UNIVERSITY OF CALIFORNIA

Los Angeles

Forming the Hematopoietic Stem Cell Niche from Pluripotent Stem Cells

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Cellular and Molecular Pathology

by

Chee Jia Chin

2016

ABSTRACT OF THE DISSERTATION

Forming The Hematopoietic Stem Cell Niche From Pluripotent Stem Cells

by

Chee Jia Chin

Doctor of Philosophy in Cellular and Molecular Pathology

University of California, Los Angeles, 2016

Professor Gay M. Crooks, Chair

Hematopoietic stem cells (HSCs) are characterized by their ability to self-renew and contribute to multi-lineage differentiation, critical functions that ensure homeostatic production of all blood cells throughout life. A major challenge in the field is the inability to culture HSCs without compromising self-renewal. My goal was to reconstruct an optimal hematopoietic niche from mesodermal differentiation of human pluripotent stem cells (hPSC) to facilitate the maintenance of functional HSCs. Initial characterization of the human fetal bone marrow and adult adipose tissues revealed that a subpopulation of CD146+ perivascular cells was capable of supporting the self-renewal of transplantable human cord blood HSCs *ex vivo*. Vector integration site-based lineage tracing technology was utilized to gain insights into the stages through which hPSC differentiate into mesodermal derivatives of the HSC niche. High throughput sequencing revealed the presence of mesodermal progenitors with trilineage (hematopoietic, endothelial and mesenchymal) potential during hPSC differentiation and uncovered the lineage bifurcation of

hematopoietic and mesenchyme early on in separated bipotent populations. The hPSC-derived mesenchymal populations were further studied in relation to their ability to promote or inhibit HSC maintenance, leading to the discovery of a population that was phenotypically, functionally and molecularly similar to primary human perivascular cells of the HSC niche. These findings elucidated lineage commitment events during early human embryogenesis and informed strategies to optimize the therapeutic development of cell lineages from hPSC.

The dissertation of Chee Jia Chin is approved.

Dinesh Subba Rao

Michael Alan Teitell

Donald Barry Kohn

Hanna K. A. Mikkola

Gay M. Crooks, Committee Chair

University of California, Los Angeles

2016

Dedication

To mother, Yuek Eng Chee,

father, Kok On Chin, and

brother, Lin Zhan Chin,

for being the cheerleaders in all of my pursuits

Table of Contents

List of figures and tables

Acknowledgements

First and foremost, I am immensely grateful to my mentor, Dr Gay Crooks, who is a passionate researcher, dedicated educator, avid learner and thoughtful leader. She has been the role model that guided me to improve continuously. When I first rotated, she challenged me to speak up my viewpoints. When I wanted to enrich my scientific toolkit, she encouraged me to learn broadly from the best minds. And when I prepared for publications, she patiently revised my writing to put forth the strongest manuscripts. Her leadership has nurtured a wonderful lab that collaborates and thrives as a team. I thank her for wholeheartedly supporting every aspect of my career and personal growth.

I would like to thank my current and past committee members, Dr. Mike Teitell, Dr. Donald Kohn, Dr. Dinesh Rao, Dr. Hanna Mikkola and Dr Benhur Lee. They have offered their time generously to guide me in my projects, to critique my manuscripts and to prepare the many letters supporting my fellowship applications.

My PhD journey was an adventurous and joyful ride because of the current and past members of the Crooks lab. I want to thank Dr Mirko Corselli, my postdoctoral mentor, who introduced me to the world of HSC niche biology and to the appreciation of Italian espresso. My deskmate/lab manager, Judy Zhu, has created a home away from home for me in Los Angeles. She taught me to learn relentlessly from everyone, to design and execute definitive experiments and to make well-rounded decisions with genuine interest in everyone's success. I want to thank Amelie Montel-Hagen, Stephanie C. de Barros, Chris Seet. The intensity of their work, the excitement they share for good data (including of others), generosity in sharing resources, makes the lab a stimulating environment to be in. I thank William Kim and Lisa Kohn, the senior graduate students in lab when I first joined. Through late experiments, weekend cell culture and even after their PhD journey, they have always been listening ears and cheering buddies. I am grateful to work with Dr. Gautam Dravid, Dr. Denis Evseenko, Dr. Batul Suterwala, Dr. Chintan Parekh, Dr Salemiz Sandoval and Dr Vanessa Bundy. I deeply appreciated the inspiration in new ideas and the warm words of encouragement they offered throughout. Last but not least, it has been amazing to work, laugh and potluck with Edward He, Nazilla Forootan, Kenneth Kim, Rebecca Chan, Su Wen Li, Priyanka Pullarkat and Brent Chick.

It was a true pleasure to collaborate with Aaron Cooper on the integration site analysis project. He was my role model as an incredible scientist, engineer and coffee connoisseur. Working with him was an intellectual osmosis, as he was always available for an in depth chalk talk. I also thank Georgia Lill for her enormous contribution to the integration site analysis and her countless revisions on the manuscript.

I would like to acknowledge the bioinformatics analysis that David Casero has performed on the RNA Seq datasets. I thank him for breaking down and explaining concepts foreign to me, for raising follow-up bioinformatics approaches, and for his thorough and detailed analyses.

I am indebted to Felicia Codrea and Jessica Scholes for the meticulous sorts that they have performed at the Broad Stem Cell Center Flow Cytomery Core. I thank them for both their

technical support and the happy chats as I dissected the heterogeneity of the mesenchymal populations.

This work will not be possible without the financial support from International Fulbright Science and Technology Award, California Institute of Regenerative Medicine Predoctoral Fellowship and Dissertation Year Fellowship.

Chapter 2 is reproduced with permission from *Blood*. Chapter 3 is reproduced with permission from *Stem Cells*.

Thank you to Olivia Sattayapiwat, Tracy Chong Lu, Iris Lam, Stacey Yeh, Ricky Jin, Kathy Tran for the 5+ years of friendship. They have celebrated my milestones, supported me through the ups and downs, cheered me up with food and let me win in Monopoly®.

Finally, I thank my parents, Kok On Chin and Yuek Eng Chee, for firmly supporting me in all circumstances. I thank them for the sacrifices they made to afford me the best education and opportunities to see the world. Their unconditional love allow me to explore, experience and conquer the unknowns. I also thank my brother, Lin Zhan Chin, for listening to my numerous practice presentations, for making good analogies on my research and for cheering through rain and shine.

Vita: Chee Jia (Julia) Chin

Education

University of California, Los Angeles (UCLA) Expected June 2016
- PhD candidate in Cellular & Molecular Pathology
- GPA: 3.8/4.0, GRE Math 800

King Abdullah University of Science and Technology(KAUST), Saudi Arabia June 2011
- M.S. in Bioscience (Spearheaded the set-up and build-out of a new research laboratory)
- GPA: 4.0/4.0 (top 1%)

University of California, Los Angeles (UCLA) June 2009
- B.S. Molecular, Cell and Developmental Biology. Minor in Biomedical research
- GPA: 3.98/4.0 (*summa cum laude*, top 5%)

Green River Community College, United States Aug 2007
- Associate of Science, Biology

Academic Honors

Fulbright International Science and Technology Award July 2011-July 2014
- Nominated by Malaysia, competed worldwide for 40 awards/year (<1% selectivity), full ride scholarship
- Most prestigious science and technology fellowship sponsored by the U.S. Department of State (scholarship worth totaling $180,000/person)

Publications

- **Chin CJ**, Corselli M, Casero D, Zhu Y, He CB, Hardy R, Péault B, Crooks GM. Pericyte-like cells generated from human pluripotent stem cells support hematopoietic stem and progenitors *ex vivo.* In preparation. May 2016.
- **Chin CJ***, Cooper AC*, Lill GR, Evseenko D, Zhu Y, He CB, Casero D, Pallegrini M, Kohn DB, Crooks GM. Genetic Tagging During Human Mesoderm Differentiation Reveals Tripotent Lateral Plate Mesodermal Progenitors. *Stem Cells.* 2016 Mar 2. doi: 10.1002/stem.2351. [Epub ahead of print] *Equal contribution
- Kim W, Zhu Y, Deng Q, **Chin CJ,** He CB, Grieco A, Dravid G, Parekh C, Hollis R, Lane TF, Bouhassira EE, Kohn DB, Crooks GM. Erythropoiesis from human embryonic stem cells through Erythropoietin-independent AKT signaling. *Stem Cells.* doi: 10.1002/stem.1677. 2014. PMID: 24677652
- Gkountela S, Li Z, **Chin CJ,** Lee SA and Clark AT. PRMT5 is required for human embryonic stem cell proliferation but not pluripotency. *Stem Cell Reviews and Reports.* 10(2):230-9.2014. PMID: 24477620 PMCID: PMC3972323
- Corselli M, **Chin CJ**, Parekh C, Sahaghian A, Wang W, Ge S, Evseenko D, Wang X, Montelatici E, Lazzari L, Crooks GM, Péault B. Perivascular support of human hematopoietic stem/progenitor cells. *Blood.* 121(15):2891-901. 2013. PMID: 23412095 PMCID: PMC3707421

Research Experience

Department of Pathology and Laboratory Medicine, UCLA July 2011-July 2016
PhD Thesis Advisor: Gay Crooks, MD
Project title: Forming the Hematopoietic Niche from Human Pluripotent Stem Cells (hPSCs)

- Characterized mesenchymal stroma cells derived from hPSC to identify positive and negative regulatory signals for *ex vivo* hematopoietic stem cell maintenance
- Investigated the clonal relationships of the hematopoietic and non-hematopoietic lineages during mesoderm differentiation of hPSCs using genetic labeling and Illumina high-throughput sequencing

Division of Bioscience, KAUST, Saudi Arabia Sept 2009-June 2011

Masters Thesis Advisor: Jasmeen Merzaban, Ph.D

Using mass spectrometry-based proteomics:

- Identified bone marrow homing receptors expressed on human hematopoietic stem cells
- Elucidated signaling pathways responsible for the blockage of differentiation in leukemia

Leadership Experience

Advancing Women in Science and Engineering (AWiSE), UCLA Sept 2014-June 2016

- Logistics Coordinator of AWiSE (250+ graduate student members); organized career panels and networking nights
- Fundraised $2,500 for AWiSE STEM Day, an outreach activity for 100+ inner city, middle-school girls

Women Initiative for Support and Development, KAUST Sept 2009-June 2010

- 1 out of 5 co-founders, led 11 science departments to raise awareness of issues facing women scientists at Saudi Arabia
- Organized a symposium (80+ attendees) to foster communications and to promote diversity inclusion

Fellowships and Training Grants

- **Dissertation Year Fellowship, UCLA Graduate Division** July 2015
- **California Institute of Regenerative Medicine Predoctoral Fellowship** July 2014

Presentations

International Society for Stem Cell Research 13th Annual Meeting, Stockholm, Sweden

- Invited speaker, oral presentation to 400+ stem cell scientists and biotech industry attendees (Jun 2015)

Keystone Symposia: Hematopoiesis, Keystone, Colorado, USA Feb 2015

- Poster presentation

11th Annual Stem Cell Symposium, UCLA Broad Stem Cell Research Center Feb 2015

- Poster presentation

Teaching Experience

Teaching assistant in Developmental Biology, UCLA Sept 2012-Mar 2013

- Taught two upper division discussions to 250+ Juniors and Seniors
- Two office hours/week; Designed 12 lectures, generated quizzes to enhance student learning in science

Academic awards

Dean's Prize for Outstanding Research, UCL June 2008

- 20 awardees among competition from over 200 undergraduate researchers

Chapter 1

Introduction

The hematopoietic system supplies our body with >100 billion mature blood cells every day that carry out functions such as oxygen transport, immunity, and tissue remodeling[1]. Hematopoietic stem cells (HSCs), located at the top of the hematopoietic hierarchy, are responsible for replenishing our pool of blood cells throughout life when the system is stressed. Intensive research efforts have begun to uncover the cellular constituents that comprise the HSC niche and the signaling pathways that regulate self-renewal and differentiation of HSC.

During steady state adult hematopoiesis, HSCs are located in specialized niches in the bone marrow. Several cell populations that regulate HSC have been described in terms of their spatial relationship to bone and vascular components of the bone marrow, termed the osteoblastic and perivascular niche, respectively. The generation of knockout and transgenic mice has been extremely useful in these experiments, as it allows researchers to selectively and systematically evaluate the functional importance of individual cell population or growth factors. By studying a Stem cell factor (SCF)-reporter mouse, perisinusoidal leptin receptor (*lepr*)-expressing stromal cells were discovered to be the major source of SCF and CXCL12 in the bone marrow. Conditional deletion of SCF with *Lepr*-Cre led to the depletion of quiescent HSCs [2]. $Nestin^{Bright}NG2^{+}$ mesenchymal progenitors, located adjacent to arterioles that run along the endosteal surface, are also necessary in maintaining HSC quiescence [3]. Studies from the murine HSC niche has established multiple mesenchymal subsets that co-regulate HSC maintenance.

On the other hand, study of the human hematopoietic niche is largely limited to the analysis of BM biopsy specimens or studying selective population isolated from the human bone marrow. The limited yield and access to tissues poses significant challenge to study the potential existence of distinct perivascular stromal cells as seen in the murine bone marrow and to characterize the mechanisms of HSC regulation underlying each human mesenchyme subsets.

I have developed my PhD thesis around the question of how mesenchymal components in the perivascular niche regulate development and self-renewal of HSC. Human adipose tissue was chosen as an alternative source for study because the microvessels in lipoaspirate provide an abundant and easily accessible source of perivascular cells. My work with Dr Mirko Corselli, a postdoctoral fellow in the Crooks lab, led to the discovery that CD146+ perivascular cells isolated from human adipose tissues and fetal bone marrow were functionally distinct from other mesenchymal cells in their ability to maintain repopulating HSCs during ex vivo culture [4]. We showed that expression of key ligands discriminated the CD146+ pericytes from the more numerous PDGFRa-expressing mesenchyme and that notch ligands such as Jagged-1 mediate pericyte support of HSC. This work is described in Chapter 2.

Using these findings from fetal and adult human tissue as a foundation, I have developed a model to study the embryonic development of components of the HSC niche as they emerge during differentiation of human pluripotent stem cells (hPSCs). hPSC have the potential to differentiate into any adult tissue, can be cultured long term, and are amenable to genetic modification. Therefore, the prospect of using hPSC-derived niche populations represents an effective, versatile way to study human HSC maintenance in vitro. Numerous cellular constituents such as osteoblastic, endothelial, adipocytic and perivascular cells in the niche are

mesodermal in origin. Our lab has previously characterized a human embryonic mesoderm progenitor (hEMP) population derived from hPSC that corresponds to the onset of mesoderm commitment and has the potential to generate broad mesodermal derivatives [5]. The differentiation of hEMP offers a platform to track the initial fate decision between hematopoietic cells and the hematopoietic-supporting stroma and to systematically examine various mesoderm derivatives for their ability to support hematopoiesis.

Specifically, my thesis research aims to (1) Define the clonal relationships of hematopoietic and non-hematopoietic lineages during mesodermal differentiation from hPSC; (2) Characterize the hPSC-derived mesenchymal subsets that provide negative and positive regulatory signals in regulating *ex vivo* maintenance of definitive HSC. Studying the embryonic origin of mesodermal commitment and modeling HSC-supportive niche from embryonic mesoderm will contribute to the understanding of early human embryogenesis and improvement of directed differentiation of hPSC for therapeutic applications.

1. Boulais PE, Frenette PS. Making sense of hematopoietic stem cell niches. **Blood**. 2015;125:2621-2629.
2. Ding L, Saunders TL, Enikolopov G et al. Endothelial and perivascular cells maintain haematopoietic stem cells. **Nature**. 2012;481:457-462.
3. Kunisaki Y, Bruns I, Scheiermann C et al. Arteriolar niches maintain haematopoietic stem cell quiescence. **Nature**. 2013;502:637-643.
4. Corselli M, Chin CJ, Parekh C et al. Perivascular support of human hematopoietic stem/progenitor cells. **Blood**. 2013;121:2891-2901.
5. Evseenko D, Zhu Y, Schenke-Layland K et al. Mapping the first stages of mesoderm commitment during differentiation of human embryonic stem cells. **Proceedings of the National Academy of Sciences of the United States of America**. 2010;107:13742-13747.

Chapter 2

Regular Article

HEMATOPOIESIS AND STEM CELLS

Perivascular support of human hematopoietic stem/progenitor cells

Mirko Corselli,[1,2] Chee Jia Chin,[3] Chintan Parekh,[4] Arineh Sahaghian,[3] Wenyuan Wang,[5] Shundi Ge,[3] Denis Evseenko,[1,2,6] Xiaoyan Wang,[7] Elisa Montelatici,[8] Lorenza Lazzari,[8] Gay M. Crooks,[2,3,7] and Bruno Péault[1,2,9]

[1]University of California at Los Angeles (UCLA) and Orthopaedic Hospital Department of Orthopaedic Surgery and the Orthopaedic Hospital Research Center, [2]Broad Stem Cell Research Center, and [3]Department of Pathology and Laboratory Medicine, UCLA, Los Angeles, CA; [4]Division of Pediatric Hematology/Oncology, Children's Hospital Los Angeles, Los Angeles, CA; [5]Molecular Biology Institute, [6]Jonsson Comprehensive Cancer Center, and [7]Department of Biostatistics, UCLA, Los Angeles, CA; [8]Cell Factory, Fondazione Istituto Di Ricovero e Cura a Carattere Scientifico (IRCCS) Ca' Granda Ospedale Maggiore Policlinico, Milan, Italy; and [9]Center for Cardiovascular Science and Center for Regenerative Medicine, University of Edinburgh, Edinburgh, United Kingdom

Key Points

- Perivascular cells maintain HSPCs ex vivo.

Hematopoietic stem and progenitor cells (HSPCs) emerge and develop adjacent to blood vessel walls in the yolk sac, aorta-gonad-mesonephros region, embryonic liver, and fetal bone marrow. In adult mouse bone marrow, perivascular cells shape a "niche" for HSPCs. Mesenchymal stem/stromal cells (MSCs), which support hematopoiesis in culture, are themselves derived in part from perivascular cells. In order to define their direct role in hematopoiesis, we tested the ability of purified human CD146$^+$ perivascular cells, as compared with unfractionated MSCs and CD146$^-$ cells, to sustain human HSPCs in coculture. CD146$^+$ perivascular cells support the long-term persistence, through cell-to-cell contact and at least partly via Notch activation, of human myelolymphoid HSPCs able to engraft primary and secondary immunodeficient mice. Conversely, unfractionated MSCs and CD146$^-$ cells induce differentiation and compromise ex vivo maintenance of HSPCs. Moreover, CD146$^+$ perivascular cells express, natively and in culture, molecular markers of the vascular hematopoietic niche. Unexpectedly, this dramatic, previously undocumented ability to support hematopoietic stem cells is present in CD146$^+$ perivascular cells extracted from the nonhematopoietic adipose tissue. (*Blood.* 2013;121(15):2891-2901)

Introduction

Blood and vasculature are indispensable to embryonic development, and are thus the first differentiated tissues produced in life. Incipient human hematopoiesis adapts to the rudimentary anatomy of the embryo and proceeds first in the yolk sac, then transiently in the placenta and liver before being stabilized in fetal bone marrow (FBM). Definitive hematopoietic stem and progenitor cells (HSPCs) first emerge in the aorta-gonad-mesonephros region of the embryo.[1] Therefore, several organs of distinct germline origins, structures, and eventual roles converge functionally to produce blood cells during development. What remains however remarkably constant through pre- and postnatal life is the physical association of incipient hematopoietic cells with blood vessels. In the yolk sac, erythroid cells emerge within intravascular blood islands.[2] It is now also well accepted that, from fish to humans, specialized blood-forming endothelial cells present in the dorsal aorta and possibly other organs supply the embryo with hematopoietic cells,[3-7] an ontogenic transition that has been modeled in human embryonic stem cells.[8] In addition to this direct developmental affiliation between embryonic endothelial cells and HSPCs, there is evidence that vascular cells nurture blood cells in pre- and postnatal life. The cellular and molecular mechanisms involved in this support can be analyzed in cocultures of stromal and hematopoietic cells.[9-11] For instance, cultured endothelial cells use angiocrine factors to regulate HSPC differentiation or self-renewal.[12-14] Mesenchymal stem/stromal cells (MSCs), the multilineage mesodermal progenitors spontaneously selected in long-term cultures of unfractionated cells from bone marrow and other tissues,[15-18] can also, to some extent, sustain hematopoiesis in vitro.[19-24] However, the relevance of this support to physiologic blood cell production in vivo has been unknown because MSCs have long eluded prospective identification.[25] Similarities between MSCs and pericytes, which ensheath capillaries and microvessels in all organs, have been described.[26-28] In an experimental approach combining stringent cell purification by flow cytometry and differentiation in culture and in vivo, we have demonstrated that human CD146$^+$ perivascular cells represent ubiquitous ancestors of MSCs.[29]

Although hematopoietic stem cells (HSCs) were originally detected in the endosteal regions of the bone marrow,[30] recent findings have suggested the existence of a distinct, perivascular niche for HSPCs.[31-34] Perivascular reticular cells expressing CXCL12 were found to play a role in murine HSC maintenance.[35] In a seminal study by Méndez-Ferrer et al,[36] the function and identity of perivascular niche cells were further defined. The authors showed the existence in murine bone marrow of perivascular nestin$^+$ MSCs associated with HSCs. Ablation of nestin$^+$ MSCs led to a significant reduction in the number and homing ability of HSCs. The direct role for perivascular cells in hematopoiesis regulation was confirmed in a recent study by Ding et al.[37] Selective shutoff of c-kit ligand expression in leptin receptor

Submitted August 24, 2012; accepted February 10, 2013. Prepublished online as *Blood* First Edition paper, February 14, 2013; DOI 10.1182/blood-2012-08-451864.

