Title
Expression of the Medial HOXA genes is indispensible for self-renewal in human hematopoietic stem cells

Author
Dou, Diana Remy

Publication Date
2017

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Expression of the Medial HOXA genes is indispensible for self-renewal in human hematopoietic stem cells

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Molecular Biology

by

Diana Remy Dou

2017

ABSTRACT OF THE DISSERTATION

Expression of the Medial HOXA genes is indispensible for self-renewal in human hematopoietic stem cells

by

Diana Remy Dou

Doctor of Philosophy in Molecular Biology

University of California, Los Angeles, 2017

Professor Hanna K.A. Mikkola, Chair

Hematopoietic stem cells (HSCs) are able to self-renew, generate all lineages of the blood, reconstitute the hematopoietic system upon bone marrow (BM) transplantation, and thereby cure diseases of the blood and immune system. BM transplantations have successfully treated a wide range of blood-related diseases (i.e., β-thalassemia, leukemia, and HIV/AIDS). However, current demand for BM transplantations outstrips the quantity of HSCs available through cord blood banks and BM registries, and further requires finding matching HLA-types between patients and donors. Given these restrictions in quantity and compatibility, there is significant interest in developing alternative sources of transplantable HSCs by differentiating HSCs *in vitro* from pluripotent stem cells (PSCs), such as embryonic stem cells (ESCs). Unfortunately, lack of understanding of the regulatory mechanisms governing HSC function has prevented their *in vitro* generation.

The focus of my thesis research is on deciphering the molecular blocks preventing the generation of HSCs from ESCs, and identifying key molecular cues to help overcome these blocks. In the course of my research, I identified the medial HOXA genes as critical regulators for human HSC self-renewal, and explored the pathways that regulate the HOXA genes that are dysregulated during ESC differentiation cultures.

Our lab developed a two-step differentiation protocol mimicking the environments a developing HSC would encounter in the embryo and successfully generated cells with the human HSC immunophenotype, $CD34^{+}CD38^{-}CD45^{+}CD90^{+}GPI80^{+}$. However, although ESC-derived cells differentiated into definitive erythroid, myeloid, and T-cells, they displayed impaired self-renewal and engraftment ability compared to hematopoietic stem/progenitor cells (HSPCs) isolated from the fetal liver (FL), the site of the most active HSC expansion and differentiation during human development. Microarray analysis witnessed the overall close molecular correlation between ESC-HSPCs and FL-HSPCs, but revealed the lack of expression of the medial HOXA genes in ESC-derived cells as a critical defect preventing their self-renewal. Knockdown of HOXA5 or HOXA7 in FL-HSPCs recapitulated the self-renewal defects observed in ESC-HSPCs. A six-day stimulation of the retinoic acid (RA) signaling pathway was sufficient to significantly upregulate HOXA gene expression in ESC-HSPCs at day 6. However, HOXA gene expression was not sustained long-term.

Microarray and ChIP data further revealed that the regulator of HOXA genes, MLL1, was expressed, but not bound to HOXA genes in ESC-derived cells, suggesting the inability to recruit MLL1 to HOXA genes prevents their activation during ESC differentiation. I identified in the human genome a region of 89% sequence homology to the mouse lincRNA Mistral, which in mice recruits MLL1 to HOXA6/HOXA7 genes. RNA-seq and microarray data showed high

expression of putative human MISTRAL in FL-HSCs but not in their differentiated progeny or in ESC-derived hematopoietic cells, and high correlation of HOXA6/A7 expression to Mistral expression in multiple cell types. These data identify insufficiency of HOXA gene expression is a developmental barrier for generating HSCs from pluripotent cells and suggest RA-signaling is a major inductive signal and MISTRAL a novel component of the fetal HSC regulatory machinery that conveys "stemness" in hematopoietic cells.

The dissertation of Diana Remy Dou is approved.

Gay M. Crooks

William E. Lowry

Kathrin Plath

Owen N. Witte

Hanna K.A. Mikkola, Committee Chair

University of California, Los Angeles

2017

DEDICATION

This dissertation is dedicated to my parents, Jim H. Dou and Christina F. Yueh, and my extended family and friends, whose unfaltering love, support and encouragement made the pursuit of my dreams a reality.

I would also like to thank the exceptional mentors who have guided me in my endeavors, in particular Dr. Hanna Mikkola, whose boundless enthusiasm for advancing scientific research and untiring dedication to the training of scientists will serve as inspiration throughout my career.

TABLE OF CONTENTS

Preliminary pages:

Dissertation Research:

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AGM	Aorta-gonado-mesonephros region
ATAC	Assay for transposase-accessible chromatin
ChIP	Chromatin immunoprecipitation
EB	Embryoid body
EHT	Endothelial-to-hematopoietic transition
ESC	Embryonic stem cell
FL	Fetal liver
HE	Hemogenic Endothelium
HSC	Hematopoietic stem cell
HSPC	Hematopoietic stem/ progenitor cell
lncRNA	long noncoding RNA
PCR	Polymerase chain reaction
PL	Placenta
PSC	Pluripotent stem cell
RA	Retinoic acid
RACE	Rapid amplification of cDNA ends
RALDH	Retinoic aldehyde dehydrogenase
RAR	Retionic acid receptor
TSS	Transcription start site
UTR	Untranslated region
YS	Yolk sac

Acknowledgements

Chapter 1 is a version of the invited review for Experimental Hematology currently in preparation.

Chapter 2 was originally published in *Nature Cell Biology*. **DR Dou**§, V Calvanese§, MI Sierra, A Nguyen, A Minasian, P Saarikoski, R Sasidharan, JA Zack, GM Crooks, Z Galic, HKA Mikkola. "Medial HOXA genes demarcate haematopoietic stem cell fate during human development." Nature Cell Biology. 2016 Jun;18(6):595-606. doi:10.1038/ncb3354. Epub 2016 May 16. PMID:27183470. §Denotes equal contribution

Chapter 3 is a version of **DR Dou,** V Calvanese, T Org, B Mishra, P Ernst, Z Galic, HKA Mikkola. A non-coding transcript in the human HOXA7 3' UTR is correlated with medial HOXA gene expression and HSC stemness. (In preparation)

Appendix is a version of the protocol originally published in the *Protocol Exchange*. **DR Dou**, V Calvanese, P Saarikoski, Z Galic, HKA Mikkola. "Induction of HOXA genes in hESC-derived HSPC by two-step differentiation and RA signalling pulse. Protocol Exchange. doi:10.1038/protex.2016.035

I would like to thank the Mikkola and the Lowry labs for discussions and help with experiments. I would like to thank the UCLA BSCRC Flow Cytometry Core for their assistance in FACS sorting. I would like to thank the UCLA Clinical Pathology Microarray Core and BSCRC

Sequencing Core for their assistance in microarray and RNA-sequencing. I would like to thank W. E. Lowry for discussions. I would like to thank A. Hossain and T. Johnson for advice for RACE experiments. I would like to thank T. Soyanova, D. Johnson and O. Witte for help with NSG mice. I would like to thank A. Montel-Hagen for help with the figures of the introduction.

I would like to acknowledge our funding sources. This work was funded by CIRM RN1-00557 and RT3-07763 and NIH RO1 DK100959. H.K.A.M. was supported by the Rose Hills Foundation Scholar Award and the Leukemia & Lymphoma Society Scholar Award (20103778). D.R.D was supported by the Ruth L. Kirschstein National Research Service Award GM007185 Training Grant and the National Science Foundation Graduate Research Fellowship Program (NSF GRFP).

Biographical Sketch:

Diana Remy Dou

EDUCATION

10/2006 – 06/2010 B.S. with honors in the majors: Biology and Business, Economics and Management, and minor: English at the California Institute of Technology (Caltech), Pasadena USA

PUBLICATIONS

DR Dou§, V Calvanese§, MI Sierra, A Nguyen, A Minasian, P Saarikoski, R Sasidharan, JA Zack, GM Crooks, Z Galic, HKA Mikkola. "Medial HOXA genes demarcate haematopoietic stem cell fate during human development." Nature Cell Biology. 2016 Jun;18(6):595-606. doi:10.1038/ncb3354. Epub 2016 May 16. PMID:27183470. §Denotes equal contribution

DR Dou, V Calvanese, P Saarikoski, Z Galic, HKA Mikkola. "Induction of HOXA genes in hESC-derived HSPC by two-step differentiation and RA signalling pulse. Protocol Exchange. doi:10.1038/protex.2016.035

T Petersen, YJ Lee, A Osinusi, VK Amorosa, C Wang, M Kang, R Matining, X Zhang, **D Dou**, T Umbelia, S Kottilil, MG Peters. "Interferon Stimulated Gene Expression in HIV/HCV Coinfected Patients Treated with Nitazoxanide/Peginterferon-Alfa-2a and Ribavarin. AIDS Research and Human Retroviruses. 2016 Jul;32(7):660-7. doi:10.1089/aid.2015.0236. PMID:26974581

A Katsounas, JJ Hubbard, CH Wang, **D Dou**, B Shivakumar, S Winter, RA Lempicki, H Masur, A Polis, S Kottilil, A Osinusi. "High Inteferon-stimulated gene ISIG-15 expression affects HCF treatment outcome in patients co-infected with HIV and HCV." Journal of Medical Virology. 2013 Jun;85(6):959-63. doi:10.1002/jmv.23576. PMID:23588721

FELLOWSHIPS & HONORS

2016 UCLA Molecular Biology Institute Dissertation Year Award

2015 UCLA Molecular Cell and Developmental Biology Department Retreat Poster Award

2015 ISEH New Investigator Dirk van Bekkum 1st Place Graduate Student Award

2014 UCLA Molecular Biology Interdisciplinary Doctoral Program Retreat Poster Award

2013 National Science Foundation Graduate Research Fellowship Program (NSF-GRFP)

2012 Ruth L. Kirschstein National Research Service Award GM007185 Training Grant

2012 American Society of Hematology 54th Annual Meeting Abstract Achievement Award

NATIONAL AND INTERNATIONAL CONFERENCE PRESENTATIONS

Dou DR*, Calvanese V, Sierra MI, Sasidharan R, Zack JA, Crooks GM, Galic Z, Mikkola HKA, "Medial HOXA gene expression is required for establishing "stemness" in human HSCs." Talk presented at the ISEH Society for Hematology and Stem Cells 44th Annual Scientific Meeting, Kyoto, Japan, September 2015.

Dou DR*, Calvanese V, Sierra MI, Minasian A, Sasidharan R, Saarikoski P, Org T, Mishra B, Ernst P, Zack JA, Crooks GM, Galic Z, Mikkola HKA, "Activation of medial HOXA genes is required for establishment of "stemness" in human HSCs." Poster presented at the 13th Annual Meeting of the International Society for Stem Cell Research, Stockholm, June 2015.

Dou DR*, Calvanese V, Sierra MI, Sasidharan R, Org T, Mishra B, Ernst P, Galic Z, Mikkola HKA, "Medial HOXA cluster and associated lncRNAs as indicators of "stemness" during human HSC development." Poster presented at the Keystone Symposia on Molecular and Cellular Biology: Long Noncoding RNAs: From Evolution to Function (C7), Keystone, CO, March 2015.

Dou DR*, Minasian A, Sierra MI, Saarikoski P, Prashad SL, Magnusson M, Xu, J, Orkin SH, Zack JA, Crooks GM, Galic Z, Mikkola HKA, "Inability to express HOXA cluster genes compromises self-renewal of hESC-derived hematopoietic cells." Poster presented at the 11th Annual Meeting of the International Society for Stem Cell Research, Boston, MA, June 2013.

Dou DR*, Minasian A, Sierra MI, Saarikoski P, Prashad SL, Magnusson M, Xu, J, Orkin SH, Zack JA, Crooks GM, Galic Z, Mikkola HKA, "Inability to express HOXA cluster and BCL11A genes compromises self-renewal and multipotency of hESC-derived hematopoietic cells." Poster presented at the Annual Meeting of the American Society of Hematology, Atlanta, GA, December 2012.

Dou DR*, Minasian A, Sierra MI, Saarikoski P, Prashad SL, Magnusson M, Xu, J, Orkin SH, Zack JA, Crooks GM, Galic Z, Mikkola HKA, "Deficiency of HOXA gene and BCL11A expression compromises self-renewal and multipotency of hESC-derived hematopoietic cells." Poster presented at the ISEH Society for Hematology and Stem Cells 41st Annual Scientific Meeting, Amsterdam, Netherlands, August 2012.

WORK & EXTRACURRICULARS

Teaching Assistant for Developmental Biology, UCLA Apr – Jun 2014
Led three weekly recitation sessions to discuss developmental biology publications and trained students to critically analyze and present journal articles.

Teaching Assistant for Human Genetics, UCLA Jan – Mar 2013
Led two weekly recitation sessions to discuss human genetics publications, prompted student discussion, and helped write exam questions.

Chapter 1:

Introduction

PREFACE

Hematopoietic stem cells (HSCs) are self-renewing cells that are the foundation of blood, and the active component in bone marrow transplantation. However, the demand for bone marrow transplants exceeds HSCs available through cord blood banks and bone marrow registries, and is complicated by the necessity of proper HLA match between patients and donors to avoid graft-versus-host disease. As such, there is significant motivation to develop methods that will either expand cord blood or bone marrow HSCs or generate HSCs *in vitro* from pluripotent cells, such as induced human pluripotent stem cells (iPSCs) or embryonic stem cells (ESCs). While the functional attributes of HSCs are well known (self-renewal capacity, multipotency and bone marrow engraftment ability), the underlying mechanisms that endow HSCs with these properties are poorly understood, especially in human.

Decades of studies on various model organisms ranging from insects to mice have elucidated the anatomical origins of hematopoietic cells in the embryo, identified markers to isolate the various blood precursors, and dissected important gene regulatory networks that govern blood production. The use of model organisms has taught important principles about the biology of hematopoietic stem and progenitor cells (HSPCs), but has also revealed critical differences between species, including those between mice and humans. In order to solve the quandary of deriving a transplantable human HSC *in vitro*, gaps in our understanding of human hematopoiesis must be filled. As it is impossible to study early human development *in vivo* with the same tools used with model organisms (e.g., gene knockouts, fluorescent reporters and lineage-tracing tools), the use of human ESCs has provided important contribution to our understanding the intricacies of human developmental hematopoiesis.

However, the hematopoietic cells created in culture have critical short-comings when it comes to generating true, self-renewing HSCs. Therefore, to fill the gaps in our understanding of human HSC generation, it is important to refocus the studies to *in vivo* niches where human HSCs develop, expand and differentiate. This introduction to the thesis will follow human hematopoiesis through development and provide with a comparative view to mouse hematopoiesis and human hematopoiesis in culture.

1.1: The three waves of mammalian hematopoiesis

Developmental hematopoiesis in mammalian embryos is segregated to multiple waves and anatomical niches, which enables both the rapid production of the differentiated blood cells for the embryos needs, as well as the generation of undifferentiated HSCs for lifelong hematopoiesis (**Figures 1.1** and **1.2**). Developmental hematopoiesis has been classically separated into two waves: primitive and definitive. Primitive hematopoietic cells are generally thought as the product of the extraembryonic tissues and provide the developing embryo oxygen and primitive immune protection, whereas the definitive wave, traditionally thought as an attribute of the embryo proper, produces the self-renewing, transplantable HSCs that can make all differentiated cell types found in the adult blood. Although the developmental hematopoietic waves were first classified by sites of production (extraembryonic vs. embryo proper), they are more appropriately distinguished by their developmental potential (differentiation primed, lineage restricted progenitors vs. multipotent, self-renewing HSCs). It has also become clear that the developmental hematopoietic hierarchy cannot be segregated in just two waves, and there are intermediate waves of progenitors that have functional characteristics that are "in between" primitive and true definitive hematopoietic cells. In the same vein, considerable overlap in the

sites of production and differentiation of hematopoietic cells from various waves has been uncovered as better markers and genetic tools to trace individual cell lineages have been developed. This has led to frequent revisions of the models of developmental hematopoiesis.

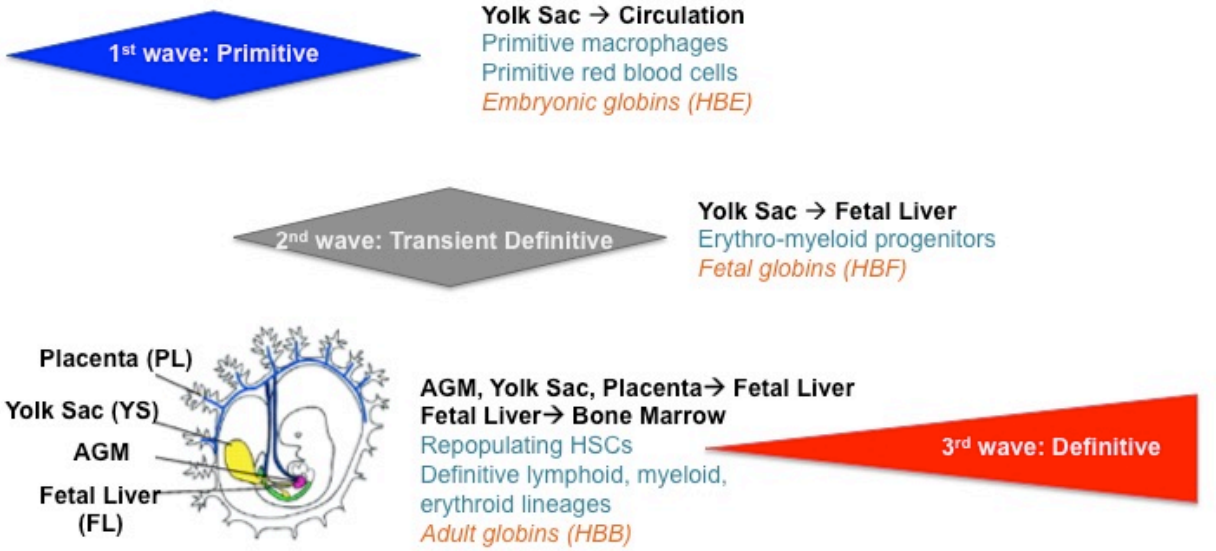


Figure 1.1: The three waves of mammalian hematopoiesis. The first wave of hematopoiesis arises from the yolk sac and generates primitive macrophages and red blood cells expressing embryonic globins. These primitive cells enter circulation to satisfy the developmental and metabolic needs of the embryo. The second wave also originates from the yolk sac and generates erythro-myeloid progenitors that give rise to erythroid cells that in human express the fetal form of beta-globin. These transient definitive cells are the first to seed the fetal liver. The final definitive wave of hematopoiesis gives rise to HSCs in the AGM, yolk sac, and placenta to seed the fetal liver, which is the main site of hematopoiesis until HSCs migrate to seed the bone marrow. These self-renewing HSCs generate lymphoid, myeloid, and erythroid cells of the definitive lineage which express the adult form of beta-globin.

1.1.1: Primitive hematopoiesis: Extra-embryonic tissues take the lead

The primitive wave of hematopoiesis is critical for meeting the immediate developmental and metabolic demands of the growing embryo. The main products of primitive hematopoiesis are nucleated erythroblasts, which are distinguished by their expression of the embryonic forms of beta-globin (epsilon (ε) in human, and beta (β) H1 in mice) (Kingsley et al., 2006). These earliest hematopoietic cells in the conceptus arise from mesodermal cell clusters in the yolk sac, known as blood islands, which can be detected as early as 16 days of human development (corresponding to Carnegie Stage (CS) 5), and embryonic day (E) 7.5 (Luckett, 1978; Moore and Metcalf, 1970) in mice. Primitive erythroblasts are also characterized by their larger size in comparison to the smaller definitive erythrocytes that have enucleated before entering circulation (Kingsley et al., 2004). The detection of circulating primitive erythroblasts does not occur until after the embryonic heart begins beating (at development day 18 (CS6) in human, and E8.5 in mice). While it was previously thought that primitive erythroblasts are incapable of enucleation, it is now known that they do enucleate later after entering in circulation. This was first documented in mice at E12.5-16.5 (Kingsley et al., 2004) and later in human (Van Handel et al., 2010). In human embryos, enucleation of primitive erythroid cells was observed in placental villous stroma at 5 weeks (CS14) of development. RBC enucleation in placental stroma was observed in macrophage-red cell clusters called erythroblast islands where macrophages digest the nucleus of the erythroblasts (Van Handel et al., 2010), similar to definitive erythropoiesis in the fetal liver and bone marrow (Yoshida et al., 2005).

Primitive hematopoiesis also produces macrophages, first detected in the precirculation yolk sac (Bertrand et al., 2005b). Studies in mice have traced the presence of myeloid progenitors to the blood islands of the yolk sac at E7.0 and the emergence of macrophages after

E9.0 (Bertrand et al., 2005a; Cline and Moore, 1972; Palis et al., 1999; Palis and Yoder, 2001) . However, fetal-derived macrophages and macrophage progenitors were also detected also in human placentae in the stroma of the chorionic plate at 3 weeks (CS7) of development, before any circulating primitive erythroid cells had reached the placenta, implying the placenta as another *de novo* source of primitive macrophages (Van Handel et al., 2010). Hence, although primitive hematopoiesis is classically associated with the yolk sac, other hematopoietic tissues also contribute to this transient blood cell production in the early embryo.

1.1.2: Transient Definitive Hematopoiesis- Co-operation between the yolk sac and the fetal liver

Although developmental hematopoiesis was initially classified into two waves, "primitive" and "definitive", as better markers and tools to assay developmental hematopoietic precursors have been developed, it has been evident that there exists an intermediary stage between primitive and true definitive hematopoiesis. In this "transient definitive wave", the yolk sac produces a large number of progenitors with erythroid and myeloid potential (EMP, erythromyeloid progenitors), which are critical for continued survival of the fetus. While EMPs and primitive hematopoietic cells both arise from the yolk sac, in contrast to primitive hematopoietic cells, EMPs have to migrate to the fetal liver to differentiate into definitive erythroid and myeloid cells. Lineage tracing in mice using Lyve1-Cre labeling demonstrated the separate lineages of yolk sac EMPs, which were Lyve1-Cre lineage traced from the mid-gestation yolk sac, and yolk sac primitive erythroid cells, which were not (Lee et al., 2016). Lyve1-Cre marked definitive HSPCs from the yolk sac were the first to initiate fetal liver

colonization and definitive erythropoiesis, demonstrating the importance of this yolk sac progenitor wave in jump-starting fetal erythropoiesis (Lee et al., 2016).

Erythroid cells generated from transient definitive progenitors in mice express the adult beta globin beta-major, hence this wave was originally labeled as "definitive". In human, the erythroid cells generated from this wave express the fetal form of beta-globin, gamma (γ) (Stamatoyannopoulos, 2005). These transient progenitors with fetal erythroid and granulopoietic potential are present in the human yolk sac as early as day 25 (CS9) of development, but become more prevalent between 4.5 and 6 weeks (CS13 and CS17) of development (Dommergues et al., 1992; Huyhn et al., 1995; Migliaccio et al., 1986; Tavian and Péault, 2005). In mice, EMPs can be detected in the yolk sac at E8.25, enter circulation at E9.0 and colonize the fetal liver by E10.5 (McGrath et al., 2011). However, while EMPs exhibit limited multipotency, they do not robustly produce lymphocytes or self-renewing HSCs and thus are not definitive hematopoietic stem cells. In mouse embryos, yolk sac EMPs can be detected with the newly described EMP marker CD16/32 co-expressed with the HSPC markers CD41 and c-kit (McGrath et al., 2015), whereas in human, no surface markers to detect specifically this wave of progenitors have been experimentally validated.

The endocardium in mice has also been shown to be hemogenic, and serve as a source of transient definitive hematopoietic progenitors (Nakano et al., 2013). So far, this finding has not been confirmed in humans. These studies suggest that the hemogenic tissues with transient definitive potential may not restricted to the yolk sac, and our understanding of the transient definitive wave is still incomplete.

1.1.3: Definitive Hematopoiesis: Multiple organs step into play

The definitive wave of hematopoiesis is distinguished from the previous primitive and transient definitive waves with the generation of multipotent, self-renewing, HSCs that can robustly produce both B and T lymphoid lineages and are capable of reconstituting recipient bone marrow in transplant setting. The intra-embryonic aorta-gonado-mesonephros (AGM) region was first proven to generate HSCs in mice, but several studies have provided evidence that HSCs can also be generated within the mouse yolk sac and the placenta (Frame et al., 2016; Gekas et al., 2005; Gekas et al., 2008; Lee et al., 2016; Ottersbach and Dzierzak, 2005; Rhodes et al., 2008; Robin et al., 2009). In human, the first multipotent myelolymphoid hematopoietic cells that are presumably the precursors for hematopoietic stem cells can be detected 3 days prior to blood circulation at day 19 (CS7) in the splanchnopleura (located in the caudal region of embryonic mesoderm that gives rise to the AGM), the suggested origin of definitive hematopoiesis (Tavian et al., 1999). $CD34^+$ intra-arterial hematopoietic cell clusters appear in the AGM beginning at day 27 and last until day 40 (CS9-16) of development, coinciding with migration of CD34+ cells to the fetal liver (Tavian et al., 1999), and corresponding to E9.5-11.5 in mice (Garcia-Porrero et al., 1995). Explant studies of human AGM at this stage show that the AGM can differentiate into definitive hematopoietic cells, including T cells (Tavian et al., 2001) while transplantation studies into NSG mice demonstrated the presence of HSCs with long-term repopulating ability at day 33 (CS14) (Ivanovs et al., 2011).

Recently, the placenta has also been identified as a source of HSCs in both the mouse and human. In mouse placentas, long-term repopulating HSCs are detected beginning from E10.5-11, and at mid-gestation quantitatively contain more HSCs than either the yolk sac or the AGM. The placental HSC pool diminishes by E15, concurrent to the migration into the fetal liver (Gekas et

al., 2005; Gekas et al., 2008), suggesting that the placental HSC pool is a significant contributor to fetal liver HSC colonization. With the onset of circulation, hematopoietic cells can spread to other sites, making it difficult to prove the origin of specific subsets of hematopoietic cells from this stage on. Nevertheless, lineage-tracing experiments in zebrafish and mice have been able to track the emergence and migration of early hematopoietic cells to specific vascular beds. In mice, studies using circulation-defective $Ncx1^{-/-}$ embryos confirmed that the placenta is not just an HSC reservoir, but a likely site of HSC emergence, as Rhodes et al. 2008 were able to document de novo production hematopoietic cells with robust myelo-lymphoid potential in $Ncx^{-/-}$ placental vasculature (Rhodes et al., 2008).

Although transplantation studies show that definitive placental and AGM HSCs emerge at the same time in the mouse embryo (Gekas et al., 2005; Ottersbach and Dzierzak, 2005), the Medvinsky group first detected long-term repopulating HSCs exclusively in the human AGM at approximately 5 weeks of development (CS 14-15), which was 5 days earlier than their assays detected HSCs in the yolk sac (CS16). Their studies did not detect transplantable HSCs in the human placenta even at 6 weeks (CS 17) (Ivanovs et al., 2011). This finding is in contrast to other studies suggesting the human placenta contains repopulating HSCs from much earlier stages, throughout gestation (Robin et al., 2009). While clonogenic progenitors were identified already in the pre-circulation placenta at 3 weeks (Van Handel et al., 2010), signs of HSCs with evidence of *in vivo* reconstitution were reported beginning from gestational week 6 and enduring until birth (Robin et al., 2009).

While the possible contribution of yolk sac to the adult HSC pool has been debated for decades, several studies in mice have now provided evidence that definitive HSCs also arise from the yolk sac (Frame et al., 2016; Lee et al., 2016). Thus, the function of the extraembryonic

tissues in developmental hematopoiesis may not be restricted to the short-lived embryonic progenitor pools.

As the nascent HSCs emerge from their sites of origin and migrate to seed the fetal liver, the fetal liver becomes the most active site of HSC expansion and differentiation during the second trimester of human development. During this period, the main beta-globin expressed in the erythroid progeny in human concepti switches from fetal (γ) to adult beta-globin (β) (Sankaran et al., 2008; Sankaran et al., 2010; Stamatoyannopoulos, 2005). Fetal liver hematopoiesis in human continues until HSCs seeding of the bone marrow is complete, after which adult hematopoiesis continues on in the medulla of the bones. In this final site of hematopoiesis, HSCs comprise the well-known "red" bone marrow while "yellow" bone marrow contains adipose and stromal cells. As the individual ages, HSCs are gradually replaced by fat cells until nearly all the marrow is "yellow" by old age.

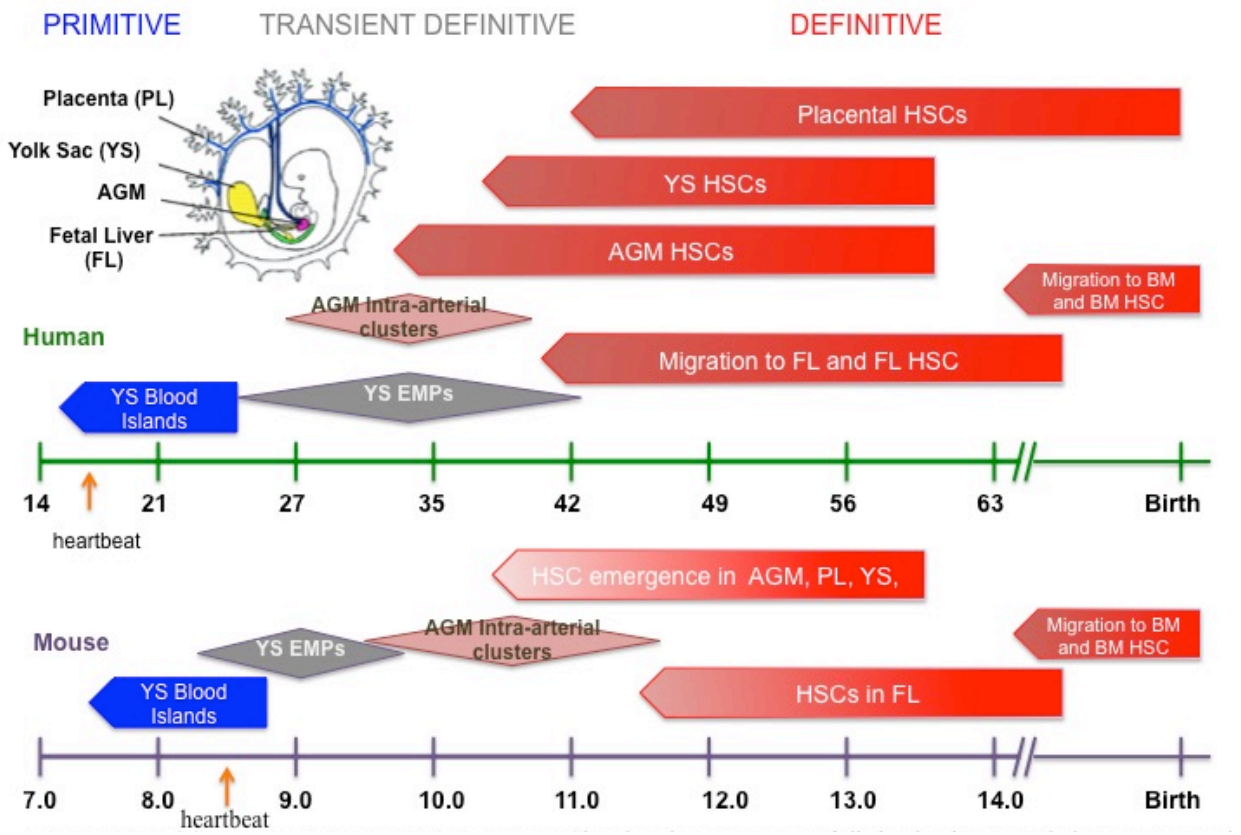


Figure 1.2: HSC emergence in developing embryo. Showing the appearance of distinctive hematopoietic structures and times of HSC emergence in human (green timeline) and mouse (purple) and coded by color to the corresponding developmental wave (blue=primitive, grey=transient definitive, red=definitive).

1.1.4: Cellular origin of definitive HSCs: Hemogenic endothelium in the spotlight

The hemogenic endothelium is the source of definitive HSCs (**Figure 1.3**). The cellular origin of the HSC, while speculated to arise from the endothelium, was only definitively traced with the advent of sophisticated imaging and lineage tracing tools (Bertrand et al., 2010; Chen et al., 2009; Eilken et al., 2009; Kissa and Herbomel, 2010; Lancrin et al., 2009; Zovein et al., 2008). An alternative model to the hemogenic endothelium was suggested from the observation that multipotent hematopoietic precursors of the AGM in mice were localized to the subaortic

patches beneath the aortic floor. Phenotypically, these AGM multipotent precursors also shared other hematopoietic markers with HSCs from the fetal liver and bone marrow but expressed low level of CD45, similar to the subaortic patch which also does not express CD45 (Bertrand et al., 2005a). The location and phenotype of these early HSCs supported the model wherein HSCs first arise from the subaortic patch and then migrate through the aortic floor to subsequently enter circulation and colonize the fetal liver. However, continuous, long-term single-cell imaging and cell-tracking of mouse ESC-derived mesodermal cells detected the generation of endothelial cells expressing hematopoietic molecular markers in an endothelial background from which hematopoietic cells emerged (Eilken et al., 2009). Furthermore, lineage-tracing studies using VE-Cadherin-Cre in both mouse and zebrafish definitively demonstrated that HSCs directly originate from aortic endothelium (Bertrand et al., 2010; Zovein et al., 2008). At this stage, the cells undergo a transition from endothelial (hemogenic endothelium) to hematopoietic (pre-HSC) characteristics as single endothelial cells bud off the aorta ventral wall to become hematopoietic cells in a progression known as the endothelial to hematopoietic transition (EHT) (Kissa and Herbomel, 2010). Mouse knockout studies showed that the transcription factor Scl is critical for the specification of hemogenic endothelium from the mesodermal precursor sometimes referred to as the hemangioblast, (Lancrin et al., 2009; Van Handel et al., 2012) while Runx1 is required specifically at the EHT transition stage (Chen et al., 2009). These studies established both the existence of the hemogenic endothelium and its critical role as the cellular origin of definitive HSCs.

1.1.5: In Culture: Progress toward recreating developmental hematopoietic hierarchy in a petri dish

As studies manipulating human embryos (i.e., knockout strains cannot be intentionally generated with human embryos and allowed to develop), studying differentiation of human ESC in culture has provided a new model to study the early stages of developmental hematopoiesis that cannot be studied *in vivo*. hESC-differentiation cultures first appear to generate the primitive erythroid lineage and closely follow the primitive and transient definitive waves of hematopoietic development as observed in the yolk sac (Zambidis et al., 2005). Additionally, hESC differentiation cultures have also generated hemogenic endothelial-like cells (Ditadi et al., 2015; Kennedy et al., 2007). However, while the primitive and transient definitive waves of hematopoiesis can be recapitulated well *in vitro*, the production of the true self-renewing and engraftment-enabled HSCs of the definitive lineage is still under development. As such, these studies have also provided hints at the importance of environmental cues to HSC development, as it has not been possible yet to recreate the conditions to generate transplantable HSCs *in vitro*.

From the clinical perspective, the high demand for bone marrow transplants far exceeding available, HLA-matched HSC donations, has resulted in significant interest in generating limitless quantities of transplantable HSCs *in vitro* from pluripotent stem cells (PSCs). However, current efforts to generate transplantable HSCs from PSCs have been unsuccessful. Studies have documented the generation of definitive precursors similar to the cells emerging from the hemogenic endothelium with multi-differentiation potential and severely limited self-renewal. The most rigorous test for HSCs is transplantation into the bone marrow of NSG mice, commonly administered through retro-orbital injection to also test for homing capability. So far, PSC-derived hematopoietic cells have not demonstrated the capacity to

function properly in transplant, irrespective of the differentiation protocol that was used for their generation (Choi et al., 2012b; Dou et al., 2016a; Doulatov et al., 2013; Kennedy et al., 2012; Ng et al., 2016; Vodyanik et al., 2006). There exist many differentiation protocols used to generate hematopoietic cells from PSC, including, but not limited to embryoid body formation and culture on stroma in the presence of specific cytokines. In this introduction, we will not compare in detail the successes and limitations of each protocol except to observe that, while all have made critical advances in studying and understanding human developmental hematopoiesis, none have resulted in a transplantable HSC.

On the other hand, since the hematopoietic cells in culture have limited exposure to critical environmental niches, such as mesenchymal stroma and other niche elements in the bone marrow, this *in vitro* shortcoming emphasizes the importance of understanding the niche environment and the signals the developing HSC encounters during its migration from sites of origin to sites of maturation, expansion and long-term maintenance.

One important step in generating the definitive wave is in identifying the stage of development to which the differentiating ESC has progressed. The Keller group showed that the generation of T-lymphocytes marks definitive hematopoietic progenitors in human PSC differentiation cultures also capable of generating myeloid and erythroid cells (Kennedy et al., 2012). While the ability to generate T lymphoid cells is not sufficient criteria for HSC generation, if the hematopoietic precursors cannot give rise to T lymphoid cells, they are unlikely to represent fully multipotent, self-renewing HSCs.

1.2: Identifying the Hematopoietic Stem Cell: The changing "face" of the HSC through development

An important facet of studying hematopoietic progenitors and stem cells is the ability to correctly identify these cells at each stage of development. The immunophenotype of the hematopoietic cell changes as development progresses and the mesoderm-derived hematovascular precursors become committed to the hematopoietic lineage, mature into self-renewing stem cells with homing ability and then, finally, lose self-renewal ability as they differentiate into progenitors of the different blood lineages. A large obstacle to distinguishing the HSC in humans is the ability to use markers unique to the human HSC that are not shared with earlier non-HSC hematopoietic progenitors or other cells from the sites of emergence.

1.2.1: Surface marker expression of HSCs

The most common method to separate hematopoietic cell populations is through the use of cell surface markers that can then be sorted or measured using flow cytometry methods. The murine HSC in the bone marrow can easily be distinguished from multipotent progenitor (MPP) cells with the SLAM family markers—$CD150^{+}$ (Slamf1), $CD48^{-/low}$ (Slamf2), $CD229^{=/-}$ (Slamf3), and $CD244^{-}$ (Slamf4)—and $lineage^{-/low}Sca1^{+}c\text{-}Kit^{+}$ (LSK) markers (MPPs are LSK^{+} $CD150^{-}$) (Oguro et al., 2013). However, these markers cannot be used in human cells, which do not express Sca1, nor do the SLAM markers distinguish human HSCs to the same degree as compared to mice (Larochelle et al., 2011). This makes the selective isolation of HSCs from non-HSC progenitor cells much more difficult in human. The currently accepted surface marker phenotype for definitive human HSCs consists of expression of CD34, CD90 (Thy1), the pan-

hematopoietic marker CD45, and negative/lo CD38 expression, collectively referred to as $Thy1^{+,}$ or CD90, HSCs (Baum et al., 1992; Hogan et al., 2002; Magnusson et al., 2013; Majeti et al., 2007; Uchida et al., 2001). However, both CD34 and CD90 expression are shared with the endothelium and therefore do not alone distinguish HSCs from their hemogenic endothelial precursors or endothelial niche cells.

CD133 (PROM1) has been used to isolate cells from bone marrow and cord blood for successful transplantation into mouse, human, and sheep, but its expression is not limited to the HSC as it is expressed on multiple stem cells as well as hematopoietic progenitors. The combination $CD34^{+}CD38^{-}CD90^{+}CD133^{+}$ marks immature progenitors (Ramshaw et al., 2001). (Author's personal observation that, in human fetal liver, CD133 is not as restrictive in HSCs as CD90) Other markers used in the identification of early hematopoietic progenitors include CD31, CD41 and CD43, the pre-panhematopoietic marker that distinguishes commitment to the hematopoietic fate and precedes the appearance of pan-hematopoietic marker CD45 that is expressed in all nucleated hematopoietic cells. CD31 is also expressed in the endothelium, while CD43 and CD45 are specific to hematopoietic cells.

The angiotensin-converting enzyme (ACE/BB9/CD143) can be detected in the earliest definitive HSCs arising from the AGM but is also in other hematopoietic cells in the fetal liver and umbilical cord blood (Jokubaitis et al., 2008; Oberlin et al., 2010). Since ACE is expressed throughout hematopoietic development, it provides a useful marker to track hematopoiesis through each stage and niche, but cannot distinguish between progenitors and HSC.

VE-Cadherin (CD144) has been shown to mark the highly self-renewing HSC population in human fetal liver, as opposed to the less self-renewing VE-Cadherin- subsets (Oberlin et al., 2010). Human HSCs are also present in the $VE\text{-}cadherin^{+}CD45^{+}$ fraction in the AGM (Ivanovs et

al., 2014). However, VE-Cadherin is also expressed with the endothelial lineages and has been shown in mice to be only transiently expressed in fetal liver HSCs from E13.5 to E16.5, and not detectable in bone marrow (Kim et al., 2005). Recent studies identified GPI80 expression in the human fetal liver $CD90^+$ ($Thy1^+$) HSPCs as marking the only self-renewing and engraftment-capable cells within the $CD90^+$ HSPCs (Prashad et al., 2014). $GPI80^+$ as a hematopoietic self-renewal indicator is applicable throughout human development; $GPI80^+$ HSCs can be detected as early as early as 8 weeks in the human fetal liver and as early as 5 weeks in the human placenta. Its expression is preserved in fetal bone marrow HSCs, but is not observed robustly in adult bone marrow HSCs. GPI80 was not expressed in endothelial cells at sites of HSC origin, but only appeared in HSPCs after their emergence. Furthermore, while $GPI80^+$ HSCs isolated from the fetal liver were able to reconstitute the hematopoietic system upon transplantation into NSG mice, isolation of a similar population in the early placenta was shown to expand *in vitro* but has not yet been tested *in vivo*. Future transplantation studies will determine whether GPI80 can also identify the engrafting, self-renewing cells from earlier hemaopoietic tissues (Prashad et al., 2014).

Identifying new markers that can be utilized to further separate the hematopoietic populations in the different stages and sites of human HSC development remains a critical goal for the field. Such advances would contribute greatly in the isolation of HSCs for clinical use, increase the safety and efficacy of bone marrow transplantations, and help document the progress toward generating definitive HSCs *in vitro*.

