10.5 due to neural and hematopoietic defects\textsuperscript{55} and components of the Megalin complex are amongst the most significantly disrupted genes in null embryos\textsuperscript{38}. Coordinated up-regulation of a receptor that facilitates folate uptake in hemogenic endothelial cells would be consistent with demand for an essential hematinic in cells that are on the threshold of a replicative phase.

Although we focused our attention on \textit{Lrp2}, as a gene without a described role in hematopoiesis, from a list of six that were up-regulated in both hemangioblasts and HE cells, there were other genes that were UP in only one or the other cell fraction. Given the overlap of these gene sets with those generated from gene knockout embryos that showed stage specific developmental defects or from sorted primary cells harvested from embryos, they will serve as a rich resource to explore and manipulate the emergence of hemangioblasts from FLK1+ mesoderm or definitive hematopoiesis in HE cells. Insights gained from these manipulations will in turn inform tissue regeneration protocols that aim to generate functional HSCs.
ACKNOWLEDGEMENTS

This work was supported by grants from the National Health and Medical Research Council of Australia, Australian Research Council and the Dr Tom Bee Stem Cell Research Fund to JEP, Cancer Research UK to VK and GL, the BBSRC, Leukaemia and Lymphoma Research, The Leukaemia and Lymphoma Society, Cancer Research UK, and core support grants by the Wellcome Trust to the Cambridge Institute for Medical Research and Wellcome Trust - MRC Cambridge Stem Cell Institute to BG.

AUTHOR CONTRIBUTIONS


DISCLOSURES

No conflicts of interest to declare.
REFERENCES


FIGURE LEGENDS

FIGURE 1. Mesoderm to hemangioblast transition is accompanied by increased Eng expression and chromatin accessibility at hemato-endothelial regulatory elements. (A) Schematic representation of the ENG locus. The transcription start site is marked with an arrow. The -8kb, +7kb and +9kb enhancers and the promoter (P) are marked in orange, exons are marked in brown and the 5’UTR in cyan. (B) Schematic representation of Bry-GFP ES cell differentiation. At day 3 of EB differentiation, Bry-GFP+/FLK1- (G+F-) and Bry-GFP+/FLK1+ (G+F+) cells were sorted and analysed by RT-PCR and ChIP. (C) Bar graph showing Eng mRNA expression levels in sorted FLK1+ve and –ve mesodermal cell populations in day 3 EBs generated from Bry-GFP ES cells. (D) A bar graph showing levels of enrichment of the active chromatin mark, H3K9Ac at Eng -8, P, +7 and +9 hematoendothelial enhancers relative to IgG in pre-hemangioblast mesoderm (G+F-; black) and in hemangioblast mesoderm (G+F+; gray). Eng -4 was included as a negative control region. **; P < 0.01, ***; p < 0.001

FIGURE 2. The Eng promoter when combined with the -8, +7 and +9 haematoendothelial enhancers targets FLK1+ mesodermal cells enriched for BL-CFC potential. (A) Schematic representation of the experimental procedure. The Eng -8/P/LacZ, Eng -8/P/LacZ/+7, Eng -8/P/LacZ/+9, Eng -8/P/LacZ/+7/+9, Eng -8/P/LacZ/+7Δ/+9, and Eng -8/P/LacZ/+7Δ/+9Δ reporter constructs were introduced by homologous recombination into the HPRT locus of HM1 ES cells. Recombinant clones were differentiated into day 3 EBs and stained for FLK1 expression and β-galactosidase activity. FLK1+/LacZ- (F+/L-; gray) and FLK1+/LacZ+ (F+/L+; blue) cells were sorted and seeded into BL-CFC assays. Fractions sorted from the Eng -8/P/LacZ/+7/+9 were further analysed by RNA-sequencing. (B) Flow cytometry profiles of Eng -8/P/LacZ/+7/+9 day 3 EBs (left). BL-CFCs from sorted F+/L-
(gray) and F+/L+ (blue) fractions. (C) (i) Flow cytometry profile of day 3 EBs derived from ES cells targeted with $Eng \,-8/P/LacZ^{+7Δ/+9} \,Δ$ (mutated GATA motifs in the +7 enhancer) is shown to the left with corresponding BL-CFCs from sorted F+/L- (gray) and F+/L+ (blue) fractions shown to the right. (ii) Flow cytometry profile of day 3 EBs derived from ES cells targeted with $Eng \,-8/P/LacZ^{+7Δ/+9Δ}$ (mutated GATA motifs in the +7 enhancer and mutated ETS motifs in the +9 enhancer) and corresponding BL-CFCs from sorted F+/L- (gray) and F+/L+ (blue) fractions. (D) Flow cytometry profiles of day 3 EBs and BL-CFCs from sorted F+/L- (gray) and F+/L+ (blue) fractions are shown for ES cells targeted with (i) $Eng \,-8/P/LacZ$ (ii) $Eng \,-8/P/LacZ^{+9}$ and (iii) $Eng \,-8/P/LacZ^{+7}$. BL-CFC counts are the total number of blast colonies generated from $2 \times 10^4$ seeded cells. Statistical analysis was using students T-test, *p<0.05, **p<0.01.