The online version of this article contains a data supplement.

There is an Inside *Blood* commentary on this article in this issue.

The publication costs of this article were defrayed in part by page charge payment. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 USC section 1734.

(Lep-R) positive cells surrounding murine bone marrow blood vessels significantly reduced the frequency of long-term reconstituting HSCs.[37]

In the present study, we demonstrate that $CD146^+$ perivascular cells express in vivo nestin, CXCL12, and Lep-R in human FBM as well as in adult adipose tissue. We also report for the first time that human $CD146^+$ perivascular cells are a subset of MSCs able to directly support the ex vivo maintenance of human HSPCs. We further demonstrate that cultured $CD146^+$ perivascular support HSPCs through cell-to-cell contact and activation of Notch signaling. Conversely, conventional unfractionated MSCs or the $CD146^-$ subset of MSCs favor differentiation at the expense of stemness. $CD146^+$ perivascular cells can therefore be considered as the bona fide human equivalents of the hematopoietic perivascular niche components recently described in the mouse.

Methods

Isolation of human primary stromal cells

Human stromal cells were derived from human lipoaspirate specimens (n = 4) and FBM (n = 2) as previously described.[17,29] Lipoaspirates were obtained as discarded specimens without identifiable information, therefore no institutional review board approval was required. Fetal bones (16-18 weeks of pregnancy) were obtained from Novogenix. One hundred milliliters of lipoaspirate were incubated at 37°C for 30 minutes with digestion solution composed by RPMI 1640 (Cellgro), 3.5% bovine serum albumin (Sigma), and 1 mg/mL collagenase type II (Sigma). Adipocytes were discarded after centrifugation while the pellet was resuspended and incubated in red blood cell lysis (eBioscience) to obtain the stromal vascular fraction (SVF). Fetal bones were split open to flush the bone marrow cavity. The bones were placed in digestion solution for 30 minutes at 37°C. Mononuclear cells (MNCs) were isolated using Ficoll-Paque (GE Healthcare). Hematopoietic cells were excluded by magnetic immunodepletion of $CD45^+$ cells as per manufacturer's instructions (Miltenyi Biotec). An aliquot of SVF or CD45-depleted MNCs was plated in tissue-culture–treated flask for the expansion of conventional MSCs.[17] Another aliquot of SVF or CD45-depleted MNCs was processed for fluorescence-activated cell sorting (FACS). Cells were incubated with the following antibodies: CD45–allophycocyanin (APC)–cy7 (BD Biosciences), CD34-APC (BD Biosciences), and CD146–fluorescein isothiocyanate (FITC; AbD Serotec). The viability dye 4,6 diamidino-2-phenylindole (DAPI; Sigma) was added before sorting, on a FACSAria III (BD Biosciences), $DAPI^-CD45^-CD34^-CD146^+$ perivascular cells, or $DAPI^-CD45^-CD34^+CD146^-$ cells, as previously described.[29,38] In some experiments, $CD146^-$ cells were purified from cultured MSCs.

For the animal studies, an animal care and use committee protocol (ARC no. 2008-175-11) was approved for the injection of human cells into immunodeficient mice and for the analysis of engraftment of transplanted cells.

Isolation of human $CD34^+$ cells from CB

Umbilical cord blood (CB) was collected from normal deliveries without individually identifiable information, therefore no institutional review board approval was required. MNCs were isolated by density gradient centrifugation using Ficoll-Paque (GE Healthcare). Enrichment of $CD34^+$ cells was then performed using the magnetic-activated cell sorting system (Miltenyi Biotec) as per the manufacturer's instructions.

Immunophenotype analysis of stromal cells

Cultured MSCs, $CD146^+$ cells, and $CD146^-$ cells (between passages 3 and 10) were analyzed on an LSR II flow cytometer (Becton Dickinson). Cells were stained with monoclonal antibodies: CD146-FITC (AbD Serotec), CD31-APC (Biolegend), CD44–phycoerythrin (PE), CD73-PEcy7, CD105-PE, CD90-APC, and CD45-FITC (all from BD Biosciences). Unstained samples were used as negative controls. Data were analyzed using FlowJo software (Tree Star).

Mesodermal lineage differentiation assays

The ability of cells to differentiate into mesodermal lineages was tested in osteogenic or adipogenic differentiation medium (Hyclone). After 3 weeks of culture in differentiation conditions, cells were stained with Alizarin red or Oil red O (Sigma) for the detection of mineral deposits or lipids as previously described.[29]

Quantitative RT-PCR

Five hundred thousand cultured cells were processed for RNA extraction using a Qiagen micro kit. An Omniscript reverse transcriptase (RT) kit was used to make complementary DNA, which was subjected to quantitative polymerase chain reaction (qPCR) using Sybr green probe–based gene expression analysis (Applied Biosystems) for 2 housekeeping genes, TBP and GAPDH, and the target genes CD146, nestin, α-SMA, and NG2. A 7500 real-time PCR system was used (ABI). Data were analyzed using the comparative C(T) method.

Western blotting

Cells were lysed in denaturing cell extraction buffer (Invitrogen) containing protease inhibitor tablets (Roche). Proteins were then separated by sodium dodecyl sulfate–polyacrylamide gel electrophoresis and analyzed using the XCell II western blot system (Invitrogen). Rat anti-human Jagged-1 (Abcam, 1:50) and monoclonal mouse anti-β actin (Sigma, 1:5000) antibodies were used. Donkey anti-rat horseradish peroxidase (HRP) and donkey anti-mouse HRP (Jackson Immunoresearch Inc; 1:5000) were used as secondary antibodies. The blots were developed using ECL Plus Western Blotting Substrate (Pierce).

Coculture of stromal cells and CB $CD34^+$ cells

Cultured stromal cells (between passages 3 and 8) were irradiated (20 Gy) and plated on 96 multi-well plates at 1.5×10^4 cells per well. Twenty-four hours later, CB $CD34^+$ cells ($5\text{-}7 \times 10^4$ per well) were plated on top of the stromal layer. Stroma-free cultures were performed seeding CB $CD34^+$ cells on recombinant retronectin (RN; Lonza) coated wells. Cocultures were performed in RPMI 1640, 5% fetal bovine serum, 1× penicillin/streptavidin. No supplemental cytokines were ever added. Cells were harvested after 1, 2, 4, and 6 weeks. Cocultures in the absence of cell-to-cell contact were performed in 96 multi-well transwell plates (Corning). For the inhibition of Notch, 10μM DAPT (N-[N-(3, 5-difluorophenacetyl-L-alanyl)]-S-phenylglycine t-butyl ester; Sigma) or 10 μg/mL of anti-human Notch 1 neutralizing antibody (Biolegend) were added to each well every 48 hours. An equal volume of dimethylsulfoxide (DMSO; Sigma) or an equal concentration of mouse unrelated IgG (Biolegend) was added to wells as negative controls for DAPT and anti-Notch1 antibody, respectively.

Flow cytometric analysis of cultured CB $CD34^+$ cells

After 1, 2, 4, and 6 weeks of coculture, cells were harvested and stained with the following antibodies: CD45-APC-cy7, CD34-PE-cy7, CD14-APC, CD10-APC, CD33-PE, CD19-FITC (all from BD Biosciences). Dead cells were identified with propidium iodide (PI; BD Biosciences).

Colony forming unit assay

After 1, 2, 4 and 6 weeks of coculture cells were harvested and 2.5×10^3 cells were plated in methylcellulose (Methocult; Stem Cell Technologies). Colonies, here reported as the sum of the progeny of colony forming unit (CFU) granulo-macrophage, burst-forming unit erythroid, and CFU mixed, were scored after 14 days.

In vivo repopulation assay

CB $CD34^+$ cells were cocultured with MSCs or $CD146^+$ cells for 2 weeks in RPMI 1640, 5% fetal bovine serum, 1× penicillin/streptavidin. An equal number of $CD45^+$ cells (10^5) obtained from the cocultures was intratibially injected in sublethally irradiated (250 cGy), 6- to 8-week-old NSG mice (The Jackson Laboratory). Mice were sacrificed 6 weeks posttransplantation. Engraftment of human hematopoietic cells was evaluated by FACS analysis after staining with anti-human specific monoclonal antibodies: CD45-APC-cy7, HLA (A/B/C)–PE, CD34-PE-cy-7, CD19-FITC, CD14-APC, CD15-APC, CD33-APC (all from BD Biosciences). For secondary transplantation, bone marrows from 2 engrafted mice were pooled and intratibially injected into a secondary host (n = 4). Engraftment was evaluated 4 weeks posttransplantation.

Immunocytochemistry and immunohistochemistry

For immunofluorescence analysis, human adipose tissue-frozen sections, cells cultured in chamber slides (Millipore) or cytospun on microscope slides, were fixed with cold methanol/acetone (1:1) for 5 minutes at room temperature prior to incubation with blocking solution (phosphate-buffered saline 5% donkey serum) for 1 hour at room temperature. Overnight incubation at 4°C was performed with unconjugated primary antibodies: mouse anti-human CD146 (BD Biosciences), mouse anti-human CD45 (eBioscience), rat anti-human Jagged-1, rabbit anti-human N1ICD, mouse anti-human nestin, rabbit anti-human CXCL12, rabbit anti-human Lep-R, rabbit anti-human CD146 (all from Abcam). Tissue sections or cells were incubated for 2 hours at room temperature with FITC-conjugated mouse anti-human von Willebrand factor (VWF; US Biological). Tissue sections or cells were incubated for 1 hour at room temperature with the following conjugated antibodies: donkey anti-rabbit–Alexa 488, donkey anti-rabbit– Alexa 647, donkey anti-rat–Alexa 594 or donkey anti-mouse–Alexa 594 (all from Jackson Immunoresearch Inc). For immunohistochemistry on human FBM, fetal bones (16-18 weeks of pregnancy) were fixed in 4% paraformaldehyde (Sigma-Aldrich). Fixed tissues were embedded in paraffin and sections were stained with the same antibodies against nestin, CXCL12, Lep-R, and CD146. Secondary HRP-conjugated IMPRESS anti-rabbit and anti-mouse antibodies and 3, 3′-diaminobenzidine (Vector Laboratories) were used for revelation. As negative controls, tissue sections or cells were incubated only with secondary antibodies. Images were acquired on an Axiovision microscope (Carl Zeiss; software version 4.8) equipped with ApoTome.2 modules for Axio Imager.2 and Axio Observer, with 10×, 20×, 40×, and 63× (1.4 NA) objectives.

Statistical analysis

Mean and SDs were used to summarize continuous variables. Bivariate cross-sectional comparisons of continuous variables were performed using paired *t* tests. Continuous outcomes such as total numbers of $CD45^+$ and $CD34^+$ cells, frequency of $CD34^+Lin^-$ cells and CFUs were collected over time. The experimental design involved 2 within-experiment factors, MSCs and pericytes, and time (week 1, 2, 4, 6), which corresponded to a strip-plot design. A mixed-model approach was used. Within the mixed-model framework, we performed hypothesis testing for the comparison of MSCs and pericytes at different time points. Pearson's correlation (r) was reported to assess the linear correlation between $CD34^+lin^-$ cells and CFUs. For the qPCR data, ΔCT values were calculated for each marker. A randomized block design model was fitted on ΔCT values. Donors were treated as random effects while stromal cells groups were treated as fixed effects. For all statistical investigations, tests for significance were 2-tailed. To account for type I error inflation due to multiple comparisons, *P* values were adjusted by Bonferroni correction. The Fisher exact test was performed to compare engraftment and not engraftment ability. Statistically significant threshold of *P* value was set at .05. Statistical analyses were carried out using SAS version 9.2 (SAS Institute).

Results

Human $CD146^+$ perivascular cells express nestin, CXCL12, and Lep-R in hematopoietic and nonhematopoietic tissues

Recent studies have described murine perivascular cells as key players for the maintenance of HSPCs. Perivascular niche cells, displaying MSC features, have been identified based on the expression of CXCL12,[34] nestin,[36] and Lep-R.[37] We have previously demonstrated that pericytes, surrounding microvessels, and capillaries, can be detected in multiple human tissues on expression of CD146.[29] Consistent with our previous findings, immunohistochemistry performed on human FBM revealed the presence of CD146-expressing perivascular cells (Figure 1A). Nestin, CXCL12, and Lep-R, markers of the perivascular niche previously described in murine studies, were also expressed in human perivascular cells in FBM (Figure 1B-D). We further investigated the expression of the same stromal cell markers in human adult adipose tissue, considered as an abundant source of MSCs and recently suggested to also be a reservoir of HSCs.[39] Nestin, CXCL12, and Lep-R were all expressed in cells immediately adjacent to VWF-positive endothelial cells (Figure 1E-G). Multicolor immunofluorescence showed that $CD146^+$ pericytes, surrounding microvessels and capillaries, coexpress nestin, CXCL12, and Lep-R (Figure 1H-S). Thus, human $CD146^+$ perivascular cells express in situ markers previously identified in murine studies to mark the perivascular hematopoietic niche.

Purified and ex vivo–expanded $CD146^+$ perivascular cells maintain expression of markers of the perivascular niche

We then analyzed the expression of the perivascular niche markers in purified and ex vivo–expanded $CD146^+$ perivascular cells as compared with unfractionated MSCs and $CD146^-$ cells. MSCs were conventionally derived from the adipose tissue SVF by plastic adherence, while $CD146^+$ perivascular cells and $CD146^-$ cells were purified by FACS as previously described (Figure 2A).[29,38] $CD146^+$ perivascular cells demonstrated expression of cell-surface markers typical of unfractionated cultured MSCs, such as CD44, CD105, CD73, and CD90 and did not express the hematopoietic and endothelial cell markers CD45 and CD31 (supplemental Figure 1a-b, available on the *Blood* website). Also similar to unfractionated MSCs, cultured $CD146^+$ cells were able to differentiate into osteoblasts and adipocytes in culture (supplemental Figure 1c-f). $CD146^+$ perivascular cells retained uniform CD146 expression in culture, as did a small fraction of MSCs, while $CD146^-$ cells remained negative for CD146 expression in culture (Figure 2B). Quantitative RT-PCR analysis of established cultures confirmed that $CD146^+$ cells expressed higher levels of the perivascular cell markers CD146, α-SMA, NG2, and nestin than did either unfractionated MSCs or $CD146^-$ cells derived from fat (Figure 3A) or FBM (Figure 3B). Furthermore, immunocytochemistry demonstrated that cultured $CD146^+$ perivascular cells isolated from fat or FBM express higher levels of nestin and CXCL12 than $CD146^-$ cells do. No significant difference in the expression of Lep-R was observed between cultured $CD146^+$ and $CD146^-$ cells (Figure 3C-N).

$CD146^+$ perivascular cells support HSPCs ex vivo

The ability of distinct stromal cells to support HSPCs ex vivo was assessed by coculturing CB-derived $CD34^+$ cells (CB $CD34^+$) in direct contact with either $CD146^+$ perivascular cells, unfractionated MSCs, or $CD146^-$ cells all obtained from both lipoaspirate

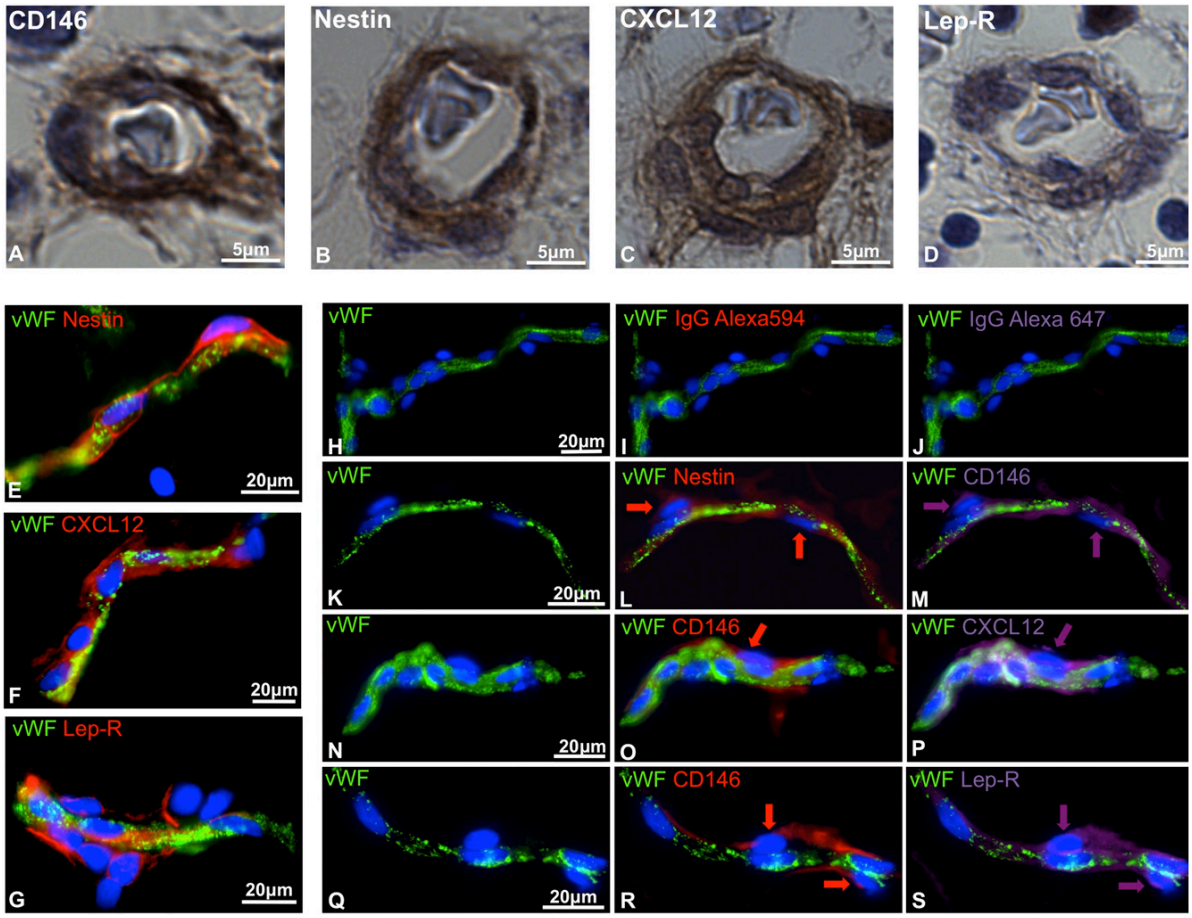


Figure 1. In situ expression of hematopoietic niche markers by human perivascular cells. (A-D) Immunohistochemistry performed on paraffin-embedded sections of 17-week-old human FBM. Pericytes surrounding microvessels express (A) CD146, (B) nestin, (C) CXCL12, and (D) leptin receptor (Lep-R) (original magnification, ×63). (E-S) Immunohistochemistry performed on human adipose tissue cryosections. (E-G) VWF-positive endothelial cells (green) are surrounded by perivascular cells expressing (E) nestin, (F) CXCL12, and (G) Lep-R. (H-S) Triple-staining immunohistochemistry performed on human adipose tissue cryosections shows coexpression of CD146 with (K-M) nestin, (N-P) CXCL12, and (Q-S) LepR. (H-J) Single staining with anti-VWF antibody followed by incubation with conjugated IgG controls revealed the lack of autofluorescence (original magnification, ×40). IgG, immunoglobulin G.