1.2.2: In Culture: Clues to hematopoietic maturity in changing cell surface identity

Identifying the exact precursor for HSC is critical for understanding how the definitive HSC lineage is initially established. The use of hPSC cultures to observe the emergence of separate hematopoietic lineages enabled the discrimination between the distinct lineages of hemogenic endothelium and arterial vascular endothelium (Ditadi et al., 2015). The hemogenic endothelium is enriched in the $CD34^+$ population and both murine hemogenic endothelium and hESC-derived hemogenic endothelium lack CD184 and CD73 expression while arterial and venous endothelium can be distinguished by varying levels of CD184 and CD43. Hemogenic endothelial cells also do not express the endothelial progenitor marker DLL4. In hPSC-differentiation cultures, the $CD34^+CD73^-CD184^-DLL4^-$ population therefore marked the hemogenic endothelium population capable of generating hematopoietic progenitors (Ditadi et al., 2015).

Studying the development of the HSC in ESC-differentiation cultures has also proven informative in identifying the first markers that emerge in definitive hematopoiesis. In mESCs, CD41 marks the first definitive hematopoietic progenitors (Mikkola et al., 2003) whereas CD43 identifies the hESC-derived hematopoietic progenitors (Vodyanik et al., 2006). Differentiation cultures have shown it is possible to generate cells that closely match immunophenotypically *in vivo* HSCs with the expression of the classical human HSC immunophenotype $CD34^+CD38^-CD45^+CD90^+$ as well as induction of the fetal HSC self-renewal marker GPI80 (Dou et al., 2016b; Prashad et al., 2014). However, none of the hPSC-differentiation cultures, no matter how stringently selective of these hematopoietic stem cell markers, was able to document the generation of a self-renewing, transplantable HSC. The field will need to continue to identify and validate a new marker or marker combination that also in vitro setting functionally correlates with definitive, mature HSCs with the engraftment capability.

1.3: HSC Regulatory Networks

Many cues, both environmental and cell intrinsic, impact the development and maturation of the emerging HSC (**Figure 1.3**). The majority of the *in vivo* studies conducted at this stage have been in mice. However, many clues about the EHT stage in humans have been discovered using PSC differentiation cultures that allow the study of human hematopoietic development at stages inaccessible in the living organism in lieu of *in vivo* studies that are possible in other model organisms. hPSC studies have been used both to identify new factors and signals important to hematopoietic specification and to corroborate and more closely study results observed in *in vivo* models. In fact, hPSC differentiation cultures mimic in vivo yolk sac hematopoietic development and successfully replicate the primitive and transient definitive waves (Qiu et al., 2005; Zambidis et al., 2005). However, although it is possible to generate hemogenic endothelial-like cells (Ditadi et al., 2015; Kennedy et al., 2012), less success has been achieved in replicating the definitive wave in culture. So far, it has not been possible to differentiate transplantation-capable HSCs with full erythroid and lymphoid potential from hPSCs (Dravid et al., 2011; Martin et al., 2008). Gaps in our understanding of HSC networks must be bridged to achieve this goal.

1.3.1: Transcription Factor Establishment of the HSC

Several transcription factor networks are critical for directing the development of hemogenic endothelium from mesoderm and then the specification of HSCs from the hemogenic endothelium. SCL/TAL1 has been shown to be a master regulator of hematopoiesis. It suppresses the cardiac fate in mice during mesoderm specification (Org et al., 2015; Van Handel

et al., 2012) and is indispensible for establishment of the hemogenic endothelium (Lancrin et al., 2009).

Studies in both mouse and human ESCs have revealed the importance of SOX17 in definitive HSC development. In mice, Sox17 expression marks the precursors to the hemogenic endothelial and definitive hematopoietic lineages and is required for generation and maintenance of fetal and neonatal HSCs from the fetal liver and yolk sac (Choi et al., 2012a; Kim et al., 2007). Studies using Sox17-GFP reporter mice showed that the DNA-binding factor SOX17 is critical for the development of the definitive hemogenic endothelium and is expressed in both the hemogenic endothelium and emerging HSC (Clarke et al., 2013). Moreover, differentiation studies in mESCs showed that Sox17 demarcates emerging hemogenic endothelial and hematopoietic progenitors, and SOX17 expression is required for both EHT and definitive hematopoiesis (Clarke et al., 2013). Sox17 is also expressed at high levels in hESC-derived endothelial cells and low levels in pre-hematopoietic progenitor cells. In culture, overexpression studies of Sox17 promoted the expansion of HE-like cells in a stage between endothelial and pre-hematopoietic progenitor cell (Nakajima-Takagi et al., 2013).

Following hemogenic endothelium patterning, Runx1 expression distinguishes hemogenic endothelium from other endothelial cells and is required specifically at the EHT stage (Chen et al., 2009). Conditional deletion of Runx1 prevents the generation of intra-arterial clusters, EMP and HSC in mice (Chen et al., 2009). Direct downstream Runx1 targets are Gfi1 and Gfi1B, which suppress the endothelial programs during EHT in mice to allow the emergence of hematopoieic cells from hemogenic endothelium (Lancrin et al., 2012) and, at a later stage, are also important for maintaining HSC quiescence (Hock et al., 2004a; Zeng et al., 2004) and

the development and function of HSCs and the blood lineages arising from definitive HSCs (reviewed in (van der Meer et al., 2010)).

Tel1/ETV6 rearrangements are frequently found in human leukemias (Bertrand et al., 2005a; Golub et al., 1994). In mice, Tel1/ETV6 is required for blood vessel development and later in hematopoiesis for adult HSC survival in the bone marrow (Hock et al., 2004b; Wang et al., 1997; Wang et al., 1998) while in zebrafish, Tel1/ETV6 is a critical for specification of the first HSCs from the dorsal aorta (Ciau-Uitz et al., 2010). Gata2 in mice is required for definitive hematopoiesis and is critical specifically in HSC production and expansion in the AGM and proliferation and survival in the adult bone marrow (Ling et al., 2004; Tsai et al., 1994; Tsai and Orkin, 1997). In hPSC differentiation cultures, overexpression of ETV2 with GATA1 or GATA2 or SCL/TAL1 with GATA2 was able to induce hemogenic endothelium with, respectively, either pan-myeloid or erythroid and megakaryocytic potential (Elcheva et al., 2015). Myb is also required during definitive hematopoiesis for HSC self-renewal and to suppress differentiation programs but is dispensable in primitive hematopoiesis in mice (White and Weston, 2000).

Following definitive HSC emergence, several transcription factors contribute to the maintenance of the HSC identity and function. For example, HOXB4 directly targets the hematopoiesis-specific protein Hemogen (HEMGN). In mice, HOXB4 overexpression was able to instigate robust ex vivo expansion in murine HSCs (Antonchuk et al., 2002), suggesting its importance in HSC self-renewal. PSC-differentiation cultures further demonstrated the importance of HOXB4 in engraftment when overexpression of HOXB4 generated engraftable HSCs in EB/ESC cultures (Chan et al., 2008; Kyba et al., 2002; Pilat et al., 2013). However, HOXB4 has not been shown to be beneficial in human ESC differentiation cultures as its

overexpression showed no detectable effect in improving repopulating ability, highlighting another major difference between mouse and human development (Lu et al., Wang et al.).

The HOXA genes have also emerged as key HSC regulators in humans, largely due to *in vitro* PSC differentiation studies (Dou et al., 2016b; Doulatov et al., 2013; Ng et al., 2016), but the importance of HoxA genes is also described in mouse definitive HSCs *in vivo*. HOXA9 knockout mice display hematopoietic defects, particularly in bone marrow hypocellularity from impaired self-renewal and proliferation defects in the LT-HSCs (Lawrence et al., 2005). Monoallelic deletion of the HoxA cluster demonstrated that HOXA cluster haploinsufficiency was sufficient to increase primitive hematopoietic cells at the cost of HSCs. Additionally, $HoxA^{+/-}$ HSCs were not as competitive in engraftment as wild type HSCs, suggesting that the HoxA cluster regulates definitive hematopoiesis (Lebert-Ghali et al., 2010). Additionally, conditional homozygous deletion of the HoxA cluster in adult murine HSCs resulted in impaired proliferation potential in HSPCs and subsequently reduced competitiveness in engraftment, suggesting HoxA genes mediate proliferation in HSPCs (Lebert-Ghali et al., 2016). In human cells, knockdown of medial HOXA genes in self-renewing HSCs of the second trimester fetal liver also led to an inability to self-renew and a molecular correlation to the more immature, earlier hematopoietic progenitors (Dou et al., 2016b) (Chapter 2). We noted the inability of hESC-derived HSPCs, which cannot self-renew or engraft, to express the medial HOXA genes and implicated the medial HOXA genes as critical to human HSC self-renewal and identity. Doulatov et al., identified the combination HOXA9 with transcription factors ERG and RORA as required and sufficient to respecify hPSC-derived myeloid-restricted precursors to multipotential hematopoietic progenitors (Doulatov et al., 2013).

The B cell CLL/lymphoma 11A (BCL11A) transcription factor is required for the switch from fetal gamma (HBG) globin to adult beta (HBB) globin (Sankaran et al., 2008; Sankaran et al., 2010), which is the marker for erythrocytes of the definitive wave, and is also critical for normal lymphopoiesis (Liu et al., 2003; Yu et al., 2012). Recent studies in BCL11A-difficent mice have shown that BCL11A is also required for normal HSC function, and lack of BCL11A produces HSCs with aged phenotype in the bone marrow (Luc et al., 2016). Unpublished preliminary data in our lab corroborates this finding in human HSCs as knockdown of BCL11A in HSCs isolated from the human fetal liver severely depletes the ability to maintain HSCs in culture and the cells display severe self-renewal defects.

The hepatic leukemia factor HLF is also and important HSC regulator, and its overexpression in HSPCs causes increased self-renewal (Gazit et al., 2013). PRDM16 is another transcription factor that is highly expressed in HSCs and is critical for HSC expansion through protection from apoptosis (Aguilo et al., 2011). MECOM, which encodes the transcription factor EVI1, regulates hematopoiesis through maintaining HSC proliferation in a dosage-dependent manner (Goyama et al., 2008). Timely expression of all these factors during HSC development is critical for generating functional HSCs, and monitoring of their expression in culture provides clues about the developmental stage to which in vitro generated hematopoietic cells have progressed.

1.3.2: Signaling in the definitive HSC

Signaling pathways co-operate with cell intrinsic factors and are also critical for specification of the definitive hemogenic endothelium, the endothelial to hematopoietic transition, and HSC emergence. In particular, Notch signaling is critical for HSC development

but is dispensable in adult HSCs following emergence from the AGM (Maillard et al., 2008; Souilhol et al., 2016). Studies using Notch1$^{-/-}$ mice showed that Notch1 signaling is especially critical for HSC generation from endothelial cells (Kumano et al., 2003). Mouse, ES cell, and zebrafish studies have shown that Notch signaling, acting in conjunction with Hedgehog and Scl, promotes the endothelial-to-hematopoietic transition. In zebrafish, Notch, together with Runx1, is required to specify adult HSCs. Notch1 signaling through the Notch targets, Hey1/Hey2 transcription factors, in mice implicated Notch signaling in developing the embryonic vascular system (Fischer et al., 2004) although studies in Notch-ligand-null mice shows Notch signaling may be dispensable for establishing arterial fate (Robert-Moreno et al., 2008).

In 2013, the Keller group showed that retinoic acid signaling initiated by RALDH2 through RAR-alpha was critical for the specification of definitive hematopoietic stem cells in early mouse embryos (Chanda et al., 2013). Indeed, our studies showed that RA signaling was able to induce the HOXA genes, prolong maintenance of the HSPC compartment, and improve clonogenic potential in hESC-HSPCs (Dou et al., 2016b) (see Chapter 2). Further highlighting the importance of signaling pathways involving the HOXA genes, Ng et al. showed that simultaneous upregulation of Wnt signaling and inhibition of Activin signaling differentiates hESCs to CD34^{+} HOXA^{+} cells that more closely resemble repopulation-competent CD34^{+} cord blood cells than CD34^{+} HOXA^{-} cells (Ng et al., 2016).

Survival cues are also important for maintaining the definitive HSC. For example, the receptor kinase FLT3 is highly expressed in early lymphoid progenitors and HSCs capable of long-term engraftment and signaling through FLT3 acts to prevent HSPCs from spontaneous apoptosis (Kikushige et al., 2008). Mouse studies of the cytokine thrombopoietin (TPO) showed TPO is important for both platelet and megakaryocyte production (de Sauvage et al., 1996;

Gurney et al., 1994) and adult HSC maintenance (Kimura et al., 1998) through maintaining quiescence (de Sauvage et al., 1996; Gurney et al., 1994; Qian et al., 2007) and expands HSCs following stress, such as transplantation (Craig et al., 1993). Both these factors are used in PSC differentiation and HSC supportive cultures.

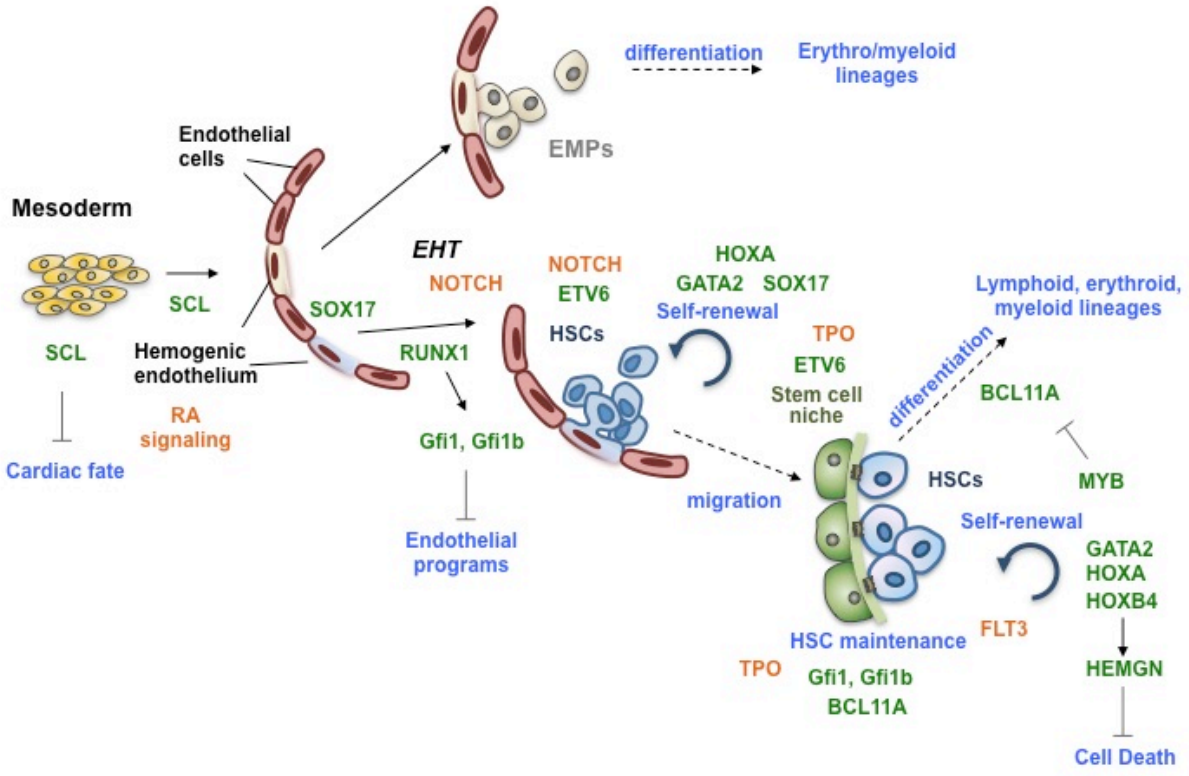


Figure 1.3: Transcription factors and signaling during HSC specification. Many transcription factors are required to suppress other fates and to drive the differentiation into HE, promote EHT, and maintain HSC identity, survival, and expansion once established. This figure shows key transcription factors (green) and signaling pathways (orange) involved in these processes.

Collectively, these studies emphasize the intricacy of transcription factors and signaling networks that are critical for definitive hematopoietic development and in maintaining HSC

identity after emergence. However, our current knowledge of definitive HSC emergence is still incomplete, as evidenced by the lack of success in generating a transplantable HSC from hPSC differentiation cultures. In addition to transcription factors and signaling pathways, other elements, particularly in the chromatin landscape, must also be considered when aiming to reconstruct HSC development. Recent studies using ATAC-seq to profile the chromatin accessibility in 13 human primary blood cell types from 21 donors showed the striking parallel between chromatin accessibility patterns unique to each cell type that reflect cell identity better than mRNA expression levels obtained through RNA-seq (Corces et al., 2016). Thus, future studies studying hematopoietic cells and lineages must take into consideration both the traditional transcriptome and noncoding aspects impacting the underlying genomic structural organization. In our own studies, we observed that RA signaling induced both HOXA gene expression as well as accessibility to the HOXA gene locus (Dou et al., 2016b) (Chapter 2), suggesting that the initial impediment to HOXA gene expression may have been due, at least in part, to structural characteristics of the region which made the HOXA genes inaccessible for transcription. To fully achieve the goal of differentiating transplantable HSCs from PSCs, elucidating the structural components of gene regulation is just as critical as understanding the genes and pathways underlying HSC specification and identity.

1.4: Long noncoding RNAs in Hematopoiesis

The intergenic regions between genes were previously regarded as "junk" and are now regarded as the "dark matter" of the genome. One example, long noncoding RNAs (lncRNAs), are only now being understood as critically important in the structural regulation of the universe/genome (**Figure 1.4**), just as physical dark matter is critically important for the

structure of the universe (Peebles, 2015). Since only 1.5% of the mammalian genome consists of protein coding potential, the complexity of organisms can be attributed to noncoding RNAs. LncRNAs are classified as RNA transcripts >200 nucleotides with little coding potential. Unlike small RNAs, such as miRNAs, shRNAs, and piRNAs, which are highly conserved as they act at the level of transcription and post-transcription silencing through sequence-specific base pairing with their targets, lncRNAs function through diverse mechanisms not restricted to base pairing with targets and are poorly conserved. Mechanistic archetypes for lncRNAs include signals, decoys, guides, and scaffolds, and some lncRNAs fit multiple archetypes at once (reviewed in (Wang and Chang, 2011; Wang et al., 2011b)). As many lncRNAs are under tight transcriptional control, they are also tissue- and cell-type specific and thus may serve as molecular signals identifying cell type, location, and developmental stage (Guttman et al., 2009; Guttman et al., 2010; Wang and Chang, 2011).

Since lncRNAs cannot often be identified through sequence conservation across species, criteria such as evolutionary conservation of synteny, microhomology, and secondary structure (Quinn et al., 2016). The tightly transcriptionally regulated HOX clusters transcribe many lncRNAs that are also expressed in temporally restricted and site-specific manner (Rinn et al., 2007). Two of the most well-studied lncRNAs of the human HOXA cluster include HOTAIRM1 and HOTTIP which work as guides, signals, and scaffolds to configure a locus of the HOXA cluster and recruit transcriptional complexes to the otherwise inaccessible genes. HOTAIRM1, first discovered in human myeloid cells, is expressed at the proximal end of the cluster between HOXA1 and HOXA2 and interacts with the UTX/MLL H3K27 demethylase or polycomb repressive complex 2 (PRC2) complexes to regulate transcription of the proximal HOXA1-HOXA4 genes (Wang and Dostie, 2016; Zhang et al., 2009). HOTTIP, expressed distally in the

HOXA cluster beside HOXA13 also shows distal anatomic localization, activates the expression of the distal HOXA9-HOXA13 genes through interaction with WDR5, the binding partner for the H3K4 methyltransferase MLL1 (Wang et al., 2011a; Yang et al., 2014). In mice, the lncRNA Mistral is expressed medially between HOXA6 and HOXA7 and also acts *in cis* to its site of transcription to regulate the expression of HOXA6 and HOXA7 in mouse ESCs through direct interaction with MLL1 (Bertani et al., 2011). While Hotairm1 has also been described in mouse cells (De Kumar et al., 2015) and Hottip in mouse and chick (Wang et al., 2011a), thus far, there are no reports of human MISTRAL or its functional equivalent situated in the medial HOXA cluster.

The interaction of HOXA-associated lncRNAs with transcriptional co-activators, such as UTX/MLL and MLL1 and co-repressors, such as PRC2, and chromosomal looping of the neighboring genetic locus to regulate HOXA gene expression provides an interesting avenue of exploration in the context of hematopoietic development. Both HOTTIP and HOTAIRM1 currently serve as biomarkers for multiple carcinomas (Díaz-Beyá et al., 2015; Quagliata et al., 2014; Wan et al., 2016; Zhang et al., 2015). With regards to hematopoiesis, HOTAIRM1 is critical for myeloid differentiation and is specifically associated with acute myeloid leukemia (Díaz-Beyá et al., 2015). Other leukemia related lncRNAs include the Notch-regulated lncRNA LUNAR1, which is required for the oncogenic proliferation of T-cell acute lymphoblastic leukemia cells (Trimarchi et al., 2014). Recent studies have also observed unique lncRNA profiles between HSCs, lymphoid progenitors, and committed B and T lymphoid cells, implicating lncRNAs in hematopoietic differentiation and lineage commitment (Casero et al., 2015). Thus, in addition to having important regulatory functions. lncRNAs could serve as molecular markers for identifying distinct subsets of hematopoietic cells.

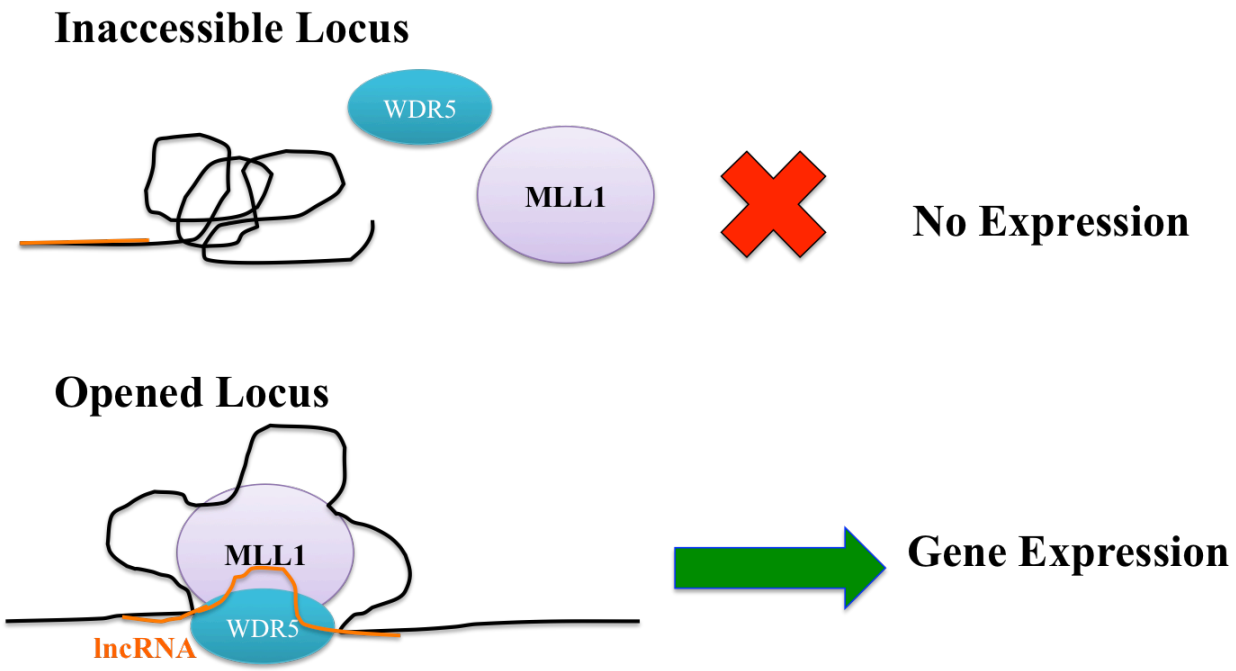


Figure 1.4: Long Noncoding RNAs can function as molecular scaffolds. An example of how lncRNAs may function in changing the chromatin landscape: an inaccessible gene locus can be transformed into an open configuration by lncRNAs. LncRNAs can also recruit co-activators, such as methyltransferases, to activate gene expression.

OVERVIEW

While the functional attributes of HSCs that make them invaluable in treating a vast catalogue of blood disorders are well documented, there remain gaps in our understanding of the molecular mechanisms underlying these functional attributes. Here, we have outlined the advances the field of hematopoiesis has made in charting human HSC development both *in vivo* and *in vitro* as well as the remaining black boxes that have yet to be clarified. However, it is clear that the HSC is a complex cell defined by more than surface identity and gene expression,

reliant on both cell intrinsic and extracellular cues, regulated by both coding and non-coding factors, and tightly governed throughout its development into and subsequence maintenance as a self-renewing stem cell capable of producing all lineages found in blood.

Aims of this Dissertation

HSCs are characterized by three unique properties: their ability to self-renew, engraft into bone marrow, and generate all lineages found in blood. There are significant gaps in our understanding of the genetic pathways underlying these functional characteristics and why these properties fail to be induced during HSPC differentiation *in vitro* from pluripotent cells (Dravid et al., 2011; Dravid and Crooks, 2011). Here, we use hESC differentiation cultures and definitive HSCs isolated from the human fetal liver to understand the molecular programs and regulatory mechanisms important for human HSC function and identity.

Aim 1: Define the functional and molecular differences between ESC-derived HSPCs and HSPCs from the human fetal liver

Previous ESC differentiation cultures successfully generated hematopoietic cells from the primitive and transitive waves of developmental hematopoiesis (Kennedy et al., 2007; Zambidis et al., 2005; Zambidis et al., 2007), but the generation of engraftable multilineage HSCs from PSCs has not yet been achieved. To define the extent of specification of definitive HSPCs progress in vitro, and identify the functional and molecular defects in PSC derived HSPCs, we developed a two-step differentiation culture intended to simulate the environments a developing HSC would encounter *in vivo*. Here we compare the phenotype, multipotency, self-renewal, and

engraftment ability of ESC-derived HSPCs to definitive HSPCs isolated from the fetal liver using FACS analysis, lymphoid and semi-solid differentiation cultures, and transplantation into NSG mice. DNA microarray analysis of HSPCs isolated from multiple stages of development and ESC-HSPC differentiation culture is used to identify programs dystregulated between *in vivo*-derived FL-HSPCs and ESC-derived HSPCs.

Aim 2: Identify the molecular blocks preventing the *in vitro* generation of functional HSCs from ESCs

Aim 1 identified the downregulation of the entire cluster of HOXA genes, in particular the medial HOXA genes, as a molecular defect in ESC-derived HSPCs. The involvement of HOXA genes in hematopoietic proliferation or self-renewal has been implicated before. HOXA9 and HOXA10 fusions with MLL have been observed in many cases of leukemia and HOXA9 is one of the key factors identified in direct reprogramming hPSC-derived myeloid precursors to multipotent progenitors (Doulatov et al., 2013), but less is known about the medial HOXA genes in hematopoiesis. Here we use knockdown in FL-HSPCs and overexpression in FL- and ESC-HSPCs of the medial HOXA genes to determine the impact of the medial HOXA genes on HSPC multipotency, self-renewal, and engraftment ability while RNA-seq and qPCR are used to assess HOXA gene regulated programs.

Aim 3: Determine the mechanisms regulating the molecular pathways critical for HSC function and identity

Since the silencing of the medial HOXA genes was identified as a key molecular block preventing ESC-HSPC self-renewal, our goal has been to identify cell intrinsic and extrinsic

mechanisms that regulate the induction and/or maintenance of the medial HOXA genes. Retinoic acid (RA) signaling is known to be critical for definitive HSC specification at the hemogenic endothelium stage (Chanda et al., 2013), and in other cell types regulates anterior and medial HOXA genes. Here we use RNA-seq, qPCR, and ATAC-seq to assess changes in gene expression and chromatin accessibility of the HOXA cluster and genes associated with hematopoietic programs upon RA signaling induction in ESC-derived HSPCs. FACS analysis and differentiation cultures is used to examine whether RA signaling improves ESC-HSPC phenotype and functionality.

MLL1 is an upstream regulator of the HOXA genes (Schuettengruber et al., 2007; Schwartz and Pirrotta, 2007; Soshnikova, 2013; Soshnikova and Duboule, 2009), and the lncRNA HOTTIP is known to function as molecular scaffold to recruit the MLL1 complex to activate the HOXA9-HOXA13 locus in other cell types (Wang et al., 2011a; Yang et al., 2014). The lncRNA, Mistral, was proposed to act in an analogous manner on HOXA6/HOXA7 in mice (Bertani et al., 2011). Here we use ChIP-PCR to determine binding of MLL1 to the HOXA gene locus and DNA microarray and RNA-seq to compare gene expression and identify epigenetic mechanisms that are involved in regulating the programs that differ between self-renewing and non-self-renewing HSPCs. Alignment tools are used to compare the sequence conservation of a postulated lncRNA peak identified from RNA-seq to mouse Mistral. Rapid amplification of cDNA ends (RACE) and ChIP-seq of methylation marks were used to determine the positioning of the putative human MISTRAL in the medial HOXA gene locus.

Chapter 2 addresses Aim 1, Aim 2, and the first half of Aim 3. Chapter 3 discusses the second part of Aim 3.

Bibliography

Aguilo, F., Avagyan, S., Labar, A., Sevilla, A., Lee, D.F., Kumar, P., Lemischka, I.R., Zhou, B.Y., and Snoeck, H.W. (2011). Prdm16 is a physiologic regulator of hematopoietic stem cells. Blood *117*, 5057-5066.

Antonchuk, J., Sauvageau, G., and Humphries, R.K. (2002). HOXB4-induced expansion of adult hematopoietic stem cells ex vivo. Cell *109*, 39-45.

Baum, C.M., Weissman, I.L., Tsukamoto, A.S., Buckle, A.M., and Peault, B. (1992). Isolation of a candidate human hematopoietic stem-cell population. Proc Natl Acad Sci U S A *89*, 2804-2808.

Bertani, S., Sauer, S., Bolotin, E., and Sauer, F. (2011). The Noncoding RNA Mistral Activates Hoxa6 and Hoxa7 Expression and Stem Cell Differentiation by Recruiting MLL1 to Chromatin. Molecular Cell *43*, 1040-1046.

Bertrand, J.Y., Chi, N.C., Santoso, B., Teng, S., Stainier, D.Y., and Traver, D. (2010). Haematopoietic stem cells derive directly from aortic endothelium during development. Nature *464*, 108-111.

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

Bertrand, J.Y., Jalil, A., Klaine, M., Jung, S., Cumano, A., and Godin, I. (2005b). Three pathways to mature macrophages in the early mouse yolk sac. Blood *106*, 3004-3011.

Casero, D., Sandoval, S., Seet, C.S., Scholes, J., Zhu, Y., Ha, V.L., Luong, A., Parekh, C., and Crooks, G.M. (2015). Long non-coding RNA profiling of human lymphoid progenitor cells reveals transcriptional divergence of B cell and T cell lineages. Nat Immunol *16*, 1282-1291.

Chan, K.M., Bonde, S., Klump, H., and Zavazava, N. (2008). Hematopoiesis and immunity of HOXB4-transduced embryonic stem cell-derived hematopoietic progenitor cells. Blood *111*, 2953-2961.

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

Chen, M.J., Yokomizo, T., Zeigler, B.M., Dzierzak, E., and Speck, N.A. (2009). Runx1 is required for the endothelial to haematopoietic cell transition but not thereafter. Nature *457*, 887-891.

Choi, E., Kraus, M.R., Lemaire, L.A., Yoshimoto, M., Vemula, S., Potter, L.A., Manduchi, E., Stoeckert, C.J., Grapin-Botton, A., and Magnuson, M.A. (2012a). Dual lineage-specific expression of Sox17 during mouse embryogenesis. Stem Cells *30*, 2297-2308.

Choi, K.D., Vodyanik, M.A., Togarrati, P.P., Suknuntha, K., Kumar, A., Samarjeet, F., Probasco, M.D., Tian, S., Stewart, R., Thomson, J.A*., et al.* (2012b). Identification of the hemogenic endothelial progenitor and its direct precursor in human pluripotent stem cell differentiation cultures. Cell Rep *2*, 553-567.

Ciau-Uitz, A., Pinheiro, P., Gupta, R., Enver, T., and Patient, R. (2010). Tel1/ETV6 specifies blood stem cells through the agency of VEGF signaling. Dev Cell *18*, 569-578.

Clarke, R.L., Yzaguirre, A.D., Yashiro-Ohtani, Y., Bondue, A., Blanpain, C., Pear, W.S., Speck, N.A., and Keller, G. (2013). The expression of Sox17 identifies and regulates haemogenic endothelium. Nat Cell Biol *15*, 502-510.

Cline, M.J., and Moore, M.A. (1972). Embryonic origin of the mouse macrophage. Blood *39*, 842-849.

Corces, M.R., Buenrostro, J.D., Wu, B., Greenside, P.G., Chan, S.M., Koenig, J.L., Snyder, M.P., Pritchard, J.K., Kundaje, A., Greenleaf, W.J.*, et al.* (2016). Lineage-specific and single-cell chromatin accessibility charts human hematopoiesis and leukemia evolution. Nat Genet *48*, 1193-1203.

Craig, W., Kay, R., Cutler, R.L., and Lansdorp, P.M. (1993). Expression of Thy-1 on human hematopoietic progenitor cells. J Exp Med *177*, 1331-1342.

De Kumar, B., Parrish, M.E., Slaughter, B.D., Unruh, J.R., Gogol, M., Seidel, C., Paulson, A., Li, H., Gaudenz, K., Peak, A.*, et al.* (2015). Analysis of dynamic changes in retinoid-induced transcription and epigenetic profiles of murine Hox clusters in ES cells. Genome Res *25*, 1229-1243.

de Sauvage, F.J., Carver-Moore, K., Luoh, S.M., Ryan, A., Dowd, M., Eaton, D.L., and Moore, M.W. (1996). Physiological regulation of early and late stages of megakaryocytopoiesis by thrombopoietin. J Exp Med *183*, 651-656.

Ditadi, A., Sturgeon, C.M., Tober, J., Awong, G., Kennedy, M., Yzaguirre, A.D., Azzola, L., Ng, E.S., Stanley, E.G., French, D.L.*, et al.* (2015). Human definitive haemogenic endothelium and arterial vascular endothelium represent distinct lineages. Nat Cell Biol *17*, 580-591.

Dommergues, M., Aubény, E., Dumez, Y., Durandy, A., and Coulombel, L. (1992). Hematopoiesis in the human yolk sac: quantitation of erythroid and granulopoietic progenitors between 3.5 and 8 weeks of development. Bone Marrow Transplant *9 Suppl 1*, 23-27.

Dou, D.R., Calvanese, V., Saarikoski, P., Galic, Z., and Mikkola, H.K.A. (2016a). Induction of HOXA genes in hESC-derived HSPC by two-step differentiation and RA signalling pulse. Nature Protocol Exchange.

Dou, D.R., Calvanese, V., Sierra, M.I., Nguyen, A.T., Minasian, A., Saarikoski, P., Sasidharan, R., Ramirez, C.M., Zack, J.A., Crooks, G.M.*, et al.* (2016b). Medial HOXA genes demarcate haematopoietic stem cell fate during human development. Nat Cell Biol *18*, 595-606.

Doulatov, S., Vo, L.T., Chou, S.S., Kim, P.G., Arora, N., Li, H., Hadland, B.K., Bernstein, I.D., Collins, J.J., Zon, L.I.*, et al.* (2013). Induction of multipotential hematopoietic progenitors from human pluripotent stem cells via respecification of lineage-restricted precursors. Cell Stem Cell *13*, 459-470.

Dravid, G., Zhu, Y., Scholes, J., Evseenko, D., and Crooks, G.M. (2011). Dysregulated gene expression during hematopoietic differentiation from human embryonic stem cells. Mol Ther *19*, 768-781.

Dravid, G.G., and Crooks, G.M. (2011). The challenges and promises of blood engineered from human pluripotent stem cells. Adv Drug Deliv Rev *63*, 331-341.

Díaz-Beyá, M., Brunet, S., Nomdedéu, J., Pratcorona, M., Cordeiro, A., Gallardo, D., Escoda, L., Tormo, M., Heras, I., Ribera, J.M.*, et al.* (2015). The lincRNA HOTAIRM1, located in the HOXA genomic region, is expressed in acute myeloid leukemia, impacts prognosis in patients in the intermediate-risk cytogenetic category, and is associated with a distinctive microRNA signature. Oncotarget *6*, 31613-31627.

Eilken, H.M., Nishikawa, S., and Schroeder, T. (2009). Continuous single-cell imaging of blood generation from haemogenic endothelium. Nature *457*, 896-900.

Elcheva, I., Brok-Volchanskaya, V., and Slukvin, I. (2015). Direct Induction of Hemogenic Endothelium and Blood by Overexpression of Transcription Factors in Human Pluripotent Stem Cells. J Vis Exp, e52910.

Fischer, A., Schumacher, N., Maier, M., Sendtner, M., and Gessler, M. (2004). The Notch target genes Hey1 and Hey2 are required for embryonic vascular development. Genes Dev *18*, 901-911.

Frame, J.M., Fegan, K.H., Conway, S.J., McGrath, K.E., and Palis, J. (2016). Definitive Hematopoiesis in the Yolk Sac Emerges from Wnt-Responsive Hemogenic Endothelium Independently of Circulation and Arterial Identity. Stem Cells *34*, 431-444.

Garcia-Porrero, J.A., Godin, I.E., and Dieterlen-Lièvre, F. (1995). Potential intraembryonic hemogenic sites at pre-liver stages in the mouse. Anat Embryol (Berl) *192*, 425-435.

Gazit, R., Garrison, B.S., Rao, T.N., Shay, T., Costello, J., Ericson, J., Kim, F., Collins, J.J., Regev, A., Wagers, A.J.*, et al.* (2013). Transcriptome analysis identifies regulators of hematopoietic stem and progenitor cells. Stem Cell Reports *1*, 266-280.

Gekas, C., Dieterlen-Lièvre, F., Orkin, S.H., and Mikkola, H.K. (2005). The placenta is a niche for hematopoietic stem cells. Dev Cell *8*, 365-375.

Gekas, C., Rhodes, K.E., and Mikkola, H.K. (2008). Isolation and visualization of mouse placental hematopoietic stem cells. Curr Protoc Stem Cell Biol *Chapter 2*, Unit 2A.8.1-2A.8.14.

Golub, T.R., Barker, G.F., Lovett, M., and Gilliland, D.G. (1994). Fusion of PDGF receptor beta to a novel ets-like gene, tel, in chronic myelomonocytic leukemia with t(5;12) chromosomal translocation. Cell *77*, 307-316.

Goyama, S., Yamamoto, G., Shimabe, M., Sato, T., Ichikawa, M., Ogawa, S., Chiba, S., and Kurokawa, M. (2008). Evi-1 is a critical regulator for hematopoietic stem cells and transformed leukemic cells. Cell Stem Cell *3*, 207-220.

Gurney, A.L., Carver-Moore, K., de Sauvage, F.J., and Moore, M.W. (1994). Thrombocytopenia in c-mpl-deficient mice. Science *265*, 1445-1447.

Guttman, M., Amit, I., Garber, M., French, C., Lin, M.F., Feldser, D., Huarte, M., Zuk, O., Carey, B.W., Cassady, J.P.*, et al.* (2009). Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. Nature *458*, 223-227.

Guttman, M., Garber, M., Levin, J.Z., Donaghey, J., Robinson, J., Adiconis, X., Fan, L., Koziol, M.J., Gnirke, A., Nusbaum, C.*, et al.* (2010). Ab initio reconstruction of cell type-specific transcriptomes in mouse reveals the conserved multi-exonic structure of lincRNAs. Nat Biotechnol *28*, 503-510.

Hock, H., Hamblen, M.J., Rooke, H.M., Schindler, J.W., Saleque, S., Fujiwara, Y., and Orkin, S.H. (2004a). Gfi-1 restricts proliferation and preserves functional integrity of haematopoietic stem cells. Nature *431*, 1002-1007.

Hock, H., Meade, E., Medeiros, S., Schindler, J.W., Valk, P.J., Fujiwara, Y., and Orkin, S.H. (2004b). Tel/Etv6 is an essential and selective regulator of adult hematopoietic stem cell survival. Genes Dev *18*, 2336-2341.

Hogan, C.J., Shpall, E.J., and Keller, G. (2002). Differential long-term and multilineage engraftment potential from subfractions of human CD34+ cord blood cells transplanted into NOD/SCID mice. Proc Natl Acad Sci U S A *99*, 413-418.

Huyhn, A., Dommergues, M., Izac, B., Croisille, L., Katz, A., Vainchenker, W., and Coulombel, L. (1995). Characterization of hematopoietic progenitors from human yolk sacs and embryos. Blood *86*, 4474-4485.

Ivanovs, A., Rybtsov, S., Anderson, R.A., Turner, M.L., and Medvinsky, A. (2014). Identification of the niche and phenotype of the first human hematopoietic stem cells. Stem Cell Reports *2*, 449-456.

Ivanovs, A., Rybtsov, S., Welch, L., Anderson, R.A., Turner, M.L., and Medvinsky, A. (2011). Highly potent human hematopoietic stem cells first emerge in the intraembryonic aorta-gonad-mesonephros region. J Exp Med *208*, 2417-2427.

Jokubaitis, V.J., Sinka, L., Driessen, R., Whitty, G., Haylock, D.N., Bertoncello, I., Smith, I., Péault, B., Tavian, M., and Simmons, P.J. (2008). Angiotensin-converting enzyme (CD143) marks hematopoietic stem cells in human embryonic, fetal, and adult hematopoietic tissues. Blood *111*, 4055-4063.

Kennedy, M., Awong, G., Sturgeon, C.M., Ditadi, A., LaMotte-Mohs, R., Zúñiga-Pflücker, J.C., and Keller, G. (2012). T lymphocyte potential marks the emergence of definitive hematopoietic progenitors in human pluripotent stem cell differentiation cultures. Cell Rep *2*, 1722-1735.

Kennedy, M., D'Souza, S.L., Lynch-Kattman, M., Schwantz, S., and Keller, G. (2007). Development of the hemangioblast defines the onset of hematopoiesis in human ES cell differentiation cultures. Blood *109*, 2679-2687.