FIGURE 3. RNA-sequencing of FLK1 mesoderm targeted by $Eng \,-8/P/LacZ^{+7/+9}$ identifies genes associated with hemangioblast activity. (A) RNA-sequencing profiles showing $Kdr$ ($Flk1$) transcripts (top panel) and $Eng$ transcripts (bottom panel) in the F+/L- and F+/L+ fractions. FPKM expression values are shown to the right. (B) (i) Heat map representation of up- and down-regulated genes in FLK1+/LacZ- (F+/L-) and FLK1+/LacZ+ (F+/L+) fractions in three independent experiments. (ii) Expression (FPKM values) levels of genes that have previously been associated with hemangioblast function. The left panel shows a subset of genes that are differentially expressed between F+/L- and F+/L+ fractions and the right panel shows a subset of genes that are not. (C) GSEA profiles showing the correspondence of genes that are differentially expressed between F+/L- and F+/L+ fractions and those that are differentially expressed in ETV2+/+ vs. ETV2/- (top panel) and LDB wt vs. LDB -/- gene sets. DEG; differentially expressed genes.
FIGURE 4. The Eng promoter in combination with the -8 endothelial enhancer targets hemogenic endothelial cells enriched for hematopoietic potential. (A) Schematic diagram outlining the experimental procedure. Recombinant ES cells generated using the Eng reporter constructs were differentiated into day 3 EBs. FLK1+ mesodermal cells were sorted from representative clones for each recombinant ES cell line and cultured in liquid blast media. At 48 hours, CD41-/TIE2+/c-KIT+ (HE) cells were sorted into lacZ+ and lacZ− fractions. The sorted fractions were re-cultured in liquid blast media for a further 48 hours followed by flow cytometry and CFU-C assays. (B) (i) CD41 and TIE2 expression in sorted c-KIT+ HE LacZ- (white) and LacZ+ (blue) fractions (after 2 days of re-culture) derived from Eng -8/P/LacZ ES cells. (ii) Percentage of CD45+ cells generated from LacZ- and LacZ+ HE fractions. (iii) Bar chart showing the number and type of hematopoietic colonies generated by each fraction. (C) (i)-(iii) Corresponding data to (B) generated from Eng -8/P/LacZ/+7/+9 ES cells. (D) (i)-(iii) Corresponding data to (B) generated from Eng -8/P/LacZ/+7 ES cells. (E) (i)-(iii) Corresponding data to (B) generated from Eng -8/P/LacZ/+9 ES cells. Primitive and definitive colonies were scored after four and nine days respectively. Statistical analysis was using students T-test, *; p<0.05, **; p<0.01.

FIGURE 5. Hematopoietic potential is highest in Eng -8/P/lacZ targeted HE cells that do not express surface ENG. (A) Schematic diagram outlining the experimental procedure. FLK1+ mesodermal cells were sorted from day 3 EBs generated from the Eng -8/P/LacZ recombinant ES cell line and cultured in liquid blast culture media. At 48 hours, CD41-/TIE2+/c-KIT+ (HE) cells were sorted into ENG+/LacZ-, ENG+/LacZ+, and ENG-/LacZ+ fractions. These fractions were either directly seeded into CFU-C assays (B) or re-cultured in liquid blast media for a further 48 hours and analysed by flow cytometry and CFU-C assays (C). (B) (i)-(ii) Flow cytometry to show the frequencies of CD41-/TIE2+/c-KIT+ (HE) cells
in ENG+/lacZ+, ENG+/lacZ- and ENG-lacZ+ fractions. (iii) CFU-C potential of each sorted fractions in (i). (C) (i) Flow cytometry analysis of CD41 and TIE2 expression in sorted HE cell fractions after two days of re-culture in liquid blast media. (ii) Bar chart showing the percentage of CD45 positive cells in sorted fraction. (iii) Bar chart showing hematopoietic colony numbers from each fraction. Primitive and definitive colonies were scored after four and nine days respectively. Statistical analysis was using student T-test, *; p < 0.05, **; p <0.01 and ***; p <0.001.