specimens and FBM. These cultures were performed in basal medium with a low concentration of serum (5%) and in the absence of any supplemental cytokines, so that the specific effect of each stromal cell subset could be assessed with minimal influence of exogenous factors. In the absence of any stromal cells or cytokines, hematopoietic cells cultured on RN died within the first 2 weeks, whereas $CD45^+$ cells survived for up to 6 weeks in the presence of either MSCs or $CD146^+$ perivascular cells (Figure 4A). The total number of $CD45^+$ cells recovered from $CD146^+$ cell cocultures remained significantly higher at any time of culture when compared with MSC cocultures (Figure 4A). A similar pattern was observed for the total number of $CD34^+$ cells (Figure 4B). CD34 expression identifies human hematopoietic cells without discriminating between HSCs and lineage-committed progenitors. The most immature progenitors present in cocultures were further defined as $CD34^+Lin^-$ cells based on expression of CD34 and lack of the early myeloid cell marker CD33 and lymphoid cell markers CD10 and CD19. $CD146^+$ cell cocultures contained a significantly higher frequency and number of $CD34^+Lin^-$ cells at all time points (Figure 4C-D). Consistent with these findings, culture in the presence of MSCs resulted in accelerated differentiation of CB $CD34^+$ cells into $CD14^+$ myeloid cells and $CD10/CD19^+$ lymphoid cells, relative to coculture with $CD146^+$ cells (Figure 4E-F). The increased frequency of myeloid and lymphoid cells was counterbalanced by the lower numbers of $CD45^+$ cells in MSC cocultures, hence no significant difference in the absolute numbers of myeloid or lymphoid cells was observed (Figure 4E-F). Furthermore, the number of clonogenic cells detected after 1, 2, 4, and 6 weeks was significantly higher when CB $CD34^+$ cells were cocultured with $CD146^+$ perivascular cells compared with MSCs (Figure 5A).

$CD146^+$ perivascular cells isolated from either FBM or adipose tissue sustained significantly more $CD34^+Lin^-$ cells and CFUs from CB $CD34^+$ cells than $CD146^-$ stromal cells did (supplemental Figure 2a-d), thus confirming that within the heterogeneous MSC population, the ability to support HSPCs is confined to the subset of $CD146^+$ perivascular cells, regardless of the tissue of origin.

$CD146^+$ perivascular cells maintain human HSPCs with repopulating ability and self-renewal potential

We next investigated whether coculture with MSCs or $CD146^+$ perivascular cells retains functional HSPCs. Sublethally irradiated

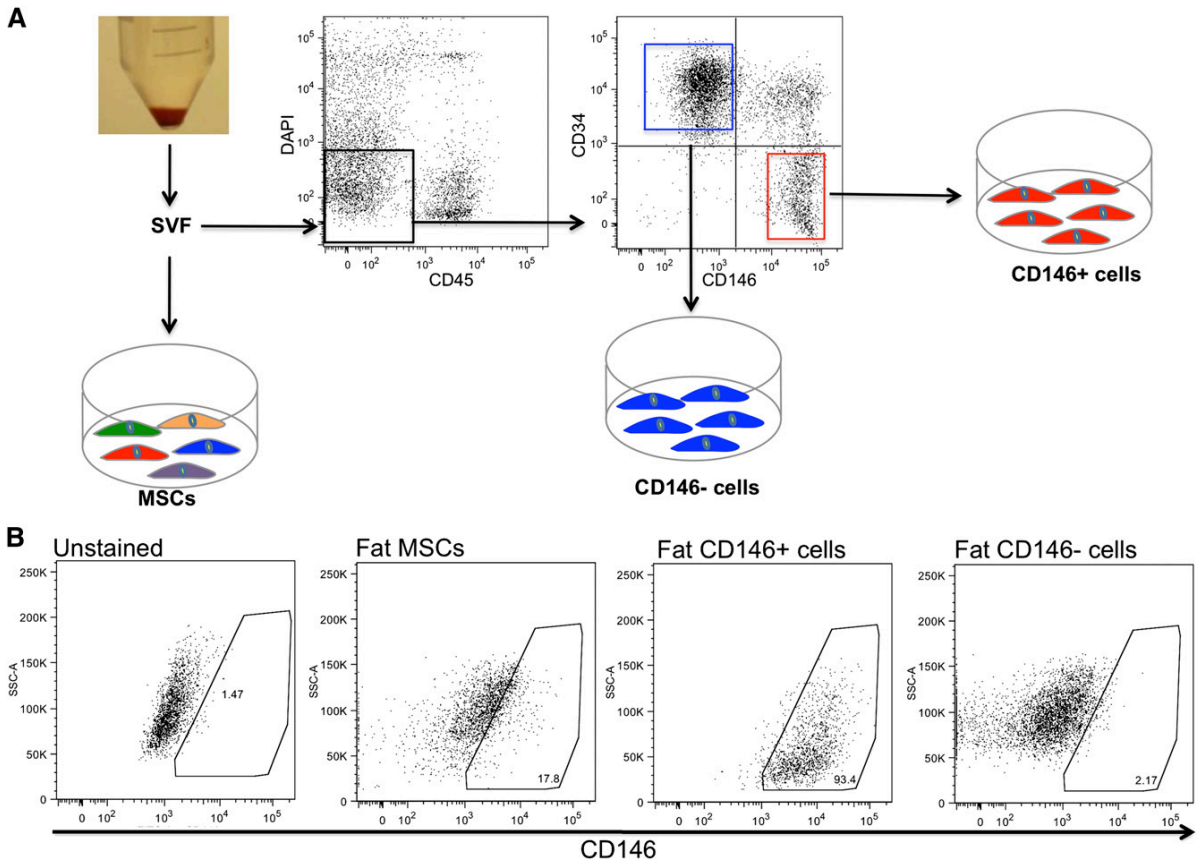


Figure 2. Isolation and culture of MSCs and stromal subsets from lipoaspirate. (A) SVF was obtained from human lipoaspirate specimens (n = 4 donors). An aliquot of SVF was directly seeded in tissue-culture plates for the isolation of conventional MSCs by plastic adherence. Another aliquot of SVF was processed for FACS sorting of DAPI$^-$CD45$^-$CD34$^-$CD146$^+$ perivascular cells and DAPI$^-$CD45$^-$CD34$^+$CD146$^-$ cells. (B) FACS analysis of cultured fat-derived MSCs, CD146$^+$ perivascular cells, and CD146$^-$ cells. After 9 passages in culture, MSCs retain a low percentage of CD146$^+$ cells, while purified CD146$^+$ perivascular cells and CD146$^-$ cells retain a stable phenotype homogeneously positive and negative for CD146, respectively.

NOD/SCID/IL-2 receptor γ-chain null (NSG) mice were injected with hematopoietic cells cocultured with CD146$^+$ perivascular cells or MSCs for 2 weeks in low-serum concentration without added cytokines. Strikingly, all mice transplanted with hematopoietic cells cocultured with perivascular cells exhibited human hematopoietic cell engraftment 6 weeks posttransplantation, whereas no engraftment was observed in any of the mice transplanted with hematopoietic cells cocultured with MSCs (n = 11 mice per group, n = 3 individual experiments; Fisher exact test, $P < .0001$) (Figure 5B-C). Human CD34$^+$ progenitors, CD19$^+$ lymphoid cells and CD14$^+$ myeloid cells were detected in the chimeric mice (Figure 5D). Human CD45$^+$HLA$^+$ cells were not only detected in the medullary site of injection, but also in the contralateral tibia, thus suggesting that HSPCs cocultured with CD146$^+$ perivascular cells maintained the ability to migrate and home to distant sites after initial engraftment (Figure 5E). To assess the self-renewal potential of HSPCs cultured in the presence of CD146$^+$ perivascular cells, bone marrow from chimeric mice was transplanted into secondary NSG mouse hosts. Lymphoid and myeloid engraftment of human cells was still detectable in secondary hosts (Figure 5F-I), demonstrating that the CD146$^+$ cell fraction of MSCs is uniquely able to sustain human HSPCs with multilineage repopulating capacity and self-renewal ability.

Contact with CD146$^+$ cells is required for HSPC maintenance

In addition to the phenotypic and functional differences described above (Figures 2 and 3), a different morphology and spatial distribution was observed between hematopoietic cells cocultured with MSCs or CD146$^+$ perivascular cells. When CD146$^+$ cells were used as a feeder layer, hematopoietic cells appeared small, rounded, and clustered (supplemental Figure 3a). In the presence of MSCs, hematopoietic cells were larger, less uniform in size, and scattered throughout the cultures, consistent with more vigorous hematopoietic differentiation (supplemental Figure 3c). Immunocytochemical analysis confirmed the presence of clusters of CD34$^+$ cells in contact with underlying CD146$^+$ cells in perivascular cell cocultures but not in MSC cocultures (supplemental Figure 3b,d). Based on these observations, we next investigated the role of cell-to-cell contact on HSPC maintenance. When direct contact between CD146$^+$ cells and CB CD34$^+$ cells was prevented in a transwell culture system, the total number of CD45$^+$ cells was dramatically reduced after 1 week of coculture (supplemental Figure 3e). In these noncontact conditions, hematopoietic cells were barely detectable after 2 weeks and the limited number of cells recovered did not allow us to perform further immunophenotypic or functional analyses.

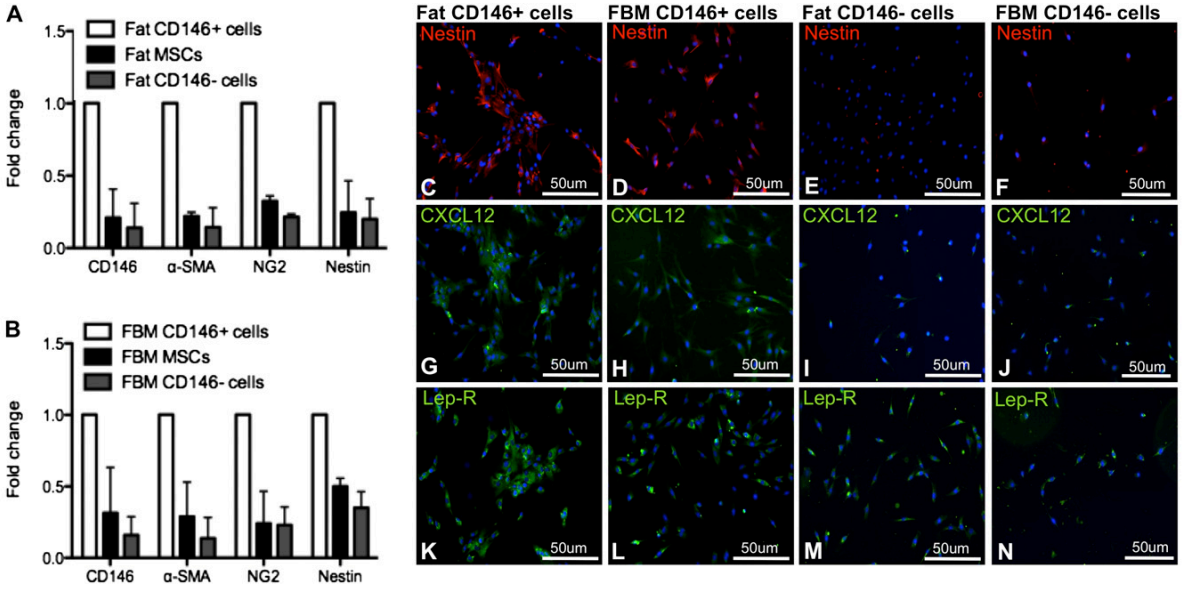


Figure 3. Cultured CD146$^+$ perivascular cells express markers of hematopoietic perivascular niche cells. (A-B) Ex vivo–expanded CD146$^+$ perivascular cells purified from fat and FBM similarly express higher levels of mRNA of perivascular cell markers when compared with MSCs and CD146$^-$ cells (n = 2 donors for each tissue). (C-N) Fat and FBM-derived CD146$^+$ perivascular cells similarly and almost exclusively express (C-F) nestin and (G-J) CXCL12 in culture compared with CD146$^-$ cells. (K-N) No difference in Lep-R expression was observed between CD146$^+$ and CD146$^-$ cells from either fat and FBM (original magnification, ×20). mRNA, messenger RNA.

CD146$^+$ perivascular cells express Notch ligands and activate Notch in hematopoietic cells

Transwell culture experiments suggested that CD146$^+$ perivascular cells sustain hematopoietic cells through cell-to-cell contact rather than by secretion of soluble factors. As Notch signaling is one of the key pathways through which the microenvironment affects growth and differentiation of HSPCs during development,[40] we investigated whether CD146$^+$ perivascular cells sustain HSPCs through the activation of Notch. Immunocytochemistry revealed that all cultured CD146$^+$ perivascular cells express high levels of the Notch ligand Jagged-1. In contrast, only rare MSCs expressed Jagged-1, the majority of which also expressed CD146 (Figure 6A). Western blot analysis detected Jagged-1 expression at high levels in CD146$^+$ perivascular cells compared with unfractionated MSCs (Figure 6B). Expression of other Notch ligands (Jagged-2, DLL-1, and DLL-4) was also detected by qPCR in MSCs, albeit at a lower level compared with CD146$^+$ perivascular cells (Figure 6C).

We used an antibody recognizing an epitope exclusively exposed after Notch 1 receptor cleavage (NICD) to measure the frequency of hematopoietic cells activating Notch in the presence of CD146$^+$ cells or MSCs (supplemental Figure 4a-b). As expected, Notch activation was not observed when direct contact between CD146$^+$ cells and hematopoietic cells was inhibited in transwell cocultures (supplemental Figure 4c-d). MSCs, which express all 4 Notch ligands tested, were able to activate Notch1 in ~50% of hematopoietic cells and progenitors (Figure 6D-E). The percentage of NICD$^+$CD45$^+$ cells was significantly higher in CD146$^+$ cell cocultures than in cocultures with total MSCs or CD146$^-$ cells, regardless of the tissue of origin (FBM or fat) (Figure 6D). Furthermore, Notch activation was significantly stronger in CD34$^+$ progenitors cocultured with CD146$^+$ cells compared with those cocultured with MSCs or CD146$^-$ cells (Figure 6E).

Notch inhibition in CD146$^+$ cell/HSPC cocultures reduces progenitor cell numbers and stimulates B-cell differentiation

To further assess the functional role of Notch activation in HSPCs, CB CD34$^+$ cells, and CD146$^+$ perivascular cells were cocultured in the presence of the γ-secretase inhibitor DAPT. Notch inhibition resulted in significantly reduced total number of CD45$^+$ cells, CD34$^+$Lin$^-$ cells, and CFUs compared with control cocultures performed in the presence of the DMSO solvent alone (Figure 7A-C). A significantly higher frequency of PI$^+$ dead cells was measured after Notch inhibition (Figure 7D). Of note, the frequency of PI$^+$ cells was not increased when CD146$^+$ perivascular cells or CB CD34$^+$ cells were treated separately with DAPT in the absence of supplemental cytokines, thus excluding nonspecific cytotoxicity from DAPT (supplemental Figure 5a). Notch inhibition also resulted in a dramatic increase in B-cell differentiation (Figure 7E). A comparable decrease in output of total CD34$^+$Lin$^-$ cells and increase in B-lymphoid cells (supplemental Figure 5b-d) was also observed when cocultures were performed in the presence of an antibody to specifically block the Notch-1 receptor. However, the effect was less pronounced when compared with DAPT treatment. To explain this difference, we determined the levels of Notch inhibition following DAPT or anti-Notch-1 blocking antibody treatment. Although Notch activation was totally abrogated by DAPT, low-level activation was still detected in a few cells after antibody treatment, confirming that the latter treatment is less efficient than chemical inactivation of Notch (supplemental Figure 4e-g).