Kikushige, Y., Yoshimoto, G., Miyamoto, T., Iino, T., Mori, Y., Iwasaki, H., Niiro, H., Takenaka, K., Nagafuji, K., Harada, M.*, et al.* (2008). Human Flt3 is expressed at the hematopoietic stem cell and the granulocyte/macrophage progenitor stages to maintain cell survival. J Immunol *180*, 7358-7367.

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

Kim, I., Yilmaz, O.H., and Morrison, S.J. (2005). CD144 (VE-cadherin) is transiently expressed by fetal liver hematopoietic stem cells. Blood *106*, 903-905.

Kimura, S., Roberts, A.W., Metcalf, D., and Alexander, W.S. (1998). Hematopoietic stem cell deficiencies in mice lacking c-Mpl, the receptor for thrombopoietin. Proc Natl Acad Sci U S A *95*, 1195-1200.

Kingsley, P.D., Malik, J., Emerson, R.L., Bushnell, T.P., McGrath, K.E., Bloedorn, L.A., Bulger, M., and Palis, J. (2006). "Maturational" globin switching in primary primitive erythroid cells. Blood *107*, 1665-1672.

Kingsley, P.D., Malik, J., Fantauzzo, K.A., and Palis, J. (2004). Yolk sac-derived primitive erythroblasts enucleate during mammalian embryogenesis. Blood *104*, 19-25.

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

Kumano, K., Chiba, S., Kunisato, A., Sata, M., Saito, T., Nakagami-Yamaguchi, E., Yamaguchi, T., Masuda, S., Shimizu, K., Takahashi, T.*, et al.* (2003). Notch1 but not Notch2 is essential for generating hematopoietic stem cells from endothelial cells. Immunity *18*, 699-711.

Kyba, M., Perlingeiro, R.C., and Daley, G.Q. (2002). HoxB4 confers definitive lymphoid-myeloid engraftment potential on embryonic stem cell and yolk sac hematopoietic progenitors. Cell *109*, 29-37.

Lancrin, C., Mazan, M., Stefanska, M., Patel, R., Lichtinger, M., Costa, G., Vargel, O., Wilson, N.K., Möröy, T., Bonifer, C.*, et al.* (2012). GFI1 and GFI1B control the loss of endothelial identity of hemogenic endothelium during hematopoietic commitment. Blood *120*, 314-322.

Lancrin, C., Sroczynska, P., Stephenson, C., Allen, T., Kouskoff, V., and Lacaud, G. (2009). The haemangioblast generates haematopoietic cells through a haemogenic endothelium stage. Nature *457*, 892-895.

Larochelle, A., Savona, M., Wiggins, M., Anderson, S., Ichwan, B., Keyvanfar, K., Morrison, S.J., and Dunbar, C.E. (2011). Human and rhesus macaque hematopoietic stem cells cannot be purified based only on SLAM family markers. Blood *117*, 1550-1554.

Lawrence, H.J., Christensen, J., Fong, S., Hu, Y.L., Weissman, I., Sauvageau, G., Humphries, R.K., and Largman, C. (2005). Loss of expression of the Hoxa-9 homeobox gene impairs the proliferation and repopulating ability of hematopoietic stem cells. Blood *106*, 3988-3994.

Lebert-Ghali, C., Fournier, M., Kettyle, L., Thompson, A., Sauvageau, G., and Bijl, J.J. (2016). Hoxa cluster genes determine the proliferative activity of adult mouse hematopoietic stem and progenitor cells. Blood *127*, 87-90.

Lebert-Ghali, C.E., Fournier, M., Dickson, G.J., Thompson, A., Sauvageau, G., and Bijl, J.J. (2010). HoxA cluster is haploinsufficient for activity of hematopoietic stem and progenitor cells. Exp Hematol *38*, 1074-1086.e1071-1075.

Lee, L.K., Ghorbanian, Y., Wang, W., Wang, Y., Kim, Y.J., Weissman, I.L., Inlay, M.A., and Mikkola, H.K. (2016). LYVE1 Marks the Divergence of Yolk Sac Definitive Hemogenic Endothelium from the Primitive Erythroid Lineage. Cell Rep *17*, 2286-2298.

Ling, K.W., Ottersbach, K., van Hamburg, J.P., Oziemlak, A., Tsai, F.Y., Orkin, S.H., Ploemacher, R., Hendriks, R.W., and Dzierzak, E. (2004). GATA-2 plays two functionally distinct roles during the ontogeny of hematopoietic stem cells. J Exp Med *200*, 871-882.

Liu, P., Keller, J.R., Ortiz, M., Tessarollo, L., Rachel, R.A., Nakamura, T., Jenkins, N.A., and Copeland, N.G. (2003). Bcl11a is essential for normal lymphoid development. Nat Immunol *4*, 525-532.

Luc, S., Huang, J., McEldoon, J.L., Somuncular, E., Li, D., Rhodes, C., Mamoor, S., Hou, S., Xu, J., and Orkin, S.H. (2016). Bcl11a Deficiency Leads to Hematopoietic Stem Cell Defects with an Aging-like Phenotype. Cell Rep *16*, 3181-3194.

Luckett, W.P. (1978). Origin and differentiation of the yolk sac and extraembryonic mesoderm in presomite human and rhesus monkey embryos. Am J Anat *152*, 59-97.

Magnusson, M., Sierra, M.I., Sasidharan, R., Prashad, S.L., Romero, M., Saarikoski, P., Van Handel, B., Huang, A., Li, X., and Mikkola, H.K. (2013). Expansion on stromal cells preserves the undifferentiated state of human hematopoietic stem cells despite compromised reconstitution ability. PLoS One *8*, e53912.

Maillard, I., Koch, U., Dumortier, A., Shestova, O., Xu, L., Sai, H., Pross, S.E., Aster, J.C., Bhandoola, A., Radtke, F.*, et al.* (2008). Canonical notch signaling is dispensable for the maintenance of adult hematopoietic stem cells. Cell Stem Cell *2*, 356-366.

Majeti, R., Park, C.Y., and Weissman, I.L. (2007). Identification of a hierarchy of multipotent hematopoietic progenitors in human cord blood. Cell Stem Cell *1*, 635-645.

Martin, C.H., Woll, P.S., Ni, Z., Zuniga-Pflucker, J.C., and Kaufman, D.S. (2008). Differences in lymphocyte developmental potential between human embryonic stem cell and umbilical cord blood-derived hematopoietic progenitor cells. Blood *112*, 2730-2737.

McGrath, K.E., Frame, J.M., Fegan, K.H., Bowen, J.R., Conway, S.J., Catherman, S.C., Kingsley, P.D., Koniski, A.D., and Palis, J. (2015). Distinct Sources of Hematopoietic Progenitors Emerge before HSCs and Provide Functional Blood Cells in the Mammalian Embryo. Cell Rep *11*, 1892-1904.

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

Migliaccio, G., Migliaccio, A.R., Petti, S., Mavilio, F., Russo, G., Lazzaro, D., Testa, U., Marinucci, M., and Peschle, C. (1986). Human embryonic hemopoiesis. Kinetics of progenitors and precursors underlying the yolk sac----liver transition. J Clin Invest *78*, 51-60.

Mikkola, H.K., Fujiwara, Y., Schlaeger, T.M., Traver, D., and Orkin, S.H. (2003). Expression of CD41 marks the initiation of definitive hematopoiesis in the mouse embryo. Blood *101*, 508-516.

Moore, M.A., and Metcalf, D. (1970). Ontogeny of the haemopoietic system: yolk sac origin of in vivo and in vitro colony forming cells in the developing mouse embryo. Br J Haematol *18*, 279-296.

Nakajima-Takagi, Y., Osawa, M., Oshima, M., Takagi, H., Miyagi, S., Endoh, M., Endo, T.A., Takayama, N., Eto, K., Toyoda, T.*, et al.* (2013). Role of SOX17 in hematopoietic development from human embryonic stem cells. Blood *121*, 447-458.

Nakano, H., Liu, X., Arshi, A., Nakashima, Y., van Handel, B., Sasidharan, R., Harmon, A.W., Shin, J.H., Schwartz, R.J., Conway, S.J.*, et al.* (2013). Haemogenic endocardium contributes to transient definitive haematopoiesis. Nat Commun *4*, 1564.

Ng, E.S., Azzola, L., Bruveris, F.F., Calvanese, V., Phipson, B., Vlahos, K., Hirst, C., Jokubaitis, V.J., Yu, Q.C., Maksimovic, J.*, et al.* (2016). Differentiation of human embryonic stem cells to HOXA(+) hemogenic vasculature that resembles the aorta-gonad-mesonephros. Nat Biotechnol *34*, 1168-1179.

Oberlin, E., El Hafny, B., Petit-Cocault, L., and Souyri, M. (2010). Definitive human and mouse hematopoiesis originates from the embryonic endothelium: a new class of HSCs based on VE-cadherin expression. Int J Dev Biol *54*, 1165-1173.

Oguro, H., Ding, L., and Morrison, S.J. (2013). SLAM family markers resolve functionally distinct subpopulations of hematopoietic stem cells and multipotent progenitors. Cell Stem Cell *13*, 102-116.

Org, T., Duan, D., Ferrari, R., Montel-Hagen, A., Van Handel, B., Kerényi, M.A., Sasidharan, R., Rubbi, L., Fujiwara, Y., Pellegrini, M.*, et al.* (2015). Scl binds to primed enhancers in mesoderm to regulate hematopoietic and cardiac fate divergence. EMBO J *34*, 759-777.

Ottersbach, K., and Dzierzak, E. (2005). The murine placenta contains hematopoietic stem cells within the vascular labyrinth region. Dev Cell *8*, 377-387.

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

Palis, J., and Yoder, M.C. (2001). Yolk-sac hematopoiesis: the first blood cells of mouse and man. Exp Hematol *29*, 927-936.

Peebles, P.J. (2015). Dark matter. Proc Natl Acad Sci U S A *112*, 12246-12248.

Pilat, S., Carotta, S., and Klump, H. (2013). Development of hematopoietic stem and progenitor cells from mouse embryonic stem cells, in vitro, supported by ectopic human HOXB4 expression. Methods Mol Biol *1029*, 129-147.

Prashad, S.L., Calvanese, V., Yao, C.Y., Kaiser, J., Wang, Y., Sasidharan, R., Crooks, G., Magnusson, M., and Mikkola, H.K. (2014). GPI-80 Defines Self-Renewal Ability in Hematopoietic Stem Cells during Human Development. Cell Stem Cell.

Qian, H., Buza-Vidas, N., Hyland, C.D., Jensen, C.T., Antonchuk, J., Månsson, R., Thoren, L.A., Ekblom, M., Alexander, W.S., and Jacobsen, S.E. (2007). Critical role of thrombopoietin in maintaining adult quiescent hematopoietic stem cells. Cell Stem Cell *1*, 671-684.

Qiu, C., Hanson, E., Olivier, E., Inada, M., Kaufman, D.S., Gupta, S., and Bouhassira, E.E. (2005). Differentiation of human embryonic stem cells into hematopoietic cells by coculture with human fetal liver cells recapitulates the globin switch that occurs early in development. Exp Hematol *33*, 1450-1458.

Quagliata, L., Matter, M.S., Piscuoglio, S., Arabi, L., Ruiz, C., Procino, A., Kovac, M., Moretti, F., Makowska, Z., Boldanova, T.*, et al.* (2014). Long noncoding RNA HOTTIP/HOXA13 expression is associated with disease progression and predicts outcome in hepatocellular carcinoma patients. Hepatology *59*, 911-923.

Quinn, J.J., Zhang, Q.C., Georgiev, P., Ilik, I.A., Akhtar, A., and Chang, H.Y. (2016). Rapid evolutionary turnover underlies conserved lncRNA-genome interactions. Genes Dev *30*, 191-207.

Rhodes, K.E., Gekas, C., Wang, Y., Lux, C.T., Francis, C.S., Chan, D.N., Conway, S., Orkin, S.H., Yoder, M.C., and Mikkola, H.K. (2008). The emergence of hematopoietic stem cells is initiated in the placental vasculature in the absence of circulation. Cell Stem Cell *2*, 252-263.

Rinn, J.L., Kertesz, M., Wang, J.K., Squazzo, S.L., Xu, X., Brugmann, S.A., Goodnough, L.H., Helms, J.A., Farnham, P.J., Segal, E.*, et al.* (2007). Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. Cell *129*, 1311-1323.

Robert-Moreno, A., Guiu, J., Ruiz-Herguido, C., López, M.E., Inglés-Esteve, J., Riera, L., Tipping, A., Enver, T., Dzierzak, E., Gridley, T.*, et al.* (2008). Impaired embryonic haematopoiesis yet normal arterial development in the absence of the Notch ligand Jagged1. EMBO J *27*, 1886-1895.

Robin, C., Bollerot, K., Mendes, S., Haak, E., Crisan, M., Cerisoli, F., Lauw, I., Kaimakis, P., Jorna, R., Vermeulen, M.*, et al.* (2009). Human placenta is a potent hematopoietic niche containing hematopoietic stem and progenitor cells throughout development. Cell Stem Cell *5*, 385-395.

Sankaran, V.G., Menne, T.F., Xu, J., Akie, T.E., Lettre, G., Van Handel, B., Mikkola, H.K., Hirschhorn, J.N., Cantor, A.B., and Orkin, S.H. (2008). Human fetal hemoglobin expression is regulated by the developmental stage-specific repressor BCL11A. Science *322*, 1839-1842.

Sankaran, V.G., Xu, J., and Orkin, S.H. (2010). Transcriptional silencing of fetal hemoglobin by BCL11A. Ann N Y Acad Sci *1202*, 64-68.

Schuettengruber, B., Chourrout, D., Vervoort, M., Leblanc, B., and Cavalli, G. (2007). Genome regulation by polycomb and trithorax proteins. Cell *128*, 735-745.

Schwartz, Y.B., and Pirrotta, V. (2007). Polycomb silencing mechanisms and the management of genomic programmes. Nat Rev Genet *8*, 9-22.

Soshnikova, N. (2013). Hox genes regulation in vertebrates. Dev Dyn.

Soshnikova, N., and Duboule, D. (2009). Epigenetic regulation of vertebrate Hox genes: a dynamic equilibrium. Epigenetics *4*, 537-540.

Souilhol, C., Lendinez, J.G., Rybtsov, S., Murphy, F., Wilson, H., Hills, D., Batsivari, A., Binagui-Casas, A., McGarvey, A.C., MacDonald, H.R.*, et al.* (2016). Developing HSCs become Notch independent by the end of maturation in the AGM region. Blood *128*, 1567-1577.

Stamatoyannopoulos, G. (2005). Control of globin gene expression during development and erythroid differentiation. Exp Hematol *33*, 259-271.

Tavian, M., Hallais, M.F., and Péault, B. (1999). Emergence of intraembryonic hematopoietic precursors in the pre-liver human embryo. Development *126*, 793-803.

Tavian, M., and Péault, B. (2005). Embryonic development of the human hematopoietic system. Int J Dev Biol *49*, 243-250.

Tavian, M., Robin, C., Coulombel, L., and Peault, B. (2001). The human embryo, but not its yolk sac, generates lympho-myeloid stem cells: mapping multipotent hematopoietic cell fate in intraembryonic mesoderm. Immunity *15*, 487-495.

Trimarchi, T., Bilal, E., Ntziachristos, P., Fabbri, G., Dalla-Favera, R., Tsirigos, A., and Aifantis, I. (2014). Genome-wide mapping and characterization of Notch-regulated long noncoding RNAs in acute leukemia. Cell *158*, 593-606.

Tsai, F.Y., Keller, G., Kuo, F.C., Weiss, M., Chen, J., Rosenblatt, M., Alt, F.W., and Orkin, S.H. (1994). An early haematopoietic defect in mice lacking the transcription factor GATA-2. Nature *371*, 221-226.

Tsai, F.Y., and Orkin, S.H. (1997). 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 *89*, 3636-3643.

Uchida, N., Fujisaki, T., Eaves, A.C., and Eaves, C.J. (2001). Transplantable hematopoietic stem cells in human fetal liver have a CD34(+) side population (SP)phenotype. J Clin Invest *108*, 1071-1077.

van der Meer, L.T., Jansen, J.H., and van der Reijden, B.A. (2010). Gfi1 and Gfi1b: key regulators of hematopoiesis. Leukemia *24*, 1834-1843.

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

Van Handel, B., Prashad, S.L., Hassanzadeh-Kiabi, N., Huang, A., Magnusson, M., Atanassova, B., Chen, A., Hamalainen, E.I., and Mikkola, H.K. (2010). The first trimester human placenta is a site for terminal maturation of primitive erythroid cells. Blood *116*, 3321-3330.

Vodyanik, M.A., Thomson, J.A., and Slukvin, I.I. (2006). Leukosialin (CD43) defines hematopoietic progenitors in human embryonic stem cell differentiation cultures. Blood *108*, 2095-2105.

Wan, L., Kong, J., Tang, J., Wu, Y., Xu, E., Lai, M., and Zhang, H. (2016). HOTAIRM1 as a potential biomarker for diagnosis of colorectal cancer functions the role in the tumour suppressor. J Cell Mol Med *20*, 2036-2044.

Wang, K.C., and Chang, H.Y. (2011). Molecular mechanisms of long noncoding RNAs. Mol Cell *43*, 904-914.

Wang, K.C., Yang, Y.W., Liu, B., Sanyal, A., Corces-Zimmerman, R., Chen, Y., Lajoie, B.R., Protacio, A., Flynn, R.A., Gupta, R.A.*, et al.* (2011a). A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. Nature *472*, 120-124.

Wang, L.C., Kuo, F., Fujiwara, Y., Gilliland, D.G., Golub, T.R., and Orkin, S.H. (1997). Yolk sac angiogenic defect and intra-embryonic apoptosis in mice lacking the Ets-related factor TEL. EMBO J *16*, 4374-4383.

Wang, L.C., Swat, W., Fujiwara, Y., Davidson, L., Visvader, J., Kuo, F., Alt, F.W., Gilliland, D.G., Golub, T.R., and Orkin, S.H. (1998). The TEL/ETV6 gene is required specifically for hematopoiesis in the bone marrow. Genes Dev *12*, 2392-2402.

Wang, X.Q., Crutchley, J.L., and Dostie, J. (2011b). Shaping the Genome with Non-Coding RNAs. Curr Genomics *12*, 307-321.

Wang, X.Q., and Dostie, J. (2016). Reciprocal regulation of chromatin state and architecture by HOTAIRM1 contributes to temporal collinear HOXA gene activation. Nucleic Acids Res.

White, J.R., and Weston, K. (2000). Myb is required for self-renewal in a model system of early hematopoiesis. Oncogene *19*, 1196-1205.

Yang, Y.W., Flynn, R.A., Chen, Y., Qu, K., Wan, B., Wang, K.C., Lei, M., and Chang, H.Y. (2014). Essential role of lncRNA binding for WDR5 maintenance of active chromatin and embryonic stem cell pluripotency. Elife *3*, e02046.

Yoshida, H., Kawane, K., Koike, M., Mori, Y., Uchiyama, Y., and Nagata, S. (2005). Phosphatidylserine-dependent engulfment by macrophages of nuclei from erythroid precursor cells. Nature *437*, 754-758.

Yu, Y., Wang, J., Khaled, W., Burke, S., Li, P., Chen, X., Yang, W., Jenkins, N.A., Copeland, N.G., Zhang, S.*, et al.* (2012). Bcl11a is essential for lymphoid development and negatively regulates p53. J Exp Med *209*, 2467-2483.

Zambidis, E.T., Peault, B., Park, T.S., Bunz, F., and Civin, C.I. (2005). Hematopoietic differentiation of human embryonic stem cells progresses through sequential hematoendothelial, primitive, and definitive stages resembling human yolk sac development. Blood *106*, 860-870.

Zambidis, E.T., Sinka, L., Tavian, M., Jokubaitis, V., Park, T.S., Simmons, P., and Péault, B. (2007). Emergence of human angiohematopoietic cells in normal development and from cultured embryonic stem cells. Ann N Y Acad Sci *1106*, 223-232.

Zeng, H., Yücel, R., Kosan, C., Klein-Hitpass, L., and Möröy, T. (2004). Transcription factor Gfi1 regulates self-renewal and engraftment of hematopoietic stem cells. EMBO J *23*, 4116-4125.

Zhang, H., Zhao, L., Wang, Y.X., Xi, M., Liu, S.L., and Luo, L.L. (2015). Long non-coding RNA HOTTIP is correlated with progression and prognosis in tongue squamous cell carcinoma. Tumour Biol *36*, 8805-8809.

Zhang, X., Lian, Z., Padden, C., Gerstein, M.B., Rozowsky, J., Snyder, M., Gingeras, T.R., Kapranov, P., Weissman, S.M., and Newburger, P.E. (2009). A myelopoiesis-associated regulatory intergenic noncoding RNA transcript within the human HOXA cluster. Blood *113*, 2526-2534.

Zovein, A.C., Hofmann, J.J., Lynch, M., French, W.J., Turlo, K.A., Yang, Y., Becker, M.S., Zanetta, L., Dejana, E., Gasson, J.C.*, et al.* (2008). Fate tracing reveals the endothelial origin of hematopoietic stem cells. Cell Stem Cell *3*, 625-636.

Chapter 2:

Medial HOXA genes demarcate haematopoietic stem cell fate during human development

ARTICLES

nature cell biology

Medial *HOXA* genes demarcate haematopoietic stem cell fate during human development

Diana R. Dou[1,2,3,8], Vincenzo Calvanese[1,2,8], Maria I. Sierra[1,2], Andrew T. Nguyen[1], Arazin Minasian[1,2], Pamela Saarikoski[1,2], Rajkumar Sasidharan[1,2], Christina M. Ramirez[4], Jerome A. Zack[2,5,6], Gay M. Crooks[1,2,7], Zoran Galic[2,5] and Hanna K. A. Mikkola[1,2,3,9]

Pluripotent stem cells (PSCs) may provide a potential source of haematopoietic stem/progenitor cells (HSPCs) for transplantation; however, unknown molecular barriers prevent the self-renewal of PSC-HSPCs. Using two-step differentiation, human embryonic stem cells (hESCs) differentiated *in vitro* into multipotent haematopoietic cells that had the CD34$^+$CD38$^{-/lo}$CD90$^+$CD45$^+$ GPI-80$^+$ fetal liver (FL) HSPC immunophenotype, but exhibited poor expansion potential and engraftment ability. Transcriptome analysis of immunophenotypic hESC-HSPCs revealed that, despite their molecular resemblance to FL-HSPCs, medial *HOXA* genes remained suppressed. Knockdown of *HOXA7* disrupted FL-HSPC function and caused transcriptome dysregulation that resembled hESC-derived progenitors. Overexpression of medial *HOXA* genes prolonged FL-HSPC maintenance but was insufficient to confer self-renewal to hESC-HSPCs. Stimulation of retinoic acid signalling during endothelial-to-haematopoietic transition induced the *HOXA* cluster and other HSC/definitive haemogenic endothelium genes, and prolonged HSPC maintenance in culture. Thus, medial *HOXA* gene expression induced by retinoic acid signalling marks the establishment of the definitive HSPC fate and controls HSPC identity and function.

Haematopoietic stem cells (HSCs) regenerate the blood system on transplantation, and can therefore cure inherited and acquired blood diseases. However, lack of HLA (human leukocyte antigen)-matched bone marrow or cord blood donors limits their therapeutic use[1]. Generation of HSCs from human embryonic stem cells (hESCs) or induced pluripotent stem cells (PSCs) could provide alternative HSC sources. Recent studies used transcription factor reprogramming to convert fibroblasts or mature blood cells[2–4] to haematopoietic cells possessing some properties of HSCs. Despite these promising approaches, clinical application of *in vitro*-generated HSCs remains unachieved.

Although hESCs can differentiate into most blood lineages[5], efforts to produce engraftable HSCs have failed[6]. The molecular barriers preventing HSC generation *in vitro* are poorly understood owing to lack of studies comparing candidate HSCs from PSC cultures and human conceptus that match by immunophenotype and developmental stage. During embryogenesis, haematopoiesis starts in the yolk sac by the generation of two distinct waves of myelo-erythroid progenitors (primitive and transient definitive) that can be distinguished by the specific globins expressed in their progeny[7]. These progenitors lack self-renewal ability and robust lymphoid potential[8,9]. Definitive HSCs possessing these properties emerge in the third haematopoietic wave from specialized haemogenic endothelium in major arteries in the AGM (aorta–gonad–mesonephros) region, yolk sac, placenta and vitelline and umbilical vessels[10]. Human haemogenic endothelial cells express CD34 and CD31[11] and upregulate CD43 on haematopoietic commitment[12,13], whereas HSCs also co-express CD45 (pan-haematopoietic), CD90 (HSC, endothelium) and GPI-80 (human fetal HSCs[14]), and typically have low CD38 expression (lineage commitment/HSC activation). Haematopoietic differentiation of mouse and human ESCs mirrors embryonic haematopoiesis[8,15] and recapitulates mesoderm and haemato-vascular commitment[16,17] followed by waves of primitive and definitive erythropoiesis[18,19]. However, hESC-derived haematopoietic cells lack reconstitution ability[6,20,21] and full lymphoid and adult-type erythroid potential[22,23], resembling yolk sac-derived lineage-restricted progenitors[24]. A long-standing goal has been to

[1]Department of Molecular, Cell and Developmental Biology, University of California, Los Angeles, Los Angeles, California 90095, USA. [2]Eli and Edythe Broad Center for Regenerative Medicine and Stem Cell Research, University of California, Los Angeles, Los Angeles, California 90095, USA. [3]Molecular Biology Institute, University of California, Los Angeles, Los Angeles, California 90095, USA. [4]Department of Biostatistics, Fielding School of Public Health, University of California, Los Angeles, Los Angeles, California 90095, USA. [5]Department of Medicine, Division of Hematology-Oncology, University of California, Los Angeles, Los Angeles, California 90095, USA. [6]Department of Microbiology, Immunology and Molecular Genetics, University of California, Los Angeles, Los Angeles, California 90095, USA. [7]Department of Pathology and Laboratory Medicine, University of California, Los Angeles, Los Angeles, California 90095, USA. [8]These authors contributed equally to this work. [9]Correspondence should be addressed to H.K.A.M. (e-mail: hmikkola@mcdb.ucla.edu)

Received 22 December 2014; accepted 8 April 2016; published online 16 May 2016; DOI: 10.1038/ncb3354

identify regulatory cues and molecular landmarks that distinguish the definitive HSC fate from the short-lived embryonic progenitors.

We used a two-step hESC differentiation to generate HSPCs with human fetal HSC surface phenotype ($CD45^+CD34^+CD38^{-/lo}CD90^+GPI\text{-}80^+$). Molecular profiling showed remarkable resemblance of hESC-HSPCs to FL-HSPCs, yet revealed distinct differences in HSC regulatory programs, including the *HOXA* genes. Knockdown and overexpression studies revealed that medial *HOXA* genes, in particular *HOXA7*, govern definitive HSPC identity and function. Rescue of retinoic acid (RA) signalling during endothelial-to-haematopoietic transition induced medial *HOXA* genes and the definitive haematopoietic program in hESC-HSPCs. These studies uncover regulatory programs that distinguish human definitive HSC lineage from embryonic progenitors and offer a blueprint for identifying missing cues required for HSC generation.

RESULTS

hESCs generate HSPCs that are unable to engraft

To identify barriers for generating HSCs from PSCs, we compared hESC-derived HSPCs and human fetal HSCs on the basis of immunophenotypic, functional and molecular criteria. Second-trimester FL-HSCs were used as controls as they are ontologically closer to PSC-derived cells than cord blood or bone marrow. A two-step differentiation protocol was employed to generate haematopoietic cells from H1 hESCs (Fig. 1a). To promote haemato-vascular differentiation, embryoid bodies (EBs) were cultured with BMP4 (days 4–10), FLT3L and SCF (days 4–14)[25]. Day-14 EBs had generated cells that co-expressed CD34, CD90 and CD43 (Fig. 1b), but not CD45. To promote the development of haemato-vascular precursors towards HSCs, EB-derived $CD34^+$ cells were plated with TPO, SCF and FLT3L on OP9-M2 stroma, which supports the expansion of multipotent human HSPCs[26]. Two-week OP9-M2 co-culture (EB-OP9) generated cells with $CD34^+CD38^{-/lo}CD90^+CD45^+$ HSPC immunophenotype (Fig. 1b) and some co-expressed the HSC marker GPI-80 (Fig. 1c). These data showed that two-step differentiation generates cells with the immunophenotype of human fetal HSCs.

To assess whether the hESC-HSPC maturation culture on OP9-M2 stroma confers functional properties of HSCs, $CD34^+$ cells from EBs (isolated directly or after two-step differentiation), and FL (isolated directly or after 2-week culture) were transplanted into sublethally irradiated *NOD–scid IL2Rγ*-null (NSG) mice (Supplementary Fig. 1A). Human $CD45^+$ chimaerism in bone marrow was measured 12 weeks post-transplantation. Although FL-HSPCs engrafted successfully before or after OP9-M2 culture, hESC-derived cells showed minimal engraftment (Fig. 1d). Human $CD45^+$ cells in the bone marrow of mice transplanted with FL contained HSPCs (Supplementary Fig. 1B), $CD19^+$ B cells, $CD3^+$ T cells and $CD13^+$ or $CD66^+$ myeloid cells, whereas the mice transplanted with hESC-derived cells harboured only rare human myeloid cells (Fig. 1e). These data show that hESC-HSPCs are severely impaired functionally.

hESC-HSPCs have poor proliferative potential

To understand the functional defects in hESC-HSPCs, hESC- and cultured FL-HSPCs ($CD34^+CD38^{-/lo}CD90^+CD45^+$) were sorted and re-plated on OP9-M2 co-culture to assess their expansion (Fig. 2a). Both FL- and hESC-HSPC cultures maintained an immunophenotypic HSPC population 1 week later (Fig. 2b,c); however, at 3 weeks, hESC-HSPCs had disappeared (Fig. 2b,c). BrdU incorporation analysis did not reveal differences in cell cycle between FL- and hESC-HSPCs (Supplementary Fig. 2A), suggesting that loss of hESC-HSPCs was not due to inability to divide.

hESC-HSPCs also produced fewer clonogenic progenitors on OP9-M2 than FL-HSPCs (Fig. 2d), although both formed erythroid (BFU-E), granulocyte-macrophage (CFU-GM), macrophage (CFU-M) and mixed myelo-erythroid (CFU-mixed) colonies (Fig. 2e). CFU-C potential in EB-OP9 cells was higher in the $CD34^+CD38^{-/lo}CD90^+CD45^+$ fraction than total $CD34^+$ cells (Supplementary Fig. 2B), which also showed minimal CFU-C expansion (Supplementary Fig. 2C). These data indicate that hESC-derived immunophenotypic HSPCs exhibit poor proliferative potential.

Previous studies documented impaired differentiation of hESCs to adult-type erythroid cells[23]. Colonies from EB $CD34^+$ cells generated erythroid cells that predominantly expressed embryonic ε-globin, whereas fetal γ-globin (*HBG1*) was low (16.3% of FL levels) and adult β-globin (*HBB*) nearly undetectable (1.8% of FL levels). However, colonies derived from EB $CD34^+$ cells after 2 weeks on OP9-M2 induced some expression of γ- and β-globins (84.9% and 54.0% of FL-derived cells, respectively; Supplementary Fig. 2D). β-globin induction suggests that hESC-HSPCs can differentiate to adult-type erythroid cells if given adequate maturation time in culture.

T-cell generation is another hallmark of definitive haematopoiesis[27]. Both EB- and FL-derived $CD34^+$ cells generated CD4 and CD8 single- and double-positive T-lymphoid cells in OP9-DL1 culture (Supplementary Fig. 2E). These results imply that the main defect in hESC-derived haematopoietic cells is not differentiation, but self-renewal.

Microarray uncovers molecular defects in hESC-HSPCs

To understand the defective self-renewal of hESC-HSPCs, we assessed their relationship to primary human fetal HSPCs using microarray analysis. Gene expression differences were compared between immunophenotypic HSPCs ($CD34^+CD38^{-/lo}CD90^+CD43^+CD45^{+/-}$) derived from hESCs through EB or two-step differentiation (EB-OP9), and from FL. To assess culture-induced effects, FL-HSPCs were cultured on OP9-M2 for 2 or 5 weeks (FL-OP9). $CD34^+CD38^{-/lo}CD90^+CD43^+$ cells from early human placenta (PL) were used as a reference for immature human HSPCs (Supplementary Table 1). Spearman rank coefficient comparisons and hierarchical clustering of samples revealed that FL-HSPCs are more similar to EB-OP9-HSPCs than EB- or PL-HSPCs (Fig. 3a,b). EB-OP9-HSPCs were most similar to cultured FL-HSPCs, implying that hESC-HSPC transcriptome was influenced by prolonged culture (Fig. 3a). Many transcription factors governing the development of definitive HSPCs (*SCL/TAL1, GATA2, RUNX1, MYB, ETV6, HOXB4, GFI1B, BCL11A* and so on) were expressed in both EB-OP9-HSPCs and FL-HSPCs (Fig. 3c). These data revealed that EB-OP9-HSPCs are remarkably similar to FL-HSPCs at the molecular level.

To identify co-regulated and differentially expressed programs between samples, K-means clustering and DAVID GO (gene ontology) analysis was performed for all differentially expressed genes (>2-fold, P value <0.05; Supplementary Tables 2 and 3). K-means clustering

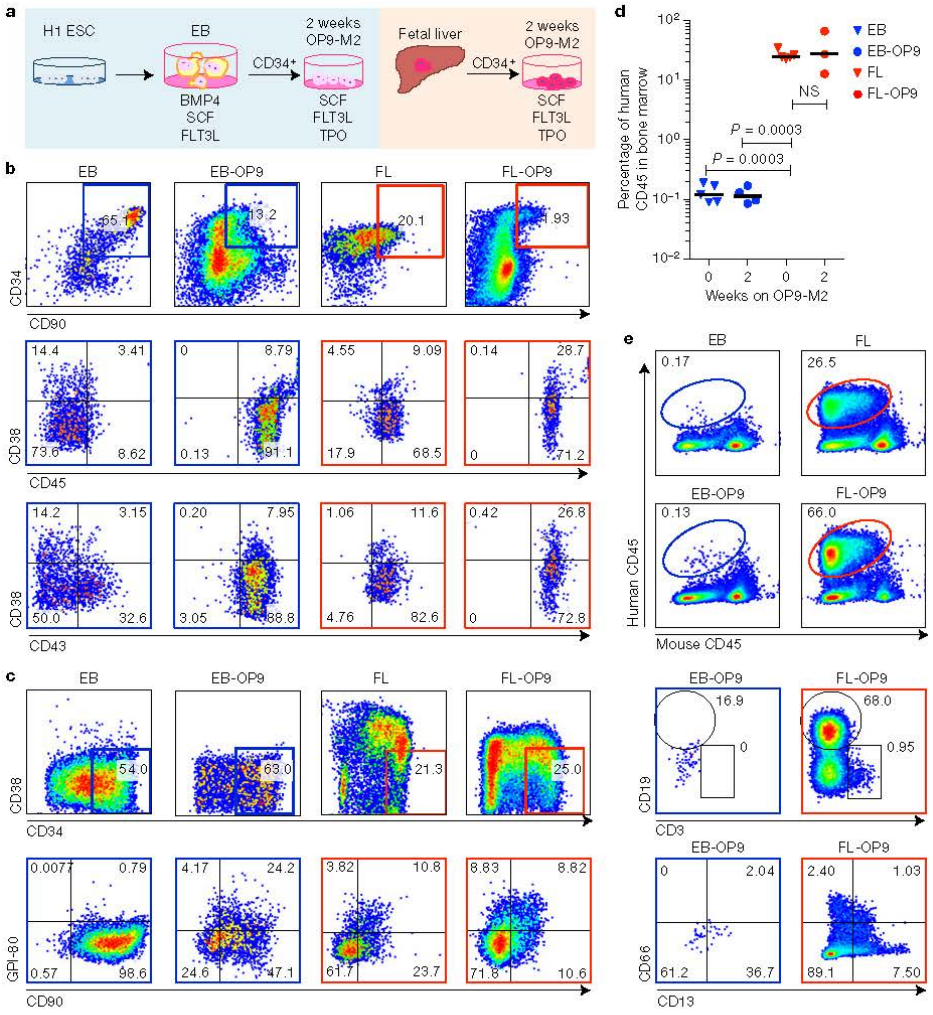


Figure 1 Two-step culture of hESCs generates immunophenotypic HSPCs that engraft poorly. (**a**) Culture and isolation strategy for differentiating H1 hESCs to HSPCs. (**b**) Representative FACS plots from 11 experiments staining for CD34, CD90, CD38, CD45 and CD43 on hESC-derived CD34$^+$ cells isolated from 2-week EBs (EB) and after 2-week maturation culture on OP9-M2 (EB-OP9), compared with cells from second-trimester fetal liver that were isolated directly (FL) or cultured on OP9-M2 (FL-OP9). (**c**) Representative FACS plots from nine experiments staining for CD38, CD34, CD90 and human fetal HSC self-renewal marker GPI-80 on hESC- and FL-derived cells. (**d**) Human engraftment in NSG mice with hESC-derived and FL-derived CD34$^+$ cells, before and after OP9-M2 co-culture (individual values and mean are shown; $n=5$ EB, $n=4$ EB-OP9 and FL, and $n=3$ FL-OP9 transplanted mice; statistical significance was calculated using the Wilcoxon rank sum test; see Supplementary Table 7 for statistics source data). (**e**) Representative FACS plots showing the human CD45$^+$ fraction in the mouse bone marrow 12 weeks post-transplantation. Multi-lineage engraftment is assessed by CD19 and CD3 (B and T lymphoid), and CD66 and CD13 (myeloid) stainings.

showed that OP9-M2 co-culture of EB-derived cells induced many FL-HSPC-associated genes (Fig. 3d, clusters 2 and 9) that encoded DNA-repair factors and transcriptional regulators (Fig. 3e), whereas vascular genes were downregulated (Fig. 3d, clusters 5 and 6, Fig. 3f).

K-means clustering also identified genes that remained dysregulated in hESC-HSPCs. The most striking differences were observed in clusters 4 and 8 that contained genes highly expressed in FL-HSPCs but suppressed in hESC-HSPCs and early PL. The most highly enriched GO category was 'Transcription' (Fig. 3d), which included many *HOXA* genes. All *HOXA* genes except for *HOXA13* (a regulator of placental vascular labyrinth specification[28]) were severely suppressed in hESC-derived HSPCs (Fig. 3g). A similar

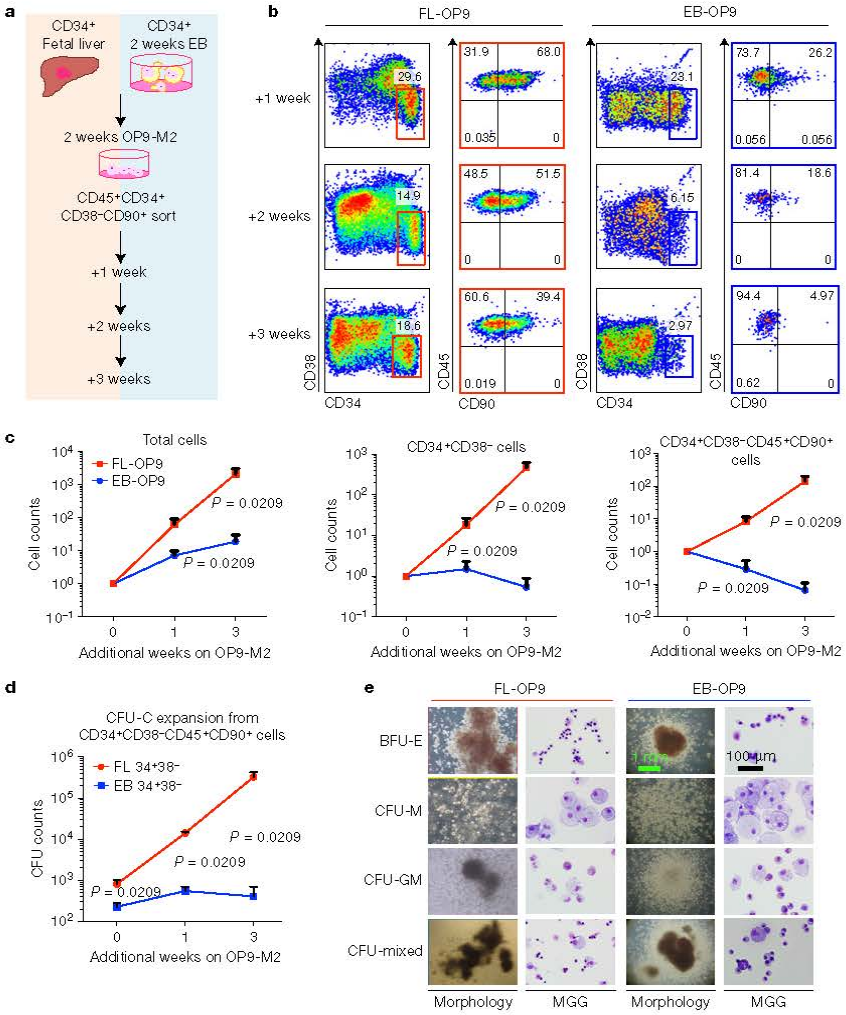


Figure 2 hESC-derived haematopoietic cells have limited proliferative potential *in vitro*. (**a**) Strategy for comparing the expansion of hESC- and FL-HSPCs. (**b**) FACS staining for the HSPC surface markers CD34$^+$, CD38$^{-/lo}$, CD45$^+$ and CD90$^+$ at various time points in OP9-M2 co-culture. (**c**) Expansion of FL- and hESC-derived haematopoietic cells sorted for HSPC phenotype after two-step differentiation, and cultured for additional weeks on OP9-M2. Mean ± s.e.m. (upward bars) from $n=4$ experiments; statistical significance was assessed using the Wilcoxon rank sum test. (**d**) CFU-C expansions from 10,000 hESC-derived or FL-derived immunophenotypic HSPCs in methylcellulose following 0, 1 and 3 additional weeks on OP9-M2 co-culture. Mean ± s.e.m. (upward bars) from $n=4$ experiments; statistical significance was assessed using the Wilcoxon rank sum test. (**e**) The morphology of myelo-erythroid colonies generated from hESC- or FL-HSPCs on methylcellulose as assessed by light microscopy and May–Grünwald–Giemsa (MGG) staining. E, erythroid; M, macrophage; GM, granulocyte-macrophage; mixed, mixed myelo-erythroid. Statistics source data used to generate graphs in **c** and **d** can be found in Supplementary Table 7.