FIGURE 6. Transcriptomic analysis of HE fractions identifies genes associated with hemogenic endothelial to hematopoietic transition. (A) RNA-sequencing profiles showing Eng transcripts (top panel) and Runx1 transcripts (bottom panel) in the E+/L-, E+/L+ and E-/L+ fractions. FPKM expression values are shown to the right. (B) (i) Heat map representation of up- and down-regulated genes in ENG+/LacZ- (E+/L-) HE, ENG+/LacZ+ (E+/L+) HE and ENG-/LacZ+ (E-/L+) HE fractions in three independent experiments. (ii) Expression (FPKM values) levels of genes that have previously been associated with hemogenic endothelium. The top panel shows a subset of genes that are differentially expressed between E+/L-, E+/L+ and E-/L+ fractions and the bottom panel shows a subset of genes that are not. (C) GSEA profiles showing the correspondence of genes that are differentially expressed between the E+/L-, E+/L+ and E-/L+ fractions and those that are differentially expressed in endothelial cells (EC) vs. hemogenic endothelial cells (HEC) (top panel) and HECs vs. hematopoietic stem cells (HSC) gene sets. (D) Transcription factors and cell surface receptors that are up- and down regulated in the Eng -8/P E-/L+ HE fraction. The log fold changes (logFC) and log false discovery rates (logFDR) are listed for each gene. DEG; differentially expressed genes.
FIGURE 7. Lrp2 is required for normal definitive hematopoiesis. (A) Venn diagram showing the overlap of genes that are UP in FLK1 mesoderm enriched for BL-CFCs and/or HE cells enriched for hemogenic potential. (B) Immunohistochemistry of E10.5 Ly6aGFP AGM shows co-expression of GFP and LRP2 in endothelial cells and hematopoietic clusters. The insets show the same sections at low magnification. (C) Homology relationships of zebrafish lrp2a and lrp2b coding sequences with that of Lrp2 in different vertebrate species. (D) ISH for the HSC marker cmyb in zebrafish at 36 hpf. (i) Low (left-side panels) and high (right-side panels) magnification images of control zebrafish (top row), lrp2 a/b morpholinos (middle row) and lrp2a/b morpholinos co-injected with hLRP2 mRNA (bottom row) zebrafish. (ii) Low (left-side panels) and high (right-side panels) magnification images of control zebrafish (top row) and lrp2b morpholinos. (E) Confocal images of flk:zsgreen reporter embryos show an intact vasculature in both control (upper panel) and lrp2a/b morphant (lower panel) embryos. DA; dorsal aorta, NC; notochord, NT; neural tube.
Figure 1

A. Genomic map showing the location of the ENG gene on the chromosome. The region from -8 to +9 is shown with regulatory elements indicated.

B. Schematic diagram showing the experimental setup with Bry-GFP ES cells producing Day 3 EBs.

C. Bar graph showing the expression of Eng relative to b-Actin in G+F- and G+F+ conditions.

D. Graph showing H3K9Ac enrichment relative to IgG for different promoter regions.

Legend:
-8, -4, P, +7, +9: Distances from the ENG gene.
G+F-, G+F+: Treatment conditions.
Figure 2
**Figure 3**

### A

- **FLK1+/lacZ- (F+L-)**
- **FLK1+/lacZ+ (F+L+)**

**Kdr (Flk1) FPKM**

**Eng FPKM**

### B

**Eng-8/P/LacZ+/7+/9**

**i**

<table>
<thead>
<tr>
<th></th>
<th>F+/L-</th>
<th>F+/L+</th>
</tr>
</thead>
<tbody>
<tr>
<td>E1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**ii**

- Genes associated with hemangioblast function:
  - CD34
  - Hoxb5
  - Fli1
  - Sox18
  - Sox7
  - Cdh5
  - Pecam1
  - Esam
  - Flt4
  - Tek

- **FPKMs**
  - Gata1
  - Sox17
  - FoxC2
  - Notch1
  - Gata2
  - Tal1
  - Bmp4
  - Brx
  - Etv2
  - Ldb1

### C

- **Etv2^{+/−} vs. Etv2^{−−}**
  - NES = 6.89
  - P < 0.001

- **Ldb1 WT vs. Ldb1^{−−}**
  - NES = 3.36
  - P < 0.001
Figure 4
Figure 5

**A**

- **HM ES+ Eng GRE**
- **Differentiation (-LIF)**
  - Day 3 EBs
  - FACS
  - FLK1+
- **Differentiation (48hrs in LBM)**
  - HE (TIE2+/- c-KIT+/CD41-)
  - FACS
  - ENG+/ixZ- HE
  - ENG+ixZ+ HE
  - ENG-ixZ+ HE
- **Differentiation (48hrs in LBM)**
  - CFU-C assays
  - RNA-sequencing
  - Flowcytometry
  - CFU-C assays

**B**

- **(i) CD41-ve**
  - CD41-/E+/L-
  - CD41-/E+/L+
  - CD41-/E-/L+
- **(ii) CD41/E+/L-**
  - CD41/E+/L+
  - CD41/E-/L+
- **(iii) CFU-C**
  - **CD41/E+/L-**
  - **CD41/E+/L+**
  - **CD41/E-/L+**
- **Colony (x 1000 cells)**
  - Primitive
  - Definitive
  - Mixed

**C**

- **(i) CD41/E+/L-**
  - CD41/E+/L+
  - CD41/E-/L+
- **(ii) CD45**
  - % positive cells
  - E+L-
  - E+L+
  - E-L+
- **(iii) CFU-C**
  - Colony (x 1000 cells)
  - Primitive
  - Definitive
  - Mixed