Altogether, these results show that CD146$^+$ perivascular cells are a subset of MSCs able to support HSPCs and regulate lineage

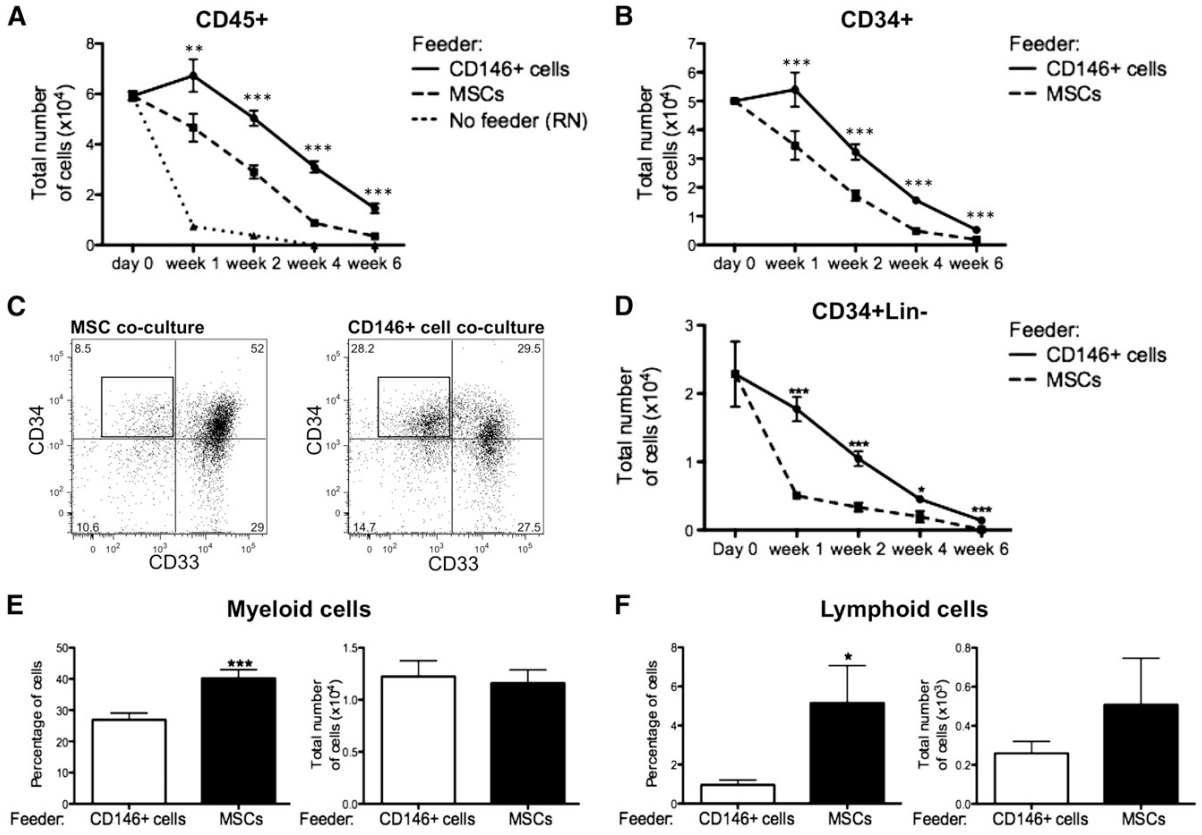


Figure 4. CD146$^+$ perivascular cells promote ex vivo maintenance of undifferentiated HSPCs. (A) In the absence of cytokines and stromal cell feeder layer (No feeder), CD45$^+$ hematopoietic cells cultured in RN-treated wells rapidly died within the first 2 weeks of culture. At any time of culture, the total number of CD45$^+$ cells recovered from CD146$^+$ cell cocultures was significantly higher when compared with MSC cocultures (n = at least 5 independent experiments for each time point, each experiment was performed in triplicate; $^{**}P < .01$, $^{***}P < .001$). (B) A similar pattern was observed for the total number of CD34$^+$ cells (n = at least 5 independent experiments for each time point, each experiment was performed in triplicate; $^{***}P < .001$). (C) Representative FACS analysis after 2 weeks of coculture of CB CD34$^+$ cells with MSCs or CD146$^+$ cell cocultures. After gating on CD45$^+$CD10$^-$CD19$^-$ cells, CD34$^+$33$^-$ cells were defined as CD34$^+$Lin$^-$ cells (black box). (D) The absolute number of CD34$^+$Lin$^-$ cells was significantly higher in CD146$^+$ cell cocultures, compared with MSC cocultures, at any time of culture (n=at least 5 independent experiments for each time point, each experiment was performed in triplicate; $^{**}P < .01$, $^{***}P < .001$). (E-F) Coculture of CB CD34$^+$ cells with MSCs led to a significantly higher frequency of CD14$^+$ myeloid cells after 2 weeks (E) (40.24% ± 2.723% vs 26.67% ± 2.075%. n = 10 independent experiments, each experiment was performed in triplicate; $^{***}P < .0001$) and a higher frequency of CD10$^+$/CD19$^+$ lymphoid progenitors or mature cells after 4 weeks of coculture (F) (5.155% ± 1.918% vs 0.9541% ± 0.2564%, n = 8 independent experiments, each experiment was performed in triplicate; $^{*}P < .05$). No difference in the absolute numbers of myeloid and lymphoid cells was observed between CD146$^+$ cell and MSC cocultures. All data are presented as mean ± SEM.

commitment in vitro through cell-to-cell interaction and partially through Notch activation.

Discussion

Blood formation in vertebrates is an opportunistic phenomenon that does not take place exclusively in specialized, hematopoiesis-restricted sites such as the bone marrow, thymus, spleen and avian bursa of Fabricius. Blood cells are also produced transiently in organs assuming other functions, such as the yolk sac, placenta, allantois and embryonic aorta-gonad-mesonephros, and liver. Moreover, extramedullary hematopoiesis can be resumed in pathologic conditions of the adult. Such anatomic diversity in blood-forming ability implies that developmentally and structurally distinct cellular environments can sustain hematopoiesis. Different blood-forming tissues may therefore share stromal cell subsets involved in blood formation. Although HSPCs have been characterized in detail and purified to homogeneity, the identity and function of the stromal cells involved in hematopoiesis have remained largely unknown. Although stroma-dependent hematopoiesis has been recapitulated in vitro for more than 3 decades using primary stromal cells or stromal cell lines,[9-11] the nature of the stromal cells involved has been elusive. As an obstacle to characterization, native stromal cells involved in supporting hematopoiesis are infrequent: Wineman et al[41] found that only a rare subpopulation of clonal fetal liver stromal cells is able to maintain HSPCs.

MSCs are cultured, multipotent adherent cells that can support hematopoiesis.[19-24] We hypothesized that MSCs contain distinct subsets of cells with different roles in the regulation of HSPCs. Based on recent descriptions of (1) a key contribution of murine perivascular cells to the medullary hematopoietic "niche,"[35-37] and (2) a pericyte ancestry for human MSCs,[29] we directly addressed whether cultured human perivascular cells can sustain human HSPCs. Conventionally derived, heterogeneous MSCs and CD146$^+$CD34$^-$CD45$^-$ perivascular cells can be obtained from virtually all human vascularized tissues.[29] In the present study, we derived MSCs and CD146$^+$ perivascular cells from FBM and human adipose tissue, which is commonly used as a convenient

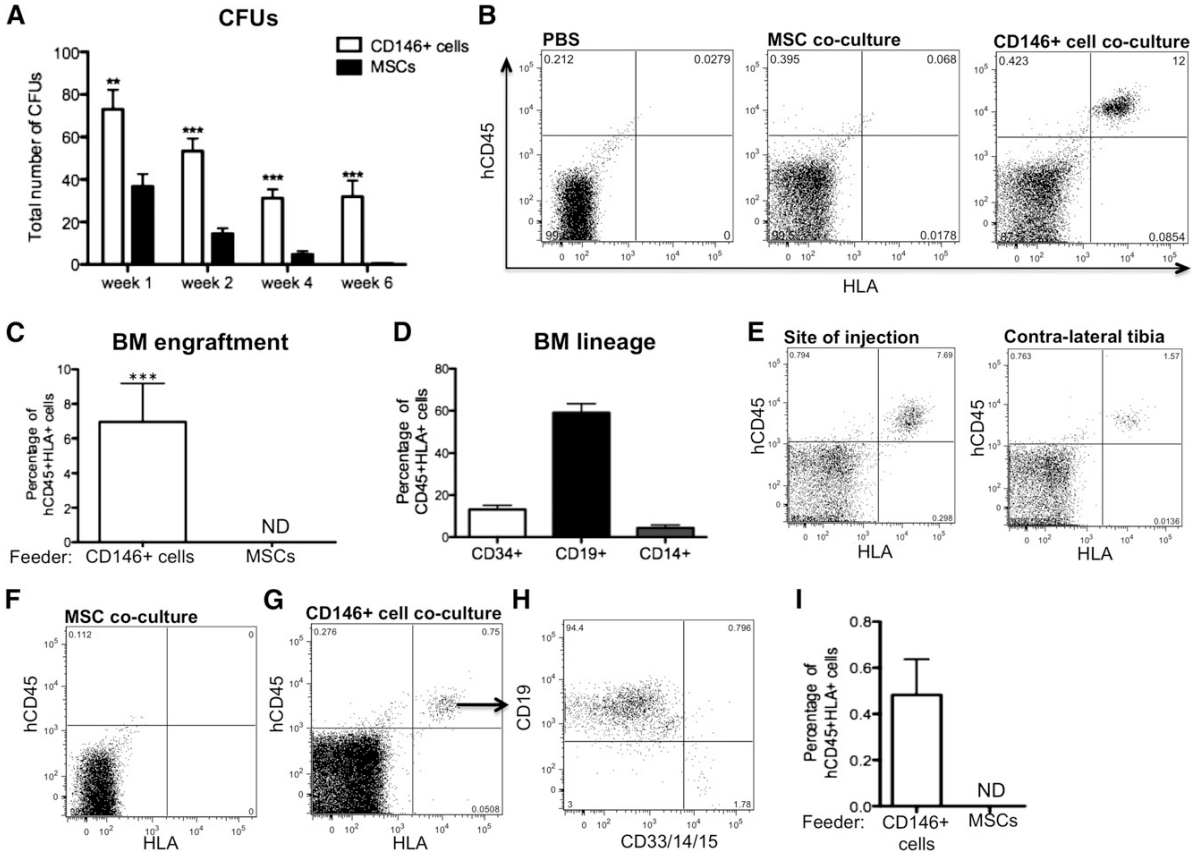


Figure 5. CD146$^+$ perivascular cells but not MSCs sustain functional HSPCs with engraftment potential and self-renewal ability. (A) CFU assay revealed significantly higher number of CFUs in CD146$^+$ cell cocultures after 1, 2, 4, and 6 weeks of coculture as compared with MSC cocultures (n = 3 independent experiments, each experiment was performed in triplicate; $^*P < .05$, $^{***}P < .001$). (B) Representative flow cytometry analysis for the detection of human CD45$^+$HLA$^+$ cells in bone marrow of NSG mice 6 weeks posttransplantation with phosphate-buffered saline, or with the same number of CD45$^+$ cells (10^5) harvested after 2 weeks of CB CD34$^+$ cell coculture with MSCs or CD146$^+$ cells. (C) All mice injected with CD45$^+$ cells obtained from CD146$^+$ cell cocultures showed human engraftment whereas no engraftment was ever detected (ND) in mice that received MSC cocultures (n = 3 independent experiments, n = 11 mice per group; $^{***}P < .0001$). (D) Frequency of CD34$^+$ progenitors, CD19$^+$ lymphoid, and CD14$^+$ myeloid cells within the CD45$^+$HLA$^+$ population of cells in the bone marrow of chimeric mice. (E) Human CD45$^+$HLA$^+$ hematopoietic cells were also detected 6 weeks posttransplantation in the contralateral tibia of mice injected with HSPCs cocultured with CD146$^+$ perivascular cells. (F-I) Representative flow cytometry analysis of secondary host bone marrow. (F) Bone marrow from primary hosts transplanted with MSC coculture was injected in secondary hosts as a negative control. (G) Human engraftment was observed 4 weeks after secondary transplantation of bone marrow from chimeric mice transplanted with CD146$^+$ cell coculture (n = 3 engrafted mice of 4). (H) Both CD19$^+$ lymphoid and CD33/CD14/CD15$^+$ myeloid cells were detectable within the human CD45$^+$ engrafted hematopoietic cells in secondary hosts. (I) Quantification of the level of chimerism in secondary mice. All data are presented as mean ± SEM.

and abundant source of MSCs. Interestingly, the sustained presence of hematopoietic cells within adipose tissue has been recently reported.[42,43]

CD146$^+$ perivascular cells expressing nestin, CXCL12, and Lep-R were found in situ in the hematopoietic FBM as well as in adipose tissue. Sorted CD146$^+$ perivascular cells homogeneously expressed in culture CD146 and higher levels of nestin, CXCL12, and Jagged-1 compared with unfractionated MSCs or to CD146$^-$ cells. CD146$^+$ perivascular cells therefore appear to represent the human counterpart of the CAR cells or nestin$^+$ cells recently described in the mouse.[35-37] A similar cell population has been documented in human bone marrow, where CD146$^+$ perivascular cells expressing CXCL12 and Jagged-1 can clonally recapitulate an ectopic hematopoietic microenvironment when implanted into mice.[44] Human bone marrow reticular stromal cells, including CD146$^+$nestin$^+$VCAM$^+$ cells, regulate HSPC homing through the secretion of CXCL12.[45] Pericyte-like cells from the human placenta have been also suggested to support hematopoietic cells in culture.[46] However, direct evidence for the ability of prospectively purified human perivascular cells to sustain primitive hematopoietic cells in long-term culture has not been provided. Several studies have investigated the ability of MSCs to maintain HSPCs in coculture systems, but these have routinely used cytokine supplementation either by direct addition or through transgene expression in MSCs.[19-24,47,48] In most cases the decisive assays, primary and secondary transplantations of cocultured hematopoietic cells into immunodeficient mice, have not been used to document the maintenance of primitive self-renewing stem cells. Most importantly, the identity of the specific subset of MSCs directly involved in the interaction with HSPCs is still unknown. In the present study, culture of CD34$^+$ cells with MSCs or CD146$^+$ perivascular cells without the addition of exogenous cytokines allowed us to define the intrinsic properties of these stromal populations in terms of hematopoietic cell support. Remarkably, unfractionated MSCs and purified CD146$^+$ perivascular cells derived from the same specimen exhibited profound differences in the ability to sustain HSPCs. Both stromal cell populations improved the survival of hematopoietic cells compared with stroma-free, cytokine-free cultures.

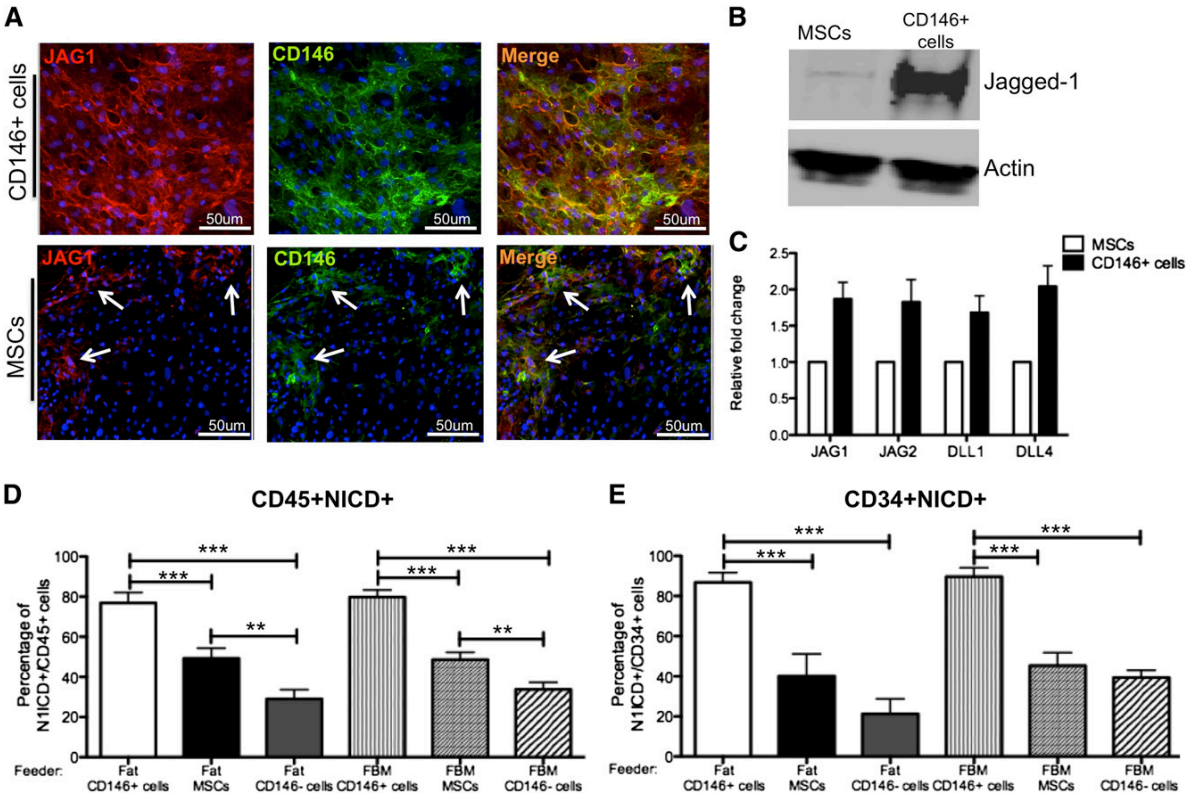


Figure 6. CD146$^+$ perivascular cells induce Notch activation in hematopoietic cells. (A) Immunocytochemical staining for Jagged-1 (JAG1, red), CD146 (green) and nuclei (DAPI, blue) performed on fat-derived CD146$^+$ perivascular cells and MSCs (original magnification, ×20). White arrows indicate clusters of cells within the MSC that coexpress JAG1 and CD146. (B) Western blot analysis showing significantly higher expression of Jagged-1 in CD146$^+$ perivascular cells compared with MSCs derived from fat. (C) qPCR analysis revealed that fat-derived MSCs and CD146$^+$ perivascular cells express multiple Notch ligands. (D-E) Quantification of (D) CD45$^+$ and (E) CD34$^+$ hematopoietic and progenitor cells with activated Notch pathway (CD45$^+$NICD$^+$) after 1 week of coculture with fat or FBM-derived CD146$^+$ perivascular cells, MSCs, and CD146$^-$ cells (n = 3 independent experiments, n = 40 random fields analyzed; ***$P < .0001$). Data are presented as mean ± SEM.

However, the total number of recovered HSPCs was consistently and significantly higher in CD146$^+$ cell cocultures. Furthermore, only CD146$^+$ perivascular cells sustained primitive HSPCs able to establish multilineage hematopoiesis in immunodeficient mice. Conversely, MSCs promoted rapid HSPC differentiation with consequent loss of engraftment ability. Our results also demonstrate that cell-to-cell contact between HSPCs and CD146$^+$ perivascular cells plays a key role in HSPCs maintenance in vitro. We show that CD146$^+$ cells express Notch ligands and that Notch inhibition in cocultures results in decreased numbers of HSPCs and increased B-cell development, as previously described.[49] Regulation of HSPCs by perivascular cells is very likely to be a multifaceted process, Notch signaling being only one of the mechanisms involved in HSPC maintenance. Although our results show an increase in PI$^+$ dead cells in the whole coculture following Notch inhibition, suggesting a role for Notch in supporting cell survival, additional studies will be required to determine whether Notch activation prevents specifically apoptosis/death in HSPCs.

De Toni et al[50] recently described the clinical-grade expansion of adipose MSCs able to support hematopoietic reconstitution in immunodeficient mice when co-injected with fresh CB CD34$^+$ cells. The authors showed higher frequency of human CD45$^+$ cells 3 weeks posttransplantation in mice that received CD34$^+$ cells and MSCs as compared with mice injected with CD34$^+$ cells alone. This difference was no longer observed 11 weeks posttransplantation, suggesting that fat MSCs support short-term progenitors. The lack of secondary transplantation assay did not allow the authors to establish whether fat MSCs can support maintenance of HSPC self-renewal. In our study, we demonstrate that unfractionated MSCs failed to ex vivo support HSPCs with reconstituting ability even in primary recipients. Conversely, we demonstrate for the first time that the CD146$^+$ subset of MSCs was uniquely able to maintain self-renewing HSPCs after 2 weeks in culture without growth factors, as shown by reconstitution of primary and secondary hosts. The identification of CD146$^+$ perivascular cells as a defined and specific subset of human stromal cells able to sustain HSPCs may therefore have a critical impact on future clinical applications based on ex vivo expansion and genetic manipulation of HSPCs. Butler et al[14] recently described a cellular platform for the expansion of human HSPCs based on the coculture with transformed endothelial cells in the presence of defined growth factors. In the present work, cocultures were performed using nontransformed stromal cells in the complete absence of added growth factors. Having identified CD146$^+$ perivascular cells as the specific subset of MSCs involved in HSPC maintenance, further studies aimed to define optimal conditions to promote HSPC expansion will be needed.