HOXA pattern was observed with early PL, raising the hypothesis that medial *HOXA* gene induction in FL-HSPCs reflects developmental maturation and acquisition of HSC properties. Clusters 4 and 8 also included the HSC regulators *HMGN, HLF, PRDM16* and *MECOM/EVI1*[29,30], and the HSC surface markers *PROM1/CD133, EMCN* and *ROBO4* (Fig. 3h). These analyses demonstrated that the two-step conditioning fails to induce transcriptional regulators highly expressed in FL-HSCs, including *HOXA* genes.

Medial *HOXA* genes govern human HSPC function

We next asked whether *HOXA* gene silencing contributes to the poor self-renewal of hESC-HSPCs. Microarray analysis of human

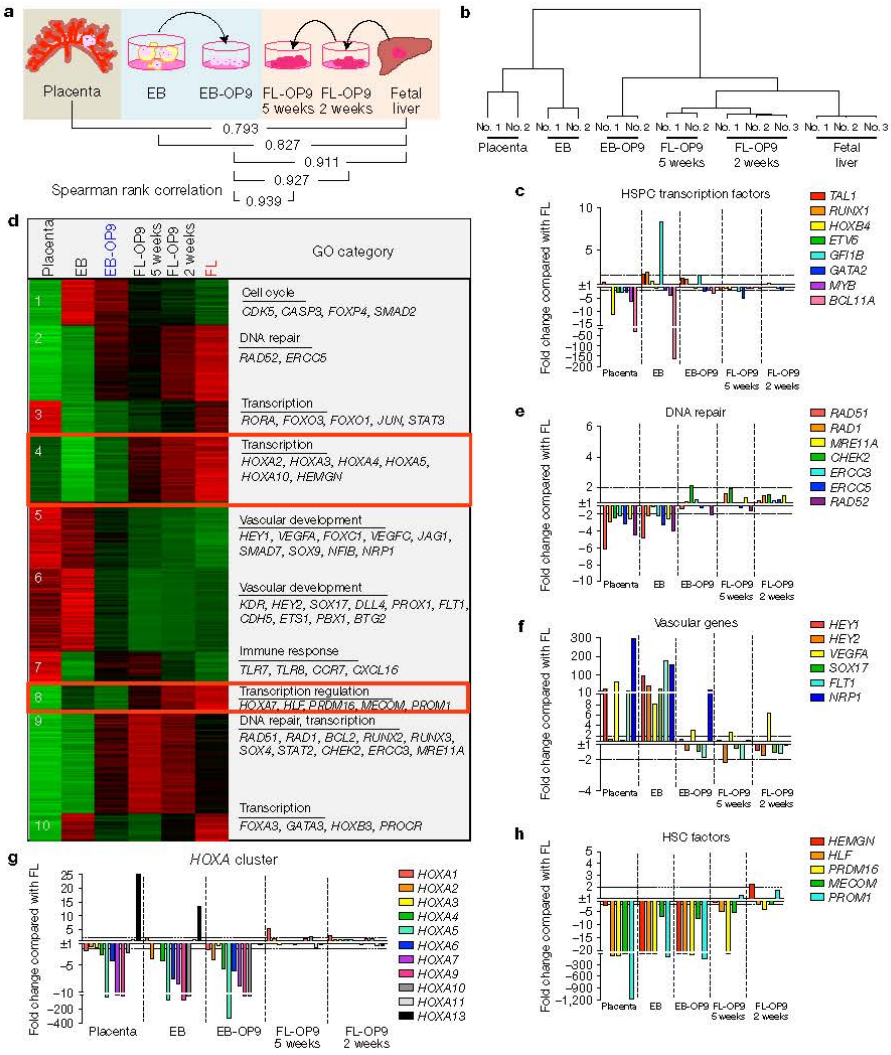


Figure 3 Identification of differentially expressed programs in hESC- and FL-HSPCs. (**a**) Spearman rank correlation of HSPCs isolated at different stages of development: 3–5-week placenta ($CD34^+CD38^{-/lo}CD90^+CD43^+$ $n=2$), hESC-HSPCs isolated from 2-week EBs (EB, $CD34^+CD38^{-/lo}CD90^+CD43^+$ $n=2$) or after two-step differentiation (EB-OP9, $CD34^+CD38^{-/lo}CD90^+CD43^+CD45^+$ $n=2$), and second-trimester FL isolated freshly (FL, $CD34^+CD38^{-/lo}CD90^+CD45^+$ $n=3$) or after 2 or 5 weeks on OP9-M2 (FL-OP9, $CD34^+CD38^{-/lo}CD90^+CD45$) ($n=3$ and $n=2$, respectively). n represents the number of tissue samples collected from separate specimens per condition. Each replicate was collected from independent experiments and analysed together. (**b**) Dendrogram showing hierarchical clustering of microarray samples. (**c**) Relative levels of haematopoietic transcription factors in different samples compared with FL-HSPCs. (**d**) K-means clustering of differentially expressed genes in HSPCs from different stages of human haematopoietic development with representative examples of gene ontology (GO) terms and genes in clusters (see Supplementary Tables 2 and 3 and GEO database GSE64865). (**e–h**) Relative levels of gene expression plotted as probe values for the indicated genes. (**e**,**f**) Levels of DNA-repair genes compared with FL-HSPCs from clusters 2 and 9 (**e**), vascular genes from clusters 5 and 6 (**f**). (**g**,**h**) *HOXA* genes (**g**) and other HSC factors (**h**) from clusters 4 and 8. See Supplementary Tables 1 and 3, and GEO database GSE64865 for values.

FL-HSPC subsets[14] documented expression of several medial *HOXA* genes in GPI-80$^+$ HSCs and their immediate progeny (Fig. 4a), and downregulation on differentiation (Fig. 4b). As *HOXA9* is a known regulator of mouse HSC proliferation[31,32], FL-HSPCs were transduced with pLKO.1 shRNA lentiviral vectors targeting *HOXA5* or *HOXA7* (Fig. 4c) to test whether other medial *HOXA* genes regulate human

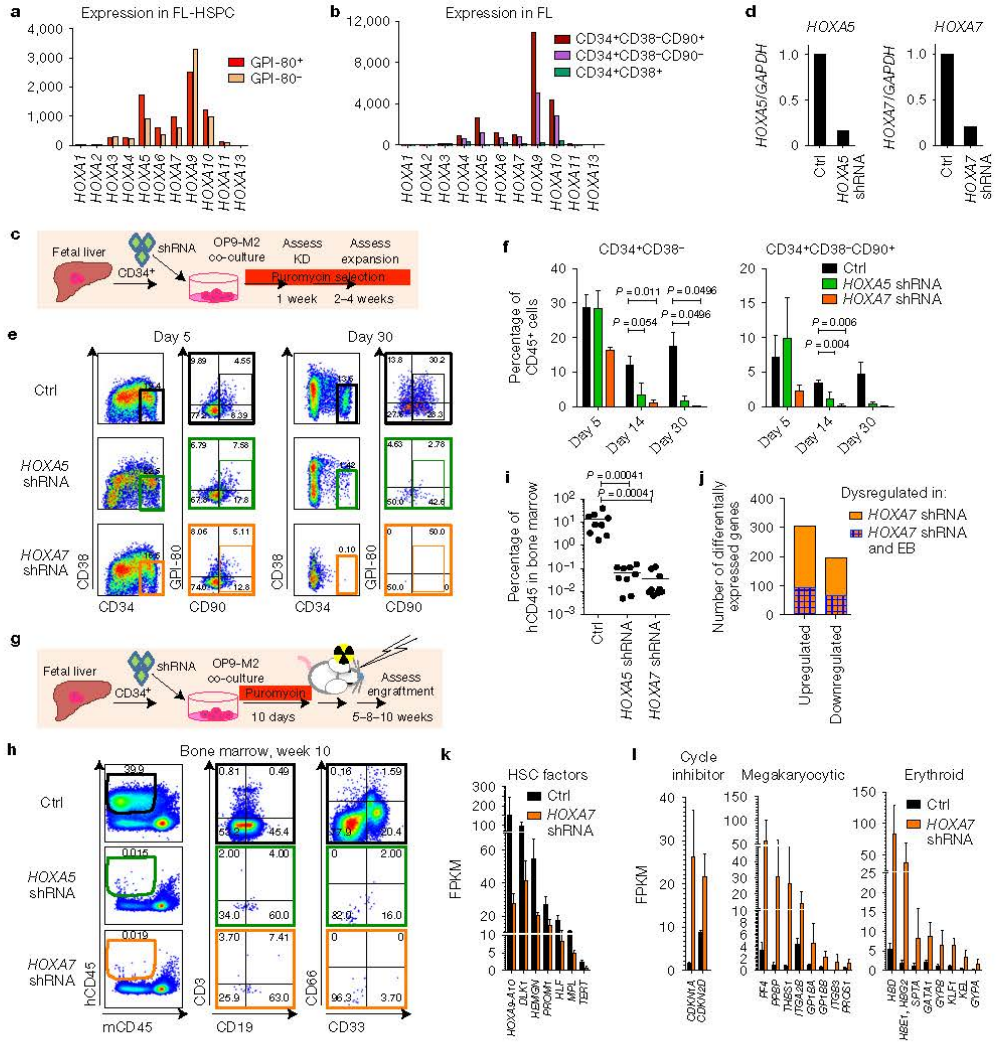


Figure 4 Medial *HOXA* genes govern the function and identity of human fetal HSPCs. (**a**,**b**) Microarray analysis of *HOXA* gene expression in CD34+CD38−/loCD90+GPI-80+ cells and their progeny (mean values are shown; left, $n=3$ samples, GEO database GSE54316; and right, $n=3$ samples (CD34+CD38−CD90+) or 2 (CD34+CD38−CD90− and CD34+CD38+), GSE34974. (**c**) A schematic showing the strategy for lentiviral shRNA knockdown of *HOXA5* or *HOXA7* in FL-HSPCs. (**d**) Knockdown is confirmed using qRT–PCR 1 week post-infection (mean ± s.d. shown from $n=3$ different FL samples). (**e**) Representative FACS plots 30 days after *HOXA5* or *HOXA7* knockdown. (**f**) Quantification of HSPC subsets in empty-vector (Ctrl)- and shRNA-infected cells (*HOXA5* shRNA or *HOXA7* shRNA) after 5, 14 and 30 days in culture (mean and s.e.m., $n=6$ independent experiments per condition for day 14 and $n=3$ for days 5 and 30). Statistical significance was assessed using the Wilcoxon signed rank test. (**g**) A schematic showing the transplantation strategy with *HOXA5*- or *HOXA7*-knockdown FL-HSPCs. (**h**) Representative FACS plots from mouse bone marrow 10 weeks post-transplantation assessing human CD45+ cells and multi-lineage engraftment (CD19 and CD3 for B and T lymphoid (middle column), and CD66 and CD33 for myeloid (right column)). (**i**) Quantification of human engraftment ($n=9$ mice per condition from three independent experiments; individual values and mean are shown). Statistical significance was assessed using the Wilcoxon rank sum test. (**j**) RNA sequencing of *HOXA7*-knockdown FL-HSPCs at day 5 post-infection. Numbers of genes up- or downregulated in *HOXA7* shRNA FL-HSPCs are shown. Genes dysregulated both in *HOXA7*-knockdown FL-HSPCs (RNA-Seq 1.8-fold change, $n=4$ independent experiments, P value < 0.05) and in EB-OP9-HSPCs compared with FL-HSPCs (microarray, twofold change, $P<0.05$) are shown in blue pattern overlay. (**k**,**l**) Bar plots showing gene expression of HSC factors downregulated in *HOXA7*-knockdown FL-HSPCs (**k**) and differentiation-associated genes upregulated (**l**) in *HOXA7*-knockdown FL-HSPCs. Mean fragments per kilobase of exon per million fragments mapped (FPKM) from $n=4$ independent specimens are shown; values used to generate graphs can be found in Supplementary Table 4 and GEO database GSE76685. See Supplementary Table 7 for statistics source data for **d**,**f**,**i**.

HSCs. Knockdown was confirmed using quantitative PCR with reverse transcription (qRT–PCR) 1 week post-transduction (Fig. 4d and Supplementary Table 6). By 2–4 weeks, *HOXA5* and *HOXA7* shRNA-treated cells were depleted of HSPCs (Fig. 4e,f). Cell cycle analysis 1 week post-transduction did not reveal significant differences in BrdU incorporation of *HOXA5*- or *HOXA7*-knockdown HSPCs, (Supplementary Fig. 3A), implying that *HOXA5*- and *HOXA7*-deficient HSPCs can divide, but cannot self-renew. Transplantation of *HOXA5* and *HOXA7* shRNA-transduced FL-HSPCs into NSG mice (Fig. 4g) showed minimal engraftment (Fig. 4h,i). These results suggest that both *HOXA5* and *HOXA7* are necessary for human FL-HSPC expansion *in vitro* and reconstitution *in vivo*. *HOXA7* showed stronger phenotypes in both assays.

To investigate how *HOXA7* regulates human HSCs, pLKO1-control- and *HOXA7*-shRNA-transduced FL-HSPCs from four different FL tissues were sorted 5 days post-infection and subjected to RNA sequencing (Fig. 4e). Five hundred significantly differentially expressed genes (306 upregulated, 194 downregulated, 1.8-fold, P value <0.05) between *HOXA7*-knockdown and pLKO1-control HSPCs were identified (Fig. 4j and Supplementary Table 4). Comparison with hESC-HSPC microarray data revealed that 30.1% of the genes upregulated on *HOXA7* knockdown were also significantly upregulated in EB-OP9-HSPCs compared with FL-HSPCs, and 34.0% of the downregulated *HOXA7*-dependent genes showed low expression in EB-OP9 HSPCs. The shared downregulated genes included HSC regulators *HOXA9-10*, *HLF* and *HMGN*, and HSC surface proteins *PROM1/CD133* and *MPL* (Fig. 4k). *HOXA7*-knockdown HSPCs and EB-OP9-HSPCs also upregulated genes associated with megakaryocytic and erythroid differentiation (Fig. 4l). The shared upregulated genes also included cell cycle inhibitors *CDKN1A* (p21Cip1)[33] and *CDKN2D* (p19Ink4d)[34], whereas other cell cycle regulators and proliferation markers were unaffected at the messenger RNA level (Supplementary Fig. 3B,C). These results imply that *HOXA7* activates factors regulating definitive HSC identity and suppresses programs associated with differentiation-primed embryonic progenitors.

HOXA gene overexpression expands FL-HSPCs

We next assessed whether *HOXA* gene overexpression improves *in vitro* expansion of FL-HSPCs and hESC-HSPCs. Tetracycline-inducible PNL lentiviral vector[35] (Supplementary Fig. 4A), which induced 30-fold overexpression of *HOXA5* or *HOXA7* in FL-HSPCs (Supplementary Fig. 4B), prolonged FL-HSPC maintenance in culture with both *HOXA5* and *HOXA7* (Supplementary Fig. 4C,D). Constitutively active FUGW vector[36], which resulted in 4–5-fold transgene overexpression, prolonged FL-HSPC maintenance with *HOXA7* (Fig. 5a–e). HSPCs overexpressing *HOXA* genes showed comparable differentiation ability in colony assays, implying that *HOXA5* or *HOXA7* expression does not prevent differentiation (Fig. 5f).

Although *HOXA5* or *HOXA7* overexpression promoted FL-HSPC expansion, their overexpression in EB-derived $CD34^+$ cells did not improve hESC-HSPC expansion *in vitro* (Fig. 5g–i and Supplementary Fig. 4E) or engraftment *in vivo* (Fig. 5j,k). Similar results were observed overexpressing *HOXA5*, *HOXA7* and *HOXA9* simultaneously. RNA-sequencing of *HOXA7*-overexpressing hESC-HSPCs showed that despite the confirmed *HOXA7* overexpression (Fig. 5l), there was minimal change in putative *HOXA7* target genes (Fig. 5m). Altogether these studies showed that medial *HOXA* genes promote HSPC expansion when expressed in a correct cellular context.

RA signalling induces *HOXA* genes in hESC-HSPCs

We next sought for upstream regulators that could specify the definitive HSPC fate and induce *HOXA* genes in a correct cellular context. Medial *HOXA* genes are developmentally regulated by RA signalling in other cell types[37,38]. RA generated by *ALDH1A2* (RALDH2) in haemogenic endothelium and its binding to RA receptor alpha (*RARA*) is necessary for generating HSCs in mouse AGM[39]. Microarray analysis showed robust expression of *RARA*, and to a lesser extent *RARB* and *RARG* in HSPCs at different stages. Notably, RALDH2 was expressed in $CD34^+CD38^{-/lo}CD90^+CD43^+$ cells in the early placenta, but not in EB. FL-HSPCs did not express RALDH2, consistent with its function in haemogenic endothelium (Fig. 6a). RALDH1 was expressed in FL-HSPCs but at low levels in both EB- and PL-derived cells. These results nominated defective RA signalling during hESC haematopoietic specification as a potential barrier for inducing *HOXA* genes.

We next asked whether RALDH2 function could be bypassed by administering all-*trans* retinoic acid (ATRA) or the *RARA* agonist AM580[40] during hESC differentiation. $CD34^+$ cells were isolated from EBs at 2 weeks and cultured for 6 days on OP9-M2 with or without ATRA (1.0 μM) or AM580 (0.2 μM) (Fig. 6b). qRT–PCR analysis showed that treatment of EB-derived $CD34^+$ cells with AM580 or ATRA induced several medial *HOXA* genes (Fig. 6c). AM580 had a more robust effect on *HOXA* genes and was used for further studies.

Fluorescence-activated cell sorting (FACS) analysis at 6 days of AM580 treatment revealed strong induction of CD38, a known RA target[41], in hESC- and FL-HSPCs (Fig. 6d). Persistent CD90 expression suggested that they had not differentiated. AM580-induced CD38 expression was reversible (Fig. 6d), and AM580-treated EB $CD34^+$ cells typically retained a higher fraction of $CD34^+CD38^{-/lo}$ cells and undifferentiated HSPCs after 23–25 days than dimethylsulfoxide (DMSO)-treated controls (Fig. 6e,f). A concomitant increase was observed with CFU-Cs, in particular with mixed myelo-erythroid colonies (Supplementary Fig. 5A,B). These data show that inducing RA signalling in the haemogenic endothelium stage prolongs hESC-HSPC maintenance.

RA signalling promotes definitive haemogenic endothelium and HSC fate

To understand how RA signalling modulates hESC-derived haemato-vascular precursors, RNA sequencing was performed to define AM580-induced genes in hESC-HSPCs ($CD45^+CD34^+CD90^+$). Six days after AM580 treatment, *RARB* and *RARG*, known *RARA* targets[42], were upregulated (Fig. 7a). The *HOXA* cluster was also induced (Fig. 7b,c), whereas *RUNX1*, which is expressed in both progenitors and HSCs, showed no difference. Altogether, AM580 induced 408 genes and repressed 562 (twofold, P value <0.05) (Supplementary Table 5). AM580-induced genes showed the highest enrichment in GO categories reflecting vasculature development, cell adhesion and migration (Fig. 7d). These include factors implicated in definitive haemogenic endothelium and HSC

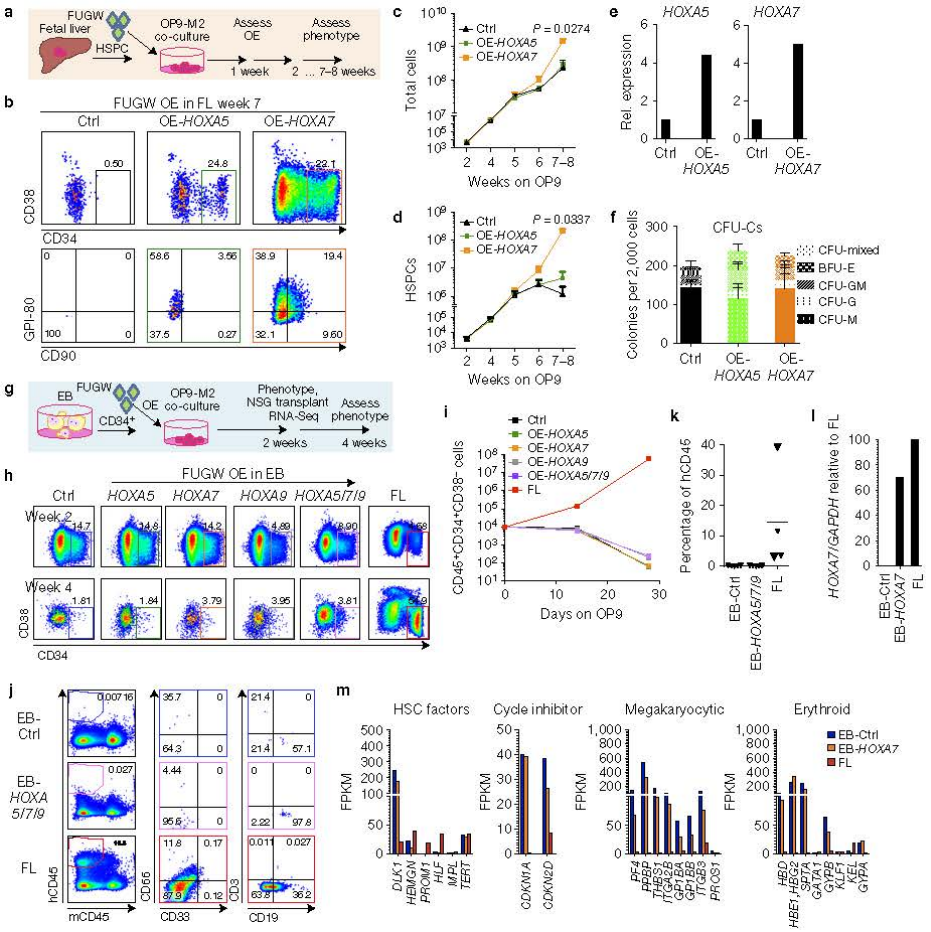


Figure 5 Overexpression of medial *HOXA* genes enhances proliferative potential in FL-HSPCs but does not confer HSC properties to hESC-HSPCs. (**a**) A schematic showing the strategy for constitutive lentiviral overexpression of *HOXA5* or *HOXA7* in FUGW vectors in FL-HSPCs. (**b**) Representative FACS plots of FUGW empty vector, *HOXA5*- or *HOXA7*-overexpressing FL-HSPCs. (**c**,**d**) Expansion of total FL cells (**c**) or HSPCs (**d**) transduced with *HOXA5*- or *HOXA7*-overexpression vectors or empty control vector (Ctrl) (mean and s.e.m. values from $n=3$ independent experiments; statistical significance was assessed using the paired Student's *t*-test). (**e**) qRT–PCR confirming overexpression in transduced HSPCs sorted 1 week post-infection ($n=1$ experiment with two pooled donors). (**f**) CFU-Cs from 2,000 HSPCs sorted after day 10 of infection with vectors overexpressing *HOXA5* or *HOXA7*, or FUGW empty control vector (mean and s.d. values shown from $n=4$ transductions from two independent experiments; *P* values shown correspond to Ctrl versus OE-*HOXA7*). (**g**) A schematic showing the strategy for lentiviral overexpression of *HOXA5* and/or *HOXA7* and/or *HOXA9* in FUGW vectors in EB CD34+ cells. (**h**) Representative examples of FACS plots of EB CD34+ cells overexpressing *HOXA5* or *HOXA7*, or a combination of *HOXA5*, *HOXA7* and *HOXA9*. Un-transduced FL is shown as a control. (**i**) Quantification of $CD34^+CD38^{-/lo}CD45^+$ haematopoietic cells from **h**; mean from $n=4$ independent experiments for Ctrl and $n=3$ for *HOXA5/7/9*, *HOXA5* and *HOXA7* at days 0 and 24, and $n=2$ at all other time points. (**j**,**k**) Transplantation assay of EB-CD34+ cells into NSG mice. Representative FACS plots (**j**) and quantification (**k**) of human CD45 cells in the bone marrow of NSG mice 12 weeks post-transplantation. Multi-lineage engraftment is assessed by CD19 and CD3 (B and T lymphoid, right column) and CD66 and CD33 (myeloid, middle column) (mean from $n=5$ mice per condition (except for FL, $n=4$) from two independent experiments). (**l**) qRT–PCR for *HOXA7* from transduced EB-OP9-HSPCs 2 weeks post-infection from one representative experiment. (**m**) Graphs representing RNA-Seq of EB-OP9 cells overexpressing *HOXA7* for genes regulated by *HOXA7* in FL-HSPCs (Fig. 4k,l) (one representative experiment, GEO database GSE76685). See Supplementary Table 7 for statistics source data in **d**–**f**,**i**,**k**.

generation such as *CDH5, NOS3* and *SOX17*; Notch ligands and targets *DLL4* and *HEY2*; and HSC surface markers *ROBO4*, *EMCN* and *PROCR*. Another highly enriched category was regulation of transcription, which included several HSC regulators (for example, the HOX cofactor *PBX1*, *MECOM/EVI1*, *ERG*, *GFI1*, *GATA3* and *HLF*; Fig. 7e).

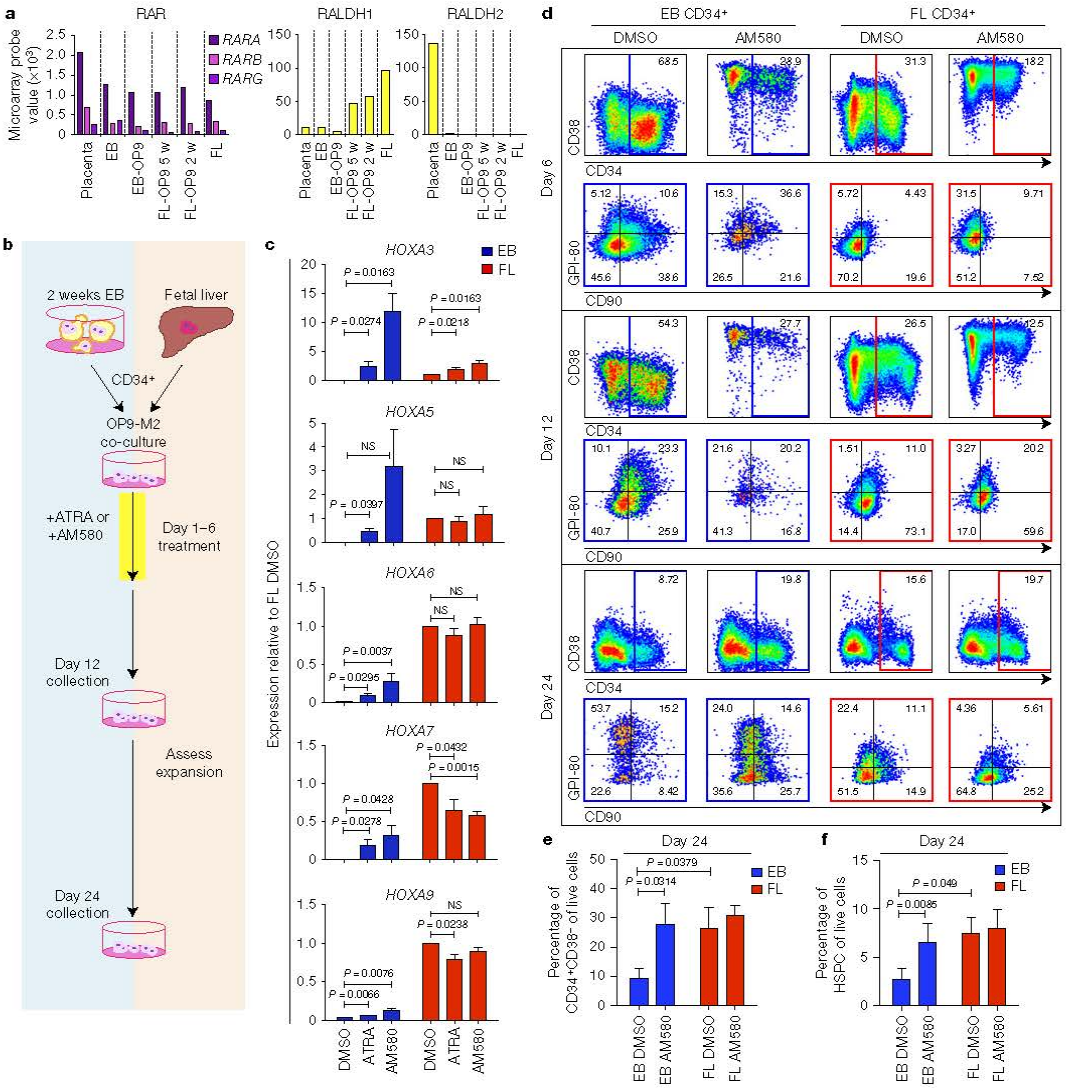


Figure 6 Retinoic acid signalling activates medial *HOXA* genes during human haematopoietic development. (**a**) Microarray analysis of gene expression of components of RA signalling pathway compared with FL-HSPCs (for *n* values see Fig. 4b; for mean values see Supplementary Table 1 and GEO database GSE64865). (**b**) A schematic showing 6-day treatment of CD34+ EB and FL cells by all-*trans* retinoic acid (ATRA) and the *RARA* agonist AM580. Cells were reseeded on OP9-M2 stroma after 12 days and analysed after an additional 12 ± 1 days (day 24 ± 1). (**c**) qRT–PCR of *HOXA3*, *HOXA5*, *HOXA6*, *HOXA7* and *HOXA9* expression in EB or FL cells treated with RA and AM580 (mean ± s.e.m. from $n=4$ independent experiments; see Supplementary Table 7 for statistics source data). (**d**) Representative FACS plots of CD45+ cells from AM580-treated EB and FL cells at 6, 12 and 24 days on OP9-M2 culture ($n=8$ independent experiments). (**e**,**f**) Quantification of CD34+CD38$^{-/lo}$ (**e**) and HSPC (**f**) fraction of EB- and FL-derived haematopoietic cells at day 24 ± 1 of OP9-M2 culture (mean ± s.e.m. from $n=8$ independent experiments). Statistical significance was assessed using the Student's paired *t*-test for **c**,**e**,**f**, one-tailed for **c** and two-tailed for **e**,**f**.

Analysis of hESC-HSPCs at day 12 of OP9-M2 culture (6 days after AM580 removal) evidenced expression of *HOXA*, albeit at lower levels, and many vascular factors and HSC regulators (Supplementary Fig. 6A,B). Interestingly, many genes upregulated in *HOXA7*-knockdown FL-HSPCs and hESC-HSPCs, including cell cycle inhibitors *CDKN1A* and *CDKN2D* and erythroid and

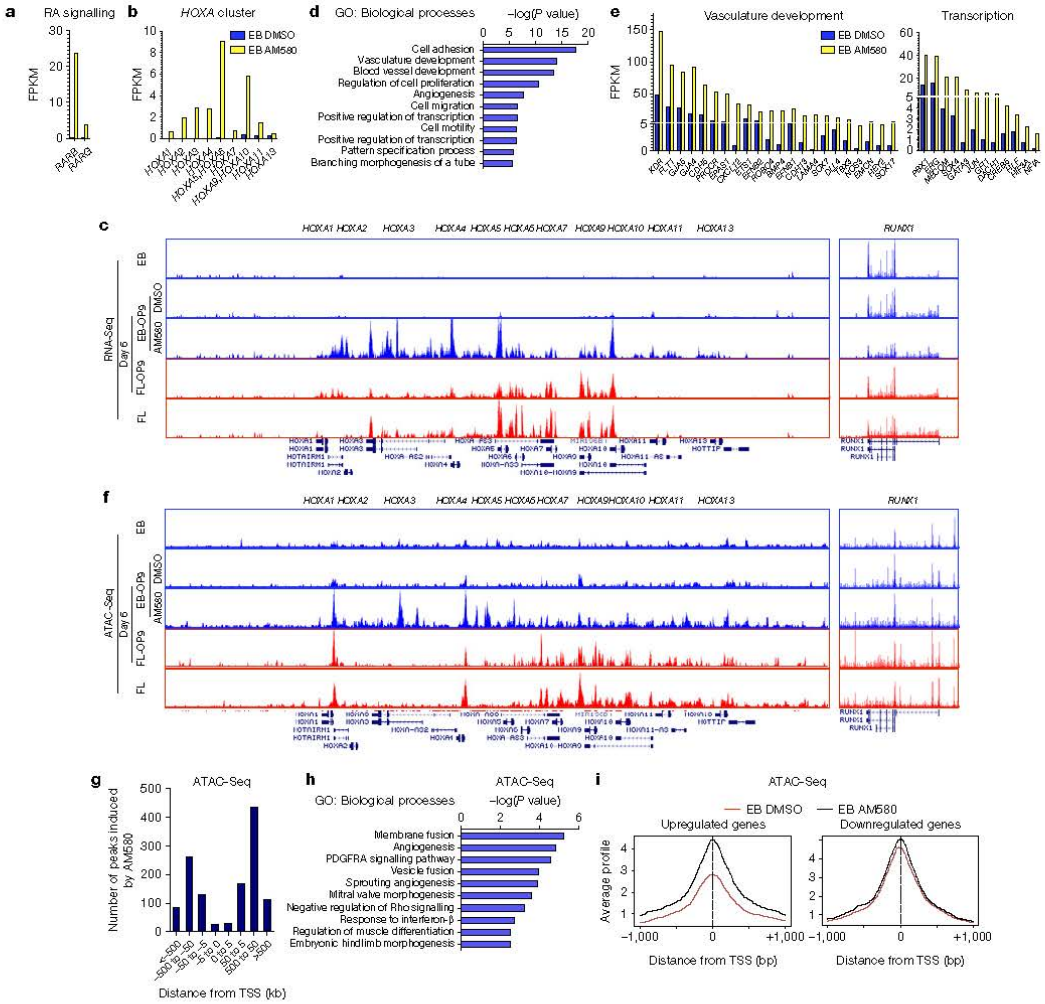


Figure 7 Retinoic acid signalling pulse in EB-derived haemato-vascular cells induces transcriptional programs associated with definitive haemogenic endothelium and HSC fate. (**a**,**b**) RNA sequencing analysis showing FPKM quantification for the *RARA* targets *RARB* and *RARG* (**a**) and *HOXA* genes (**b**) in sorted AM580-treated and DMSO control hESC-HSPCs (mean from two independent experiments). (**c**) RNA-Seq genome browser screen shot for the *HOXA* cluster and *RUNX1* in hESC- and FL-HSPCs after 6 days of AM580 treatment. (**d**) GO categories of biological processes significantly upregulated in hESC-HSPCs by AM580 treatment at day 6. (**e**) FPKM quantification values from representative genes from vasculature development and transcription GO categories from genes significantly upregulated in hESC-HSPCs by 6-day AM580 treatment (twofold or greater change, $P < 0.05$, mean from two independent experiments). (**f**) ATAC sequencing genome browser shot for the *HOXA* cluster and *RUNX1* assessing change in accessibility of regulatory regions in hESC- and FL-HSPCs on AM580 treatment. (**g**) Peaks significantly induced by AM580 treatment grouped on the basis of the distance from the TSS. (**h**) GO categories enriched among genes showing significant difference in accessibility after AM580 treatment. (**i**) ATAC-Seq signal proximal to the TSS of genes up- or downregulated by AM580 treatment. (ATAC-seq data show one representative data set from two independent experiments that showed comparable results.) See Supplementary Table 5 and GEO database GSE76685 for values used to generate graphs in **b**,**e**,**g**–**i**.

megakaryocytic genes, started to decline in AM580-treated hESC-HSPCs (Supplementary Fig. 6C). These data suggest that *RARA* signalling during endothelial-to-haematopoietic transition induces *HOXA* genes and other regulators that establish the definitive HSC fate, while suppressing embryonic progenitor programs.

To investigate whether RA signalling induces genes by modulating chromatin accessibility, ATAC sequencing was performed for $CD45^+CD34^+CD90^+$ hESC-HSPCs (DMSO- or AM580-treated for 6 days) and FL-HSPCs. AM580 stimulation increased chromatin accessibility throughout the *HOXA* cluster (Fig. 7f). Unbiased analysis

of differentially accessible peaks (q value <0.05) between DMSO- and AM580-treated hESC-HSPCs showed that most ATAC sequencing signal was 50–500 kilobases from the transcriptional start site (TSS), which are likely to represent enhancers (Fig. 7g). GREAT analysis identified vasculature-related categories as top differentially accessible gene groups (Fig. 7h). The average ATAC sequencing signal around the TSS of AM580-induced genes increased twofold on AM580 treatment (Fig. 7i), whereas genes repressed on AM580 treatment did not show a difference. These data suggest that RA signalling facilitates chromatin opening in the regulatory regions of the genes it activates.

DISCUSSION

Inability to replicate the microenvironment where HSCs develop has prevented the generation of clinically valuable HSCs from PSCs. However, the molecular barriers underlying the dysfunction of PSC-HSPCs are unknown. Here we document the generation of hESC-HSPCs that possess the immunophenotype of human fetal HSCs and differentiate into adult-type erythroid cells and T-lymphoid cells, but lack self-renewal ability. We pinpointed defective medial *HOXA* gene activation as a major developmental barrier preventing the establishment of self-renewing HSCs from hESCs, and identified RA signalling during endothelial-to-haematopoietic transition as a key inducer of *HOXA* genes and HSC fate.

Cells obtained from EBs possessed a surface phenotype of haemogenic endothelium and immature haematopoietic precursors comparable to the early placenta ($CD34^+CD38^{-/lo}CD90^+CD43^+CD45^-$GPI-$80^-$)[13,14]. Following culture on HSC-supportive stroma, hESC-HSPCs acquired surface expression of CD45 and GPI-80[14] and a closer molecular resemblance to FL-HSPCs. However, specific developmentally regulated genes, including the *HOXA* genes, remained silenced in hESC-HSPCs. Low expression of *HOXA* genes can also be found in published transcriptome analysis of hESC-derived $CD34^+$ cells[2,43–45], implying that *HOXA* gene silencing is independent of the differentiation protocol or PSC line used.

HOXA genes are dysregulated in leukaemias[46,47] and *HOXA9* has been previously implicated in mouse fetal HSC self-renewal[32,48]. Knockout mice for other *Hoxa* genes have not revealed strong HSC phenotypes, presumably owing to compensation by other Hox genes[49,50]. Deletion of the entire *Hoxa* cluster in adult mice severely reduced HSC activity and even *Hoxa* cluster haploinsufficiency compromised HSC function[51,52]. *HOXA9* overexpression enhanced haematopoietic specification from hESCs, but did not confer HSC function[44]. Our finding that *HOXA5* and *HOXA7* are critical for *in vitro* expansion and engraftment of FL-HSPCs, and their overexpression promotes FL-HSPC maintenance, documents a broader requirement for medial *HOXA* genes in human HSC regulation.

As *HOXA7* had the strongest loss- and gain-of-function phenotype, it was chosen for molecular analysis. Many genes dysregulated in *HOXA7*-deficient FL-HSPCs were also dysregulated in EB-OP9-HSPCs. Similar to hESC-HSPCs, *HOXA7*-deficient FL-HSPCs upregulated embryonic and fetal globins and other erythroid and megakaryocyte differentiation genes, implying that they acquire embryonic progenitor-like identity. Although the cell cycle was unaffected in *HOXA7*-deficient FL-HSPCs and hESC-HSPCs, they upregulated cell cycle inhibitors *CDKN1A* (p21Cip1) and *CDKN2D* (p19Ink4d). These inhibitors are low in human FL-HSPCs, but become induced after prolonged culture[26]. They are also highly expressed in FL $CD34^+CD38^+CD90^-$ progenitors that proliferate, but cannot self-renew[14]. Thus, upregulation of these CDKNs in haematopoietic cells associates with poor proliferative potential and onset of differentiation.

Although the overexpression of *HOXA7*, and to a lesser extent *HOXA5*, improved FL-HSPC maintenance in culture, their overexpression, even together with *HOXA9*, was insufficient to rescue hESC-HSPC function. *HOXA7* overexpression in hESC-HSPCs did not rescue the genes regulated by *HOXA7* in FL-HSPCs. The context-dependent regulation of HOXA target genes may depend on their epigenetic state and/or complementary regulatory factors present only in properly specified HSPCs.

We identified RA signalling as an upstream regulator for *HOXA* genes and the definitive HSPC transcriptional program. RA production by RALDH2 regulates HSC development during endothelial-to-haematopoietic transition in mouse AGM[39]. EB-derived haemato-vascular cells showed low RALDH2 expression compared with developmentally matched PL-HSPCs, suggesting that hESC-derived haemato-vascular precursors cannot produce RA. Supplementing *RARA* agonist to EB-derived haemato-endothelial cells for 6 days induced the expression of medial *HOXA* genes, and increased chromatin accessibility in the *HOXA* cluster and other RA-induced genes. Although many RA-regulated vascular genes are not highly expressed in FL-HSPCs, they promote definitive haemogenic endothelium[17] and HSC fate. RA stimulation also induced the HSC regulators *MECOM, HLF, GFI1* and *GATA3* and the HSC surface markers *ROBO4, EMCN* and *PROCR*. Strikingly, *RARA* stimulation partially suppressed the *HOXA7*-controlled differentiation programs and cell cycle inhibitors in hESC-HSPCs. These results suggest that *RARA*-induced medial *HOXA* gene activation is critical for HSC fate.

The level of *HOXA* expression in hESC-HSPCs attenuated after RA stimulation ended, implying that other regulators contribute to maintaining high *HOXA* expression in HSCs. Moreover, prolonged RA signalling can be detrimental for HSC development and maintenance[53,54] and RA effects vary with experimental conditions and developmental stage[54–56]. Therefore, the timing and dosage of RA stimulation must be optimized. Although other genetic programs are needed to impart full self-renewal capacity in HSCs, RA signalling and medial *HOXA* genes emerge as key regulators governing the patterning of haemogenic endothelium to definitive HSC fate. □

METHODS

Methods and any associated references are available in the online version of the paper.

Note: Supplementary Information is available in the online version of the paper

ACKNOWLEDGEMENTS

We thank BSCRC FACS Core at UCLA, the UCLA Clinical Pathology Microarray Core, BSCRC Sequencing Core, UCLA Tissue and Pathology Core and CFAR Gene and Cell Therapy Core (NIH grant AI028697-21) and Novogenix LLC. We thank H. Coller for discussions, T. Bolan for assistance with experiments, Y. Xing and Y.-T. Tseng for consultation on RNA sequencing analysis, and T. Stoyanova, D. Johnson and O. Witte for help with NSG mice. This work was supported by CIRM RN1-00557 and RT3-07763, NIH RO1 DK100959, LLS Scholar award and Rose Hills Foundation Scholar Award to H.K.A.M; Broad Stem Cell Research Center at UCLA and JCC Foundation; NIH P01 GM081621 to J.A.Z. and Z.G.; and NIH PO1 HL073104 and CIRM RB3-05217 to G.M.C. D.R.D. was supported by the NSF GRFP

and Ruth L. Kirschstein National Research Service Award GM007185, V.C. by an LLS Special Fellow Award and a BSCRC post-doctoral fellow award, M.I.S. by Ruth L. Kirschstein National Research Service Award HL086345, A.T.N. by a Beckman Scholarship, and P.S. by the Eugene V. Cota-Robles fellowship.

AUTHOR CONTRIBUTIONS

D.R.D., V.C., M.I.S., P.S., Z.G., J.A.Z., G.M.C. and H.K.A.M. designed experiments and interpreted data. D.R.D., V.C., M.I.S., A.T.N., A.M. and P.S. performed experiments. R.S. performed bioinformatics analysis of the microarray data and C.M.R. assisted with statistical analysis, D.R.D., V.C. and H.K.A.M. wrote the manuscript, which all authors edited and approved.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at http://dx.doi.org/10.1038/ncb3354
Reprints and permissions information is available online at www.nature.com/reprints

1. Bordignon, C. Stem-cell therapies for blood diseases. *Nature* **441,** 1100–1102 (2006).
2. Doulatov, S. *et al.* Induction of multipotential hematopoietic progenitors from human pluripotent stem cells via respecification of lineage-restricted precursors. *Cell Stem Cell* **13,** 459–470 (2013).
3. Pereira, C. F. *et al.* Induction of a hemogenic program in mouse fibroblasts. *Cell Stem Cell* **13,** 205–218 (2013).
4. Riddell, J. *et al.* Reprogramming committed murine blood cells to induced hematopoietic stem cells with defined factors. *Cell* **157,** 549–564 (2014).
5. Olsen, A. L., Stachura, D. L. & Weiss, M. J. Designer blood: creating hematopoietic lineages from embryonic stem cells. *Blood* **107,** 1265–1275 (2006).
6. Slukvin, I. I. Hematopoietic specification from human pluripotent stem cells: current advances and challenges toward de novo generation of hematopoietic stem cells. *Blood* **122,** 4035–4046 (2013).
7. McGrath, K. E. & Palis, J. Hematopoiesis in the yolk sac: more than meets the eye. *Exp. Hematol.* **33,** 1021–1028 (2005).
8. Kyba, M. & Daley, G. Q. Hematopoiesis from embryonic stem cells: lessons from and for ontogeny. *Exp. Hematol.* **31,** 994–1006 (2003).
9. Mikkola, H. K. & Orkin, S. H. The journey of developing hematopoietic stem cells. *Development* **133,** 3733–3744 (2006).
10. Swiers, G., Rode, C., Azzoni, E. & de Bruijn, M. F. A short history of hemogenic endothelium. *Blood Cells Mol. Dis.* **51,** 206–212 (2013).
11. Vodyanik, M. A., Bork, J. A., Thomson, J. A. & Slukvin, I. I. Human embryonic stem cell-derived CD34+ cells: efficient production in the coculture with OP9 stromal cells and analysis of lymphohematopoietic potential. *Blood* **105,** 617–626 (2005).
12. Vodyanik, M. A., Thomson, J. A. & Slukvin, I. I. Leukosialin (CD43) defines hematopoietic progenitors in human embryonic stem cell differentiation cultures. *Blood* **108,** 2095–2105 (2006).
13. Van Handel, B. *et al.* The first trimester human placenta is a site for terminal maturation of primitive erythroid cells. *Blood* **116,** 3321–3330 (2010).
14. Prashad, S. L. *et al.* GPI-80 defines self-renewal ability in hematopoietic stem cells during human development. *Cell Stem Cell* **16,** 80–87 (2015).
15. Keller, G. Embryonic stem cell differentiation: emergence of a new era in biology and medicine. *Genes Dev.* **19,** 1129–1155 (2005).
16. Pick, M., Azzola, L., Mossman, A., Stanley, E. G. & Elefanty, A. G. Differentiation of human embryonic stem cells in serum-free medium reveals distinct roles for bone morphogenetic protein 4, vascular endothelial growth factor, stem cell factor, and fibroblast growth factor 2 in hematopoiesis. *Stem Cells* **25,** 2206–2214 (2007).
17. Ditadi, A. *et al.* Human definitive haemogenic endothelium and arterial vascular endothelium represent distinct lineages. *Nat. Cell Biol.* **17,** 580–591 (2015).
18. Zambidis, E. T., Peault, B., Park, T. S., Bunz, F. & Civin, C. I. Hematopoietic differentiation of human embryonic stem cells progresses through sequential hematoendothelial, primitive, and definitive stages resembling human yolk sac development. *Blood* **106,** 860–870 (2005).
19. Wang, L. *et al.* Endothelial and hematopoietic cell fate of human embryonic stem cells originates from primitive endothelium with hemangioblastic properties. *Immunity* **21,** 31–41 (2004).
20. Dravid, G., Zhu, Y., Scholes, J., Evseenko, D. & Crooks, G. M. Dysregulated gene expression during hematopoietic differentiation from human embryonic stem cells. *Mol. Ther.* **19,** 768–781 (2011).
21. Shojaei, F. & Menendez, P. Molecular profiling of candidate human hematopoietic stem cells derived from human embryonic stem cells. *Exp. Hematol.* **36,** 1436–1448 (2008).
22. Martin, C. H., Woll, P. S., Ni, Z., Zuniga-Pflucker, J. C. & Kaufman, D. S. Differences in lymphocyte developmental potential between human embryonic stem cell and umbilical cord blood-derived hematopoietic progenitor cells. *Blood* **112,** 2730–2737 (2008).
23. Qiu, C. *et al.* Differentiation of human embryonic stem cells into hematopoietic cells by coculture with human fetal liver cells recapitulates the globin switch that occurs early in development. *Exp. Hematol.* **33,** 1450–1458 (2005).
24. Tian, X. & Kaufman, D. S. Differentiation of embryonic stem cells towards hematopoietic cells: progress and pitfalls. *Curr. Opin. Hematol.* **15,** 312–318 (2008).
25. Wang, L., Cerdan, C., Menendez, P. & Bhatia, M. Derivation and characterization of hematopoietic cells from human embryonic stem cells. *Methods Mol. Biol.* **331,** 179–200 (2006).
26. Magnusson, M. *et al.* Expansion on stromal cells preserves the undifferentiated state of human hematopoietic stem cells despite compromised reconstitution ability. *PLoS ONE* **8,** e53912 (2013).
27. Kennedy, M. *et al.* T lymphocyte potential marks the emergence of definitive hematopoietic progenitors in human pluripotent stem cell differentiation cultures. *Cell Rep.* **2,** 1722–1735 (2012).
28. Shaut, C. A., Keene, D. R., Sorensen, L. K., Li, D. Y. & Stadler, H. S. HOXA13 is essential for placental vascular patterning and labyrinth endothelial specification. *PLoS Genet.* **4,** e1000073 (2008).
29. Aguilo, F. *et al.* Prdm16 is a physiologic regulator of hematopoietic stem cells. *Blood* **117,** 5057–5066 (2011).
30. Klimmeck, D. *et al.* Transcriptome-wide profiling and posttranscriptional analysis of hematopoietic stem/progenitor cell differentiation toward myeloid commitment. *Stem Cell Rep.* **3,** 858–875 (2014).
31. Thorsteinsdottir, U. *et al.* Overexpression of the myeloid leukemia-associated Hoxa9 gene in bone marrow cells induces stem cell expansion. *Blood* **99,** 121–129 (2002).
32. Lawrence, H. J. *et al.* Loss of expression of the Hoxa-9 homeobox gene impairs the proliferation and repopulating ability of hematopoietic stem cells. *Blood* **106,** 3988–3994 (2005).
33. Wang, Y., Schulte, B. A., LaRue, A. C., Ogawa, M. & Zhou, D. Total body irradiation selectively induces murine hematopoietic stem cell senescence. *Blood* **107,** 358–366 (2006).
34. Hilpert, M. *et al.* p19 INK4d controls hematopoietic stem cells in a cell-autonomous manner during genotoxic stress and through the microenvironment during aging. *Stem Cell Rep.* **3,** 1085–1102 (2014).
35. Pluta, K., Luce, M. J., Bao, L., Agha-Mohammadi, S. & Reiser, J. Tight control of transgene expression by lentivirus vectors containing second-generation tetracycline-responsive promoters. *J. Gene Med.* **7,** 803–817 (2005).
36. Lois, C., Hong, E. J., Pease, S., Brown, E. J. & Baltimore, D. Germline transmission and tissue-specific expression of transgenes delivered by lentiviral vectors. *Science* **295,** 868–872 (2002).
37. Marshall, H., Morrison, A., Studer, M., Pöpperl, H. & Krumlauf, R. Retinoids and Hox genes. *FASEB J.* **10,** 969–978 (1996).
38. Gavalas, A. & Krumlauf, R. Retinoid signalling and hindbrain patterning. *Curr. Opin. Genet. Dev.* **10,** 380–386 (2000).
39. Chanda, B., Ditadi, A., Iscove, N. N. & Keller, G. Retinoic acid signaling is essential for embryonic hematopoietic stem cell development. *Cell* **155,** 215–227 (2013).
40. Delescluse, C. *et al.* Selective high affinity retinoic acid receptor alpha or beta-gamma ligands. *Mol. Pharmacol.* **40,** 556–562 (1991).
41. Kishimoto, H. *et al.* Molecular mechanism of human CD38 gene expression by retinoic acid. Identification of retinoic acid response element in the first intron. *J. Biol. Chem.* **273,** 15429–15434 (1998).
42. Balmer, J. E. & Blomhoff, R. Gene expression regulation by retinoic acid. *J. Lipid Res.* **43,** 1773–1808 (2002).
43. Salvagiotto, G. *et al.* Molecular profiling reveals similarities and differences between primitive subsets of hematopoietic cells generated *in vitro* from human embryonic stem cells and in vivo during embryogenesis. *Exp. Hematol.* **36,** 1377–1389 (2008).
44. Ramos-Mejía, V. *et al.* HOXA9 promotes hematopoietic commitment of human embryonic stem cells. *Blood* **124,** 3065–3075 (2014).
45. Wang, L. *et al.* Generation of hematopoietic repopulating cells from human embryonic stem cells independent of ectopic HOXB4 expression. *J. Exp. Med.* **201,** 1603–1614 (2005).
46. Beachy, S. H. *et al.* Isolated Hoxa9 overexpression predisposes to the development of lymphoid but not myeloid leukemia. *Exp. Hematol.* **41,** 518–529 (2013).
47. Alharbi, R. A., Pettengell, R., Pandha, H. S. & Morgan, R. The role of HOX genes in normal hematopoiesis and acute leukemia. *Leukemia* **27,** 1000–1008 (2013).
48. McKinney-Freeman, S. *et al.* The transcriptional landscape of hematopoietic stem cell ontogeny. *Cell Stem Cell* **11,** 701–714 (2012).
49. Chen, F., Greer, J. & Capecchi, M. R. Analysis of Hoxa7/Hoxb7 mutants suggests periodicity in the generation of the different sets of vertebrae. *Mech. Dev.* **77,** 49–57 (1998).
50. Boucherat, O. *et al.* Partial functional redundancy between Hoxa5 and Hoxb5 paralog genes during lung morphogenesis. *Am. J. Physiol. Lung Cell. Mol. Physiol.* **304,** L817–L830 (2013).
51. Lebert-Ghali, C. E. *et al.* HoxA cluster is haploinsufficient for activity of hematopoietic stem and progenitor cells. *Exp. Hematol.* **38,** 1074–1086 (2010).
52. Lebert-Ghali, C. E. *et al.* Hoxa cluster genes determine the proliferative activity of adult mouse hematopoietic stem and progenitor cells. *Blood* **127,** 87–90 (2016).
53. Muramoto, G. G. *et al.* Inhibition of aldehyde dehydrogenase expands hematopoietic stem cells with radioprotective capacity. *Stem Cells* **28,** 523–534 (2010).
54. Szatmari, I., Iacovino, M. & Kyba, M. The retinoid signaling pathway inhibits hematopoiesis and uncouples from the Hox genes during hematopoietic development. *Stem Cells* **28,** 1518–1529 (2010).
55. Rönn, R. E. *et al.* Retinoic acid regulates hematopoietic development from human pluripotent stem cells. *Stem Cell Rep.* **4,** 269–281 (2015).
56. Cano, E., Ariza, L., Muñoz-Chápuli, R. & Carmona, R. Signaling by retinoic acid in embryonic and adult hematopoiesis. *J. Dev. Biol.* **2,** 18–33 (2014).

METHODS

Human ESC culture. The human ESC line H1, obtained from WiCell (WA01), was maintained on 50-Gy-irradiated (CF1) murine embryonic fibroblast feeders, in DMEM/F12 medium (Invitrogen) containing 20% knockout serum replacement (Invitrogen), 0.5% penicillin/streptomycin (50 U ml^{-1}, Invitrogen), 1 mM L-glutamine, 0.5% MEM non-essential amino acids (NEA, Invitrogen), 10 ng ml^{-1} basic FGF (R&D Systems) and 0.11 mM BME (Invitrogen). hESCs were passaged weekly using 1 mg ml^{-1} collagenase IV (Invitrogen) in DMEM/F12 for 10 min, detached by gentle pipetting or cell scraper to maintain cells in small clumps, washed twice and split to new 6-well plates containing irradiated CF1 mouse embryonic fibroblasts. Cells were mycoplasma-tested and authenticated by short-term repeat analysis. hESC work was approved by the UCLA Embryonic Stem Cell Research Oversight committee. WA01 is not listed in the ICLAC or NCBI Biosample as a commonly misidentified cell line.

Differentiation of embryoid bodies. hESCs were treated with 0.5 mg ml^{-1} Dispase (Invitrogen) in DMEM/F12 for 5 min at 37 °C. Clumps were mechanically detached and transferred to low-attachment plates (Corning) in IMDM (Invitrogen), 15% FBS (Hyclone), 1% MEM-NEA, 0.3 μM BME (Invitrogen) and 1% penicillin/streptomycin (100 U ml^{-1}, Invitrogen)/1 mM L-glutamine (P/S/G). Medium was changed every 2 days and supplemented with 10 ng ml^{-1} BMP4 (R&D Systems), 50 ng ml^{-1} FLT3L (Peprotech) and 300 ng ml^{-1} SCF (Peprotech) at day 4 of EB differentiation. BMP4 was removed from the culture medium at day 12. EBs were collected at 2 weeks and dissociated with 0.25% trypsin–EDTA (Invitrogen) with 2% chick serum (Sigma-Aldrich) for 30 min at 37 °C with periodic agitation and filtered with a 70 μm cell strainer. A step-by-step protocol detailing EB generation and AM580/ATRA treatment can be found at Nature Protocol Exchange[57].

Human tissue collection. Placenta and fetal liver were de-identified, discarded material obtained from elective termination of first- and second-trimester pregnancies following informed consent. As these tissues are discarded material with no personal identifiers, this research does not constitute research on human subjects. This protocol was reviewed by the UCLA IRB committee, who determined such studies can be performed without further IRB review. Specimen age for this study is denoted as developmental age, 2 weeks less than gestational age, and was determined by ultrasound or estimated by the date of the last menstrual period. Tissues were collected into PBS with 5% FBS (Hyclone), ciprofloxacin HCl (10 ng ml^{-1}, Sigma), amphotericin B (2.50 μg ml^{-1}, Invitrogen) and 1% penicillin/streptomycin, transported on ice and processed the same day.

Human tissue processing. Single-cell suspensions were prepared from FL at 14–17 weeks of developmental age. Tissues were mechanically dissociated using scalpels and syringes. Mononuclear cells were enriched on a Ficoll layer according to the manufacturer's protocol (GE Healthcare Biosciences AB) and strained through a 70 μm mesh.

Single cell suspensions were prepared from placenta at 3–5 weeks of developmental age. Tissues were mechanically dissociated and digested in 2.5 U Dispase (Gibco), 90 mg collagenase (Worthington) and 0.075 mg DNase I (Sigma) per gram of tissue for 90 min at 37 °C with agitation. Cells were then filtered through a 70 μm cell strainer.

Selection of CD34$^+$ cells by magnetic beads. Single-cell suspensions obtained from 2-week EB differentiation cultures, human fetal livers or human placentae were magnetically isolated with anti-CD34 microbeads (Miltenyi Biotech) according to the manufacturer's protocol.

OP9-M2 stromal co-culture for HSPC maturation and expansion. OP9-M2 cells[26] were irradiated (20 Gy) and pre-plated (50,000 cells cm^{-2}) onto tissue-culture-treated wells 24 h before the start of co-culture in OP9 medium (α-MEM (Invitrogen), 20% FBS (Hyclone), P/S/G). Haematopoietic cells derived from hESCs or haematopoietic tissues were plated on a stromal layer in OP9 medium supplemented with SCF (25 ng ml^{-1}, Peprotech), Flt3 (25 ng ml^{-1}, Peprotech) and TPO (25 ng ml^{-1}, Peprotech) (HSC medium). Cells were co-cultured at 37 °C and 5% CO_2 and re-plated or analysed/sorted by flow cytometry every 7–14 days. Half of the HSC medium was replaced every other day.

Flow cytometry and cell sorting. FACS analysis was performed using single-cell suspensions prepared as described. Cells were stained with mouse anti-human monoclonal antibodies against human CD45–PE (cl. J.33, Beckman Coulter IM2078U; diluted 1:100), CD45–APC–H7 (cl. 2D1, Biolegend 368512 and 368516; 1:100) and CD45–BV711 (cl. HI30, Biolegend 304050; 1:100), and mouse CD45–APC–H7 (cl. 30-F11, BD557659; 1:100), CD34–APC (cl. 581, BD 555824; 1:20), CD90–FITC (cl. 5E10 BD 555595; 1:100), CD38–PE–Cy7 (cl. HIT2, BD 560677; 1:100), CD19–PE (cl. 1D3 or HIB19, eBiosciences 12-0193 and 12-0199; 1:50), CD43–PE (cl. MT1, SCB 51727; 1:25) or –FITC (cl.1G10, BD 555475; 1:20), GPI-80–PE (cl. 3H9, MBL D087-5; 1:50), CD3–PE–Cy7 (cl. SK7, BD 557851; 1:50), CD4–APC (cl. S3.5, Invitrogen MHCD0405; 1:50), CD8–PE (cl. HIT8A, BD 555635; 1:50), CD13–APC (cl. WM15, BD 557454; 1:500), CD66b–FITC (cl. G10F5, BD 555724; 1:50) and CD33–PE (cl. WM53, BD 561816; 1:100).

Dead cells were excluded with 7-amino-actinomycin D (7AAD) (BD Biosciences, used at 1:50). Cells were assayed on a BD-LSRII flow cytometer and data were analysed with FlowJo software (Tree Star). Cell sorting was performed using a BD FACS Aria II.

Methylcellulose colony-forming assays. Myelo-erythroid progenitor potential was assessed on methylcellulose (MethoCult GF$^+$ H4435, SCT) supplemented with TPO (10 ng ml^{-1}, Peprotech), 1% penicillin/streptomycin (100 U ml^{-1}, Invitrogen) and 1% amphotericin B (2.50 μg ml^{-1}, Invitrogen). Cultures were incubated at 37 °C and 5% CO_2 for 14–16 days and colonies were scored on the basis of morphological characteristics. Images were taken using an Olympus BX51 microscope with a DP72 camera.

Cytospins. Representative myelo-erythroid colonies from methylcellulose assays were picked and resuspended in PBS with 40% FBS (Hyclone). The cell suspension was cytospun on slides using a Shandon Cytospin 4 (Thermo Electron Corporation) spun at 10*g* under medium acceleration. Slides were air-dried overnight and stained using May–Grünwald–Giemsa (MGG) stain (Sigma-Aldrich).

Imaging. Bright-field images of individual colonies and images of MGG-stained colonies were taken using the Zeiss Axovert 40 CFL microscope under the ×10 objective with an attached Canon PC 1089 camera at ×4 zoom for a total magnification of ×40.

T-lymphoid differentiation. hESC- or FL-derived CD34$^+$ cells were plated on non-irradiated OP9-DL1 stroma (25,000 cells cm^{-2}) in OP9 medium supplemented with SCF (25 ng ml^{-1}, Peprotech), FLT3L (10 ng ml^{-1}, Peprotech) and IL-7 (20 ng ml^{-1}, Peprotech). Cells were lifted and reseeded on OP9-DL1 every week until FACS analysis.

Cell cycle analysis. Cultured cells were pulse-labelled with 10 μM BrdU for 35 min in culture. Cells were sorted for the indicated surface phenotypes and processed according to the FITC-BrdU flow kit (BD) instructions.

RNA isolation, cDNA synthesis and quantitative reverse transcriptase PCR. RNA isolation was performed using the RNeasy Mini kit (Qiagen) with an additional DNase (Qiagen) step using the manufacturer's protocol. cDNAs were prepared using the Quantitech reverse transcription kit (Qiagen) and quantitative polymerase chain reaction (qPCR) for *GAPDH*, glycophorin A (*GYPA*), haemoglobin subunit epsilon (*HBE*), haemoglobin subunit gamma (*HBG*) and beta globin (*HBB*) was performed with the LightCycler 480 SYBR Green I Master Mix (Roche) on the Lightcycler 480 (Roche); qPCR for *GAPDH* and *HOXA7* with the SYBR Select Master Mix (Life Technology, LT) using the ViiA 7 Real-Time PCR System (LT); and qPCR for *GAPDH, HOXA3, HOXA5, HOXA6* and *HOXA9* was performed with the TaqMan Gene Expression Master Mix (LT) and analysed using the ViiA 7 Real-Time PCR System (LT). SYBR Green-compatible primers were obtained from the PrimerBank database or literature[58,59]. Taqman primers were purchased from LT. Primers were tested against OP9-M2 cDNA to rule out amplification of murine genes, and gDNA or water as negative controls. Primers are presented in Supplementary Table 6.

Microarray analysis. RNA isolation was performed using RNEasy Mini kit (>50,000 cells) or RNeasy Micro kit (<50,000 cells) (Qiagen) with DNase digestion using the manufacturer's protocol. RNA was amplified by the NuGen amplification kit and hybridized on Affymetrix arrays (Human U133plus2.0 Array). Samples were quantile normalized.

K-means clustering of differentially expressed genes. Pairwise differential expression analysis was performed for the various populations by independently comparing immunophenotypic HSPCs derived from hESCs (EB and EB-OP9), placenta (PL), cultured fetal liver (FL-OP9 2 weeks and FL-OP9 5 weeks), with freshly isolated FL-HSPCs (FL). The following criteria were used to filter probes and identify differentially expressed genes: fold change value of 2 or higher; a *P* value less than 0.05; probes that are called 'Absent' in all replicates in all samples were excluded; and probes that have an absolute expression level of 50 or less were excluded. The union of all probes that met these criteria was chosen for further analysis, after which normalized expression values, which were replicate-averaged and standardized, were obtained. The standardized expression values were clustered using the K-means method. GO enrichment analysis was performed for genes in the various clusters. Differential expression assessment was performed using the

DOI: 10.1038/ncb3354

Limma R package[60] through Bioconductor[61]. To account for multiple comparisons, *P* values were adjusted using the Benjamini–Hochberg correction method to control the false discovery rate (FDR). A 5% FDR was used as the cutoff. K-means clustering of RMA-normalized, replicate-averaged and standardized expression values was done using Cluster 3.0 software[62] and the resulting clusters were viewed using Java Treeview software[63].

RNA sequencing analysis. Total RNA from 2,000 to 20,000 sorted cells was extracted using the RNeasy Mini kit (Qiagen) and a library was constructed using Ovation Rna-seq system System v2 (Nugen), followed by KAPA LTP Library Preparation Kit. Libraries were sequenced using HIseq-2000 (Illumina) to obtain single-end 50-bp-long reads. Demultiplexing of the reads based on the barcoding was performed using an in-house Unix shell script. Splice junction mapping to the human genome (hg19) was performed using TopHat v2.0.9 or v2.0.14[64] with the parameters –no-coverage-search -M -T -x 1. Coverage files were created with the Genomecov tool from Bedtools[65] with the parameters -bg -split –ibam. For abundance estimations (FPKMs) the aligned read files were further processed with Cufflinks v2.2.1[66] –compatible-hits-norm -N -u –GTF gencode.v19.annotation.gtf. Assemblies for all samples were merged using Cuffmerge and differential expression was determined using Cuffdiff.

Production of lentiviral shRNA and overexpression vectors. shRNA experiments were performed with pLKO lentiviral vectors from the TRC library (TRCN000017529 for *HOXA5*, TRCN000015084 for *HOXA7*) containing puromycin resistance gene. Human *HOXA5, HOXA7* and *HOXA9* were cloned from human FL full-length cDNA, into either the constitutive pFUGW lentiviral vector (Addgene plasmid no. 14883, from D. Baltimore, California Institute of Technology, Pasadena, California, USA), downstream and in frame with the GFP sequence with the synthetic addition of a P2A sequence between the two ORFs, or the inducible lentiviral overexpression system pNL-EGFP/TREPittdU3 (Addgene plasmid no. 18659, from J. Reiser, Louisiana State University Health Sciences Center New Orleans, New Orleans, Louisiana, USA), between the sites BamHI and NheI. pNL–TREpitt vectors were co-transduced with the constitutive pNL–EF1α–rTTA–M2 lentiviral vector to provide *in trans* the Tet transactivator. For lentiviral vector production, 20 million 293T cells were transfected with 12.5 μg deltaR8.2, 5 μg VSV-G, and 12.5 μg DNA/shRNA and 90 μl of Lipofectamine 2000 (Invitrogen) in OPTI-MEM and incubated for 5–6 h at 37 °C. After incubation for 48 h in complete medium, supernatant was filtered and concentrated using an ultracentrifuge (Beckman Coulter rotor SW32 Ti) at 20,200 r.p.m. for 1.5 h at 4 °C. Following centrifugation, pelleted viruses were resuspended in 125 μl of SFEM and stored at −80 °C.

Lentiviral transduction of $CD34^+$ cells. FL-HSPCs were prestimulated for 6–18 h in fresh SFEM culture medium supplemented with 100 ng ml^{-1} SCF, 100 ng ml^{-1} TPO and 100 ng ml^{-1} FLT3L. Wells were treated with 40 μg ml^{-1} RetroNectin (Takara) and seeded with prestimulated FL-HSPCs in 300 μl SFEM culture medium. Lentivirus (5 μl) was added twice during 24 h incubation. Transduced cells were washed and seeded on OP9-M2 with HSC medium. For shRNA lentiviral vectors, puromycin (1.0 μg ml^{-1}) treatment was used for selection of transduced HSPCs and maintained throughout culture. EB-derived $CD34^+$ cells were infected in SFEM for 2 h on a rocking platform at 37 °C and then moved to OP9-M2 co-culture on HSC medium. For overexpression experiments with PNL vectors, doxycycline was added immediately at the start of OP9-M2 culture and maintained at 1 μg ml^{-1}.

Transplantation into NOD–scid IL2Rγ-null mice. Female NOD–scid IL2Rγ-null (NSG, Jackson Laboratories) mice, 8–12 weeks old, were sublethally irradiated (325 rads) and intra-tibially injected with FL or hESC $CD34^+$ cells in a volume of 35 μl. Mice were transplanted with either 1.2×10^6 EB-derived $CD34^+$ cells (60% $CD34^+$) or with 10,000 FL $CD34^+$ cells in 50 μl PBS. A second batch of mice was transplanted with 2.5×10^6 EB-$CD34^+$-derived cells (25% $CD34^+$) or 3×10^6 FL-$CD34^+$-derived cells (35% $CD34^+$) obtained after 2 weeks of culture on OP9-M2.

For shRNA-treated FL cells, 50,000 cells were infected with shRNA lentivirus and expanded on OP9-M2 for 9 days under puromycin selection. Transduced haematopoietic cells were retro-orbitally injected into sublethally irradiated NSG mice.

For engraftment of EB-derived cells overexpressing *HOXA* genes, 300,000 cells were infected with the indicated lentivirus. After infection, cells were seeded on OP9-M2, expanded for 2 weeks, collected and injected in the right tibia of sublethally irradiated (275 rads) female NSG mice. FL $CD34^+$ cells were injected as a positive control of engraftment.

Mice were euthanized at 10–12 weeks to collect bone marrow. Collected cells were FACS-analysed to evaluate human engraftment (human CD45 and mouse CD45), differentiation into myelo-lymphoid lineages (CD13, CD33, CD66, CD3 or CD19) and preservation of the HSPC compartment ($CD34^+$, $CD38^-$, $CD90^+$).

All studies and procedures involving mice were approved by the UCLA Animal Research Committee (Protocol 2005-109).

Induction of retinoic acid signalling. ATRA (all-*trans* retinoic acid, Sigma-Aldrich) was dissolved in DMSO at 25 mM and applied at a final concentration of 1 μM; AM580 (Sigma-Aldrich) was dissolved in DMSO at 10 mM at a final concentration of 0.2 μM; DMSO was used at 1:25,000 final dilution. Treatments were performed on FL- or hESC-derived $CD34^+$ cells in HSC medium. Cells were collected at day 6 for qRT–PCR and RNA sequencing analysis. For longer cultures, RA stimulation media were removed at day 6 from the wells avoiding disruption of the cell layer and replaced with HSC medium with no treatment. Cells were then collected at day 12 reseeded on OP9-M2 for another 12–13 days, and then assayed for HSPC expansion by FACS or colony-forming potential in methylcellulose. A step-by-step protocol detailing EB generation and AM580/ATRA treatment can be found at Nature Protocol Exchange[57].

ATAC sequencing. Cells (12,000 to 60,000) were sorted in PBS and processed according to the protocol indicated[67], with minor adjustments. Nuclei were purified by the addition of 250 μl of lysis buffer (10 mM Tris-HCl, pH 7.4, 10 mM NaCl, 3 mM $MgCl_2$, 0.1%IGEPAL), pelleted and resuspended in the transposition reaction mix (Nextera DNA Library Prep Kit, Illumina) and incubated at 37 °C for 30 min. Transposed DNA was column purified and used for library amplification with custom-made adaptor primers (see Supplementary Table 6) using NEBNext High-Fidelity 2x PCR Master Mix (NEB). The amplification was interrupted after five cycles and a SyBR green qPCR was performed with 1/10 of the sample to estimate for each sample the additional number of cycles to perform, before saturation was achieved. Total amplification was between 10 and 15 cycles. Purified libraries were sequenced using HIseq-2000 (Illumina) to obtain paired-end 50-bp-long reads. Demultiplexing of the reads and creation of the fastq files was performed using an in-house Unix shell script. Read mapping to the genome (hg19) was performed using Bowtie2 or v2.2.5[68] with parameters –local -X 2000 -N 1 –no-mixed. The Bamcoverage tool from Deeptools was used to create the coverage .bw files for visualization[69]. Samtools v1.2 was used to remove duplicates and reads aligned to chrM. MACS2[70] was used to call the differentially accessible peaks between the AM580-treated and -untreated EB-derived cells, using the parameters –broad –broad-cutoff 0.1.

Statistics and reproducibility. Graphs were generated with GraphPad PRISM software. Statistical significance was calculated in R version 3.2.3. Statistical significance was assessed using the Wilcoxon rank sum test for unpaired data sets and the Student's paired *t*-test for paired data sets. All tests are two-tailed with the exception of that in Fig. 6c, where a one-tailed distribution was used because it is an established fact that RA treatment enhances HOX gene expression. Data sets were considered paired when treatments (that is, DMSO versus AM580 or LKO versus shRNA) were performed in parallel on the same batch of cells (EB) or from the same donor (FL) for the same time periods. Data were considered unpaired when cells from different batches, donors or cell types were compared. The null hypothesis of the medians/means being equal was rejected at $\alpha = 0.05$ and exact significant *P* values are shown in each graph. The comparisons to be made were decided a priori with the intention of limiting the overall number of comparisons and are indicated by lines in the figures and adjustments for multiple comparisons would be applied to $k > 4$ comparisons. To account for multiple comparisons in microarray analysis, *P* values were adjusted using the Benjamini–Hochberg correction method to control the FDR. The non-parametric Wilcoxon rank sum test is robust to outliers. We tested for differences in sample dispersion using Ansari–Bradley and we could not reject the null hypothesis of equal dispersions. A Satterthwaite's approximation was used to account for unequal variances when using the *t*-test. The investigators were not blinded to allocation during experiments and outcome assessment. For mouse engraftment studies, the experiments were not randomized, no statistical method was used to predetermine sample size and no animal was excluded from analysis.

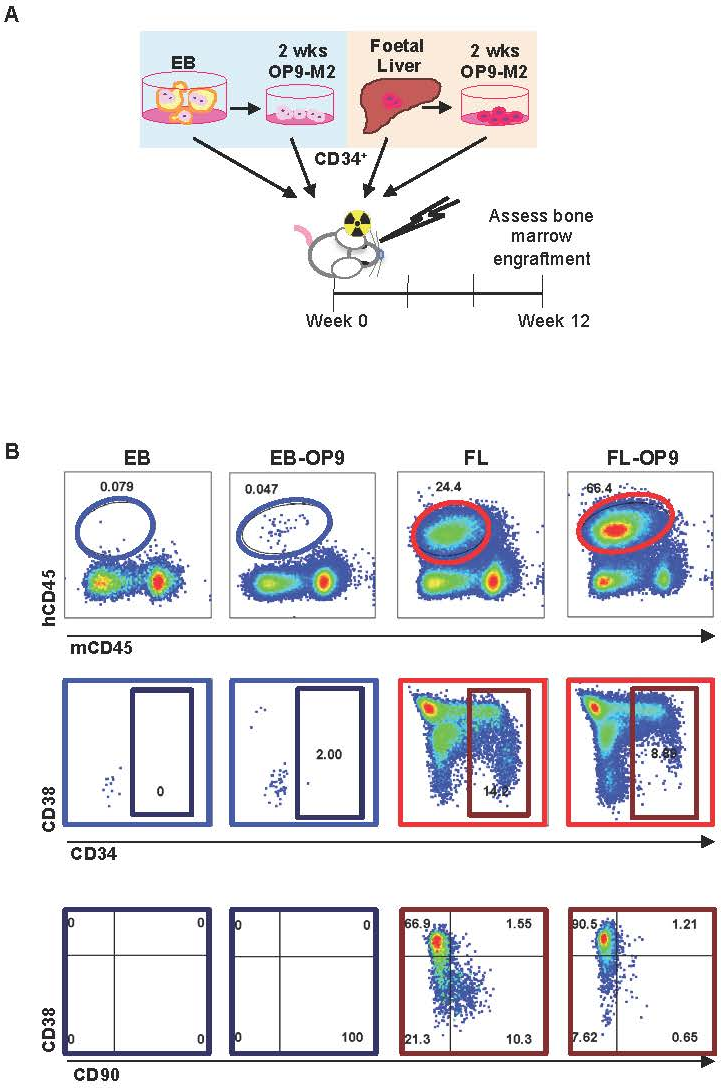
Accession numbers deposited in GEO database. Primary accession: GSE76685.
Referenced accessions: GSE54316[14] and GSE34974[26].

57. Dou, D. R., Calvanese, V., Saarikoski, P., Galic, Z. & Mikkola, H. K. A. Induction of HOXA genes in hESC-derived HSPC by two-step differentiation and RA signalling pulse. *Protocol Exchange* http://dx.doi.org/10.1038/protex.2016.035 (2016).
58. Thoma, S. J., Lamping, C. P. & Ziegler, B. L. Phenotype analysis of hematopoietic CD34+ cell populations derived from human umbilical cord blood using flow cytometry and cDNA-polymerase chain reaction. *Blood* **83**, 2103–2114 (1994).
59. Bauchwitz, R. & Costantini, F. Developmentally distinct effects on human epsilon-, gamma- and delta-globin levels caused by the absence or altered position of the human beta-globin gene in YAC transgenic mice. *Hum. Mol. Genet.* **9**, 561–574 (2000).
60. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).
61. Gentleman, R. C. *et al.* Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol.* **5**, R80 (2004).

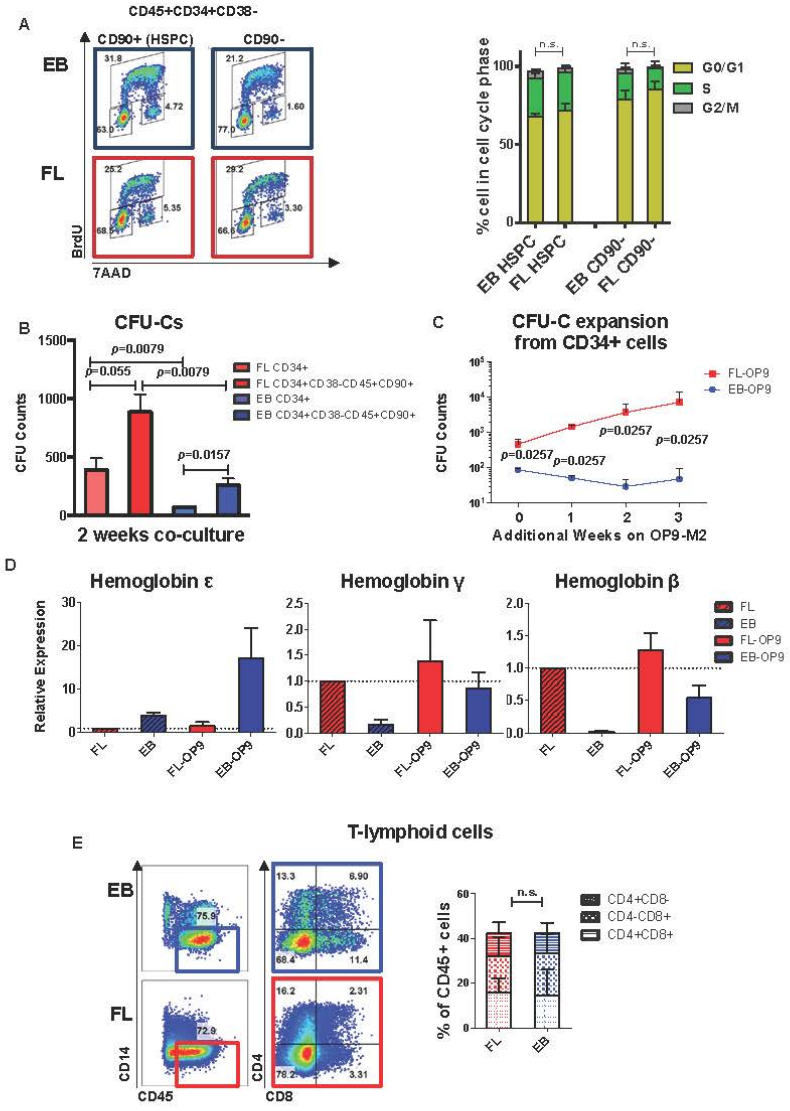
62. de Hoon, M. J., Imoto, S., Nolan, J. & Miyano, S. Open source clustering software. *Bioinformatics* **20,** 1453–1454 (2004).
63. Saldanha, A. J. Java Treeview–extensible visualization of microarray data. *Bioinformatics* **20,** 3246–3248 (2004).
64. Trapnell, C., Pachter, L. & Salzberg, S. L. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25,** 1105–1111 (2009).
65. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26,** 841–842 (2010).
66. Trapnell, C. *et al.* Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* **28,** 511–515 (2010).
67. Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y. & Greenleaf, W. J. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* **10,** 1213–1218 (2013).
68. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9,** 357–359 (2012).
69. Ramírez, F., Dündar, F., Diehl, S., Grüning, B. A. & Manke, T. deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res.* **42,** W187–W191 (2014).
70. Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9,** R137 (2008).