For the first time to our knowledge, we document the direct role of an anatomically and phenotypically defined subset of human stromal cells—the CD146$^+$ perivascular cells—in maintaining cultured HSPCs. A fraction of native and all cultured pericytes

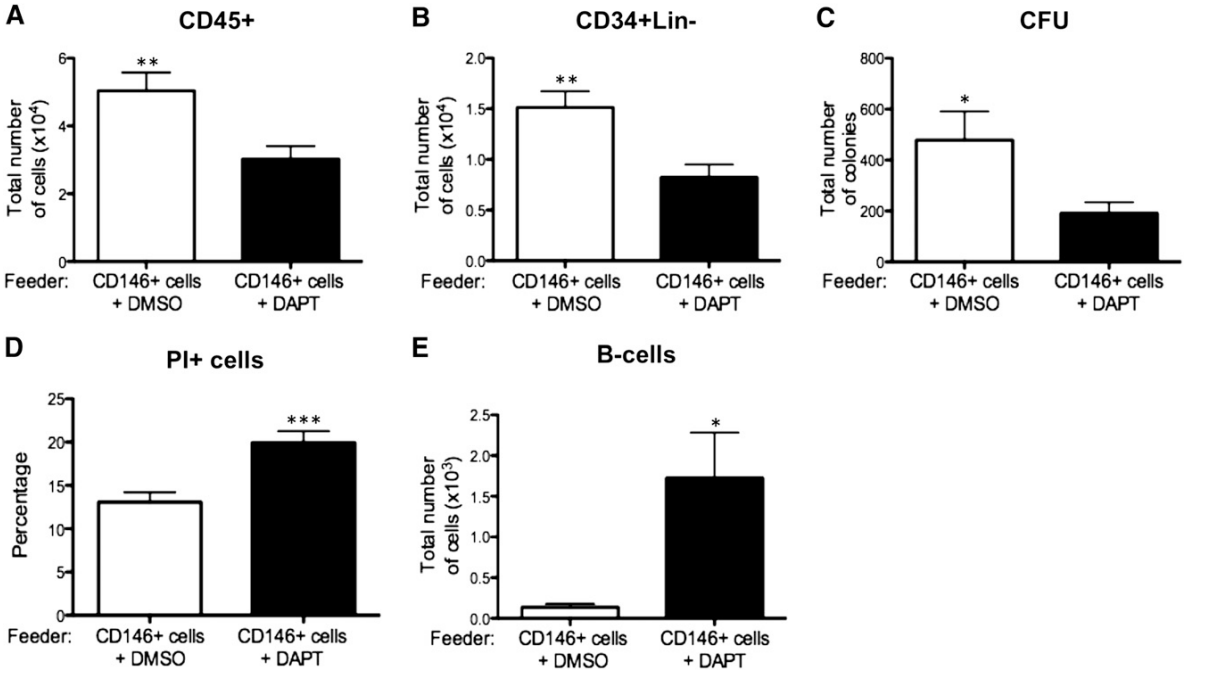


Figure 7. Notch inhibition affects survival and B-cell differentiation of HSPCs. Inhibition of Notch was achieved by addition of 10μM DAPT to CD146$^+$ perivascular cells and CB CD34$^+$ cell coculture every other day. Vehicle (DMSO) was added to control cocultures. (A-B) Total number of CD45$^+$ cells and CD34$^+$Lin$^-$ cells was significantly reduced after 2 weeks of coculture with DAPT (5.03 ± 0.54 × 10^4 vs 3.02 ± 0.37 × 10^4 CD45$^+$ cells, n = 4 independent experiments, each experiment was performed in triplicate, $^{**}P < .01$; 1.5 ± 0.16 × 10^4 vs 0.82 ± 0.12 × 10^4 CD34$^+$Lin$^-$ cells, n = 4 independent experiments, each experiment was performed in triplicate, $^{**}P < .01$). (C) Similarly, the total number of CFUs was significantly reduced after 4 weeks of coculture with DAPT (478.3 ± 112.4 vs 191.0 ± 43.28, n = 3 independent experiments, each experiment was performed in triplicate; $^{*}P < .05$). (D) Flow cytometry viability analysis revealed a significantly higher frequency of PI$^+$ dead cells in coculture performed in the presence of DAPT (13.08% ± 1.13% vs 19.94 ± 1.31, n = 4 independent experiments, each experiment was performed in triplicate, $^{***}P < .0001$). (E) Notch inhibition also significantly increased B-cell development (0.13 ± 0.04 × 10^3 vs 1.72 ± 0.55 × 10^3 of lymphoid cells, n = 3 individual experiments, each experiment performed in triplicate; $^{**}P < .01$). Data are presented as mean ± SEM.

express α-SMA,[29] therefore these findings also support a myofibroblastic identity for human hematopoietic stromal cells.[51] Besides a functional ability to support hematopoietic cells following dissociation, purification, and culture, human CD146$^+$ perivascular cells from nonhematopoietic tissues share a similar phenotype with the perivascular niche cells recently described in murine bone marrow.[35-37] Perivascular cells are ubiquitous[52] and may therefore represent the key stem cell support shared by all blood-forming organs. It remains to be determined whether and how this ability to sustain HSCs is repressed in situ in nonhematopoietic tissues, and may be reactivated in pathologic conditions, as in the course of extramedullary hematopoiesis or leukemic dissemination.

Acknowledgments

The authors thank Yuhua Zhu and Rebecca Chan (Department of Pathology, UCLA) for their valuable technical assistance, Dr Ling Wu (Orthopaedic Hospital Department of Orthopaedic Surgery and the Orthopaedic Hospital Research Center) for the assistance with immunohistochemistry on FBM, and Dr David Stoker (Marina Plastic Surgery Associates, Marina del Rey, CA) for the procurement of human tissues. The authors are also grateful for the remarkable flow cytometry and sorting assistance provided by Jessica Scholes and Felicia Codrea (Eli and Edythe Broad Stem Cell Center, UCLA).

This work was supported by funds from UCLA, Orthopaedic Hospital Department of Orthopaedic Surgery, and Eli and Edythe Broad Stem Cell Center at UCLA. This research was also made possible by a grant from the California Institute for Regenerative Medicine (grant no. RB3-05217) (G.M.C.). M.C. and C.P. acknowledge the support of a California Institute for Regenerative Medicine training grant (TG2-01169). The research was also partially founded by the European Community FP7 program, through the Reborne project (grant agreement no. 241879).

Authorship

Contribution: M.C. designed and performed research, analyzed and interpreted data, and wrote the manuscript; C.P., A.S., W.W., S.G., D.E., and C.J.C. performed research; X.W. performed statistical analysis; E.M. and L.L. analyzed and interpreted data and contributed to writing the manuscript; and G.M.C and B.P. designed research, analyzed and interpreted data, and wrote the manuscript.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

Correspondence: Bruno Péault, 615 Charles E. Young Dr South, Room 410, Los Angeles, CA 90095; bpeault@mednet.ucla.edu.

References

1. Tavian M, Péault B. The changing cellular environments of hematopoiesis in human development in utero. *Exp Hematol.* 2005;33(9): 1062-1069.
2. Sabin FR. Preliminary note on the differentiation of angioblasts and the method by which they produce blood-vessels, blood-plasma and red blood-cells as seen in the living chick. 1917. *J Hematother Stem Cell Res.* 2002;11(1):5-7.
3. Oberlin E, Tavian M, Blazsek I, et al. Blood-forming potential of vascular endothelium in the human embryo. *Development.* 2002;129(17): 4147-4157.
4. Kissa K, Herbomel P. Blood stem cells emerge from aortic endothelium by a novel type of cell transition. *Nature.* 2010;464(7285): 112-115.
5. Bertrand JY, Chi NC, Santoso B, et al. Haematopoietic stem cells derive directly from aortic endothelium during development. *Nature.* 2010;464(7285):108-111.
6. Boisset JC, van Cappellen W, Andrieu-Soler C, et al. In vivo imaging of haematopoietic cells emerging from the mouse aortic endothelium. *Nature.* 2010;464(7285):116-120.
7. Zovein AC, Hofmann JJ, Lynch M, et al. Fate tracing reveals the endothelial origin of hematopoietic stem cells. *Cell Stem Cell.* 2008; 3(6):625-636.
8. Zambidis ET, Park TS, Yu W, et al. Expression of angiotensin-converting enzyme (CD143) identifies and regulates primitive hemangioblasts derived from human pluripotent stem cells. *Blood.* 2008; 112(9):3601-3614.
9. Dexter TM, Wright EG, Krizsa F, et al. Regulation of haemopoietic stem cell proliferation in long term bone marrow cultures. *Biomedicine.* 1977; 27(9-10):344-349.
10. Whitlock CA, Witte ON. Long-term culture of B lymphocytes and their precursors from murine bone marrow. *Proc Natl Acad Sci U S A.* 1982; 79(11):3608-3612.
11. Collins LS, Dorshkind K. A stromal cell line from myeloid long-term bone marrow cultures can support myelopoiesis and B lymphopoiesis. *J Immunol.* 1987;138(4):1082-1087.
12. Butler JM, Nolan DJ, Vertes EL, et al. Endothelial cells are essential for the self-renewal and repopulation of Notch-dependent hematopoietic stem cells. *Cell Stem Cell.* 2010;6(3):251-264.
13. Kobayashi H, Butler JM, O'Donnell R, et al. Angiocrine factors from Akt-activated endothelial cells balance self-renewal and differentiation of haematopoietic stem cells. *Nat Cell Biol.* 2010; 12(11):1046-1056.
14. Butler JM, Gars EJ, James DJ, et al. Development of a vascular niche platform for expansion of repopulating human cord blood stem and progenitor cells. *Blood.* 2012;120(6):1344-1347.
15. Pittenger MF, Mackay AM, Beck SC, et al. Multilineage potential of adult human mesenchymal stem cells. *Science.* 1999; 284(5411):143-147.
16. Friedenstein AJ, Chailakhjan RK, Lalykina KS. The development of fibroblast colonies in monolayer cultures of guinea-pig bone marrow and spleen cells. *Cell Tissue Kinet.* 1970;3(4): 393-403.
17. Zuk PA, Zhu M, Ashjian P, et al. Human adipose tissue is a source of multipotent stem cells. *Mol Biol Cell.* 2002;13(12):4279-4295.
18. da Silva Meirelles L, Chagastelles PC, Nardi NB. Mesenchymal stem cells reside in virtually all post-natal organs and tissues. *J Cell Sci.* 2006; 119(Pt 11):2204-2213.
19. Gan OI, Murdoch B, Larochelle A, et al. Differential maintenance of primitive human SCID-repopulating cells, clonogenic progenitors, and long-term culture-initiating cells after incubation on human bone marrow stromal cells. *Blood.* 1997;90(2):641-650.
20. Fei XM, Wu YJ, Chang Z, et al. Co-culture of cord blood CD34(+) cells with human BM mesenchymal stromal cells enhances short-term engraftment of cord blood cells in NOD/SCID mice. *Cytotherapy.* 2007;9(4):338-347.
21. Huang GP, Pan ZJ, Jia BB, et al. Ex vivo expansion and transplantation of hematopoietic stem/progenitor cells supported by mesenchymal stem cells from human umbilical cord blood. *Cell Transplant.* 2007;16(6):579-585.
22. Corre J, Barreau C, Cousin B, et al. Human subcutaneous adipose cells support complete differentiation but not self-renewal of hematopoietic progenitors. *J Cell Physiol.* 2006; 208(2):282-288.
23. Flores-Guzmán P, Flores-Figueroa E, Montesinos JJ, et al. Individual and combined effects of mesenchymal stromal cells and recombinant stimulatory cytokines on the in vitro growth of primitive hematopoietic cells from human umbilical cord blood. *Cytotherapy.* 2009;11(7): 886-896.
24. Wagner W, Roderburg C, Wein F, et al. Molecular and secretory profiles of human mesenchymal stromal cells and their abilities to maintain primitive hematopoietic progenitors. *Stem Cells.* 2007;25(10):2638-2647.
25. da Silva Meirelles L, Caplan AI, Nardi NB. In search of the in vivo identity of mesenchymal stem cells. *Stem Cells.* 2008;26(9):2287-2299.
26. Shi S, Gronthos S. Perivascular niche of postnatal mesenchymal stem cells in human bone marrow and dental pulp. *J Bone Miner Res.* 2003;18(4): 696-704.
27. Zannettino AC, Paton S, Arthur A, et al. Multipotential human adipose-derived stromal stem cells exhibit a perivascular phenotype in vitro and in vivo. *J Cell Physiol.* 2008;214(2):413-421.
28. Covas DT, Panepucci RA, Fontes AM, et al. Multipotent mesenchymal stromal cells obtained from diverse human tissues share functional properties and gene-expression profile with CD146+ perivascular cells and fibroblasts. *Exp Hematol.* 2008;36(5):642-654.
29. Crisan M, Yap S, Casteilla L, et al. A perivascular origin for mesenchymal stem cells in multiple human organs. *Cell Stem Cell.* 2008;3(3): 301-313.
30. Calvi LM, Adams GB, Weibrecht KW, et al. Osteoblastic cells regulate the haematopoietic stem cell niche. *Nature.* 2003;425(6960):841-846.
31. Kiel MJ, Yilmaz OH, Iwashita T, et al. SLAM family receptors distinguish hematopoietic stem and progenitor cells and reveal endothelial niches for stem cells. *Cell.* 2005;121(7):1109-1121.
32. Levesque JPN. N(o)-cadherin role for HSCs. *Blood.* 2012;120(2):237-238.
33. Ehninger A, Trumpp A. The bone marrow stem cell niche grows up: mesenchymal stem cells and macrophages move in. *J Exp Med.* 2011;208(3): 421-428.
34. Kincade PW. Plasticity of supporting cells in a stem cell factory. *Immunity.* 2010;33(3): 291-293.
35. Sugiyama T, Kohara H, Noda M, et al. Maintenance of the hematopoietic stem cell pool by CXCL12-CXCR4 chemokine signaling in bone marrow stromal cell niches. *Immunity.* 2006;25(6): 977-988.
36. Méndez-Ferrer S, Michurina TV, Ferraro F, et al. Mesenchymal and haematopoietic stem cells form a unique bone marrow niche. *Nature.* 2010; 466(7308):829-834.
37. Ding L, Saunders TL, Enikolopov G, et al. Endothelial and perivascular cells maintain haematopoietic stem cells. *Nature.* 2012; 481(7382):457-462.
38. Corselli M, Chen CW, Sun B, et al. The tunica adventitia of human arteries and veins as a source of mesenchymal stem cells. *Stem Cells Dev.* 2012;21(8):1299-1308.
39. Poglio S, De Toni F, Lewandowski D, et al. In situ production of innate immune cells in murine white adipose tissue. *Blood.* 2012;120(25):4952-4962.
40. Bigas A, Robert-Moreno A, Espinosa L. The Notch pathway in the developing hematopoietic system. *Int J Dev Biol.* 2010;54(6-7):1175-1188.
41. Wineman J, Moore K, Lemischka I, et al. Functional heterogeneity of the hematopoietic microenvironment: rare stromal elements maintain long-term repopulating stem cells. *Blood.* 1996;87(10):4082-4090.
42. Cousin B, André M, Arnaud E, et al. Reconstitution of lethally irradiated mice by cells isolated from adipose tissue. *Biochem Biophys Res Commun.* 2003;301(4):1016-1022.
43. Han J, Koh YJ, Moon HR, et al. Adipose tissue is an extramedullary reservoir for functional hematopoietic stem and progenitor cells. *Blood.* 2010;115(5):957-964.
44. Sacchetti B, Funari A, Michienzi S, et al. Self-renewing osteoprogenitors in bone marrow sinusoids can organize a hematopoietic microenvironment. *Cell.* 2007;131(2):324-336.
45. Schajnovitz A, Itkin T, D'Uva G, et al. CXCL12 secretion by bone marrow stromal cells is dependent on cell contact and mediated by connexin-43 and connexin-45 gap junctions. *Nat Immunol.* 2011;12(5):391-398.
46. Robin C, Bollerot K, Mendes S, et al. Human placenta is a potent hematopoietic niche containing hematopoietic stem and progenitor cells throughout development. *Cell Stem Cell.* 2009;5(4):385-395.
47. Khoury M, Drake A, Chen Q, et al. Mesenchymal stem cells secreting angiopoietin-like-5 support efficient expansion of human hematopoietic stem cells without compromising their repopulating potential. *Stem Cells Dev.* 2011;20(8):1371-1381.
48. Xie C, Jia B, Xiang Y, et al. Support of hMSCs transduced with TPO/FL genes to expansion of umbilical cord CD34+ cells in indirect co-culture. *Cell Tissue Res.* 2006;326(1):101-110.
49. Nie L, Perry SS, Zhao Y, et al. Regulation of lymphocyte development by cell-type-specific interpretation of Notch signals. *Mol Cell Biol.* 2008;28(6):2078-2090.
50. De Toni F, Poglio S, Youcef AB, et al. Human adipose-derived stromal cells efficiently support hematopoiesis in vitro and in vivo: a key step for therapeutic studies. *Stem Cells Dev.* 2011;20(12): 2127-2138.
51. Galmiche MC, Koteliansky VE, Brière J, et al. Stromal cells from human long-term marrow cultures are mesenchymal cells that differentiate following a vascular smooth muscle differentiation pathway. *Blood.* 1993;82(1):66-76.
52. Andreeva ER, Pugach IM, Gordon D, et al. Continuous subendothelial network formed by pericyte-like cells in human vascular bed. *Tissue Cell.* 1998;30(1):127-135.

Corselli et al. Supplemental Data

Supplemental Figure Legends

Supplemental Figure S1. CD146+ perivascular cells share an immunophenotypical profile and differentiation potential typical of MSC. a-b) Flow cytometry analysis of cultured MSC and cultured CD146+ perivascular cells isolated from adipose tissue. Both cell types express stromal markers but not endothelial (CD31) or hematopoietic (CD45) markers. (Red histogram= unstained (no Antibody) control. Blue histogram= antibody as shown on x axis) (representative of 3 independent experiments). c-f) Both MSC and CD146+ perivascular cells were able to differentiate into osteoblasts and adipocytes as shown by Alizarin red (c,e) and Oil red O staining (d,f), respectively (magnification 10x).

Supplemental Figure S2. CD146+ perivascular cells from fat and fetal bone marrow (FBM) provide superior support of HSPCs *ex vivo* than MSCs and CD146- cells: a,b) The absolute number of CD34+Lin- cells and CFUs after two and four weeks of co-culture, respectively, was significantly higher in the presence of fat-derived CD146+ perivascular cells compared to MSCs and CD146- cells. c,d) Similar differences in the total number of CD34+Lin- cells and CFUs were observed when cord blood-derived CD34+ cells were co-cultured with CD146+ perivascular cells , MSCs, and CD146- cells derived from FBM. All data are presented as mean+/- SEM. *** $p<0.0001$, * $p<0.05$.