DOI: 10.1038/ncb3354

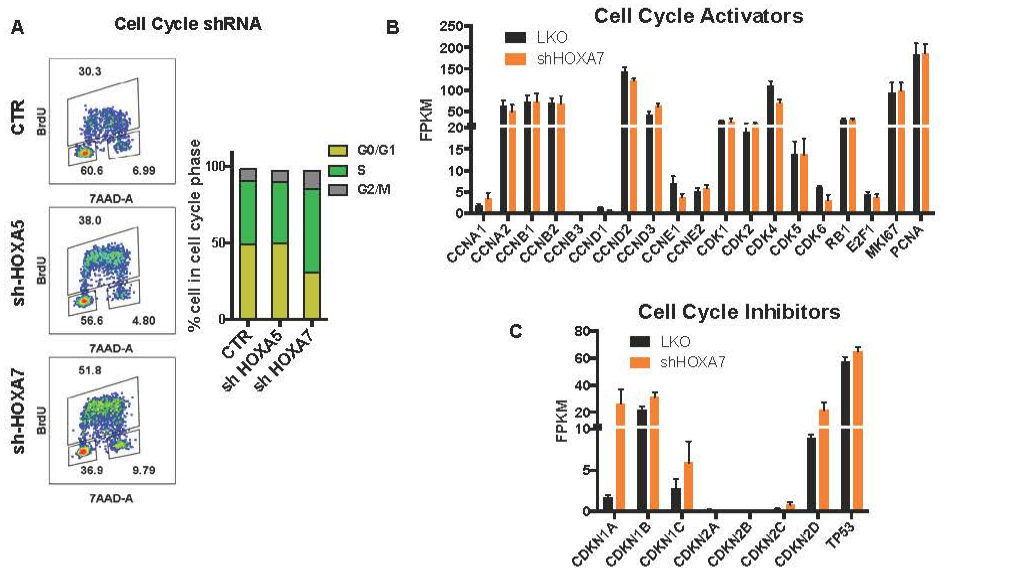
nature cell biology



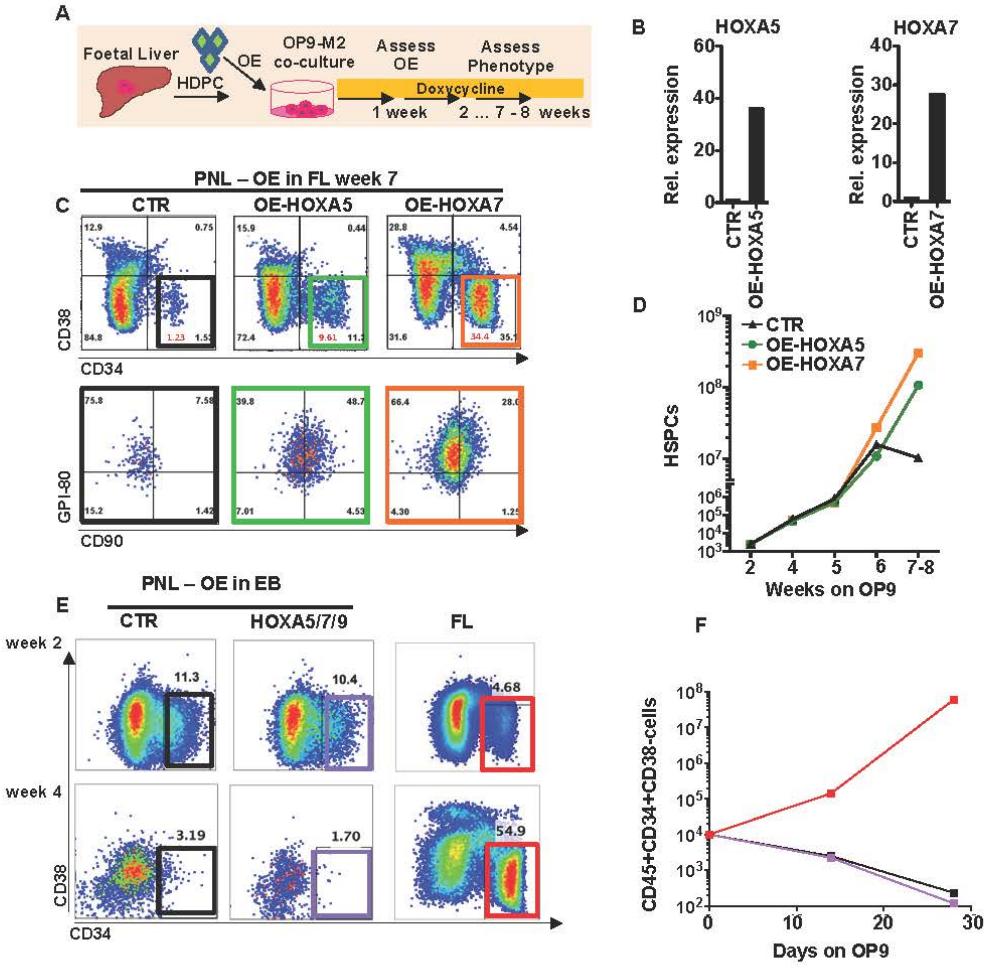
Supplementary Figure 1 FL derived, but not hESC derived haematopoietic cells can reconstitute human HSPC compartment in recipient BM. (A) Schematic of transplantation of CD34+ cells into irradiated NSG mice. (B) NSG mice were transplanted with CD34+ cells from hESCs (EB and EB-OP9) and fetal liver (FL or FL-OP9) and human engraftment in the BM assessed at 12 weeks for $CD45^+CD34^+CD38^-CD90^+$ immunophenotypic HSPCs. Results shown are from representative animals for each group of transplanted mice (5 mice transplanted with EB cells, 4 with EB-OP9 cells, 4 with FL cells, and 3 with FL-OP9 cells).



Supplementary Figure 2 hESC-derived haematopoietic cells can upregulate adult haemoglobin-beta (HBB) and differentiate into T-lymphoid cells (A) Representative FACS plots and quantification of BrdU incorporation and 7-AAD to determine cell cycle distribution in EB and FL CD90$^+$ immunophenotypic HSPCs and CD90$^-$ cells is shown (mean +/- SEM from n=3 independent experiments). (B) Comparison between CD34$^+$ haematopoietic cells and immunophenotypic (CD34$^+$CD38$^-$CD90$^+$CD45$^+$) HSPCs, all seeded at an initial density of 10,000 cells per sample, from FL and hESC-derived cells (mean +/- SEM of n=5 independent experiments). (C) CFU-C expansions from 10,000 hESC-derived or FL-derived CD34$^+$ cells in methylcellulose following 0, 1, 2 and 3 additional weeks on OP9-M2 co-culture (mean +/- SEM of n=3 independent experiments). (D) Haemoglobin levels (expression measured from colonies derived from CD34$^+$ cells) of embryonic epsilon (HBE), fetal gamma (HBG), and adult beta (HBB) measured through qRT-PCR and normalized to Glycophorin A levels (mean +/- SEM shown from n=5 independent experiments). (E) FACS staining of hESC- and FL-derived CD34$^+$ haematopoietic cells grown on OP9-DL1 stroma for 4 weeks is shown. Cells were stained for CD45, the myeloid exclusion marker CD14, and T-cell markers CD4 and CD8 (mean +/- SEM shown from n=3 independent experiments. Statistics source data for graphs shown in A, B, C, and E can be found in Supplementary Table 7. Statistical significance was assessed using the Wilcoxon Rank Sum test for A, B, C, and E.

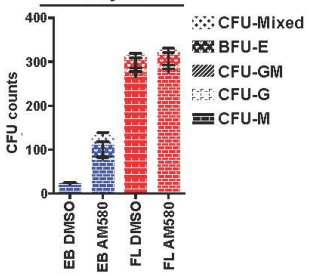


Supplementary Figure 3 Knockdown of HOXA5 or HOXA7 does not lead to changes in BRDU incorporation in FL immunophenotypic HSPCs. (A) Representative FACS plots and quantification of cell cycle analyses based on BrdU incorporation (mean from one experiment with 2 independent donors, statistics source data can be found in Supplementary Table 7) of control vector and HOXA5 and HOXA7 shRNA vector transduced FL-HSPCs. (B, C) Examples of cell cycle activators (B) and inhibitors (C) from RNA-seq analyses of FL immunophenotypic HSPCs with HOXA7 knockdown compared to empty vector controls (showing mean from 4 independent experiments, values used to generate graphs can be found in Supplementary Table 4 and GEO database GSE76685).



Supplementary Figure 4 Lentiviral overexpression of HOXA5, HOXA7 and HOXA9 in EB-derived CD34+ cells is not sufficient for rescuing HSC function. (A) Schematic showing the strategy for tet-inducible overexpression of HOXA5 or HOXA7 in FL-HSPCs using a PNL vector (B) q-RT-PCR showing induction of HOXA5 or HOXA7 expression in FL-HSPCs overexpressing HOXA5 or HOXA7, compared to empty vector control 1 week post-transduction (plotting one representative experiment). (C, D) Representative FACS plots (C) and quantification (D) of FL-HSPCs overexpressing HOXA5 or HOXA7 (mean from 3 independent experiments, except for 2 independent experiments for 7-8 weeks timepoint). (E) Representative FACS plots assessing concurrent overexpression of HOXA5, HOXA7 and HOXA9 using PNL vector. EB and FL CD34+ cells transduced with empty-vector were used as controls (mean from 2 independent experiments, except for EB-control and HOXA5/7/9 at day 14, 1 independent experiment). Statistics source data for values used to generate graphs shown in B, D, and E can be found in Supplementary Table 7.

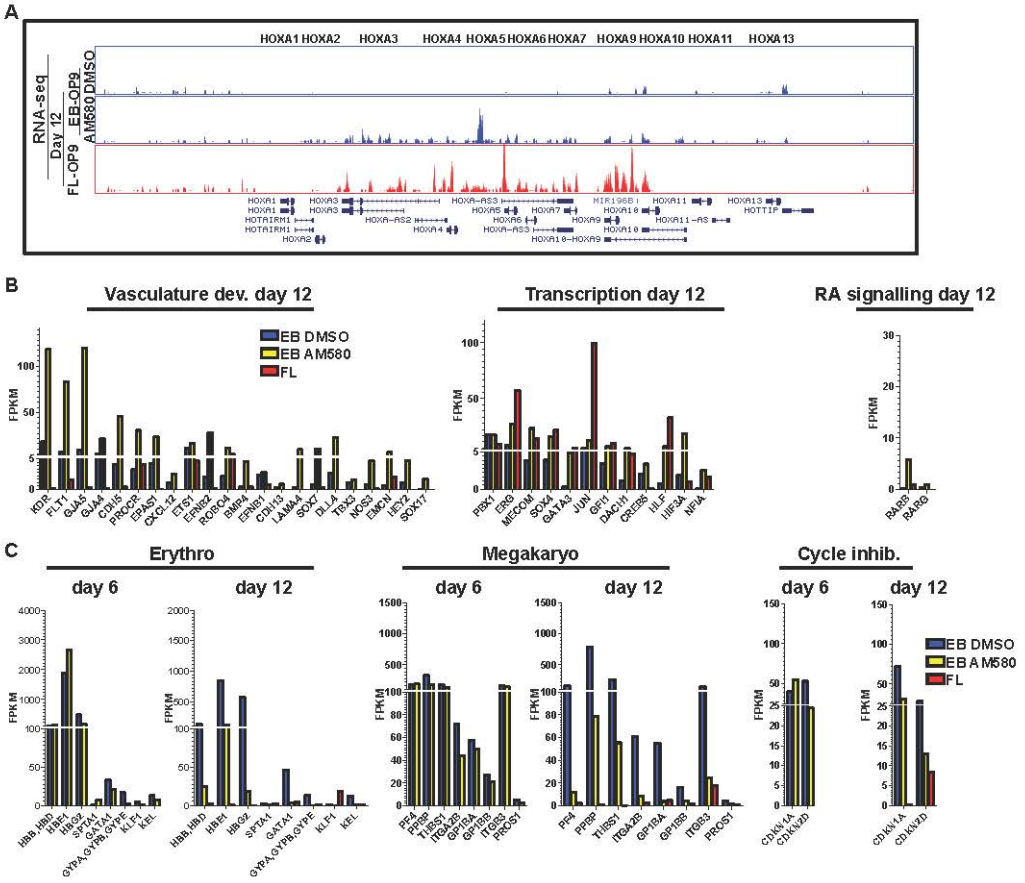
A



B

Exp.	CFU-Tot		CFU-Mix		BFU-E		CFU-GM		CFU-G		CFU-M	
EB	DM	AM	DM	AM	DM	AM	DM	AM	DM	AM	DM	AM
#1	12	260	0	28	2	24	0	4	0	6	10	198
#2	41	118	3	32	2	39	2	3	2	3	34	43
#3	32	102	2	27	1	41	0	2	2	5	27	27
#4	13	49	0	1	1	2	0	2	0	2	12	42
FL	DM	AM	DM	AM	DM	AM	DM	AM	DM	AM	DM	AM
#1	276	254	6	10	20	10	6	8	8	8	236	218
#2	347	470	7	3	7	17	7	0	0	10	327	440
#3	218	280	20	24	20	36	8	10	12	10	158	200
#4	422	304	18	12	30	34	16	8	14	8	344	242

Supplementary Figure 5 AM580 treatment prolongs CFU-C potential in hESC-derived cells. (A) Quantification of CFU-Cs generated from 10,000 EB- or FL-derived haematopoietic cells at day 24 ± 1 of OP9-M2 culture (mean +/- SEM from n=4 independent experiments, statistics source data can be found in Supplementary Figure 5B). (B) Table showing CFU counts for the indicated samples. Counts were rounded to the closest integer value (DM=DMSO, AM=AM580).



Supplementary Figure 6 Analysis of gene expression changes in hESC-HSPCs upon AM580 treatment shows partial conversion to definitive HSC transcriptome. (A) RNA-seq genome browser screenshot of the HOXA cluster of day 12 EB and FL derived immunophenotypic HSPCS that were treated with AM580 for 6 days (6 days of treatment and 6 additional days in culture). (B) Representative genes upregulated by AM580 treatment at day 6, shown at day 12 as compared to FL-HSPCs (C) Representative genes upregulated by HOXA7 shRNA knockdown (see Figure 4L) shown in day 12 EB derived cells treated with AM580 (6 days of treatment and 6 additional days in culture) as compared to FL-HSPCs (showing mean from 2 independent experiments, values used to generate graphs in B and C can be found in Supplementary Table 5 and GEO database GSE76685).

Supplementary Table Legends

Supplementary Table 1 Average probe values for all genes in HSPCs isolated at different stages of human developmental hematopoiesis in vivo and in vitro. Average probe values for all genes are shown for each of the analyzed samples on Affymetrics Human U133plus Microarray. All populations were sorted for HSPC phenotype ($CD34^+CD38^-CD90^+CD43^+CD45^{+/-}$) (FL = second trimester fetal liver; EB = 2 week embryoid bodies; OP9 = sorted cells after culture on OP9-M2 stroma for the indicated number of weeks (2 or 5); and PL= first trimester placenta).

Supplementary Table 2 Gene Ontology categories of genes differentially expressed between immunophenotypic HSPCs at different stages of human developmental haematopoiesis *in vitro* and *in vivo*. Genes differentially expressed (>2 fold, p-value < 0.05) between any of the samples are included in the clustering (see Figure 3C).

Supplementary Table 3 Differentially expressed gene list from microarray analysis with raw expression values and K means cluster organization. Genes differentially expressed (>2 fold, p-value < 0.05) between any of the samples are included in the clustering (see Figure 3C).

Supplementary Table 4 RNA sequencing analysis of HOXA7 shRNA in FL HSPCs showing differentially expressed gene list. Genes that were significantly up- or down-regulated (>1.8 fold, p-value < 0.05, 4 independent fetal liver tissues) between HOXA7 shRNA and control HSPCs are shown.

Supplementary Table 5 RNA sequencing analysis of hESC-HSPCs after AM580 treatment showing differentially expressed genes. Genes up-or down-regulated (>2 fold, p-value < 0.05) after 6 days of AM580 treatment are listed. For comparison, the average FKPM values for the same treatment after 6 additional days in culture and uncultured FL are shown.

Supplementary Table 6 Primers used for qPCR Quantification and Cloning. (A) SYBR Green Primers used in qPCR quantification studies. (B) Taqman® Primers used in qPCR quantification studies. (C) Primers used in cloning.

Supplementary Table 7 Statistics source data file. In experiments where n < 5, individual values for each replicate and independent experiment used to generate graphs in the figures are shown. Data corresponding to each Figure and Supplementary Figure are shown in separate spreadsheets.

Chapter 3:

A non-coding transcript in the human HOXA7 3' UTR is correlated with medial HOXA gene expression and HSC stemness

A non-coding transcript in the human HOXA7 3' UTR is correlated with medial HOXA gene expression and HSC stemness

Diana R. Dou[1,2], Vincenzo Calvanese[1], Bibhu Mishra[2], Tonis Org[1], Patricia Ernst[2], Zoran Galic[1], and Hanna K.A. Mikkola[1,2]

[1]Broad Stem Cell Research Center at UCLA

[2]Molecular Biology Institute, UCLA

[3]Department of Pediatrics and Pharmacology, University Colorado Denver

[4]Department of Chemistry and Biochemistry, UCLA

ABSTRACT

Previous studies established the requirement of medial HOXA gene expression for definitive human HSC identity and identified the lack of HOXA gene expression in hESC-derived HSPCs as the molecular block to self-renewal. Here, we examine a pathway by which the medial HOXA genes are regulated. The regulator of medial HOXA genes, MLL1, was expressed, but not bound to HOXA genes in EB-derived cells, suggesting the inability to recruit MLL1 to HOXA genes prevents their activation during ESC differentiation. The lncRNA HOTTIP has been shown to act as a molecular scaffold to facilitate expression of HOXA9-HOXA13 in human cells, but less is known about the regulators of the medial HOXA genes. We identified in the human genome a unique 5' transcription start site in a region of 89% sequence homology and conserved location to the mouse lncRNA Mistral, which in mice recruits MLL1 to HOXA6/HOXA7 genes. Moreover, we found high expression of putative human MISTRAL in FL-HSCs but not in their differentiated progeny or ESC-derived hematopoietic cells and observed high correlation of Mistral expression to medial HOXA gene expression in several cell types. These data nominate MISTRAL as a previously unidentified transcript in the human

HOXA gene cluster and novel regulator of the fetal HSC regulatory machinery and raise the hypothesis that MISTRAL functions in regulating the medial HOXA genes involved in conveying "stemness" in hematopoietic cells.

INTRODUCTION

Hematopoietic stem cells (HSCs) are instrumental in bone marrow transplantations and are defined by their ability to self-renew, engraft and generate all blood lineages. Unfortunately, the shortage in HLA-matched HSCs from bone marrow and cord blood currently limits our ability to use HSC transplantation to treat blood and immune diseases. As such, there is significant motivation to develop protocols to generate alternative sources of HSCs, either by expanding the available pool of HSCs or differentiating HSCs from pluripotent stem cells (PSCs, such as embryonic stem cells (ESCs). Both strategies require an inherent understanding of the regulatory pathways and molecular programs underlying the self-renewal and proliferation ability of HSCs.

Previously, we identified the medial HOXA genes as critical for self-renewal in human HSCs, and cells that lack HOXA gene expression, such as ESC-derived HSPCs, also display impaired self-renewal. In that study, we introduced a 2-step differentiation protocol to generate immunophenotypic $CD34^{+}CD38^{-}CD45^{+}CD90^{+}$ HSPCs from H1 ESCs. In this protocol, CD34+ cells were generated through embryoid body differentiation, and subsequently cultured for 2 weeks on OP9-M2 mesenchymal bone marrow stroma to simulate the environments a developing HSC would encounter *in vivo*. Stimulating the retinoic acid signaling pathway was able to significantly induce expression of the HOXA cluster genes, possibly through increased chromatin accessibility in the locus. However, RA signaling only transiently induced HOXA

gene expression and other elements must be involved in regulating the HOXA cluster during definitive hematopoietic specification (Dou et al., 2016b).

The methyltransferase MLL1 is a component of the TrxG/MLL complex responsible for establishing H3K4me3 marks to activate transcription (Schuettengruber et al., 2007; Schwartz and Pirrotta, 2007; Soshnikova, 2013; Soshnikova and Duboule, 2009), is an upstream regulator of the HOXA genes. Recent literature suggests that, although the SET1 domain of MLL1 has H3K4 methyltransferase activity, MLL1 may be globally critical for the H3K4 trimethylation of only 5% of promoters in mouse embryonic fibroblasts (MEFs) (Wang et al., 2009). While MLL1 knockout in MEFs did not change H3K4me3 or HOXA9 activity, MLL1 knockout combined with the depletion of WDR5 was able to reduce HOXA9 expression (Mishra et al., 2014) Additionally, we previously found that RA acid signaling also improved chromatin accessibility to the HOXA cluster, suggesting that locus inaccessibility may impede HOXA gene transcription in ESC-HSPCs and the requirement (Dou et al., 2016b).

Longnoncoding RNAs (lncRNAs) are defined as RNA transcripts > 200 nucleotides and are often polyadenylated and some may function as molecular scaffolds to improve gene accessibility and in recruiting transcriptional machinery, such as MLL, to otherwise inaccessible target genes. However, while smaller ncRNAs, such as miRNAs, shRNAs, and piRNAs, are highly sequence-conserved and function through sequence-specific base pairing with their targets, lncRNAs are not confined to sequence-specific interactions and thus are poorly conserved and often need the added criteria of synteny—conservation of landmark genomic elements around the lncRNA—, microhomology, and secondary structure similarity to be identified across species (Quinn et al., 2016). Several lncRNAs are transcribed within the human HOX clusters (Rinn et al., 2007). The lncRNAs HOTAIRM1 and HOTTIP are well characterized

in the human HOXA cluster. HOTAIRM1 was first identified as important in myelopoiesis, is expressed between HOXA1 and HOXA2, and regulates expression of HOXA1-HOXA4 through interaction with MLL/UBX and PRC2 (Wang and Dostie, 2016; Zhang et al., 2009). HOTTIP, encoded beside HOXA13, regulates expression of HOXA9-HOXA13 through recruitment of the MLL1 complex and interaction with WDR5 (Wang et al., 2011; Yang et al., 2014). The lncRNA Mistral interacts in an analogous way to activate expression of HOXA6 and HOXA7 in murine ESCs (Bertani et al., 2011), but an equivalent has not been identified in humans.

Here, we find that MLL1 displays defective binding to the medial HOXA genes in hESC-derived hematopoietic cells. Using RNA-seq and rapid amplification of cDNA ends (RACE), we identify a novel lncRNA within the medial HOXA gene locus that is highly correlated to HOXA6 and HOXA7 expression and restricted in human hematopoietic cells to the self-renewing faction. In this study, we propose the existence of human MISTRAL and further hypothesize a function for MISTRAL in the activation and maintenance of the medial HOXA gene locus that may be critical for definitive human HSC development.

RESULTS

MLL1 does not bind to medial HOXA genes in ESC-derived hematopoietic stem/progenitor cells

Our previous studies revealed that lack of expression of the HOXA genes in ESC-derived HSPCs differentiated in vitro is a major barrier for hESC-HSPC self-renewal. Transcriptome profiling using a microarray comparing self-renewing second trimester fetal liver (FL) HSPCs) to ESC-derived HSPCs showed absent or severely downregulated expression of the HOXA cluster genes, particularly the medial HOXA4-HOXA10 genes, both before (EB) and after the

maturation culture on OP9-M2 stroma (EB-OP9) (**Fig 1A**). In contrast, FL-HSPCs that also underwent OP9_M2 co-culture (FL-OP9) did not display significant changes in HOXA gene expression compared to freshly isolated FL-HSPCs, indicating that the dysregulated expression is not induced by OP9 culture (**Fig 1A**).

To understand why HOXA gene expression could not be induced, we assessed the expression and binding of the known upstream regulators of HOXA gene expression. The methyltransferase MLL1 and its binding partner, WDR5, add covalent modifications to histones. To test whether there were differences in MLL1 binding in 34^+ cells isolated from FL and EBs, MLL1-ChIP was peformed and compared with a Gal4-ChIP negative control. In the control regions examined, as expected, no differences were observed in the negative Gal4 baseline and MLL1-ChIP in the intergenic and HOXC8 regions whereas, in the positive control gene EYA1, MLL1 binding was clear in both the 34^+ FL and EB cells. However, although robust MLL1 binding to HOXA7 and HOXA9 was observed in the 34^+ FL cells, there was little MLL1 binding above the Gal4 baseline in the 34^+ ESC-derived HSPCs, suggesting that MLL1 does not bind to the medial HOXA genes in ESC-derived HSPCs (**Fig 1B)**. Nevertheless, transcriptome profiling between FL-HSPCs and ESC-derived immunophenotypic HSPCs showed that MLL1 and WDR5 mRNA was expressed at close-to-physiological levels also in the *in vitro* derived HSPCs. Since these levels remain virtually unchanged before (FL, EB) and after culture on OP9 stroma (FL-OP9, EB-OP9), MLL1/WDR5 expression levels were unlikely to represent a bottleneck for HOXA gene induction (**Fig 1C**). Taken together, these data suggests that, while MLL1 and WDR5 are both expressed in HOXA-deficient ESC-derived HSPCs, defective MLL1 binding in the HOXA cluster contributes to poor HOXA gene expression.

Syntenic conservation and sequence homology suggest the presence of a human equivalent to the lncRNA Mistral

Several studies have documented lncRNAs associating with the MLL1 methyltransferase complex to recruit the complex to target genes (Bertani et al., 2011; Yang et al., 2014), and thus aid in gene activation. The lncRNAs HOTAIRM1 and HOTTIP are both known regulators within the HOXA cluster, acting proximally and distally, respectively. HOTAIRM1 is expressed between HOXA1 and HOXA2 and impacts the expression of HOXA1-HOXA4 to regulate differentiation of myeloid and granulocytic cells and is also highly expressed in leukemia and colorectoral cancer (Díaz-Beyá et al., 2015; Wan et al., 2016; Wei et al., 2016; Zhang et al., 2009; Zhang et al., 2014). HOTTIP has been proposed to regulate HOXA9-HOXA13 expression in distal anatomic sites, such as foreskin fibroblasts and the prostate, as well as in hepatocarcinomas (Wang et al., 2011). However, no expression of HOTAIRM1 or HOTTIP was observed in either FL-HSPCs or ESC-derived HSPCs, suggesting that these lncRNAs do not activate the HOXA genes important in human HSCs. As HOXA cluster associated lncRNAs have been shown to activate in cis directly next to their target genes (Dasen, 2013; Wang et al., 2011; Zhang et al., 2009), we examined the medial HOXA gene neighborhood in RNA sequencing data, and identified a peak in the non-coding, 3' UTR region of HOXA7 that is prominent in the FL-derived HSPCs, but absent from ESC-HSPCs (orange arrow, **Fig 2A**). This peak in the 3' UTR region is covered by the Affymetrix probe "235753_at". RNA microarray comparison verified very low expression of that region in ESC-derived HSPCs compared to FL-HSPCs, which express this region stably also post-culture (**Fig 2B**). Taken together, both RNA-seq and microarray data suggest the peak in the HOXA7 3' UTR associates closely with medial HOXA gene expression and HSC self-renewal properties.

The annotated length of the HOXA7 3' UTR is much longer in humans than in mice, suggesting that the peak annotated to the 3' UTR region of human HOXA7 is localized outside of the murine HOXA7 UTR. In mice, there is a reported lncRNA named Mistral between HOXA6 and HOXA7, closer to HOXA7 than HOXA6. Mouse Mistral is an unspliced, single exon expressed transcript and has been reported to work in *cis* to regulate the expression of HOXA6 and HOXA7 (Bertani et al., 2011). To test the hypothesis that the FL-HSPC associated peak next to the human HOXA7 gene is the human equivalent of Mistral, we compared the sequence conservation with both UCSC BLAT (**Fig 2C**) and the pairwise sequence alignment tool LALIGN on the EMBL-EBI server (Li et al., 2015) (**Fig 2D**). The UCSC BLAT query of sequence alignment of the annotated mouse Mistral was found to share high conservation (89.1%) with a region of the human HOXA7 UTR (**Fig 2C**). Analysis of pairwise alignment using LALIGN software showed 75.6% similarity with the 798 basepair murine Mistral and a corresponding 820 nt region of in the human genome (**Fig 2D**).

In order to further define the identity of the putative human MISTRAL, we first conducted strands-specific RNA-sequencing to identify the strand that expresses the MISTRAL transcript. Consistent with murine MISTRAL's expression from the same strand as the HOXA genes, strand-specific RNA-seq in human $CD34^{+}CD38^{-}90^{+}$ FL-HSPCs showed a distinct and separate peak in the HOXA7 3' UTR in the same sense strand as the HOXA genes (**Fig 2E**). No signal was detected in the immediate vicinity of the HOXA6 and HOXA7 genes on the opposite strand.

As non-coding RNAs are often not sequence-conserved across species, the finding of high sequence homology in addition to conservation of synteny suggests an important role for MISTRAL that is retained through evolutionary branching.

HOXA6/HOXA7 expression is correlated with MISTRAL expression

Since Mistral is shown to regulate HOXA6 and HOXA7 during mouse ESC differentiation (Bertani et al., 2011), we used a microarray of RNA expression to examine the expression of HOXA6 and HOXA7 and putative MISTRAL in human cells. The probe "235753_at", described above, which encompasses a short span in the HOXA7 3' UTR overlapping the predicted MISTRAL ortholog in humans, was used to gauge putative human MISTRAL expression. Among hematopoietic progenitor and stem cells, HOXA6 and HOXA7 were expressed at highest levels in the self-renewing GPI80^{+} HSPCs, and decreased as the cells become progenitor-like and first lose GPI80 expression, then CD90 (Thy1) expression, and finally elevate the differentiation marker, CD38. The probe designated as MISTRAL also gave off the highest signal in GPI80+ self-renewing HSPCs, followed by the same drastic loss in expression as the cells lose stemness and become CD38^{hi} progenitors **(Fig 3A)**. RNA-sequencing of highly self-renewing GPI80^{+} FL-HSPCs compared to non-self-renewing CD38^{hi} hematopoietic progenitors showed the high expression of the HOXA genes, in particular HOXA4-HOXA10, and MISTRAL in the self-renewing HSCs, whereas the non-self-renewing CD38^{hi} hematopoietic progenitors do not express the HOXA genes or MISTRAL (**Fig 3B**), further supporting the notion that the medial HOXA genes and MISTRAL are linked to self-renewal.

We next assessed if the pattern linking MISTRAL expression to HOXA6 and HOXA7 expression emerges also in other cell types than hematopoietic cells. In a dataset comprised of 117 samples from 89 unique sources ranging from adipocytes, iPSCs, neurons, ESCs, HUVECs, and other human primary cells and cell lines (Germanguz et al., 2016), the R^2 correlation of

MISTRAL to HOXA6 expression was 0.678 and to HOXA7 expression was 0.9069, while correlation to HOXA1 and HOXA13 were 0.3345 and 0.01051, respectively (**Fig 3C**). MISTRAL expression was not correlated to randomly selected genes that are not expected to associate with the HOXA cluster, such as ACTN3 (R^2=0.2692), APOB (R^2=0.104), and TAL1 (R^2=0.04289) (**Fig 3C**). This observation suggests a correlation, in *cis*, between MISTRAL and the neighboring medial HOXA genes, that is not restricted to only cells of the hematopoietic lineage.

The human HOXA7 3' UTR harbors a unique 5' transcription start site

HSCs and even primary 34^+ FL cells and 34^+ EBs are difficult to obtain in large quantity, which necessitated the use of cell lines for assays requiring large quantities of cells. Since MISTRAL and HOXA6/HOXA7 correlation was observed in a variety of non-hematopoietic cells as described above (**Fig 3C**), a cell line with high signal for the Affymetrix microarray probe corresponding to the MISTRAL region was chosen in lieu of HSCs for experiments requiring several millions of cells. A cell line that highly expressed this probe region as determined through a search on the BioGPS gene annotation portal (Wu et al., 2016; Wu et al., 2013; Wu et al., 2009) was DLD1 (**Fig 4A**), a colorectoral cancer cell line that also had high expression of HOXA6 and HOXA7, as observed in RNA-seq (**Fig 4B**).

We employed 5' RACE (Rapid Amplification of cDNA ends), a technique used to determine where transcript ends are located, to test for the presence of a unique 5' transcription start site in the HOXA7 3' UTR. Four different primers were designed to amplify in the rightwards direction (towards 5' of the HOXA7 gene) in the UTR region (**Fig 5A**). The results from all four primers' reactions were two distinct products that matched the sizes of a full

HOXA7 transcript and a smaller product close to the predicted size of MISTRAL (**Fig 5B**). Sequencing of the products showed that the much larger product terminated at the beginning of the annotated HOXA7 exon 1 (Chromosome 7, 27196228) (**Fig 5C**) and the smaller product returned a second, previously undescribed transcription start site (Chromosome 7, 27194325) (**Fig 5D**) in the HOXA7 3' UTR.

Since genes actively transcribed by PolII have activating H3K4me3 marks at the promoter and the PolII elongation mark H3K36me3 on the transcribed region, we looked for K4-K36 domains (Guttman et al., 2009) outside the annotated genes in the vicinity of the medial HOXA cluster. A strong peak was observed in both RNA-seq and the H3K36me3 ChIP-seq that is separate from the HOXA7 exons in the HOXA7 3' UTR corresponding to the predicted MISTRAL transcript (**Fig 5E**). Upon aligning the novel transcription start site from 5' RACE (Chromosome 7, 27194325) in the HOXA7 3' UTR to the sequencing data, we observed that the start site tightly corresponded with the initiation of the H3K36me3 elongation peak (**Fig 5E**). The H3K4me3 mark in HOXA7 was too broad to determine whether there was also a separate, distinct peak that could correspond to MISTRAL (**Fig 5E**), which does not rule out either possibility that 1) MISTRAL could share a promoter with HOXA7 or 2) MISTRAL and HOXA7 have distinct promoters. However, Cap Analysis of Gene Expression (CAGE) data, which utilizes high throughput sequencing of 5' mRNA ends to profile transcription start sites (Shiraki et al., 2003), from the FANTOM5 hub (Forrest et al., 2014; Lizio et al., 2015), shown as the bottom-most track, indicate possible sites of initiation around the predicted beginning of MISTRAL in the vicinity of the start site identified in 5' RACE.

3' RACE identifies multiple candidate 3' end sites

3’ RACE was used to assay whether or not this unique transcript also possessed a unique termination site. Six different primers amplifying in the leftwards direction (towards the 3’ end of the HOXA7 UTR) (**Fig 6A**) were used in nested combinations that, upon sequencing of the excised band products (from gel in **Fig 6B**), returned 3 distinct end sites. One site corresponded to the annotated 3’ end of the HOXA7 UTR (Chromosome 7, 27193335), another approximately 260 base pairs from the end of the 3’ UTR (Chromosome 7, 27193596), and the third even further from the end (Chromosome 7, 27193794). Since the last product ends before the region amplified by the 5’ RACE primer, Right 619 (**Fig 5A, Table 1**), to the novel 5’ start site in the 3’ UTR (5’ RACE product: Chromosome 7, 27193756-27194325 (**Fig 5D**)), it can be eliminated as a possible 3’ termination site (**Fig 6C**). These analyses thus identified two candidate 3’ end sites, as we cannot rule out the possibility that MISTRAL does not terminate in the same position as the HOXA7 transcript.

DISCUSSION

In this study, we found that MLL1 and WDR5 display defective binding to the medial HOXA genes in ESC-derived hematopoietic cells. We then sought to confirm the existence of a human equivalent to Mistral and found the presence of a unique transcript in the noncoding 3’ UTR region of HOXA7. Putative MISTRAL is highly expressed in self-renewing HSPCs and is absent in non-self-renewing hematopoietic cells and is also highly correlated to the expression of HOXA6 and HOXA7 in multiple cell lines. In terms of conservation, putative human MISTRAL and mouse Mistral share high sequence homology and conservation of synteny. It remains to be seen if secondary structures are conserved as well and whether the two are orthologous (Quinn et al., 2016).

Bertani et al., 2011, suggested that in mice, Mistral directly interacts with MLL1 to activate the expression of HOXA6 and HOXA7. Knockdown of Mistral led to the inability to induce HOXA6 and HOXA7 gene expression during mouse ESC differentiation even in the presence of retinoic acid (Bertani et al., 2011). Yang et al., 2014, showed that the lncRNA HOTTIP directly binds with WDR5 to activate gene expression (Yang et al., 2014). In order to test whether MISTRAL in human cells directly binds MLL1, WDR5, or both, future CLIP-Seq experiments involving both MLL1 and WDR5 will be conducted.

Alternative methods to isolate the full-length sequence of human MISTRAL, such as Northern Blot, are necessary to determine which of the transcription end sites identified in 3' RACE is the true end. The full length of MISTRAL is required for functional studies using CRISPR to delete MISTRAL in cells expressing high levels of Mistral and the medial HOXA genes. These studies will evaluate the impact of MISTRAL on medial HOXA gene expression in humans. In FL-HSPCs, MISTRAL deletion studies can also be used to determine whether MISTRAL is also a critical element of the definitive hematopoietic landscape or merely an incidental, non-functional indicator of hematopoietic cells with self-renewal potential.

However, traditional overexpression methods are unlikely to succeed in *cis*-activating lncRNAs as they are distance dependent (*cis*-restricted) and must be expressed immediately proximal to the genes of interest. Ectopic expression of HOTTIP delivered by retroviral infection failed to activate HOXA gene expression and required the addition of the BoxB RNA element for recruitment (Wang et al., 2011). Retroviral/lentiviral overexpression of MISTRAL is also unlikely to succeed without the identification of an additional targeting adaptor to ensure proper localization and tethering to the medial HOXA gene cluster. Thus, future studies will also focus

on determining the elements required to recruit MISTRAL to its appropriate gene locus as well as the regulatory pathways that may induce MISTRAL expression endogenously.

Preliminary data suggests that human MISTRAL first appears after HOXA7 expression in the earliest HSPCs arising from the AGM. This result would suggest that MISTRAL may not be required to induce medial HOXA gene expression in human HSCs and might, instead play a critical role in maintaining HOXA gene expression. If such is the case, then the RA signaling pathway, shown to be required for HOXA cluster expression induction (Dou et al., 2016b), and MISTRAL may play complementary roles in regulating medial HOXA gene expression in human definitive hematopoiesis.

Future studies to test the impact of manipulating MISTRAL expression, and ways to induce expression of MISTRAL at its locus, on HSPC self-renewal will provide an interesting link in the overall pathways involved in HOXA gene regulation and definitive hematopoietic specification. This study has the broader ranging impact of furthering our understanding of lncRNA conservation across species and the interplay between coding genes and structural elements in regulating gene expression in underlying human hematopoiesis. In conclusion, we have identified a novel HOXA-associated lncRNA with high similarity to murine Mistral that may also be used as an indicator of "stemness" in human HSPCs.

ACKNOWLEDGEMENTS

We thank the BSCRC FACS core at UCLA, the UCLA Clinical Pathology Microarray Core, BSCRC Sequencing Core, UCLA Tissue and Pathology Core and CFAR Gene and Cell Therapy Core (NIH grant AI029697-21). We thank A. Hussein and T. Johnson for consultation on RACE experiments and W.E. Lowry for discussions. This work was supported by (SAME GRANTS AS

HOXA?). H.K.A.M. was supported by the LLS Scholar award and Rose Hills Foundation Scholar Award; D.R.D. was supported by the National Science Foundation Graduate Research Fellowship Program (NSF GRFP) and Ruth L. Kirschstein National Research Service Award GM007185; and V.C. by an LLS Special Fellow Award and a BSCRC post-doctoral fellow award.

AUTHOR CONTRIBUTIONS

D.R.D., V.C., B.M., J.G., G.C., P.E., Z.C., and H.K.A.M. designed experiments and interpreted data. D.R.D., V.C., B.M., and J.G. performed experiments. D.R.D. and H.K.A.M. wrote the manuscript.

CONFLICTS OF INTEREST

The authors declare no competing conflicts of interest.

MATERIALS AND METHODS

Tissue Culture

Human H1 ESCs were maintained and passaged weekly as previous described (Dou et al., 2016a; Dou et al., 2016b) and a step-by-step method of the differentiation protocol of embryoid bodies for two weeks in the presence of cytokines can be found at Nature Protocol Exchange (Dou et al., 2016a). 14-17 week human fetal liver collection and processing and culture of ESC-derived and FL cells on OP9-M2 stroma were all performed as previously described (Dou et al.,

2016a; Dou et al., 2016b). All hESC work was approved by the UCLA Embryonic Stem Cell Research Oversight committee.

Cell Lines

The DLD1 colorectal cancer cell line was obtained from ATCC (CCL-221) and cultured in RPMI Media 1640, 1X (Corning Cellgro Ref 15-016-CV) supplemented with 10% fetal bovine serum (Corning Cellgro), 1% penicillin/streptomycin (Life Technologies Ref 15140-122), 1 uM sodium pyruvate (Gibco Ref 11360-070) and 10 uM HEPES (Gibco Ref 15630-080) and passaged when cells reached 80% confluency. The K-562 chronic myelogenous leukemia cell line (ATCC CL-243) was cultured in Iscove's DMEM, 1X (Gibco Ref 11875-093) supplemented with 10% fetal bovine serum (Corning Cellgro) and 1% penicillin/streptomycin (Life Technologies Ref 15140-122) and passaged when cells reached 80% confluency.

Microarray Analysis

Array information for the HSPC graphs can be found in NIH GEO dataset Series GSE34974 (Magnusson et al., 2013), GSE 54316 (Prashad et al., 2014), GSE76685 (Dou et al., 2016b).

Array information for the graphs showing the correlation of multiple cell lines and primary cells (Fig 4C) can be found in NIH GEO dataset Series GSE47796 (Germanguz et al., 2016).

RNA isolation, cDNA synthesis, and quantitative RT-PCR

RNA isolation, cDNA synthesis, qPCR for GAPDH, HOXA6, and MISTRAL were performed using the TaqMan Gene Expression Master Mix (LT), qPCR for GAPDH and HOXA7 were performed using the SYBR Select Master Mix (Life Technology, LT), and analyzed as

previously described (Dou et al., 2016b). The MISTRAL primer was custom ordered from TaqMan (Assay Name: HMISTRAL, Assay ID: AI39SEO) with the Forward Primer 5'-GTGCCAAGCCAGAGAAGGA-3', Reverse Primer 5'- CAGGGTATTGGCATCTATTTAAA TCGAAAAATAATATATTTAT-3', and Reporter 5'-CTTGCACCCTAGAATCA-3'. All other primer information are previously described (Dou et al., 2016b).

Fluorescence Activated Cell Sorting (FACS)

Flow cytometry and cell sorting were performed as previously described (Dou et al., 2016b). Cell sorting was performed using the BD FACS Aria II (for RNA-seq experiments) or BD FACS Aria III (for ChIP-seq experiments).

RNA-sequencing analysis

Strand-specific analysis for the tracks EB, EB-OP9, FL, and FL-OP9 used in Fig 2A: Library was prepared using the Illumina TruSeq Stranded mRNA kit and data was analyzed using the command options: --library-type fr-firststrand with Tophat v2.0.8, Bowtie2 2.1.0.0, and Samtools 0.1.19.

Single end analysis for the tracks $GPI80^{+}$ FL and $CD38^{hi}$ FL used in Fig 3A was prepared using the. Nugen Ovation RNA-seq System V2 kit and data was analyzed using the command options: --no-coverage-search with Tophat v2.0.8, Bowtie2 2.1.0, and Samtools 0.1.19.

The DLD1 total RNA RNA-seq data was obtained from the DDBJ database DRA000308 (Yamashita et al., 2011) and analyzed using the command options: --no-coverage-search with Tophat v2.0.14, Bowtie2 2.2.9, and Samtools 0.1.19. The K-562 track was obtained from ENCODE/Caltech (Consortium, 2012).

ChIP-sequencing

Antibodies for H3K4me3 (abcam ab8580), H3K36me3 (abcam ab9050), and Rabbit IGG (Santa Cruz Antibodies sc-2027) were used at 1 ug/sample. Each sample contained 250,000 cells. Cell fixation in 1% formaldehyde, library preparation using the Nugen Ovation Ultralow Kit and sequencing through HiSeq 2000 (Illumina), and ChiP-sequencing analysis using Bowtie and peak calling with MACS were performed as previously described (Org et al., 2015).