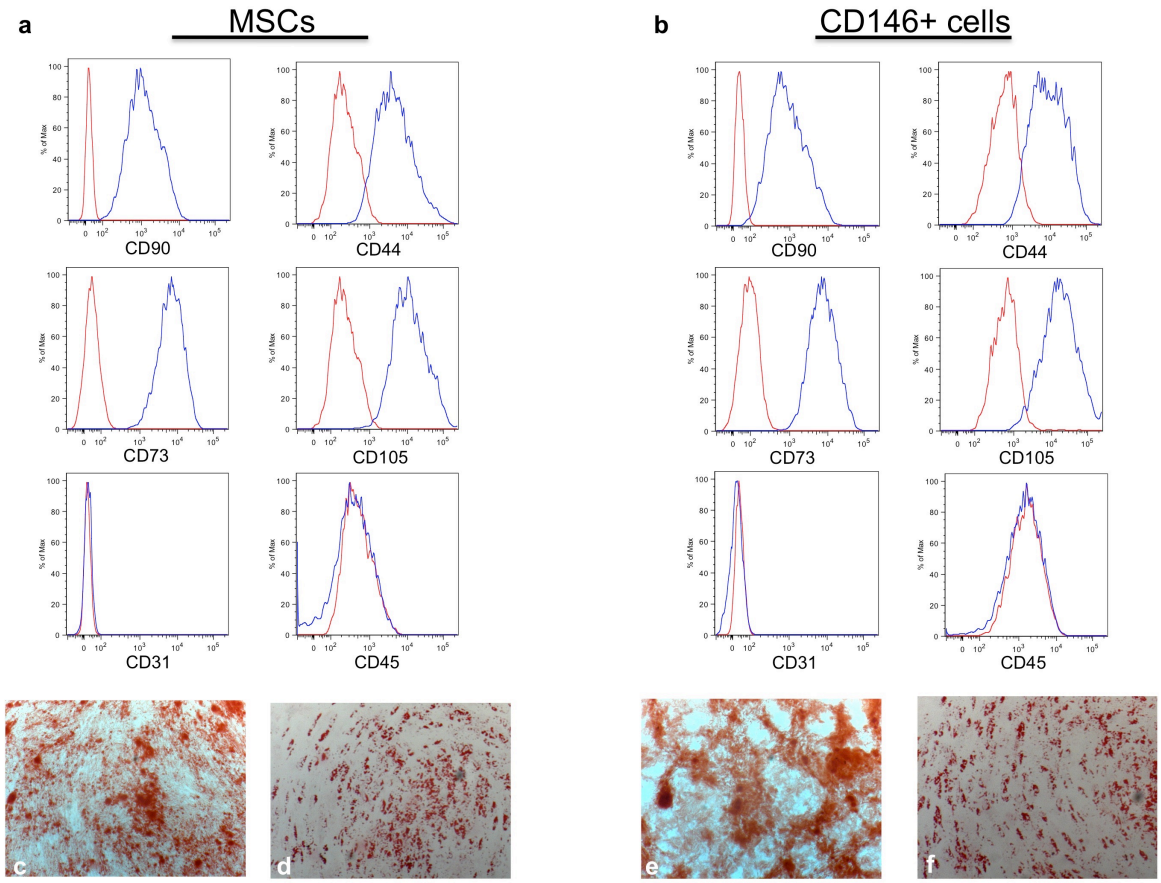
Supplemental Figure S3. Cell-to-cell contact is required for the maintenance of hematopoietic cells on CD146+ stroma: a-d) Phase contrast (magnification 10x) and immunofluorescence (magnification 20x) images showing different morphology and distribution of CB CD34+ cells co-cultured with either CD146+ perivascular cells or

MSCs. (a) Red-dotted and (b) white-dotted circles highlight the formation of hematopoietic clusters observed exclusively in CD146+ cell co-cultures. c) Transwell plates were used to prevent cell-to-cell contact between CD146+ perivascular cells and hematopoietic cells. In this condition, the total number of CD45+ hematopoietic cells was significantly decreased after 1 week of co-culture ($4.93 \pm 0.5779 \times 10^4$ vs. $0.91 \pm 0.12 \times 10^4$. n=5 independent experiments, each experiment was performed in triplicate. Contact= CD146+ perivascular cells and CB CD34+ both plated on the bottom well. No Contact= CD146+ perivascular cells plated on the bottom well, CB CD34+ cells plated on the top well. Data presented as mean +/-SEM. *** $p<0.0001$

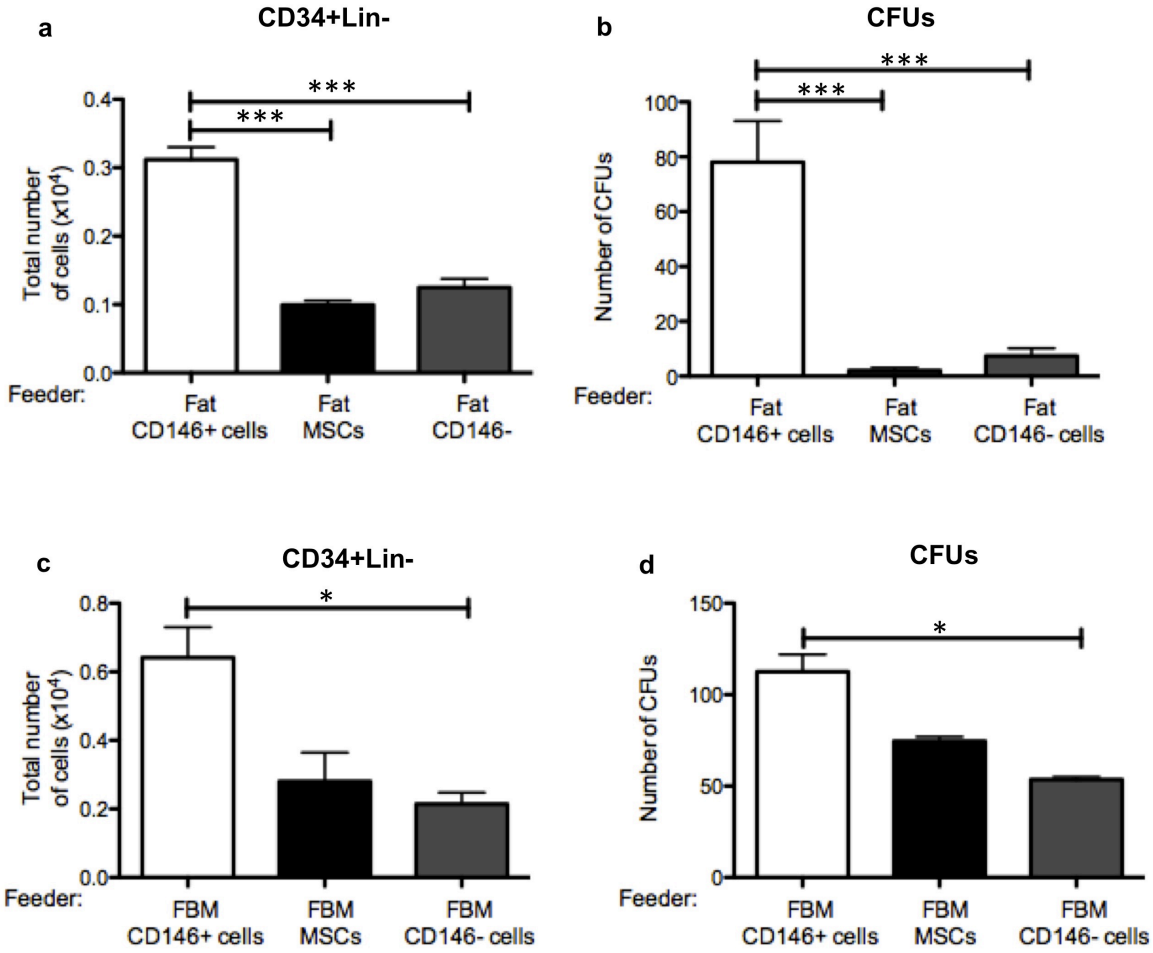
Supplemental Figure S4. Detection of Notch1 intracellular domain (NICD) in hematopoietic cell cultures: a,b) Detection of NICD+CD45+ hematopoietic cells in CD146+ cell and MSC co-cultures in chamber slides. c) N1ICD was detected when direct contact between CD146+ perivascular cells and hematopoietic cells was allowed but d) not when direct contact between CD146+ cells and HSPCs was inhibited in transwell co-cultures (no contact). e-f) Detection of NICD+CD45+ hematopoietic cells in CD146+ cell and MSC co-cultures in cytospin slides. NICD in HSPCs co-cultured with CD146+ and after treatment with f) DAPT or g) anti-Notch1 blocking antibody. Treatment with DAPT resulted in a complete inhibition of Notch activation. NICD+CD45+ cells were still detected after treatment with anti-Notch1 blocking antibody, albeit at a lower frequency compared to controls (magnification 20x)

Supplemental Figure S5. Evaluation of DAPT cytotoxicity and specificity: a) CD146+ perivascular cells and CB CD34+ cells were cultured separately in medium with DMSO or DAPT in the absence of added growth factors. Retronectin coated plates were

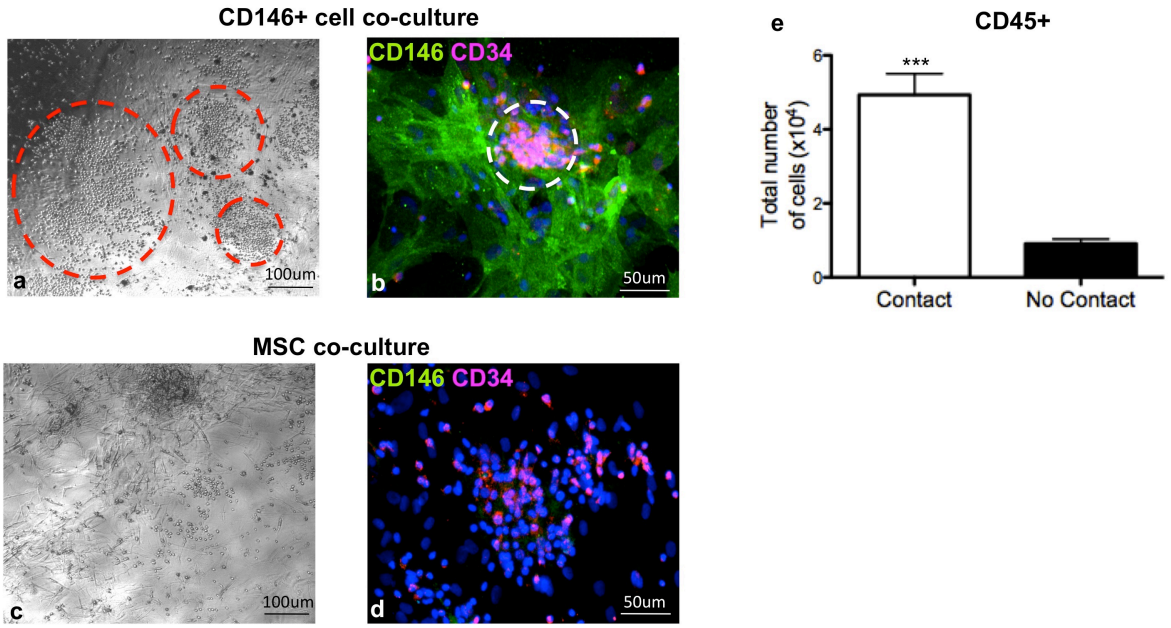
used for the culture of CB CD34+ cells. Neither CD146+ perivascular cells nor CB CD34+ cells showed increase in the percentage of PI+ (dead) cells after DAPT treatment (n=4 independent experiments, each experiment was performed in triplicate, ns= not statistically significant). b-d) Addition of a specific anti-Notch1 neutralizing antibody to HSPC/CD146+ cell co-cultures produced a similar effect in terms of (b) number of CD45+ (c) number of CD34+Lin- cells and (d) increase in B-cell development compared to DAPT treated co-cultures. All data are presented as mean+/- SEM. * $p<0.05$, ** $p<0.001$.



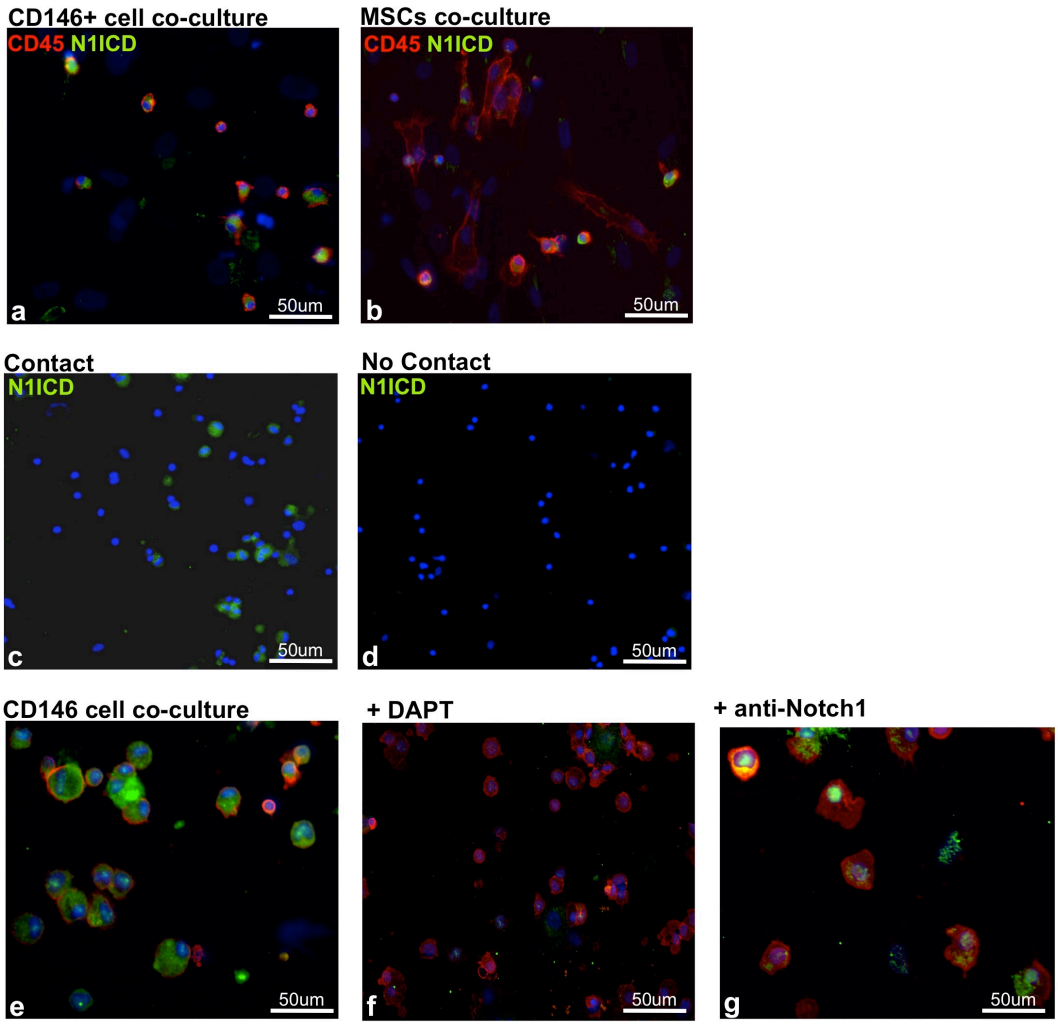
Supplemental Figure S1



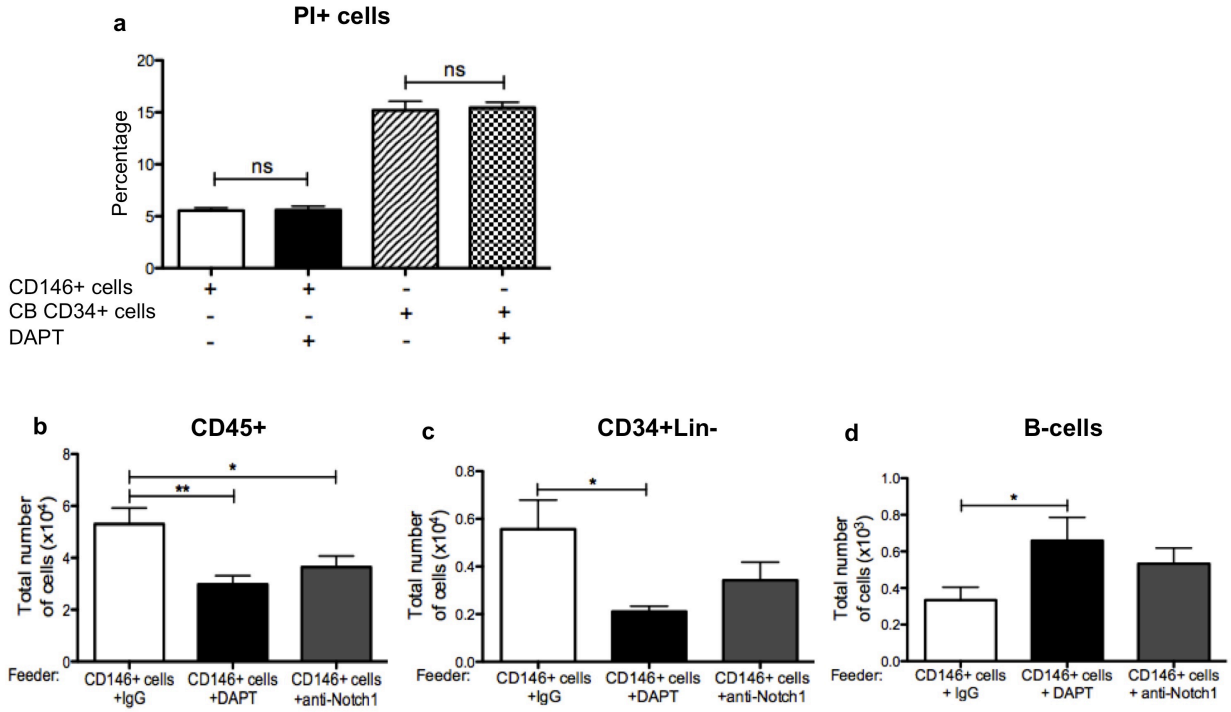
Supplemental Figure S2



Supplemental Figure S3



Supplemental Figure S4



Supplemental Figure S5

Chapter 3

Stem Cells. 2016 Mar 2. doi: 10.1002/stem.2351.

Genetic Tagging During Human Mesoderm Differentiation Reveals Tripotent Lateral Plate Mesodermal Progenitors

Chin CJ[1], Cooper AR[2], Lill GR[3], Evseenko D[4], Zhu Y[1], He CB[1], Casero D[5], Pellegrini M[6], Kohn DB[7], Crooks GM[8].

* Chin and Cooper share first authorship

[1]Department of Pathology & Laboratory Medicine, David Geffen School of Medicine (DGSOM), UCLA.

[2]Molecular Biology Interdepartmental PhD Program, UCLA.

[3]Department of Microbiology, Immunology and Molecular Genetics, DGSOM, UCLA.

[4]Department of Orthopedic Surgery, Keck School of Medicine of USC.

[5]Department of Pathology & Laboratory Medicine, DGSOM, UCLA.

[6]Department of Molecular, Cell, and Developmental Biology, Molecular Biology Institute, UCLA.

[7]Department of Microbiology, Immunology and Molecular Genetics, and Pediatrics, DGSOM, BSCRC, Jonsson Comprehensive Cancer Center (JCCC), Molecular Biology Institute, UCLA.

[8]Department of Pathology & Laboratory Medicine, and Pediatrics, DGSOM, BSCRC, JCCC, UCLA.

ABSTRACT

Although clonal studies of lineage potential have been extensively applied to organ specific stem and progenitor cells, much less is known about the clonal origins of lineages formed from the germ layers in early embryogenesis. We applied lentiviral tagging followed by vector integration site analysis (VISA) with high-throughput sequencing to investigate the ontogeny of the hematopoietic, endothelial and mesenchymal lineages as they emerge from human embryonic mesoderm. In contrast to studies that have used VISA to track differentiation of self-renewing

stem cell clones that amplify significantly over time, we focused on a population of progenitor clones with limited self-renewal capability. Our analyses uncovered the critical influence of sampling on the interpretation of lentiviral tag sharing, particularly among complex populations with minimal clonal duplication. By applying a quantitative framework to estimate the degree of undersampling we revealed the existence of tripotent mesodermal progenitors derived from pluripotent stem cells, and the subsequent bifurcation of their differentiation into bipotent endothelial/hematopoietic or endothelial/mesenchymal progenitors.

INTRODUCTION

During the earliest stages of embryogenesis, a morphologic region called the primitive streak (PS) contains cells committed to form either mesoderm or definitive endoderm. Fate-mapping experiments in vertebrates show that mesoderm patterning in the PS strictly correlates with the place and time of mesoderm induction, specifying posterior PS (extraembryonic mesoderm, lateral plate mesoderm), anterior PS (cardiac mesoderm, definitive endoderm, axial mesoderm) and late PS (presomitic mesoderm) [1]. This dramatic period of morphogenesis, and the dynamic transcriptional and signaling events that shape mesoderm patterning, suggest that the PS is a rapidly differentiating population in which self-renewal may be limited or non-existent.

Given the inaccessibility of early human embryonic tissues, modeling with human pluripotent stem cells (hPSC) has become an essential tool for studying the complex cellular events of human germ layer commitment [2]. An early PS-like population has been identified during hPSC differentiation using transcriptional reporters [3] and by expression of cell surface markers of epithelial to mesenchymal transition [4]. Recently, cardiac and paraxial mesoderm, subtypes

derived from the anterior PS and late PS respectively, were shown to be specified through distinct *BRACHYURY*$^{+}$ mesoderm progenitors during exit from pluripotency [1].