5'RACE

5 ug of total DLD RNA was used in the SMARTer RACE 5'/3' Kit (Clontech Cat# 624858). Primers were designed according to manufacturer's instructions and listed in Table 1. The final PCR amplified reactions were run on 1.5% agarose gels with 0.005% ethidium bromide. Bands were extracted and gel-purified using the Zymoclean™ Gel DNA Recovery Kit (Zymo Research Cat# D4001) and products cloned into One Shot® TOP10 Chemically Competent *E. coli*, minipreps prepared and collected using the Invitrogen™ PureLink™ Quick Plasmid Miniprep Kit (Cat# K210011). Products were sent for sequencing to Laragen using the company's M13 Forward primer and aligned to the hg19 genome using the UCSC BLAT tool.

3'RACE

50 ng of polyA-purified DLD1 RNA was used in the FirstChoice® RLM-RACE Kit (Ambion AM1700) and prepared according to manufacturer's instructions, including the optional step of the "nested" or inner RACE reaction. Primers were also designed according to manufacturer's instructions and listed in Table 1. The final PCR amplified reactions were run on 1.5% agarose

gels with 0.005% ethidium bromide. Bands were extracted and gel-purified using the Zymoclean™ Gel DNA Recovery Kit (Zymo Research Cat# D4001) and products cloned into One Shot® TOP10 Chemically Competent *E. coli*, minipreps prepared and collected using the Invitrogen™ PureLink™ Quick Plasmid Miniprep Kit (Cat# K210011). Products were sent for sequencing to Laragen using the company's M13 Forward primer and aligned to the hg19 genome using the UCSC BLAT tool.

Graphs and Statistical Methods

Graphs and R^2 correlation data were generated using the GraphPad Prism 6 software.

Figure 1: **MLL1 fails to bind to the medial HOXA cluster genes in non-self-renewing ESC-HSPCs. A** Fold change expression as determined by microarray analysis of the HOXA genes compared to human HSPCs isolated from definitive fetal liver (FL)-HSPCs in cultured FL-HSPCs (FL-OP9) and ESC-derived HSPCs, both cultured (EB-OP9) and un-cultured (EB). **B** Graphs plotting signal compared to input from ChIP of MLL and GAL4 (control) in 34+ FL and 34+ EBs in control (intergenic, eya1) and HOXA7 and HOXA9 regions. **C** Graphs showing the expression of MLL and WDR5 from microarray in human HSPCs isolated from FL, FL-OP9 as well as ESC-derived HSPCs. Values used to generate graphs from microarray data can be found in GEO database GSE76685.

Figure 2: A region of the human HOXA7 3' UTR shares high sequence identity with the murine lncRNA MISTRAL. **A** RNA-seq data of FL- and ESC-HSPCs, both pre- (FL, EB) and post- OP9-M2 co-culture (FL-OP9, EB-OP9), shows low expression of the HOXA genes and "MISTRAL" in (indicated with orange arrow) in non-self-renewing human ESC-HSPCs. **B** Expression of the probe "235753_at" encompassing human "MISTRAL" from DNA microarray in human HSPCs isolated from FL and ESC-derived, both un-cultured and cultured. Microarray data used in 2A and 2B can be found in GEO database GSE 76685. **C** UCSC BLAT comparison of sequence alignment shows high conservation between human (i, red box) and mouse (ii, blue box) MISTRAL and in the HOXA7 region (iii, conservation tracks). **D** Sequence alignment using the LALIGN tool shows highly shared identity between the annotated MISTRAL from the mouse genome (bottom strand sequence) and an 820 nt region in the human (top strand sequence) HOXA7 3' UTR. ":" indicates a base pair match in that position and "-" indicates a

base that does not exist in that position for the corresponding genome. **E** Strand-specific RNA-seq isolates a 3'UTR peak (orange box) in the same sense strand as the HOXA genes.

Figure 3: Expression of HOXA6 and HOXA7 is correlated to MISTRAL expression in several cell types. A Expression of the Affymetrix probe "235753_at" corresponding to MISTRAL region and HOXA6 and HOXA7 genes in self-renewing GPI80^{+} and non-self-renewing GPI80- HSPCs and in CD90^{+} HSPCs and CD90- hematopoietic progenitors. Graphs generated from microarray data from the GEO database GSE4316 and GSE34974. **B** RNA-seq data showing the HOXA cluster gene region in GPI80+ highly self-renewing FL-HSPCs (red), CD38hi FL hematopoietic progenitors (green), and CD90+ ESC-derived immunophenotypic HSPCs (blue). **C** (i) Plots of HOXA6 vs MISTRAL and HOXA7 vs. MISTRAL with R^2 correlation in pink. (ii) R^2 correlation of expression of the HOXA cluster genes to MISTRAL expression and expression of control genes to MISTRAL expression. Graphs generated from microarray data found in GSE47796.

Figure 4: Expression of HOXA7 and MISTRAL region is high in the DLD1 colorectal cancer cell line and absent in K562 cells. **A** Search using the BioGPS database of Probe 235753_at "MISTRAL" shows the expression of this location of multiple cell lines, including DLD1 (purple arrow) and K562 (light blue arrow). **B** RNA-seq of the DLD1 and K562 cell lines of the medial HOXA gene region and the control gene ACTNB shows high expression of ACTNB in both cell lines. K562 is used as a negative control due to lack of HOXA gene expression. DLD1 shows high expression of HOXA6 and HOXA7 as well as the peak for MISTRAL, in keeping with HOXA6/HOXA7 expression correlating with MISTRAL expression

Figure 5: **5' RACE identifies a second TSS distinct from the annotated HOXA7 TSS located in the 3' UTR region of HOXA7**. **A** Locations of primers designed for 5' RACE in the predicted region of MISTRAL transcribing in the rightwards direction of the HOXA7 strand. **B** Gel of products from 5' RACE experiments (primers labeled in blue) in DLD1 cells with two distinct 5' start sites as shown by bands (larger product predicted HOXA7, smaller product predicted MISTRAL). **C** Longer sequence amplified from MISTRAL Right 619 primer (shown in yellow box) from heavier band extraction extends to the 5' UTR region of HOXA7 (Chromosome 7:27196228). **D** Shorter sequence amplified from MISTRAL Right 619 primer (shown in orange box). Three other primers extended to same 5' start region in the HOXA7 3' UTR (Chromosome 7:27184325), indicated by the arrow. **E** RNA-seq, H3K36me3 and H3K4me3 assay of CD90+ FL-HSPCs zoomed in to the HOXA6/HOXA7 region shows the possible existence of a second transcript in the HOXA7 3' UTR (boxed in orange). 5' RACE of peak region identifies a unique 5' start site in the HOXA7 3' UTR, in compliance with predicted TSS from the FANTOM5 CAGE data (blue peak=plus strand, red peak=minus strand as oriented by HOXA gene transcription).

Figure 6: **3' RACE identifies 3 possible termination sites for predicted MISTRAL. A** Locations and of primers used for 3' RACE in the HOXA7 3' UTR region corresponding to putative human MISTRAL transcribing in the leftwards direction. **B** Gel showing products from nested reactions using two primers sequentially, outer reactions using only one primer, and the positive mouse b-actin control. All band products were sequenced to determine 3' termination

sites shown in **C**. There were three 3' termination sites in the 3' UTR of HOXA7 identified through 3' RACE.

Table 1: Sequences of primers used in 5' and 3' RACE designed in the human HOXA7 3' UTR region corresponding to predicted MISTRAL.

Figure 1

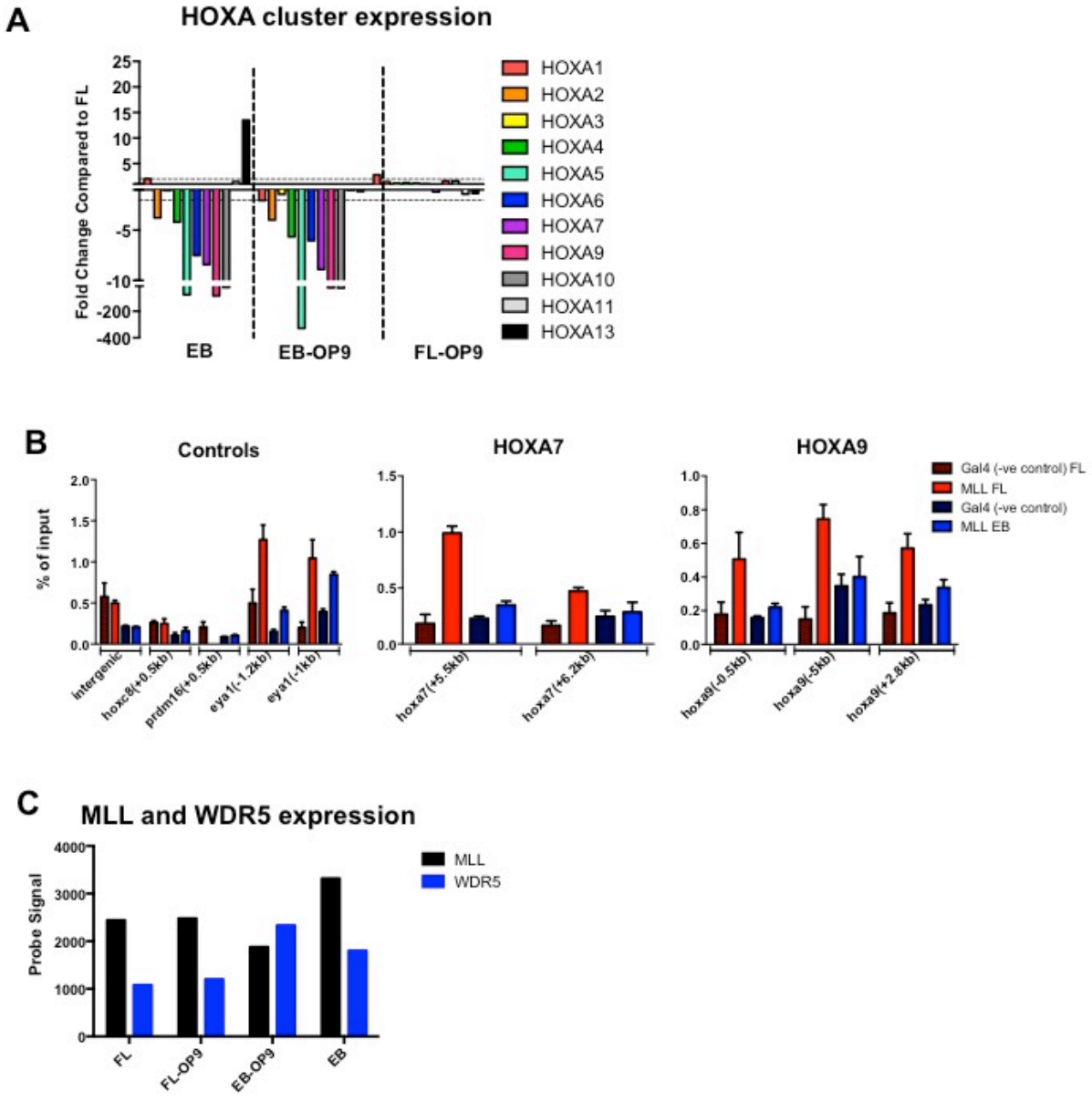

A
HOXA cluster expression
Fold Change Compared to FL
25
20
15
10
5
-5
-10
-200
-400
EB
EB-OP9
FL-OP9
HOXA1
HOXA2
HOXA3
HOXA4
HOXA5
HOXA6
HOXA7
HOXA9
HOXA10
HOXA11
HOXA13
B
Controls
HOXA7
HOXA9
% of input
Gal4 (-ve control) FL
MLL FL
Gal4 (-ve control)
MLL EB
intergenic
hoxc8(+0.5kb)
prdm16(+0.5kb)
eya1(-1.2kb)
eya1(-1kb)
hoxa7(+5.5kb)
hoxa7(+6.2kb)
hoxa9(-0.5kb)
hoxa9(-5kb)
hoxa9(+2.8kb)
C
MLL and WDR5 expression
Probe Signal
MLL
WDR5
FL
FL-OP9
EB-OP9
EB

Figure 2

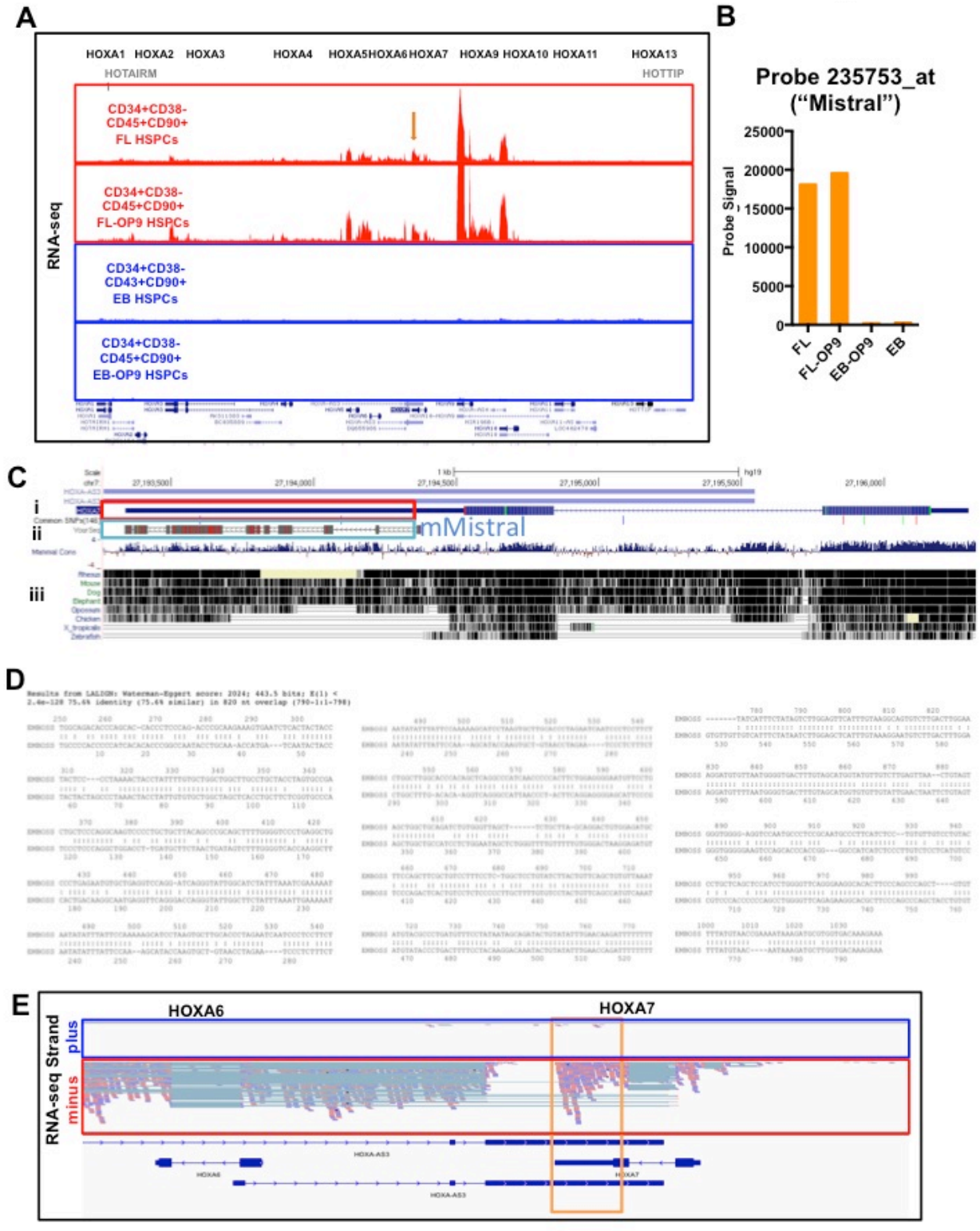

A
HOXA1 HOXA2 HOXA3 HOXA4 HOXA5 HOXA6 HOXA7 HOXA9 HOXA10 HOXA11 HOXA13
HOTAIRM
HOTTIP
RNA-seq
CD34+CD38-CD45+CD90+ FL HSPCs
CD34+CD38-CD45+CD90+ FL-OP9 HSPCs
CD34+CD38-CD43+CD90+ EB HSPCs
CD34+CD38-CD45+CD90+ EB-OP9 HSPCs
B
Probe 235753_at ("Mistral")
Probe Signal
25000
20000
15000
10000
5000
0
FL
FL-OP9
EB-OP9
EB
C
i
ii
iii
mMistral
D
E
HOXA6
HOXA7
RNA-seq Strand
plus
minus

Figure 3

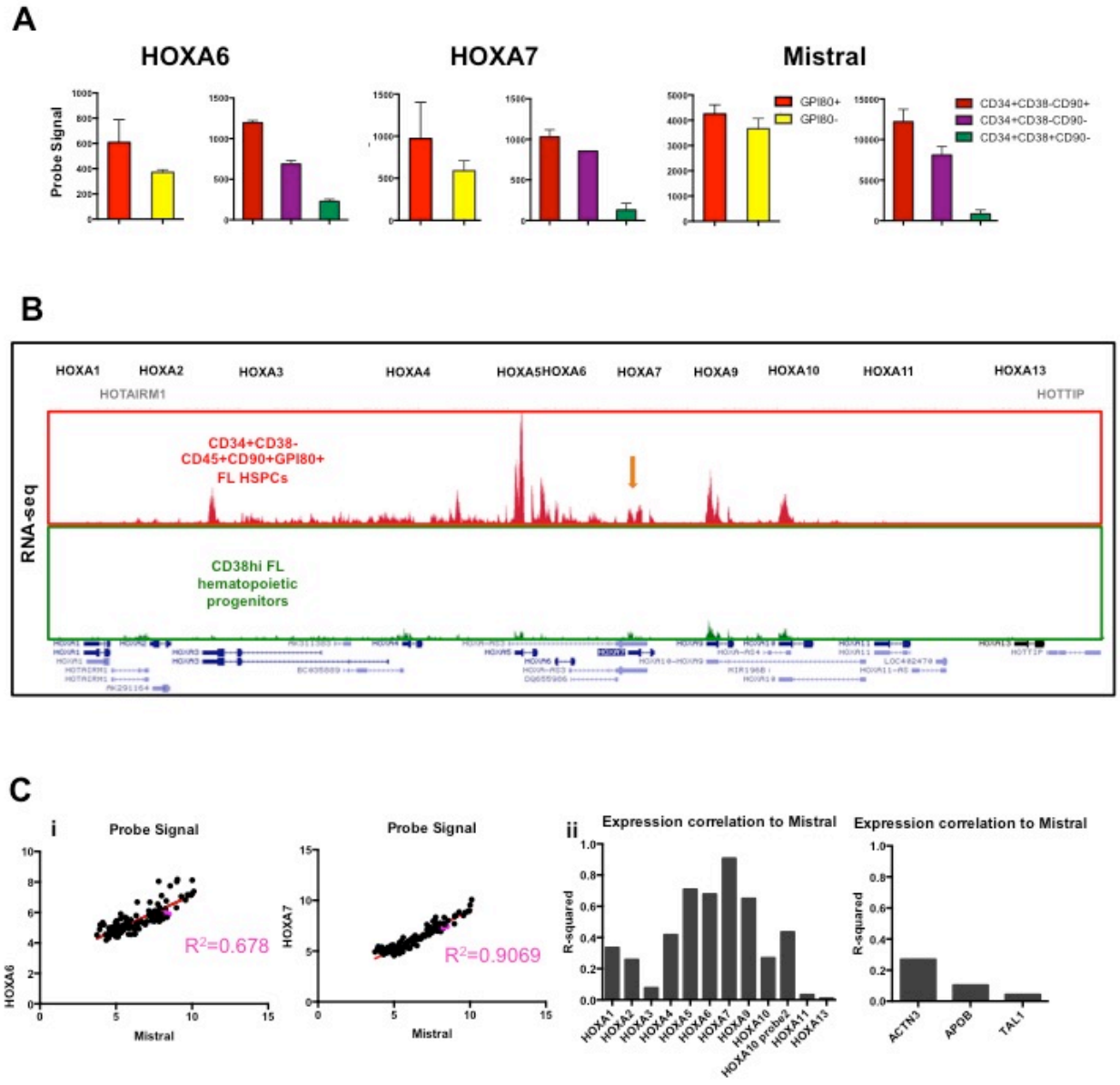

A
HOXA6
HOXA7
Mistral
Probe Signal
GPI80+
GPI80-
CD34+CD38-CD90+
CD34+CD38-CD90-
CD34+CD38+CD90-
B
HOXA1
HOXA2
HOXA3
HOXA4
HOXA5HOXA6
HOXA7
HOXA9
HOXA10
HOXA11
HOXA13
HOTAIRM1
HOTTIP
CD34+CD38-
CD45+CD90+GPI80+
FL HSPCs
CD38hi FL
hematopoietic
progenitors
RNA-seq
C
i
Probe Signal
HOXA6
Mistral
R2=0.678
Probe Signal
HOXA7
Mistral
R2=0.9069
ii
Expression correlation to Mistral
R-squared
HOXA1
HOXA2
HOXA3
HOXA4
HOXA5
HOXA6
HOXA7
HOXA9
HOXA10
HOXA10 probe2
HOXA11
HOXA13
Expression correlation to Mistral
R-squared
ACTN3
APOB
TAL1

Figure 4

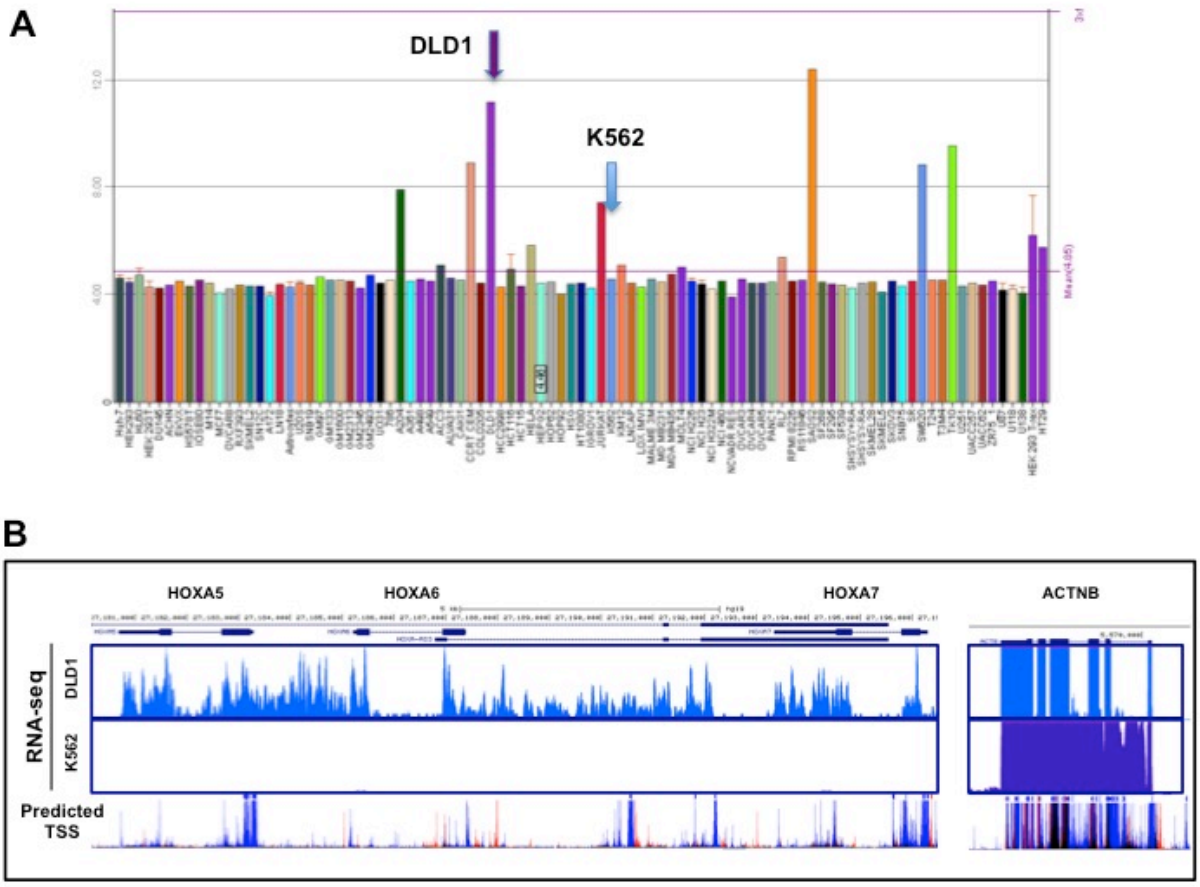

A
DLD1
K562
B
HOXA5
HOXA6
HOXA7
ACTNB
RNA-seq
DLD1
K562
Predicted
TSS

Figure 5

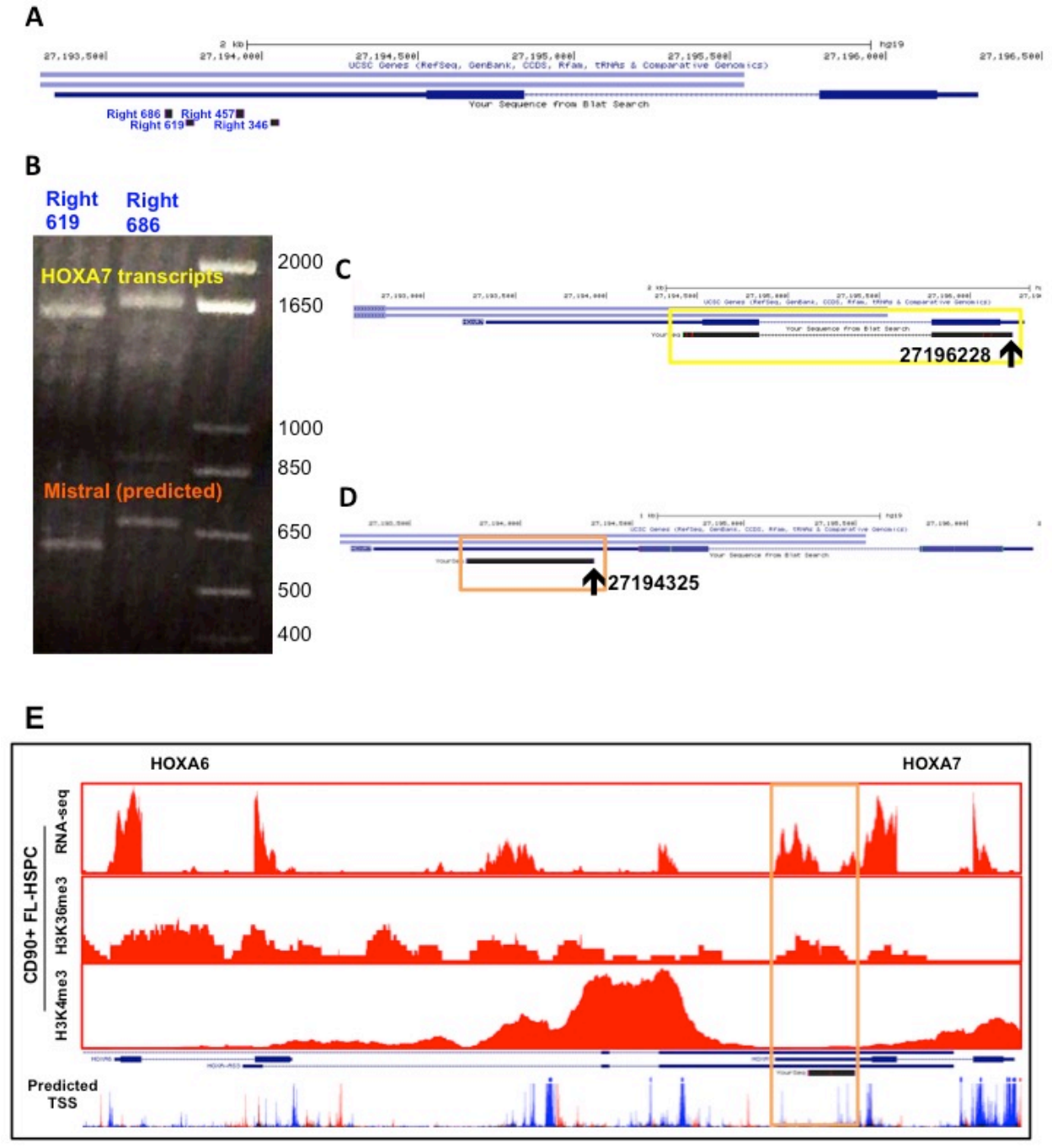

A
Right 686
Right 457
Right 619
Right 346
Your Sequence from Blat Search
B
Right 619
Right 686
HOXA7 transcripts
Mistral (predicted)
2000
1650
1000
850
650
500
400
C
27196228
D
27194325
E
HOXA6
HOXA7
CD90+ FL-HSPC
RNA-seq
H3K36me3
H3K4me3
Predicted TSS

Figure 6

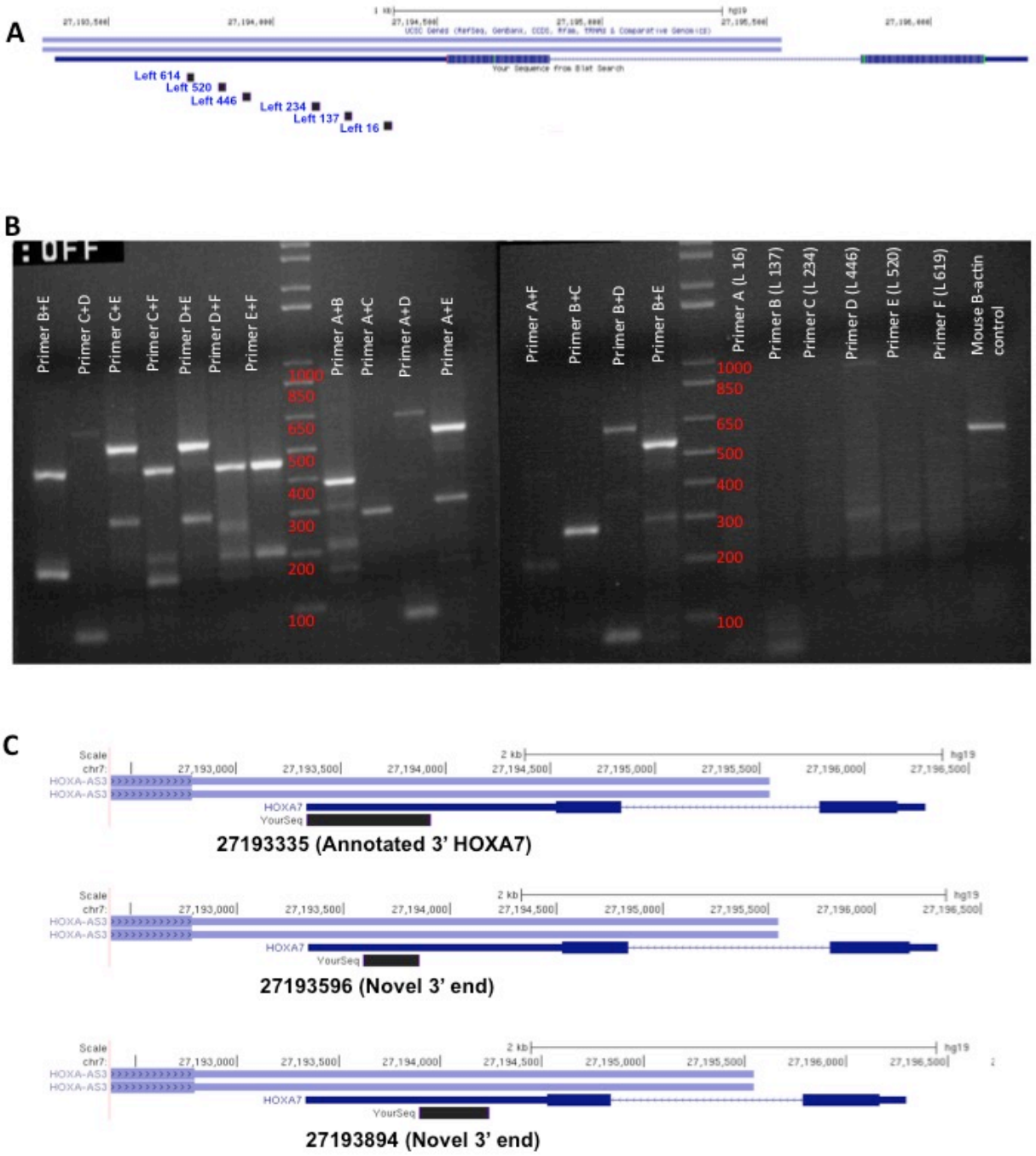

A
Left 614
Left 520
Left 446
Left 234
Left 137
Left 16
B
Primer B+E
Primer C+D
Primer C+E
Primer C+F
Primer D+E
Primer D+F
Primer E+F
Primer A+B
Primer A+C
Primer A+D
Primer A+E
Primer A+F
Primer B+C
Primer B+D
Primer B+E
Primer A (L 16)
Primer B (L 137)
Primer C (L 234)
Primer D (L 446)
Primer E (L 520)
Primer F (L 619)
Mouse B-actin control
1000
850
650
500
400
300
200
100
C
Scale
chr7:
27,193,000
27,193,500
27,194,000
27,194,500
27,195,000
27,195,500
27,196,000
27,196,500
2 kb
hg19
HOXA-AS3
HOXA7
YourSeq
27193335 (Annotated 3' HOXA7)
27193596 (Novel 3' end)
27193894 (Novel 3' end)

Table 1: Primers used for 5’ and 3’ RACE

5' RACE Primers	
Mistral Right 346	GCAAGCCAGCCAGCACAAAATAGGT
Mistral Right 457	TACCCTGATCCTGGACCTCAGCACA
Mistral Right 686	CAGATCTGCAGCCAGCTCAGGAACA
Mistral Right 619	GGAGCCAGAGGAAAGGACAGCGAAG
3' RACE Primers	
Mistral Left 16	TACTCCAGCTCCCAACTTTTGT
Mistral Left 137	CGCGTCCAAAATACTACCTAGC
Mistral Left 234	TATGTATCCCAACACTGGCAGA
Mistral Left 446	AGGATCAGGGTATTGGCATCTA
Mistral Left 520	CCCTAGAATCAATCCCTCCTTC
Mistral Left 614	GATCTGTGGGTTAGCTTCTGCT

Bibliography

Bertani, S., Sauer, S., Bolotin, E., and Sauer, F. (2011). The Noncoding RNA Mistral Activates Hoxa6 and Hoxa7 Expression and Stem Cell Differentiation by Recruiting MLL1 to Chromatin. Molecular Cell *43*, 1040-1046.

Consortium, E.P. (2012). An integrated encyclopedia of DNA elements in the human genome. Nature *489*, 57-74.

Dasen, J.S. (2013). Long noncoding RNAs in development: solidifying the Lncs to Hox gene regulation. Cell Rep *5*, 1-2.

Dou, D.R., Calvanese, V., Saarikoski, P., Galic, Z., and Mikkola, H.K.A. (2016a). Induction of HOXA genes in hESC-derived HSPC by two-step differentiation and RA signalling pulse. Nature Protocol Exchange.

Dou, D.R., Calvanese, V., Sierra, M.I., Nguyen, A.T., Minasian, A., Saarikoski, P., Sasidharan, R., Ramirez, C.M., Zack, J.A., Crooks, G.M.*, et al.* (2016b). Medial HOXA genes demarcate haematopoietic stem cell fate during human development. Nat Cell Biol *18*, 595-606.

Díaz-Beyá, M., Brunet, S., Nomdedéu, J., Pratcorona, M., Cordeiro, A., Gallardo, D., Escoda, L., Tormo, M., Heras, I., Ribera, J.M.*, et al.* (2015). The lincRNA HOTAIRM1, located in the HOXA genomic region, is expressed in acute myeloid leukemia, impacts prognosis in patients in the intermediate-risk cytogenetic category, and is associated with a distinctive microRNA signature. Oncotarget *6*, 31613-31627.

Forrest, A.R., Kawaji, H., Rehli, M., Baillie, J.K., de Hoon, M.J., Haberle, V., Lassmann, T., Kulakovskiy, I.V., Lizio, M., Itoh, M.*, et al.* (2014). A promoter-level mammalian expression atlas. Nature *507*, 462-470.

Germanguz, I., Listgarten, J., Cinkornpumin, J., Solomon, A., Gaeta, X., and Lowry, W.E. (2016). Identifying gene expression modules that define human cell fates. Stem Cell Res *16*, 712-724.

Guttman, M., Amit, I., Garber, M., French, C., Lin, M.F., Feldser, D., Huarte, M., Zuk, O., Carey, B.W., Cassady, J.P.*, et al.* (2009). Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. Nature *458*, 223-227.

Li, W., Cowley, A., Uludag, M., Gur, T., McWilliam, H., Squizzato, S., Park, Y.M., Buso, N., and Lopez, R. (2015). The EMBL-EBI bioinformatics web and programmatic tools framework. Nucleic Acids Res *43*, W580-584.

Lizio, M., Harshbarger, J., Shimoji, H., Severin, J., Kasukawa, T., Sahin, S., Abugessaisa, I., Fukuda, S., Hori, F., Ishikawa-Kato, S.*, et al.* (2015). Gateways to the FANTOM5 promoter level mammalian expression atlas. Genome Biol *16*, 22.

Magnusson, M., Sierra, M.I., Sasidharan, R., Prashad, S.L., Romero, M., Saarikoski, P., Van Handel, B., Huang, A., Li, X., and Mikkola, H.K. (2013). Expansion on stromal cells preserves the undifferentiated state of human hematopoietic stem cells despite compromised reconstitution ability. PLoS One *8*, e53912.

Mishra, B.P., Zaffuto, K.M., Artinger, E.L., Org, T., Mikkola, H.K., Cheng, C., Djabali, M., and Ernst, P. (2014). The histone methyltransferase activity of MLL1 is dispensable for hematopoiesis and leukemogenesis. Cell Rep *7*, 1239-1247.

Org, T., Duan, D., Ferrari, R., Montel-Hagen, A., Van Handel, B., Kerényi, M.A., Sasidharan, R., Rubbi, L., Fujiwara, Y., Pellegrini, M.*, et al.* (2015). Scl binds to primed enhancers in mesoderm to regulate hematopoietic and cardiac fate divergence. EMBO J *34*, 759-777.

Prashad, S.L., Calvanese, V., Yao, C.Y., Kaiser, J., Wang, Y., Sasidharan, R., Crooks, G., Magnusson, M., and Mikkola, H.K. (2014). GPI-80 Defines Self-Renewal Ability in Hematopoietic Stem Cells during Human Development. Cell Stem Cell.

Quinn, J.J., Zhang, Q.C., Georgiev, P., Ilik, I.A., Akhtar, A., and Chang, H.Y. (2016). Rapid evolutionary turnover underlies conserved lncRNA-genome interactions. Genes Dev *30*, 191-207.

Rinn, J.L., Kertesz, M., Wang, J.K., Squazzo, S.L., Xu, X., Brugmann, S.A., Goodnough, L.H., Helms, J.A., Farnham, P.J., Segal, E.*, et al.* (2007). Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. Cell *129*, 1311-1323.

Schuettengruber, B., Chourrout, D., Vervoort, M., Leblanc, B., and Cavalli, G. (2007). Genome regulation by polycomb and trithorax proteins. Cell *128*, 735-745.

Schwartz, Y.B., and Pirrotta, V. (2007). Polycomb silencing mechanisms and the management of genomic programmes. Nat Rev Genet *8*, 9-22.

Shiraki, T., Kondo, S., Katayama, S., Waki, K., Kasukawa, T., Kawaji, H., Kodzius, R., Watahiki, A., Nakamura, M., Arakawa, T.*, et al.* (2003). Cap analysis gene expression for high-throughput analysis of transcriptional starting point and identification of promoter usage. Proc Natl Acad Sci U S A *100*, 15776-15781.

Soshnikova, N. (2013). Hox genes regulation in vertebrates. Dev Dyn.

Soshnikova, N., and Duboule, D. (2009). Epigenetic regulation of vertebrate Hox genes: a dynamic equilibrium. Epigenetics *4*, 537-540.

Wan, L., Kong, J., Tang, J., Wu, Y., Xu, E., Lai, M., and Zhang, H. (2016). HOTAIRM1 as a potential biomarker for diagnosis of colorectal cancer functions the role in the tumour suppressor. J Cell Mol Med *20*, 2036-2044.

Wang, K.C., Yang, Y.W., Liu, B., Sanyal, A., Corces-Zimmerman, R., Chen, Y., Lajoie, B.R., Protacio, A., Flynn, R.A., Gupta, R.A.*, et al.* (2011). A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. Nature *472*, 120-124.

Wang, P., Lin, C., Smith, E.R., Guo, H., Sanderson, B.W., Wu, M., Gogol, M., Alexander, T., Seidel, C., Wiedemann, L.M.*, et al.* (2009). Global analysis of H3K4 methylation defines MLL family member targets and points to a role for MLL1-mediated H3K4 methylation in the regulation of transcriptional initiation by RNA polymerase II. Mol Cell Biol *29*, 6074-6085.

Wang, X.Q., and Dostie, J. (2016). Reciprocal regulation of chromatin state and architecture by HOTAIRM1 contributes to temporal collinear HOXA gene activation. Nucleic Acids Res.

Wei, S., Zhao, M., Wang, X., Li, Y., and Wang, K. (2016). PU.1 controls the expression of long noncoding RNA HOTAIRM1 during granulocytic differentiation. J Hematol Oncol *9*, 44.

Wu, C., Jin, X., Tsueng, G., Afrasiabi, C., and Su, A.I. (2016). BioGPS: building your own mash-up of gene annotations and expression profiles. Nucleic Acids Res *44*, D313-316.

Wu, C., Macleod, I., and Su, A.I. (2013). BioGPS and MyGene.info: organizing online, gene-centric information. Nucleic Acids Res *41*, D561-565.

Wu, C., Orozco, C., Boyer, J., Leglise, M., Goodale, J., Batalov, S., Hodge, C.L., Haase, J., Janes, J., Huss, J.W.*, et al.* (2009). BioGPS: an extensible and customizable portal for querying and organizing gene annotation resources. Genome Biol *10*, R130.