Transcriptome profiling of single cells often reveal a surprising level of heterogeneity within purified populations that contain apparently identical cells. The differential expression of lineage specific genes between individual cells is often inferred as evidence that those cells intrinsically possess different types of lineage potential. Such analyses provide a snapshot of the transcriptional status of individual mesoderm progenitors, but do not definitively prove the clonal relationship of the lineages that will be ultimately produced from each progenitor. The functional interrogation of lineage output at a clonal level is crucial to understand the process of lineage commitment, and can provide valuable insights in how to guide stepwise generation of therapeutically relevant tissues and organs from hPSC. Clonal analyses of lineage potential have been extensively applied to adult tissues and to fetal cells isolated from tissues after germ layer differentiation. Although PS-like populations can be differentiated into various lineages [3, 4], little is known about the clonal relationship of cells that initiate early embryonic development. We have previously published a differentiation and isolation strategy that captures the earliest stage of mesodermal commitment from hPSC, comprising a population that displays broad lateral plate and cardiac mesoderm potential [4]. In the current study, we utilized these human embryonic mesoderm progenitors (hEMP) as the starting population for clonal tracking of mesoderm derivatives, specifically examining the hematopoietic, endothelial and mesenchymal lineage potential of single cells by high throughput sequencing of lentiviral tags. A fundamental challenge in applying this approach to heterogeneous progenitors undergoing rapid differentiation is that clonal output must be detected within a highly complex population with limited clonal duplication. To meet this challenge, we applied a “mark-recapture” statistical

approach to estimate clonal abundance [5-7] and developed a mathematical model to accurately interpret datasets from high complexity samples that cannot be exhaustively characterized.

Through high throughput sequencing and analysis of lentiviral integration sites combined with attention to the impact of sampling and the examination of essential control populations, we demonstrated the presence of a tripotent mesodermal progenitor and uncovered the bifurcation of the hematopoietic and mesenchyme lineages early in lateral plate mesoderm commitment. We propose that the critical concept of undersampling and the mathematical approaches used here are highly relevant to other studies of transitional populations, including those that attempt to uncover the role of genetically diverse malignant subclones during disease evolution.

MATERIALS AND METHODS

Vector constructs and production

FUGW (carrying the EGFP reporter)[8] was used in all transductions except for the negative control experiment (Fig. 4D) in which mCitrine, mCerulean and mStrawberry were expressed in pCCLc-UBC-reporter-PRE-FB-2xUSE [9]. Vector production and titer were as described [10].

Mesoderm differentiation from hPSC and hEMP

The hESC line H1 (WiCell, Madison, WI) was maintained and expanded on irradiated primary mouse embryonic fibroblasts (EMD Millipore, Billerica, MA). Mesoderm commitment was induced as previously described [4]. $CD326^{-}CD56^{+}$ embryonic mesoderm progenitors were isolated by flow cytometry at day 3.5 (Fig. 1A) and co-cultured on OP9 stroma for trilineage (hematopoietic, endothelial and mesenchymal) differentiation over the next 13 days (see supplemental methods).

Transduction of hPSC, hEMP and K562 cells

A full description of transduction methods for each cell type is found in Supplemental Methods. In brief, hPSC were transduced in mTESR on matrigel for 24 hours, yielding ~50% GFP+ cells. Isolated hEMP were transduced for 12 hours, yielding 80% GFP+ cells. To limit further viral integration during hEMP differentiation, integrase inhibitor (raltegravir, Merck & Co, White House Station, NJ) was added at 1 μM at 12 hours when OP9 coculture was initiated and was supplemented throughout the 13 days of differentiation. K562 cells (American Type Culture Collection) were transduced at MOI 4 and raltegravir added at various time points between 6 h and 3d. After 3 weeks of further expansion, DNA was extracted to measure vector copy number/cell by droplet digital PCR (see supplemental methods).

Flow cytometry and cell sorting

Identification and isolation of lineages was performed using the following gating sequence: murine CD29-APC-Cy7 was used to exclude murine cells. CD45-PE-Cy7 identified the hematopoietic population ($CD45^+$). From the non-hematopoietic compartment ($CD45^-$), the CD31-APC and CD73-PE-Cy7 coexpressing cells were first gated, and CD144-PerCP-Cy5.5 positivity was then used to define the endothelial population ($CD45^-CD31^+CD73^+CD144^+$). From the non-hematopoietic, non-endothelial compartment ($CD45^-CD31^-CD144^-$), the mesenchymal population was identified based on CD73-PE-Cy7 positivity ($CD45^-CD31^-CD144^-CD73^+$). Lineage negative cells were $mCD29^-$ cells that could not be assigned to a lineage based on immunophenotype ($mCD29^-CD45^-CD31^-CD73^-$). See Supplemental Methods for details.

Lentiviral tag sequencing

Genomic DNA was isolated from cells using the PureLink Genomic DNA Mini kit (Invitrogen) or NucleoSpin Tissue XS kit (Clontech, Mountain View, CA), depending on starting cell number. 5µl of DNA (or a maximum of 75,000 lentiviral tags/reaction) was used as starting input for non-restrictive linear amplification-mediated PCR (nrLAM-PCR) [11]. Four to ten independent nrLAM-PCR reactions were performed on the genomic DNA from each population. PCR products were mixed and quantified by probe-based droplet digital qPCR and appropriate amounts were used to load Illumina v3 flow cells. See Supplemental Methods for PCR primer sequences and conditions.

Paired-end 50- or 100-bp sequencing was performed on an Illumina HiSeq 2000 using a custom read 1 primer (GAGATCTACACTGATCCCTCAGACCCTTTTAGTC). Sequence reads were required to begin with the end of the vector LTR sequence and have no more than two mismatches with the LTR sequence. To map lentiviral tags within the human genome, LTR sequences and Illumina adapter sequences were trimmed from the reads, which were then aligned to the hg19 build of the human genome with Bowtie2[12]. Alignments were condensed and annotated by a custom Python wrapper script. Conservative cutoffs were set for calling integration sites and demultiplexing (i.e. calling lentiviral tags as detected within samples marked by different indexes) to control for spurious tag sharing due to sequencing artifacts or FACS impurity. Four reads were required to call a lentiviral tag detected, and the power law function (threshold) = $0.0874 * (\text{total readcount among samples})^{0.7025}$ was used to establish a readcount threshold for calling a tag detected in a sample. The result of this processing was a set of lentiviral tags (integration sites denoted by chromosome number, strand and nucleotide position) that were detected in each sampling of DNA from the various differentiated lineages.

The numbers of sequence reads obtained for the samples analyzed are detailed in Supplementary Fig. 6

Estimation of total tag count and calculation of detection scores

From the set of lentiviral tags detected among all samplings from all lineages in an experiment, the total number of tags (the union) was taken, along with the number of tags detected in only one sample and the number of tags detected in exactly two samples. These three values were used to calculate a Chao2 lower bound on the total lentiviral tags among all lineages. This estimate was taken to reflect the number of tags generated during hEMP transduction at the beginning of the experiment that survived through the end of the experiment (see Supplementary Fig. 3). These calculated lower bounds were used in the model described in Fig. S4 to calculate the expected number of shared lentiviral tags between all possible pairs and trios of lineages, assuming that all tagged cells were multipotent and that the probabilities of a tag being detected in the lineages are independent. We model lineage commitment of a clone and detection of a tag within that clone together as a random draw of a lentiviral tag from the set of tags inferred in the transduced hEMP population.

Graphical and Statistical Analysis

Graphs were generated and statistics analyzed using GraphPad Prism software. Student's two-tailed t-tests were used to calculate *p*-values, except in Fig. 5C, where an extreme value test (using the cumulative distribution function) was performed using a normal distribution with mean and standard deviation calculated from the three hEMP experiment replicates. $p < 0.05$ was considered statistically significant.

RESULTS

The CD326-CD56+ population marks early mesoderm commitment from human PSC

Our previous studies described an early stage of mesoderm commitment during hPSC differentiation that corresponds to the onset of epithelial-mesenchymal transition (EMT) in the primitive streak, and is marked by loss of CD326 (EpCAM) and acquisition of CD56 (NCAM) expression (Fig. 1A) [4]. The CD326-CD56+ human embryonic mesoderm progenitor (hEMP) population partially overlaps with the APLNR+ population [13] and first emerges as early as day 2 of differentiation [4] before cell surface expression of more lineage-specific markers (CD43, CD34, VE-Cadherin, CD235) [14-17]. Transcriptome profiling of hPSC and day 3.5 hEMP by RNA-Seq demonstrated the onset of mesoderm commitment with significant upregulation (FDR<0.01, >2-fold changes) of genes known to be involved in primitive streak formation (*MIXL1, EOMES, T*, and *MESP1*) and EMT (*CDH2, FN1, TWIST1* and *SNAI2*), with concomitant downregulation of cell-cell adhesion molecules such as *CDH1* (which encodes E-Cadherin) and claudins (Fig. 1B, Supplemental Methods). The marked downregulation of pluripotency factors in CD326-CD56+ cells (Fig. 1B) matched the functional loss of teratoma-forming ability previously seen *in vivo* [4]. Using established differentiation conditions, we have previously shown the ability of CD326-CD56+ cells to give rise to all mesodermal lineages tested including hematopoietic, endothelial, mesenchymal (bone, cartilage, fat, fibroblast), smooth muscle, and cardiomyocyte [4]. The lack of endoderm and ectoderm gene expression (Fig. 1C) and the inability to generate these germ layers *in vitro* [4] further confirmed that the hEMP population is specifically committed to mesoderm fate.

While multiple mesoderm lineages can be generated from the hEMP population, it is not known whether hEMP represent a homogenous group of multipotent progenitors or a heterogeneous

mixture of more lineage-restricted bipotent and/or unipotent progenitors. In more committed cell types (e.g. hemangioblasts and hematopoietic progenitors) lineage potential has been examined by cloning single cells and examining the lineage composition of the progeny [13, 18-20]. However, we found that a rigorous and quantitative assignment of clonal lineage potential from single hEMP in culture was not feasible due to technical limitations (e.g. temporal variability in proliferation and differentiation from single cells, the need for stromal co-cultivation obscuring readout and unreliable lineage discrimination from low frequency events).

We therefore turned to a lentiviral genetic labeling strategy to trace cellular genealogy in the mixture of clones present in differentiating bulk cultures. Lentiviral vectors integrate in a semi-random fashion throughout the cellular genome and the resulting integration sites are replicated along with the cellular genome. These integration sites can therefore be treated as unique sequence tags (hereafter referred to as "lentiviral tags") marking all progeny of a tagged clone (Fig. 1D). We developed differentiation conditions that could generate hematopoietic, endothelial and mesenchymal cells from bulk populations of hEMP in one culture vessel, so that clones would not be disturbed after lentiviral tagging and would therefore be free to proliferate and populate multiple lineages. Under these conditions, tri-lineage output was reliably detected based on cell surface marker expression after two weeks of differentiation (Fig. 1A).

Anticipating that multipotent progenitors may represent a rare subset of the total hEMP population, we chose to maximize the complexity of integrations by labeling a high starting number of cells and by using high titer vectors, achieving 60-80% transduction efficiency. To retrieve lentiviral tags from the differentiated cells with high efficiency, we amplified the vector-genome junction sequences via non-restrictive linear amplification-mediated PCR (nrLAM-PCR)

[11, 21, 22] and sequenced them on an Illumina HiSeq 2000 (Supplementary Fig. 1). nrLAM-PCR circumvents the restriction digest required in standard LAM-PCR, allowing for less biased and more comprehensive integration site amplification [23].

Efficient detection of lentiviral tagging of tri-lineage output from monoclonal hPSC

As an initial proof of concept, we assessed lentiviral tag sharing in cells differentiated from a single transduced and expanded hPSC (Fig. 2A). The undifferentiated monoclonal transduced hPSC line contained 25 distinct lentiviral tags, as determined by high-throughput lentiviral tag sequencing (Fig 2B). The tagged hPSC clone was subjected to mesoderm induction to generate a population of hEMP, which was then differentiated into hematopoietic, endothelial and mesenchymal lineages (Fig. 2A). After two weeks, each of the three lineages was isolated by flow cytometry based on cell surface markers (Fig. 1A).

HTS of nrLAM-PCR products from each lineage was performed on an Illumina HiSeq. The same twenty-five tags were detected in DNA from the original expanded, undifferentiated PSC clone and from all three lineages that were generated from isolated hEMPs derived from the hPSC clone; no additional lentiviral tags were detected (Fig. 2B). This experiment demonstrated the ability of the methodology to reliably detect lentiviral tags shared among multiple lineages differentiated from a tagged population of minimum complexity.

Lentiviral tagging of polyclonal pluripotent populations and estimation of undersampling of shared tags

We anticipated that even if the lentiviral system were applied to study a population of progenitor cells that were all multipotent, it would fail to fully detect all lentiviral tags in the multiple

differentiated lineages of interest due to undersampling arising from practical limitations inherent in our experimental system, such as the infeasibility of collecting the genomic DNA of all cells in the culture vessel, of preparing sequencing libraries containing all of the lentiviral tags present in the isolated genomic DNA, and of sequencing all of the lentiviral tags present in the sequencing libraries. In experiments attempting to detect the same clones in multiple cell populations, the impact of undersampling is amplified multiplicatively, leading to a dramatic underestimate of the frequency of sites shared between lineages.

To illustrate this issue, we present a theoretical situation in which cellular populations of three target lineages all contain the same 100 lentiviral tags (Fig. 3). Using an arbitrary scenario in which only 25% of the lentiviral tags were recovered and sequenced at random from each lineage (from compounded undersampling during processing), we can predict the expected number of shared tags between E, H, and M assuming independence in the detection events in the three lineages (Fig. 3). Each tag has a probability of 0.25 of being detected in a single lineage, and only a probability of 0.25^2, or 0.0625, of being detected in two lineages. Therefore, the most likely result is that 6 tags would be detected in both the H and E lineages. This issue compounds with each population that is added to the analysis; in the case of three populations, the probability falls to 0.25^3, or only ~0.016; thus the most likely result is to detect only one to two shared tags of the 100 that are actually present. Interpreting such a result without considering the impact of sampling would lead one to conclude that most of the tagged cells were not multipotent, when in reality, all of them were. In this theoretical setting, the expectation can be calculated because of the initial assertion that 100 clones were present, but in an experimental setting, the total number of clones is not known. Repeated sampling will discover a larger proportion of lentiviral tags in the populations and thus increase the chance of detecting shared

tags, but hundreds of sequencing runs would be needed to uncover the entire pool of lentiviral tags present in a highly complex population (Supplementary Fig. 2). This requirement is both technically and financially challenging to satisfy.

The issue of sampling has been well studied in the field of ecology, and has stimulated the development of various mark-recapture based statistical methods to address the "unseen species" problem. We chose to use one of these methods, the Chao2 estimator, to estimate the size of the sampling problem described above in our experiments. Whereas this estimator has been used previously in the context of viral tagging to estimate the number of tagged cells [5-7], we go further by using these estimates to determine the degree of undersampling and calculate dataset-specific expectations. The Chao2 estimator was chosen as it requires only presence/absence information for species in multiple samples taken from a population [24] and is thus appropriate for use with sequencing of nrLAM-PCR based products. The Chao2 formula uses the frequency of observing rare tags detected only once (f_1) or twice (f_2) to estimate the number of unsequenced tags (Supplementary Fig. 3). We applied the Chao2 estimator to estimate the number of lentiviral tags in the parental population of the E, H and M lineages by repeatedly sampling DNA from the sorted lineages as well as from cells that could not be assigned to a lineage based on immunophenotype ("lineage negative"), which were also captured during FACS sorting. Estimates of the total tags in each hEMP transduction were subsequently used in our model to calculate expectations for tag sharing between the lineages (Supplementary Fig. 4).

We first examined how important the consideration of sampling would be in a real experimental situation by genetically tagging a large population of hPSC, a cell population with known lineage potential (Fig. 4A). In this experiment, the pluripotent hPSC clones were tagged before

mesodermal commitment, and we therefore reasoned that they should individually be capable of producing cells of all three mesodermal lineages irrespective of whether the hEMP stage through which they differentiate is multipotent. Any inability to detect lentiviral tags shared between lineages in this experiment could therefore be attributed to technical limitations of the experimental system rather than an inherent restriction of lineage potential of the target cells.

A pool of approximately 7 million hPSC was transduced and expanded briefly without selection (Fig. 4A). The polyclonal pool of hPSC was then placed into mesoderm induction conditions from which hEMP were isolated at day 3.5, and then replated onto OP9 stroma for trilineage differentiation. After 14 days, hematopoietic (H), endothelial (E), mesenchymal (M) and lineage negative ($CD45^-CD31^-CD73^-$) cells were isolated as separate populations by FACS, from which genomic DNA was isolated and subjected to nrLAM-PCR and HTS. Importantly, for each lineage, ten independent samples of the genomic DNA were taken for nrLAM-PCR, and each of these sequencing libraries incorporated a distinct sequencing index that allowed for post hoc determination of which lentiviral tags were detected in which samples. The estimated sample recovery after cell collection, lineage isolation by FACS, DNA isolation and nrLAM-PCR library preparation was 2-9% of the initial input (Table 1), demonstrating the effect of compounded undersampling (as modeled in Fig. 3).

A total of 68,772 lentiviral tags were recovered from HTS of all samples from all three lineages. 142 of these were found in all three lineages, and 3,532 were found in only two lineages. The vast majority of tags (65,098) were found in only one lineage (Fig. 4B). The Chao2 estimator was applied to these data and the resulting “lower bound” for total lentiviral tags in the transduced hEMP population was used to calculate expected tag sharing between lineages (Fig.

4A, Supplementary Fig. 3 and 4). We compared the observed number of shared lentiviral tags to these expected values and henceforth refer to these observed-to-expected ratios as "detection scores" (Fig. 4A). The average detection score among all of the two- and three-lineage combinations was 0.7. The detection scores for tags shared between endothelial-hematopoietic lineages (EH), endothelial-mesenchymal lineages (EM) and hematopoietic-mesenchymal lineages (HM) were 0.5, 1.0 and 0.5 respectively, and the detection score for tags shared among endothelial-hematopoietic-mesenchymal lineages (EHM) was 0.7 (Fig. 4C). These results from the transduction of polyclonal hPSC define the maximum expectation of the system and set a standard for comparison with datasets from our experiments with transduced mesoderm progenitors in which the existence of multipotent cells is unknown.

Spurious tag sharing is rare with lentiviral tagging and FACS isolation strategy

We next designed an assay to estimate the potential for erroneously identifying common lentiviral tags between two lineages that do not share a common origin. Such false positives could occur if the vector inserted into the same site in the genome in two or more independent events [6, 7], or if cells from one lineage contaminated cells of another lineage due to errors during FACS isolation. For example, if one hematopoietic cell contaminated the endothelial population, it would yield a shared lentiviral tag incorrectly indicating a bipotent hematopoietic-endothelial clone.

To assess the magnitude of these two effects in our experimental system, three separate pools of hPSC were transduced, each with one of three lentiviral vectors expressing distinct fluorescent proteins (Fig. 4D). Mesoderm progenitors were generated and isolated from each transduced pool and differentiated in parallel cultures into hematopoietic, endothelial and mesenchymal

lineages. After 2 weeks, cells from the three differentiated cultures were mixed, and specific lineages were isolated by FACS with the added requirement that each lineage express a distinct fluorescent vector marker: mCitrine+CD31+CD73+CD144+CD45- endothelial cells; mCerulean+CD73+CD31-CD45- mesenchymal cells; mStrawberry+CD45+CD31-CD73- hematopoietic cells. Because the hPSC that generated each lineage were transduced separately prior to differentiation and isolation, the differentiated lineages would not be expected to share any common tags.