Yamashita, R., Sathira, N.P., Kanai, A., Tanimoto, K., Arauchi, T., Tanaka, Y., Hashimoto, S., Sugano, S., Nakai, K., and Suzuki, Y. (2011). Genome-wide characterization of transcriptional start sites in humans by integrative transcriptome analysis. Genome Res *21*, 775-789.

Yang, Y.W., Flynn, R.A., Chen, Y., Qu, K., Wan, B., Wang, K.C., Lei, M., and Chang, H.Y. (2014). Essential role of lncRNA binding for WDR5 maintenance of active chromatin and embryonic stem cell pluripotency. Elife *3*, e02046.

Zhang, X., Lian, Z., Padden, C., Gerstein, M.B., Rozowsky, J., Snyder, M., Gingeras, T.R., Kapranov, P., Weissman, S.M., and Newburger, P.E. (2009). A myelopoiesis-associated regulatory intergenic noncoding RNA transcript within the human HOXA cluster. Blood *113*, 2526-2534.

Zhang, X., Weissman, S.M., and Newburger, P.E. (2014). Long intergenic non-coding RNA HOTAIRM1 regulates cell cycle progression during myeloid maturation in NB4 human promyelocytic leukemia cells. RNA Biol *11*, 777-787.

Chapter 4:

Summary and Discussion

The hematopoietic stem cell is critical for the survival of the developing fetus and is required to continuously maintain the hematopoietic system, essential for both healthy metabolic and immune functions, throughout adult life. Understanding the development of HSCs and the genes and pathways underlying the functional traits of the HSCs are a requirement both in understanding the basis of and developing treatments for blood and immune diseases. HSC transplantation treatments have already helped many patients with blood and immune diseases and methods to increase accessibility and availability of engraftable HSCs will improve the prospects of many more. However, the supply of HSCs for transplantation is limited by lack of HLA-matched donors. Generation of HSCs from embryonic stem cells (ESCs) or induced pluripotent stem cells (iPSCs) would enable a more widespread utilization of HSCs in the clinic; however, *in vitro* generated hematopoietic stem/progenitor cells (HSPCs) are unable to self-renew due to as yet unknown molecular defects that compromise their function. This thesis investigated the molecular programs and regulatory mechanisms critical for human HSC-self-renewal in ESC-derived HSPCs.

In this work, we have:

1. Established a differentiation protocol to generate from ESCs immunophenotypic HSPCs with multilineage differentiation ability and high molecular correlation to *in vivo* isolated FL-HSPCs
2. Identified the expression of medial HOXA genes as critical for human HSC self-renewal and dysregulated in ESC-derived HSPCs
3. Discovered that transient induction of retinoic acid signaling is sufficient to induce, but not maintain HOXA gene expression in ESC-HSPCs

4. Identified a putative lncRNA homologous to mouse Mistral in the human HOXA cluster and shown that it associates with HOXA6/HOXA7 expression and possibly regulates their expression

Chapter 2: Medial HOXA genes demarcate haematopoietic stem cell fate during human development

Although ESC differentiation cultures have successfully generated hematopoietic cells from the earlier waves of developmental hematopoiesis that closely mirrors yolk sac blood development (Qiu et al., 2008; Wang et al., 2004; Zambidis et al., 2005), less success has been achieved in replicating the definitive hematopoietic wave in culture (Dravid et al., 2011; Martin et al., 2008). Lineage tracing and knockout studies in model organisms, such as mice and zebrafish, have enabled the study of the critical EHT stage in vertebrate embryos (Bertrand et al., 2010; Chen et al., 2009; Lancrin et al., 2009; Zovein et al., 2008), but are not possible in humans. The search for important regulatory cues during HSC specification and markers of the emerging definitive lineage in human is still ongoing and is necessary for the improvement of differentiation cultures geared towards generating a transplantable human HSC.

In this study, we document the generation of immunophenotypic HSPCs from ESCs (**Fig 1A-C**). We used HSCs isolated from the second trimester FL as the *in vivo* comparison since the FL is the site of the most active expansion and differentiation of HSCs, and FL-HSCs are ontogenically closer to ESCs than HSCs from the cord blood or bone marrow. We further showed these ESC-derived HSPCs were capable of differentiating into myeloid, T-lymphoid, and erythroid cells that express the adult form of β-globin (HBB) and thus are not restricted to the embryonic progenitor stage (**Fig 2E, Supplementary Fig D**). However, immunophenoytpic

ESC-HSPCs still displayed severe curtailment in self-renewal ability (**Fig 2A-D**) and were unable to engraft (**Fig 1D,E**), in contrast to FL-HSPCs.

Additionally, while immunophenotypic ESC-derived HSPCs displayed high molecular correlation to FL-HSPCs, a subset of transcription factors that remained downregulated, reminiscent of the more primitive HSPCs found in the EBs or first trimester PL (**Fig 3**). We discovered that the medial HOXA genes—HOXA7 in particular—are highly expressed in definitive HSPCs and required for HSPC self-renewal and engraftment (**Fig 4**). Not only are the medial HOXA genes highly expressed in self-renewing FL-HSPCs and progressively decrease in expression as the hematopoietic cells lose self-renewal ability and become more progenitor-like (**Fig 4A,B**), the medial HOXA genes are not expressed in the non-self-renewing ESC-derived immnophenotypic HSPC (**Fig 3G**). The importance of the medial HOXA genes was confirmed when knockdown of HOXA5 or HOXA7 in FL-HSPCs re-capitulated the self-renewal and engraftment defect observed in ESC-HSPCs (**Fig 4C-I).** In particular, knockdown of HOXA7 dysregulated many of the same genes dysregulated in ESC-HSPCs and specifically induced several genes associated with the embryonic progenitor lineage, such as expression of the epsilon (HBE1) and gamma globins (HBG2), while downregulating many HSC-related factors, such as HEMGN1, PROM1, and HLF (**Fig 4J-L**). Taken together, these data suggest that the expression of the medial HOXA genes is critical for definitive human HSPC identity and function.

However, while overexpression of the medial genes in FL-HSPCs robustly prolonged the maintenance of the HSPC compartment (**Fig 5A-F**), the same effect was not observed in ESC-HSPCs. Defects in self-renewal and engraftment remained resolutely in place even in the forced expression of three HOXA genes (HOXA5, HOXA7, and HOXA9) (**Fig 5G-L**). These data suggest that other components critical to the HSC pathway are missing in ESC-HSPCs, or that

the HOXA genes, which are under notoriously tight spatio-temporal regulation, must be expressed in the correct cellular context.

HOXA genes harbor many retinoic acid response elements (RAREs) (Frasch et al., 1995; Langston and Gudas, 1992), and retinoic acid is a known inducer of HOXA genes (Marshall et al., 1996), Studies in mice have determined that RALDH2 signaling through RAR- α is critical for specification of the hemogenic endothelium (HE) (Chanda et al., 2013)., Since we observed RALDH2 expression lacking in the ESC-derived HE stage (EB) compared to PL (**Fig 6A**), we stimulated the retinoic acid pathway to activate HOXA gene expression. Because prolonged application of retinoic acid stimulus is harmful to cells and drives differentiation, we used a carefully timed window of treatment lasting six days at the beginning of maturation culture on OP9-M2 stroma (**Fig 6B**). Stimulation of ESC-HSPCs using the RAR-α agonist, AM580, for 6 days was sufficient to significantly induce HOXA cluster gene expression, prolong the maintenance of the HSPC compartment, increase clonogenicity and generate multipotent colonies that were not present in the control DMSO treatment group (**Fig 6C-F**, **Supplementary Fig 5**, **Fig 7B,C**). However, the release from external RA stimulation caused a rapid decline of HOXA gene expression back to baseline only 6 days after removal of the drugs, and did not resolve the engraftment defect or completely resolve the self-renewal defect (**Fig 6D-F**, **Supplementary Fig 6A**). Despite the inability to maintain HOXA gene expression after removal of RA stimulation, several genes associated with the embryonic program (i.e., HBE1, HBG2, GP1B, ITGB3) remained suppressed and key HSC factors (ERG, MECOM, SOX4, HLF) remained expressed in ESC-HSPCs treated with AM580, indicating a partial retention of the definitive HSC transcriptome (**Supplementary Fig 6B,C**).

Collectively, we documented the generation of cells of the definitive hematopoietic lineage, demonstrated the critical role the medial HOXA cluster genes play in human HSC function and identity, and identified the RA signaling pathway as a key inducer of the HOXA genes and specification of the definitive HSC program. It is clear that a cell incapable of self-renewal also cannot engraft. Future studies to identify other regulators of the HOXA cluster and self-renewal and determine the tightly regulated spatio-temporal window in which RA signaling must occur will be critical to our understanding of human HSC development. We have filled in a gap of understanding of what underlies self-renewal capacity in human HSCs, but in order for any culture-derived cell to be clinically acceptable, the use of serum-free cultures and either the elimination of reliance on mouse stroma co-culturing or the development of comparable human stroma must occur.

Chapter 3: A non-coding transcript in the human HOXA7 3' UTR is correlated with medial HOXA gene expression and HSC stemness

Long noncoding RNAs can function in multiple different ways to regulate gene expression both as repressors or activators. The well-studied lncRNAs of the HOXA cluster, HOTAIRM1 and HOTTIP, act *in cis* to the HOXA locus in which they are transcribed and regulate, respectively, the proximal HOXA1-HOXA4 genes and the distal HOXA9-HOXA13 genes (Wang et al., 2011; Zhang et al., 2009). HOTTIP has been shown to act as molecular scaffolds in the recruitment of MLL1 and its binding partner, WDR5, to allow accessibility of the transcription machinery to target genes (Wang et al., 2011; Yang et al., 2014).. In Chapter 2, we observed the importance of the medial HOXA genes to human HSC function and identity. However, while the existence of a lncRNA essential to HOXA6 and HOXA7 regulation in

mouse ESCs has been reported (Bertani et al., 2011), less is known about the noncoding regulation of the medial HOXA genes in humans.

In this study, we found that the low expression of the medial HOXA genes in non-self-renewing ESC-derived HSPCs may be correlated with the inability of the MLL1 complex to bind to the medial HOXA genes, rather than lack of MLL1 or WDR5 expression (**Fig 1A-C**). Comparatively, we observed robust binding of MLL1 to HOXA7 and HOXA9 in MLL-ChIP pulldown using FL-derived cells (**Fig 1B**) that also express high levels of the medial HOXA genes (**Fig 1A**, **Fig 2A**). In our previous study, ATAC-seq and RNA-seq data showed the opening of the HOXA cluster and induction of HOXA gene expression with stimulation of the RA signaling pathway, suggesting that inaccessibility to the HOXA locus may explain the inability to express HOXA genes in ESC-HSPCs. Thus, molecular scaffolds, such as lncRNAs, may be required to provide locus accessibility and/or recruitment of co-activators. However, as neither HOTAIRM1 nor HOTTIP were expressed in self-renewing HSPCs (**Fig 2A**, **Fig 3B**), our search turned to finding possible noncoding elements in the region of the medial HOXA genes. We identified a distinct peak within the noncoding 3' UTR of HOXA7 that is highly expressed in self-renewing HSC,s and missing in progenitors and non-self-renewing ESC-HSPCs (**Fig 2A-B**, **Fig 3A-B**).

The lncRNA Mistral is required for expression of HOXA6 and HOXA7 in mouse ESCs and is expressed in cis to HOXA6 and HOXA7. In contrast to coding genes and smaller noncoding RNAs, lncRNAs are not as consistently sequence-conserved across species. However, sequence alignment of murine Mistral showed high sequence homology to the peak observed in the human HOXA7 3' UTR (**Fig 2C,D**). Another criterion for lncRNA identification across species is conservation of synteny—the elements surrounding the prospective lncRNA. The

prospective 3' UTR element satisfies this criterion as it, like the annotated mouse Mistral, is also located between HOXA6 and HOXA7 on the same strand as the HOXA cluster genes (**Fig 2E**). Future studies assessing the similarity of the secondary structures between mouse Mistral and the peak in the noncoding HOXA7 3'UTR will further elucidate whether these two elements are orthologous (Quinn et al., 2016). Taking into account the evidence for conservation, we labeled the peak as putative human Mistral.

The expression of HOXA6 and HOXA7 is highly correlated to putative Mistral expression across multiple cell lines and primary cells in addition to hematopoietic cells (**Fig 3**). This expression pattern suggests a possible function for human Mistral in HOXA6 and HOXA7 expression similar to the role Mistral plays in mouse. Future studies manipulating Mistral expression, such as through targeted CRISPR deletion, will evaluate the impact of Mistral on HOXA6 and HOXA7 expression. These studies cannot be completed without confirmation of the full sequence of Mistral.

Thus far, the existence of a unique transcript corresponding to the observed HOXA7 3' UTR peak has only been postulated. Confirmation that other cell lines show the same correlation of HOXA6/HOXA7 and Mistral expression (**Fig 3C**) justified the use of the DLD1 cell line in assays requiring high starting materials (**Fig 4**). Using 5' RACE, we were able to definitively show that there is a second 5' start site that is not only distinct from the HOXA7 5' start site, but also located in the predicted 5' start region of putative Mistral based on the peak, FANTOM5 CAGE TSS predictions, and further confirmed through the H3K36me3 PolII elongation signal (**Fig 5**). Due to the density of the HOXA region and the proximity of the predicted noncoding gene to the HOXA6 and HOXA7 genes, it was not possible using H3K4me3 of permissive promoters to determine whether there existed a separate promoter unique to putative Mistral or

whether Mistral shares a promoter with HOXA7 (**Fig 5E**). Efforts to find the 3' termination end of putative Mistral using 3' RACE returned 3 unique termination sites, one of which was ruled out as unsuitable as it was located before the predicted 5' start site (**Fig 6**, **Fig 5D**). Alternative methods to determine the true, full sequence of Mistral, such as Northern Blots, are currently in progress.

The mechanisms by which lncRNAs function are varied and still being explored (Wang and Chang, 2011). This study identified within the medial HOXA gene region a novel transcript with high probability of conservation to the mouse lncRNA Mistral, and further established the correlation of medial HOXA gene expression to expression of the putative human Mistral. Future studies using CLIP of components of the MLL complex and genome-wide association studies coupled with functional tests will shed further light on how, and if, the novel noncoding RNA transcript we have identified interacts with the human HOXA gene cluster. In the field of hematopoiesis, finding a Mistral homolog in humans can provide a method by which to induce or maintain medial HOXA expression, particularly HOXA7, in ESC-derived HSPCs. Future studies to test the impact of manipulating Mistral expression, and ways to induce expression of Mistral at its locus, on HSPC self-renewal will provide an interesting link in the overall pathways involved in HOXA gene regulation and definitive hematopoietic specification. These studies also have the broader ranging impact of furthering our understanding of lncRNA conservation, structure, and function in regulating gene expression.

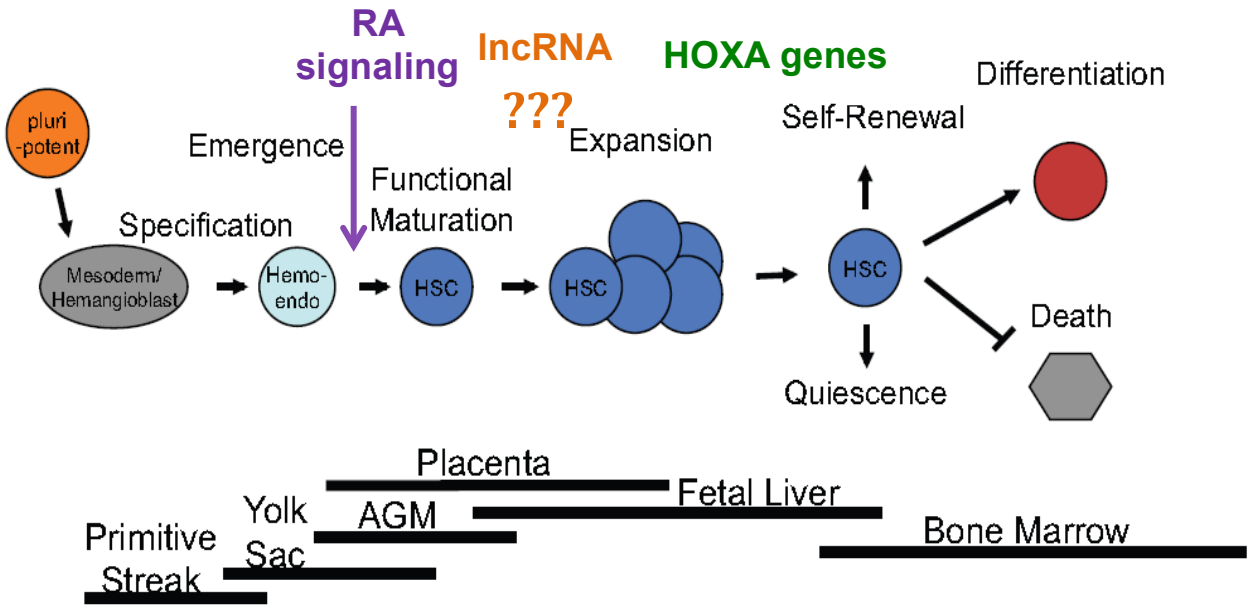


Figure 4.1: Proposed model of action for human HSC specification. Retinoic acid signaling is required for hemogenic endothelium specification to definitive HSC fate and to initiate HOXA gene expression. LncRNAs may be involved in maintaining HOXA gene expression after induction by RA signaling. HOXA gene expression is required for HSC self-renewal and definitive identity.

Bibliography

Bertani, S., Sauer, S., Bolotin, E., and Sauer, F. (2011). The Noncoding RNA Mistral Activates Hoxa6 and Hoxa7 Expression and Stem Cell Differentiation by Recruiting MLL1 to Chromatin. Molecular Cell *43*, 1040-1046.

Bertrand, J.Y., Chi, N.C., Santoso, B., Teng, S., Stainier, D.Y., and Traver, D. (2010). Haematopoietic stem cells derive directly from aortic endothelium during development. Nature *464*, 108-111.

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

Chen, M.J., Yokomizo, T., Zeigler, B.M., Dzierzak, E., and Speck, N.A. (2009). Runx1 is required for the endothelial to haematopoietic cell transition but not thereafter. Nature *457*, 887-891.

Dou, D.R., Calvanese, V., Sierra, M.I., Nguyen, A.T., Minasian, A., Saarikoski, P., Sasidharan, R., Ramirez, C.M., Zack, J.A., Crooks, G.M.*, et al.* (2016). Medial HOXA genes demarcate haematopoietic stem cell fate during human development. Nat Cell Biol *18*, 595-606.

Dravid, G., Zhu, Y., Scholes, J., Evseenko, D., and Crooks, G.M. (2011). Dysregulated gene expression during hematopoietic differentiation from human embryonic stem cells. Mol Ther *19*, 768-781.

Frasch, M., Chen, X., and Lufkin, T. (1995). Evolutionary-conserved enhancers direct region-specific expression of the murine Hoxa-1 and Hoxa-2 loci in both mice and Drosophila. Development *121*, 957-974.

Lancrin, C., Sroczynska, P., Stephenson, C., Allen, T., Kouskoff, V., and Lacaud, G. (2009). The haemangioblast generates haematopoietic cells through a haemogenic endothelium stage. Nature *457*, 892-895.

Langston, A.W., and Gudas, L.J. (1992). Identification of a retinoic acid responsive enhancer 3' of the murine homeobox gene Hox-1.6. Mech Dev *38*, 217-227.

Marshall, H., Morrison, A., Studer, M., Pöpperl, H., and Krumlauf, R. (1996). Retinoids and Hox genes. FASEB J *10*, 969-978.

Martin, C.H., Woll, P.S., Ni, Z., Zuniga-Pflucker, J.C., and Kaufman, D.S. (2008). Differences in lymphocyte developmental potential between human embryonic stem cell and umbilical cord blood-derived hematopoietic progenitor cells. Blood *112*, 2730-2737.

Qiu, C., Olivier, E.N., Velho, M., and Bouhassira, E.E. (2008). Globin switches in yolk sac-like primitive and fetal-like definitive red blood cells produced from human embryonic stem cells. Blood *111*, 2400-2408.

Quinn, J.J., Zhang, Q.C., Georgiev, P., Ilik, I.A., Akhtar, A., and Chang, H.Y. (2016). Rapid evolutionary turnover underlies conserved lncRNA-genome interactions. Genes Dev *30*, 191-207.

Wang, K.C., and Chang, H.Y. (2011). Molecular mechanisms of long noncoding RNAs. Mol Cell *43*, 904-914.

Wang, K.C., Yang, Y.W., Liu, B., Sanyal, A., Corces-Zimmerman, R., Chen, Y., Lajoie, B.R., Protacio, A., Flynn, R.A., Gupta, R.A.*, et al.* (2011). A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. Nature *472*, 120-124.

Wang, L., Li, L., Shojaei, F., Levac, K., Cerdan, C., Menendez, P., Martin, T., Rouleau, A., and Bhatia, M. (2004). Endothelial and hematopoietic cell fate of human embryonic stem cells originates from primitive endothelium with hemangioblastic properties. Immunity *21*, 31-41.

Yang, Y.W., Flynn, R.A., Chen, Y., Qu, K., Wan, B., Wang, K.C., Lei, M., and Chang, H.Y. (2014). Essential role of lncRNA binding for WDR5 maintenance of active chromatin and embryonic stem cell pluripotency. Elife *3*, e02046.

Zambidis, E.T., Peault, B., Park, T.S., Bunz, F., and Civin, C.I. (2005). Hematopoietic differentiation of human embryonic stem cells progresses through sequential hematoendothelial, primitive, and definitive stages resembling human yolk sac development. Blood *106*, 860-870.

Zhang, X., Lian, Z., Padden, C., Gerstein, M.B., Rozowsky, J., Snyder, M., Gingeras, T.R., Kapranov, P., Weissman, S.M., and Newburger, P.E. (2009). A myelopoiesis-associated regulatory intergenic noncoding RNA transcript within the human HOXA cluster. Blood *113*, 2526-2534.

Zovein, A.C., Hofmann, J.J., Lynch, M., French, W.J., Turlo, K.A., Yang, Y., Becker, M.S., Zanetta, L., Dejana, E., Gasson, J.C.*, et al.* (2008). Fate tracing reveals the endothelial origin of hematopoietic stem cells. Cell Stem Cell *3*, 625-636.

Appendix

PROTOCOL EXCHANGE | *COMMUNITY CONTRIBUTED*

Induction of HOXA genes in hESC-derived HSPC by two-step differentiation and RA signalling pulse

Diana Dou, Vincenzo Calvanese, Pamela Saarikoski, Zoran Galic & Hanna Mikkola

Mikkola Lab (UCLA)

Abstract

The generation of definitive, transplantable haematopoietic stem cells (HSC) from pluripotent stem cells (PSC) is an unmet challenge. Although previous protocols have generated clonogenic progenitors and various differentiated hematopoietic cells, there is no confirmation of their definitive identity or functional HSC properties. We describe a two-step differentiation protocol that starts with differentiation of H1 human embryonic stem cells (hESCs) into embryoid bodies (EBs) in mesoderm-inducing conditions followed by maturation of EB derived CD34+ cells on OP9M2 stroma on HSC supportive conditions. This protocol generates cells that resemble human foetal liver hematopoietic stem/progenitor cells (HSPC) by immunophenotype (CD34+CD38loCD90+CD45+GPI80+/-) and molecular profile. Treatment with a retinoic acid receptor alpha agonist (AM580) during the first six days of stroma co-culture induces the expression of developmental regulators that demarcate the definitive HSC genetic program, including the HOXA genes. AM580 treatment also prolongs the maintenance of phenotypic HSPC population and clonogenic progenitors in OP9M2 co-culture[1].

Subject terms: Cell biology Cell culture Developmental biology Isolation, Purification and Separation Tissue culture

Keywords: Embryonic stem cells Embryoid bodies Haematopoietic Stem/Progenitor Cells Retinoic Acid signalling SCF FLT3-l BMP4 TPO HOXA genes

Introduction

The attempts to generate definitive, transplantable hematopoietic stem cells (HSC) from human pluripotent stem cells have not been successful. Many protocols have been described that

achieve the generation of clonogenic progenitors and various differentiated hematopoietic cells. However, these cells do not have confirmed definitive identity, or self-renewal and engraftment potential.

Definitive HSCs are generated in a narrow time window of development in the major arteries in the embryo and extraembryonic tissues from a specialized type of haemogenic endothelium. This endothelium possesses a unique signalling environment that promotes HSC development. Prior to the generation of HSCs, large quantities of myelo-erythroid progenitors are generated in the yolk sac. However, these progenitors cannot self-renew, and have limited lymphoid potential. It has been proposed that hematopoietic differentiation from pluripotent stem cells is biased toward the generation of the yolk sac like differentiation-primed progenitors. A major challenge in the field has been to define conditions that enable the generation of definitive HSPCs rather than differentiation primed embryonic progenitors from endothelium.

Here, we describe a two-step differentiation protocol for H1 hESC that combines 2 weeks of embryoid body differentiation in mesoderm inducing conditions followed by OP9M2 stroma co-culture in HSC supportive conditions, which enables the generation of hematopoietic cells that resemble foetal liver HSPCs by immunophenotype (CD34+CD38loCD90+CD45+GPI80+)[2] and molecular profile. While the embryoid body differentiation generates immature haematovascular precursors, co-culture on mouse bone marrow OP9M2 stroma[3] enables subsequent maturation step of the CD34+ haemato-vascular cells to immunophenotypic human HSPCs. These cells have the ability to generate progeny associated with definitive haematopoiesis (adult globin expressing erythroid cells and T-lymphocytes) but have limited expansion ability in vitro and cannot engraft in vivo. Despite their close molecular correlation to foetal liver HSCs, they are unable to induce the expression of key developmental regulators associated with the HSC fate, such as the HOXA genes.[1]

Retinoic acid signalling has been recently identified as a key signalling pathway needed at the haemogenic endothelium stage to generate definitive HSCs, while the production of embryonic progenitors was unaffected[4]. We show that treatment of EB-derived CD34+ cells with a pulse of retinoic acid receptor alpha agonist (AM580) during the first days of stroma co-culture promotes the definitive HSC fate by inducing the expression of developmental regulators of definitive HSC genetic program, while suppressing genes associated with differentiation primed progenitors. The AM580 induced genes include the HOXA cluster genes and other HSC transcriptional regulators and HSC surface markers. RA signalling pulse also prolongs the maintenance of phenotypic HSPC and clonogenic progenitors in OP9M2 co-culture[1]. Although future studies will be needed to identify the conditions that induce and maintain full HSC functional properties, this protocol enables the generation of hematopoietic cells that have switched from embryonic

progenitor program toward the definitive HSPC lineage.

Reagents

1. H1 hESCs confluent, but not over-crowded. Recommend starting with 60 10cm tissue culture plates.
2. OP9M2 mouse bone marrow mesenchymal stromal cells.
3. Iscove's DMEM, 1X (Corning Cellgro, cat. no. 15-016-CV)
4. Fetal Bovine Serum (Hyclone, cat. no. SV30014.03)
5. Fetal Bovine Serum (Omega, cat. no. FB-11, Lot #101331)
6. Penicillin/Streptomycin (Invitrogen, cat. no. 15140-122)
7. Glutamax-1 (Invitrogen, cat. no. 22050-061)
8. Glutamine (Invitrogen, cat. no. 2503-081)
9. 2-mercaptoethanol (Sigma-Aldrich, cat. no. M7522)
10. Collagenase Type IV (Invitrogen, cat. no. 17104-019)
11. Phosphate Buffered Saline, 1X (Corning Cellgro, cat. no. 21-031-CV)
12. Human SCF (Invitrogen, cat. no. PHC2113)
13. Human FLT3-l (Peprotech, cat. no. 300-18)
14. Human BMP4 (Invitrogen, cat. no. PHC9534)
15. Human TPO (Peprotech, cat. no. 300-18)
16. Collagenase Type A (Worthington, cat. no. LS004176)
17. Dispase (Invitrogen, cat. no. 17105-041)
18. DNase (Sigma, cat. no. D4513)
19. StemPro ® Accutase ® Cell Dissociation Reagent (Invitrogen, cat. no. A1110501)
20. Cell Culture Grade Water (Corning Cellgro, cat. no. 25-055-CM)
21. Dead Cell Removal Kit (Miltenyi Biotec, cat. no. 130-090-101)
22. CD34 Microbead Kit Ultrapure, human (Milteny Biotec, cat. no. 130-100-453)
23. MEM-alpha, 1X (Gibco, cat. no. 12571-048)
24. DMSO (Corning Cellgro, cat. no. 29-950-CQC)
25. AM580 (Tocris, cat. no. cat. no. 0760)

Equipment

1. Cell culture hood (i.e., biosafety cabinet)
2. Incubator at 37C
3. Water bath at 37C
4. Refrigerator at 4-10C
5. Sterile aspirator system
6. Centrifuge (refrigerated at 10C)
7. Sterile filter pipette tips (2-20, 20-200, 200-1000 µL)

8. Pipettors
9. Sterile serological pipettes (5-mL, 10-mL, 25-mL)
10. Automatic pipettor (for serological pipettes)
11. StemPro EZ Passage Disposable Stem Cell Passaging Tool (Invitrogen, cat. no. 23181-010)
12. 15 and 50 mL conical tubes
13. 10 cm tissue culture-treated dishes
14. Ultralow attachment 6-well plates or 175 mL flasks (Corning, cat. nos. 3471 and 3814)
15. 24-well tissue culture-treated plates
16. 70 µm sterile nylon mesh cell strainers (Fisher Scientific, cat. no. 22363548)
17. LS Columns (Miltenyi biotec, cat. no. 130-042-401)
18. Inverted microscope
19. Hemocytometer
20. Freezer at -20C and -80C
21. Liquid nitrogen tank
22. Cryovial storage rack
23. Shaker at 37C
24. Irradiator

Procedure

Embryoid body (EB) medium:
To prepare 600 mL of EB medium:
488 mL Iscove's DMEM, 1X (Corning Cellgro)
90 mL (15%) FBS (Hyclone, or Omega Lot #101331)
6 mL (1%) Pen/Strep (Invitrogen)
6 mL (1%) Glutamax-1 (Invitrogen)
9 µL (0.0015%) 2-mercaptoethanol (Sigma-Aldrich)

As outlined in the EB Generation stepwise protocol below, for days 0-4, use EB medium without cytokines.
At day 4 and 7 of EB differentiation, supplement EB medium with 300 ng/mL of SCF (Invitrogen), 50 ng/mL of FLT3-l (Peprotech), and 10 ng/mL BMP4 (Invitrogen)
At day 10 of EB differentiation, supplement EB medium with 300 ng/mL of SCF (Invitrogen) and 50 ng/mL of FLT3-l (Peprotech)
EB Media without cytokines can be stored refrigerated at 4C for 1 months, and 2 weeks with cytokines added.

hESC Collagenase Solution:
To prepare a 1X solution of collagenase:

For every mL of DMDM/F12 1X (Corning Cellgro)
Dissolve 1 mg of Collagenase Type IV (Invitrogen)

EB Dissociation Medium:
To prepare 10 mL of a 2X solution of EB dissociation media:
PBS 1X to volume
2 mL (10%) FBS (Omega)
20 mg Collagenase Type A (Worthington)
5 mg Dispase (Gibco)
83.3 ug DNase (Sigma)
Media can be stored refrigerated at 4C for 1 week.

Human Hematopoietic Stem Cell (HSC) Medium:
To prepare 50 mL of HSC medium:
39 mL MEM alpha 1X (Gibco)
10 mL (20%) FBS (Omega)
500 µL (1%) Pen/Strep (Invitrogen)
500 µL (1%) Glutamine (Invitrogen)
25 ng/mL each of hSCF (Invitrogen), hTPO (Peprotech), and hFLT3-l (Peprotech)
HSC medium can be stored refrigerated at 4C for 2 weeks.

Stock Dilutions of AM580:
To prepare 100X dilutions:
AM580 (Tocris) was dissolved first in DMSO (Corning Cellgro) and diluted 1:500 in PBS 1X (Corning Cellgro) to make a 100X dilution and applied to wells for a final concentration of 0.2 uM. Addition of DMSO to control wells was at a final dilution of 1:25000.

EB Generation:
1. Grow 60 10 cm plates of H1 ESCs to ~80% confluency
2. Aspirate media from plates.
3. Wash once with 5 ml of PBS 1X (Corning Cellgro) per plate
4. Incubate each plate for 10 minutes with 3 mL of hESC collagenase solution in a 37C incubator
5. Following incubation, run the StemPro EZ Passage Disposable Stem Cell Passaging tool once gently, but firmly, around the entire surface areas of the plate to detach cells.
6. Collect the cell clumps by washing the plates three times each with 3 mL of 5% FBS (Omega) in PBS 1X (Corning Cellgro), scraping gently to help cells detach further.
7. Collect washes with cells in 50 mL conical tubes and spin at 1300 rpm for 5 min. at 10C.

8. Aspirate supernatant after the spin and resuspend pellet in 15 mL 5% FBS in PBS 1X per conical collection tube.
9. Pool the pellets into as many conical tubes as necessary and spin again to pellet at 1300 rpm for 5 min. at 10C.
10. Aspirate supernatant and resuspend the pellet in EB medium.
11. Allocate cells from up to 8 of the 10 cm H1 ESC plates per Ultralow attachment flask (Corning) or 10 plates per Ultralow 6-well plate (Corning) and add to volume of 40-50 mL of EB medium per flask or 4 mL per well of 6-well plate
12. On day 4, lift cells into 50 mL conical tubes and allow them to settle by gravity for 30 min. at room temperature.
13. Carefully aspirate as much of the supernatant from the loose pellet as possible.
14. Resuspend in fresh EB media supplemented with 300 ng/mL of SCF (Invitrogen), 50 ng/mL of FLT3 (Peprotech), and 10 ng/mL BMP4 (Invitrogen) and redistribute with appropriate volumes (as in Step 11) to each flask or well of 6-well plate.
15. On day 7, repeat steps 12-14.
16. On day 10, repeat steps 12-14, with the exception that BMP4 is not added to the fresh EB media.
17. Leave in incubator until day 14.

EB Processing and CD34+ cell Isolation

1. On day 14, lift cells into 50 mL conical tubes.
2. Wash with 10 mL/flask or 1 mL/well of PBS 1X and collect wash in same tubes.
3. Spin at 1300 rpm for 5 min. at 10C.
4. Carefully aspirate supernatant and resuspend pellet in a 1:1 ration by volume in PBS 1X with the EB Dissociation Media (2X).
5. Shake at 200 rpm at 37C for 15 min.
6. Filter cells through a 70 µm nylon mesh sterile cell strainer (Fisher Scientific) into a 50 mL conical tube and count.
7. Place the mesh strainer in a sterile 6-well plate and add 6 mL of accutase to cover. Incubate for 5 min. in the 37C incubator and wash the filter into the same 50 mL conical tube with the accutase followed by 10 mL PBS 1X to further dissociate any other cell clumps.
8. Pass the strained cells through the Dead Cell Removal Kit (Miltenyi Biotec) to collect live cells.
9. Isolate CD34+ cells obtained from the live cell fraction of the Dead Cell Removal Kit (Miltenyi Biotec) following instructions of the CD34 MicroBead Kit UltraPure, human (Miltenyi Biotec) and count.

AM580 treatment of EB derived CD34+ cells on stroma co-culture

1. Irradiate OP9-M2 cells (2000 rad)
2. Plate irradiated OP9-M2 at a density of 50,000 cells per well of 24-well tissue-culture treated plate in 500 uL of HSC media.
3. Allow stromal cells to settle for at least 4 hours and use irradiated plates within 3 days of irradiation.
4. Following at least 4 hours post-stromal cell plating, add 50,000 to 100,000 EB derived CD34+ cells in 500 uL of HSC media to each stroma-layered well.
5. To control wells, add 10 uL of 100X DMSO dilution. To AM580 wells, add 10 uL of 100X AM580 dilution.
6. Every 2-3 days, carefully remove 500 uL of the media from each well without disturbing cell layer on the bottom, and replace with 500 uL of fresh HSC media containing DMSO at a final dilution of 1:25000 or AM580 at 1:500000 per well.
7. On day 6, carefully remove 900 uL of media from each well without disturbing cell layer on the bottom of the wells. Replace with 900 uL of fresh HSC media.
8. On day 7, carefully remove 900 uL of media from each well without disturbing cell layer on the bottom of the wells. Replace with 900 uL of fresh HSC media.
9. Continue with half media changes (removing 500 uL of media from each well and replacing with 500 uL fresh HSC media) every 2-3 days.
10. To lift cells at any time for analysis, carefully remove all media from each well and strain through 70 uM cell strainer into 15 mL conical collection tubes. Incubate 5 min. at 37C with 200 uL/well of accutase. Add 500 uL of PBS 1X and strain into collection tubes. Wash with 1 mL/well of PBS 2X and strain washes through cell strainer into collection tubes.
11. Spin at 1300 rpm for 5 min. at 10C.
12. Aspirate supernatant and continue with tissue culture in same conditions described above but without the AM580 treatment, to assess HSPC expansion, or proceed immediately to FACS analysis or further experiments to assess HSC molecular and functional properties.

Timing

The total procedure starting from confluent plates of hESCs can be completed in 3-6 weeks, which includes 2 weeks of EB differentiation 1-4 weeks of HSPC expansion culture with initial RA pulse.

Troubleshooting

If cells do not appear to cluster into EBs by day 4 or do not display a robust CD34+ population at day 14:
1) Check that the starting hESC are morphologically undifferentiated. Our cultures normally display very low rates of spontaneous differentiation.
2) Avoid disaggregating the hESC colonies into single cells or tiny clumps. Pipet gently at every

step. Disaggregate the clumped pellet obtained after each wash step by pipetting up and down gently for a maximum of 4-6 times with a 1 ml pipet.
3) Screen FBS lots used for EB formation for best-performing lot. Test each lot for total cell yield, CD34+ cell percentage in the final EB suspension, and the ability of CD34+ cells to expand on OP9M2 co-culture and to respond to AM580 treatment.
4) Test different cytokine lots for best performance. Although commercial cytokine are regularly tested for activity by the vendors, some lots may underperform. Resuspend the cytokines according to vendor's instructions. We normally store both the stock solution of each cytokine and the working dilutions at -80C until the first use and use every working aliquot within 15 days of storage at 4C. Never re-freeze working aliquots.

If CD34+ CD45+ cells are not observed after OP9M2 co-culture:
1) Check that OP9M2 cells are morphologically uniform and cells are not crowded. If signs of cell stress or lipid droplets appear in the cytoplasm, do not use for EB-derived CD34+ co-culture. Passage at around 80% confluency and do not grow to full confluency.
2) Screen FBS lots used for HSC culture for best-performing lot. Test each lot for the ability of human foetal liver or cord blood CD34+ cells to expand in OP9M2 co-culture, and afterward validate on EB derived CD34+ cells. Test the FBS also for the ability to respond to AM580 treatment.
3) Test different cytokine lots for best performance.

If AM580 does not induce the expression of CD38 and/or other RA signalling targets, perform ar activity optimization with your specific lot by testing different concentrations.

Anticipated Results

By day 4 of EB differentiation, the first media change, small cell aggregates should be visible with bare eyes. The media should be cleared of most single cells and cell debris after the first media change. The EBs will grow in size and acquire a hollow aspect by day 10. At the end of the EB differentiation, the EBs will look like balloon-like cell clumps that float in the medium.

The haematovascular cells in 2 week EBs are CD34+, CD90+, CD43+/-, and CD38low/-. After 12 days of replating EB CD34+ cells on stroma co-culture, robust CD45 and GPI80 surface expression can be detected. Treatment with AM580 will elevate CD38 expression by 3 days after treatment initiation. The upregulation of CD38 can be reversed after the removal of AM580 treatment. Following 6 days of RA-signalling pulse with AM580, induction of HOXA genes and other HSC regulators, such as HLF, ERG, GATA3, GFI1, MECOM etc. can be detected by qPCR and RNA-sequencing[1]. RA-stimulation in CD34+ EBs also induces vascular program associated with arteries and HSC development (SOX7, SOX17, EFNB2, NOS3 etc.). Opening of

chromatin in regulatory regions associated with AM580 activated genes can be detected by ATAC-Sequencing.

References

1. Dou, D.R. et al. Medial HOXA genes demarcate haematopoietic stem cell fate during human development. Nature Cell Biology (2016).
2. Prashad, S.L. et al. GPI-80 Defines Self-Renewal Ability in Hematopoietic Stem Cells during Human Development. Cell Stem Cell (2014).
3. Magnusson, M. et al. Expansion on stromal cells preserves the undifferentiated state of human hematopoietic stem cells despite compromised reconstitution ability. PLoS One 8, e53912 (2013).
4. Chanda, B., Ditadi, A., Iscove, N.N. & Keller, G. Retinoic acid signaling is essential for embryonic hematopoietic stem cell development. Cell 155, 215-227 (2013).

Acknowledgements

This work was supported by CIRM RN1-00557, NIH RO1 DK100959 and LLS Scholar award for H.K.A.M; the Broad Stem Cell Center at UCLA and the JCC Foundation; and NIH P01 GM081621 for Z.G. D.R.D. was supported by the NSF GRFP and Ruth L. Kirschstein National Research Service Award GM007185. V.C. was supported by the LLS Special Fellow Award and BSCRC post-doctoral fellow award.

Associated Publications

This protocol is related to the following articles:

- Medial HOXA genes demarcate haematopoietic stem cell fate during human development
  Diana R. Dou, Vincenzo Calvanese, Maria I. Sierra, Andrew T. Nguyen, Arazin Minasian, Pamela Saarikoski, Rajkumar Sasidharan, Christina M. Ramirez, Jerome A. Zack, Gay M. Crooks, Zoran Galic, and Hanna K. A. Mikkola

Author information

Affiliations

University of California, Los Angeles

Diana Dou, Vincenzo Calvanese, Pamela Saarikoski & Zoran Galic

Mikkola Lab (UCLA)

Hanna Mikkola

Competing financial interests

The authors declare no conflicting financial interests.

Corresponding author

Correspondence to: Hanna Mikkola (hmikkola@mcdb.ucla.edu)

Readers' Comments

Comments on this thread are vetted after posting.

Protocol Exchange ISSN 2043-0116