When the lentiviral tags were identified from each lineage, the numbers of shared tags between endothelial-hematopoietic (EH), endothelial-mesenchymal (EM) and hematopoietic-mesenchymal (HM) lineages were 1, 6 and 23 out of a total of 5,464 tags sequenced, and no sites were shared between all three lineages (Fig. 4E). These events produced an average detection score of 0.1 (Fig. 4F). These negative control data demonstrate a clear distinction between false positive and true positive lentiviral tag sharing. False positives were predominantly seen in tags shared by the hematopoietic and mesenchymal lineages, with no false positive trilineage events (Fig. 4E,F). Hereafter, we use the detection scores from this experiment as a background cutoff/threshold for considering lentiviral tag sharing to be the result of biological events rather than technical artifacts.

Lentiviral tagging reveals tripotent and bipotent progenitors within hEMP

With the upper and lower limits of the experimental system defined, we applied the lentiviral tagging approach to interrogate the lineage potential of hEMP. hEMP were isolated at day 3.5 of mesoderm induction and then transduced with a single addition of lentiviral vector. After 12 hours, cells were washed and replated into trilineage differentiation conditions (Fig. 5A).

Because lentiviral vectors continue to integrate into host DNA for at least 48 hours in culture [25-27] and because multipotent progenitors within the hEMP population might make fate decisions soon after transduction, we chose to demarcate the window of lentiviral integration in our system. To this end, we performed experiments with or without the HIV-1 integrase inhibitor raltegravir added after 12 hours of transduction, the timing of which was established using the K562 erythroleukemia cell line (Supplementary Fig. 5).

When lentiviral tags were sequenced from hematopoietic, endothelial and mesenchymal cells isolated from tagged and differentiated hEMPs (n=3 experiments), the detection scores were analyzed in comparison to the hPSC tagging experiments in Figure 4. Using an extreme value test, the negative control result (Fig. 5C hollow dot) was found to deviate significantly from the detection scores from hEMP tagging (Fig. 5C black dots). This indicates that tag sharing in the experimental samples was not detected by chance.

Specifically, the detection scores of tag sharing in EHM, EH and EM in the negative control experiment were significantly lower than in the hEMP transduction experiments (p-values $6x10^{-93}$, $3x10^{-6}$ and 0.04). Shared tags among all three lineages had an average detection score of 0.3 (n=3), compared with a detection score of 0.7 for hPSC. The detection scores between EH lineages and between EM lineages were 0.6 and 0.9, similar to the upper limits defined from tagged hPSC (0.5 and 1.0, Fig. 5C yellow dots). This indicates that the hEMP differentiates through an EH and EM bipotent stage. On the other hand, the HM detection score in the negative control experiment was indistinguishable from the hEMP experiment scores (p=0.1), indicating that differentiation does not proceed through an HM intermediate stage.

Addition of raltegravir to hEMP transduction decreased the total number of lentiviral tags but did not influence detection scores. By combining these methods of lentiviral tagging, HTS and quantitative analysis, we conclude that the day 3.5 hEMP population contains tripotent (EHM) progenitors that rapidly differentiate into bipotent EH and EM progenitors, and that strictly bipotent HM progenitors are not generated in this system (Fig. 5D).

DISCUSSION

Genetic labeling via heritable genetic barcodes, transposon insertions or retroviral integrations has long been informative in systems with low or moderate clonal diversity and extensive clonal amplification. Typically, these studies have been used to track long-term hematopoietic stem cell behavior after transplantation in either experimental models or clinical gene therapy trials [21, 28-33]. The capacity of hematopoietic stem cells to self-renew *in vivo* after transplantation allows hematopoiesis to be sustained by relatively few clones, particularly after the initial burst of progenitor output has disappeared. Recent intriguing studies using transposon insertions and inducible genetic labeling to track endogenous murine hematopoiesis have revealed that a far greater clonal contribution during steady state is attributable to lineage-restricted progenitors than was inferred from transplantation models [34-36]. The framework we have presented extends the capabilities of these approaches to cell populations that have much greater clonal complexity and/or little clonal amplification, due to either temporal or biological constraints. Unlike barcoding methods, the high complexity labeling strategy used here does not provide quantitative information on clone size [28, 30, 37], but can nonetheless be used to interrogate a population of unknown multipotency.

As a proof of concept, we examined a transient population of mesodermal progenitors derived from hPSC and demonstrated by lentiviral tagging that at least some of this population possess trilineage (hematopoietic, endothelial and mesenchymal) potential. The finding that trilineage detection scores from labeling of mesoderm progenitors were lower than those from labeling of pluripotent cells, presumably reflects the shorter timeframe for clonal amplification from the progenitors and/or the possibility that tripotent progenitors are only a rare subset of the total progenitor population.

Unlike genetic tagging of the pluripotent hESC, in which shared lentiviral tags were readily detected between any two given downstream lineages, genetic tagging of mesoderm progenitors revealed significant shared tags only between all three lineages (EHM) or between either endothelial and hematopoietic (EH), or endothelial and mesenchymal (EM) lineages. The level of lentiviral tag sharing between the hematopoietic and mesenchymal (HM) lineages was indistinguishable from the negative control (i.e. false positive events). This data lead us to conclude that hematopoietic and mesenchymal cells, while sharing a common tripotent progenitor with endothelial cells, do not branch off from a common bipotent (HM) progenitor but rather arise from mutually exclusive pathways downstream from the tripotent EHM.

Studies based on dual-lineage colony formation assays from others independently confirm the lineage bifurcation pattern observed in our data. Modification of standard hematopoietic progenitor colony forming assays has led to identification of a common precursor for hematopoietic and endothelial cell (hemangioblast) [18, 19, 38] as well as for mesenchymal stem and endothelial cells (mesenchymangioblast) [13]. However, these reports did not identify the

earlier tripotent stage of mesoderm because of the technical limitations of the clonal culture systems.

An additional insight from our findings is that endothelial specification occurs through two mutually exclusive pathways. It was recently shown that the hemogenic endothelium and the arterial vascular endothelium derived from hESC represent non-overlapping populations [39]. However, the developmental origin of these two types of endothelium was not explored. Further studies will be needed to elucidate if endothelium from different clonal origins revealed in our system inherit distinct functional fates.

We observed fewer shared lentiviral tags than initially expected from genetically tagged hPSC and found that this discrepancy can be explained by undersampling in this experimental system. Various mark-recapture approaches have been adopted in the gene therapy setting to estimate the size of the gene-corrected cell pool [5-7]. We applied one of these methods, the Chao2 estimator, in a novel way not only to estimate the number of genetically tagged cells, but also to estimate and correct expectations for the magnitude of undersampling inherent in our HTS-based clonal tracking studies. Of note, even after use of the Chao2 method to estimate the number of undetected lentiviral tags within these large populations, the calculation of expected shared tags will most likely still be higher than observations. The Chao2 estimator for unseen species only guarantees a lower bound for the total number of species, meaning that the true number of species is always greater than or equal to the estimate. The lower bound of expected events also assumes a situation where species have a perfectly uniform abundance distribution, and in most experimental systems, clones amplify in a non-uniform fashion. Furthermore, even though the nrLAM-PCR strategy used for HTS library preparation is less biased than previous methods, it

still cannot amplify all lentiviral tags uniformly, and current technologies cannot sequence and map all lentiviral tags with uniform efficiency. The more different the species abundance distribution is from uniform, the lower the Chao2 estimate will be relative to the true species count. In our model, underestimating the total number of species ($\widehat{N}_2$) leads to an over-estimation of our sampling ($N_2/\widehat{N}$), which in turn yields a greater expected number of shared lentiviral tags ($\widehat{N}_{1,2}$) in our calculations and a lower detection score.

Additionally, our model assumes that clones proliferate sufficiently to populate all three lineages with their progeny. The number of cellular divisions that occur between genetic labeling and clonal fate decision is not known, and it is therefore possible that some clones do not populate all target lineages even though they have the biological potential to do so, or that a clone that initially populates a lineage is extinguished before the end of the experiment. Notably, this assumption is also problematic when analyzing the lineage output from of single cell cultures. A more thorough understanding of these technical and biological effects would extend the improvements in the interpretation of genetic labeling datasets made in the present study.

As genetic labeling strategies and transplantation studies continue to become more efficient and sensitive, we anticipate that the consideration of sampling will become increasingly essential for the accurate interpretation of results. We believe that the quantitative framework presented here represents significant progress towards defining and addressing this need. Moreover, our strategy allows current clonal tracking methods to be extended to experimental systems in which only limited clonal amplification is possible, whether because of limited time for clonal expansion or because clones are actively transitioning and committing into different cellular fates at the time of genetic labeling. We propose that this approach also has potential applications beyond vector

integration studies, for example in the tracking of subclones during the evolution of leukemia and other malignancies.

ACKNOWLEDGEMENTS

This work was supported by the UCLA BSCRC and by the following external grants: CIRM Basic Biology Award (RB3-05217) (GMC, DBK and DE); CIRM Training Grant (TG2-01169) (CJC); International Fulbright Science and Technology Award (CJC); Philip J. Whitcome Predoctoral Fellowship from the UCLA Molecular Biology Institute (ARC); Ruth L. Kirschstein National Research Service Award (GM007185) (ARC); NIH K01AR061415, DOD grant OR120161 and CIRM Basic Biology Award (RB5-07230) (DE). We thank Felicia Codrea and Jessica Scholes (UCLA BSCRC Flow Cytometry core), the BSCRC High Throughput Sequencing Core and the UCLA Clinical Microarray Core for their technical support, and Dr. Jerome A. Zack for providing raltegravir.

REFERENCES

1. Mendjan S, Mascetti VL, Ortmann D et al. NANOG and CDX2 pattern distinct subtypes of human mesoderm during exit from pluripotency. **Cell Stem Cell**. 2014;15:310-325.
2. Murry CE, Keller G. Differentiation of embryonic stem cells to clinically relevant populations: lessons from embryonic development. **Cell**. 2008;132:661-680.
3. Davis RP, Ng ES, Costa M et al. Targeting a GFP reporter gene to the MIXL1 locus of human embryonic stem cells identifies human primitive streak-like cells and enables isolation of primitive hematopoietic precursors. **Blood**. 2008;111:1876-1884.
4. Evseenko D, Zhu Y, Schenke-Layland K et al. Mapping the first stages of mesoderm commitment during differentiation of human embryonic stem cells. **Proceedings of the National Academy of Sciences of the United States of America**. 2010;107:13742-13747.
5. Wang GP, Berry CC, Malani N et al. Dynamics of gene-modified progenitor cells analyzed by tracking retroviral integration sites in a human SCID-X1 gene therapy trial. **Blood**. 2010;115:4356-4366.
6. Aiuti A, Biasco L, Scaramuzza S et al. Lentiviral hematopoietic stem cell gene therapy in patients with Wiskott-Aldrich syndrome. **Science**. 2013;341:1233151.
7. Biffi A, Montini E, Lorioli L et al. Lentiviral hematopoietic stem cell gene therapy benefits metachromatic leukodystrophy. **Science**. 2013;341:1233158.

8. Lois C, Hong EJ, Pease S et al. Germline transmission and tissue-specific expression of transgenes delivered by lentiviral vectors. **Science**. 2002;295:868-872.
9. Baldwin K, Urbinati F, Romero Z et al. Enrichment of Human Hematopoietic Stem/Progenitor Cells Facilitates Transduction for Stem Cell Gene Therapy. **Stem Cells**. 2015.
10. Cooper AR, Patel S, Senadheera S et al. Highly efficient large-scale lentiviral vector concentration by tandem tangential flow filtration. **J Virol Methods**. 2011;177:1-9.
11. Paruzynski A, Arens A, Gabriel R et al. Genome-wide high-throughput integrome analyses by nrLAM-PCR and next-generation sequencing. **Nat Protoc**. 2010;5:1379-1395.
12. Langmead B, Schatz MC, Lin J et al. Searching for SNPs with cloud computing. **Genome Biol**. 2009;10:R134.
13. Vodyanik MA, Yu J, Zhang X et al. A mesoderm-derived precursor for mesenchymal stem and endothelial cells. **Cell Stem Cell**. 2010;7:718-729.
14. Choi KD, Vodyanik MA, Togarrati PP et al. Identification of the hemogenic endothelial progenitor and its direct precursor in human pluripotent stem cell differentiation cultures. **Cell Rep**. 2012;2:553-567.
15. Slukvin II. Hematopoietic specification from human pluripotent stem cells: current advances and challenges toward de novo generation of hematopoietic stem cells. **Blood**. 2013;122:4035-4046.
16. Vodyanik MA, Thomson JA, Slukvin II. Leukosialin (CD43) defines hematopoietic progenitors in human embryonic stem cell differentiation cultures. **Blood**. 2006;108:2095-2105.
17. Kennedy M, Awong G, Sturgeon CM et al. T lymphocyte potential marks the emergence of definitive hematopoietic progenitors in human pluripotent stem cell differentiation cultures. **Cell Rep**. 2012;2:1722-1735.
18. Choi K, Kennedy M, Kazarov A et al. A common precursor for hematopoietic and endothelial cells. **Development**. 1998;125:725-732.
19. Kennedy M, D'Souza SL, Lynch-Kattman M et al. Development of the hemangioblast defines the onset of hematopoiesis in human ES cell differentiation cultures. **Blood**. 2007;109:2679-2687.
20. Kohn LA, Hao QL, Sasidharan R et al. Lymphoid priming in human bone marrow begins before expression of CD10 with upregulation of L-selectin. **Nat Immunol**. 2012;13:963-971.
21. Candotti F, Shaw KL, Muul L et al. Gene therapy for adenosine deaminase-deficient severe combined immune deficiency: clinical comparison of retroviral vectors and treatment plans. **Blood**. 2012;120:3635-3646.
22. Romero Z, Urbinati F, Geiger S et al. β-globin gene transfer to human bone marrow for sickle cell disease. **J Clin Invest**. 2013.
23. Gabriel R, Eckenberg R, Paruzynski A et al. Comprehensive genomic access to vector integration in clinical gene therapy. **Nat Med**. 2009;15:1431-1436.
24. Chao A. Estimating the population size for capture-recapture data with unequal catchability. **Biometrics**. 1987;43:783-791.
25. Brussel A, Sonigo P. Analysis of early human immunodeficiency virus type 1 DNA synthesis by use of a new sensitive assay for quantifying integrated provirus. **J Virol**. 2003;77:10119-10124.
26. Butler SL, Hansen MS, Bushman FD. A quantitative assay for HIV DNA integration in vivo. **Nat Med**. 2001;7:631-634.

27. Munir S, Thierry S, Subra F et al. Quantitative analysis of the time-course of viral DNA forms during the HIV-1 life cycle. **Retrovirology**. 2013;10:87.
28. Naik SH, Perié L, Swart E et al. Diverse and heritable lineage imprinting of early haematopoietic progenitors. **Nature**. 2013;496:229-232.
29. Cartier N, Hacein-Bey-Abina S, Bartholomae CC et al. Hematopoietic stem cell gene therapy with a lentiviral vector in X-linked adrenoleukodystrophy. **Science**. 2009;326:818-823.
30. Gerrits A, Dykstra B, Kalmykowa OJ et al. Cellular barcoding tool for clonal analysis in the hematopoietic system. **Blood**. 2010;115:2610-2618.
31. Brenner S, Ryser MF, Choi U et al. Polyclonal long-term MFGS-gp91phox marking in rhesus macaques after nonmyeloablative transplantation with transduced autologous peripheral blood progenitor cells. **Mol Ther**. 2006;14:202-211.
32. Kim S, Kim N, Presson AP et al. Dynamics of HSPC repopulation in nonhuman primates revealed by a decade-long clonal-tracking study. **Cell Stem Cell**. 2014;14:473-485.
33. McCracken MN, Gschweng EH, Nair-Gill E et al. Long-term in vivo monitoring of mouse and human hematopoietic stem cell engraftment with a human positron emission tomography reporter gene. **Proc Natl Acad Sci U S A**. 2013;110:1857-1862.
34. Sun J, Ramos A, Chapman B et al. Clonal dynamics of native haematopoiesis. **Nature**. 2014;514:322-327.
35. Busch K, Klapproth K, Barile M et al. Fundamental properties of unperturbed haematopoiesis from stem cells in vivo. **Nature**. 2015.
36. Naik SH, Schumacher TN, Perié L. Cellular barcoding: a technical appraisal. **Exp Hematol**. 2014;42:598-608.
37. Lu R, Neff NF, Quake SR et al. Tracking single hematopoietic stem cells in vivo using high-throughput sequencing in conjunction with viral genetic barcoding. **Nat Biotechnol**. 2011;29:928-933.
38. Huber TL, Kouskoff V, Fehling HJ et al. Haemangioblast commitment is initiated in the primitive streak of the mouse embryo. **Nature**. 2004;432:625-630.
39. Ditadi A, Sturgeon CM, Tober J et al. Human definitive haemogenic endothelium and arterial vascular endothelium represent distinct lineages. **Nat Cell Biol**. 2015;17:580-591.

SUPPLEMENTAL METHODS

hPSC culture and mesoderm induction conditions

The hPSC line H1 (WiCell, Madison, WI) was maintained and expanded on irradiated primary mouse embryonic fibroblasts (EMD Millipore, Billerica, MA). To induce mesoderm differentiation, hPSCcolonies were cut into uniform-sized pieces using the StemProEZPassage tool (Invitrogen, Thermo Fisher Scientific, Waltham, MA), transferred into 6-well plates pre-coated for 1 hour with Matrigel (growth factor reduced, no phenol red; BD Biosciences, San Jose, CA), and cultured initially in mTESR medium (Stem Cell Technologies, Vancouver, BC)

until 50–60% confluent (typically 3 days). To initiate differentiation, mTESR medium was replaced with basal induction medium X-Vivo 15 (Lonza, Walkersville, MD). Basal induction medium was supplemented with human growth factors activin A, BMP4, VEGF and bFGF for 3.5 days (all at 10 ng/mL; R&D Systems, Minneapolis, MN), with the inclusion of activin A only on day 1 ("A-BVF" condition [1]). After mesoderm induction, $CD326^{-}CD56^{+}$ cells were isolated by fluorescence activated cell sorting (FACS) at day 3.5 (typically 5-10% of total population, Figure 1A) and placed onto OP9 stroma for further differentiation over the next 13 days.

Mesoderm progenitor differentiation in OP9 coculture

To support simultaneous trilineage differentiation (hematopoietic, endothelial and mesenchymal), mesoderm progenitors ($CD326^{-}CD56^{+}$ cells) generated in the A-BVF mesodermal induction conditions were seeded at 25,000 cells/well into 6-well tissue culture plates, which had been pre-seeded with the murine bone marrow cell line OP9 (American Type Culture Collection, Manassas, VA) 1 day beforehand. Differentiation was carried out in EGM-2 complete medium (Lonza) with ALK 4/5/7 blocker SB-431542 (10μm, R&D Systems) during the first 7 days and then further supplemented with the following cytokines in the next 6 days: stem cell factor (SCF, 50 ng/mL); FMS-like tyrosine kinase-3 ligand (Flt-3, 5 ng/mL); thrombopoietin (TPO, 5 ng/mL); interleukin-3 (IL-3, 5 ng/mL) (PeproTech, Rocky Hill, NJ). During this 13-day differentiation period, half of the medium was changed every other day. At day 13, hematopoietic, endothelial and mesenchymal populations were isolated by FACS.

Transduction of hPSC

To transduce hPSC, cells were cultured in mTESR on Matrigel until 40-50% confluent. Vectors were diluted to $4x10^{7}$ TU/mL in mTESR and 1 mL/well was added to the culture. Cultures were

placed at 37 °C for 6 h. After 6 h, each well was further supplemented with 1.5 mL of fresh mTESR. At 24h post transduction, mTESR containing vectors was fully replaced with fresh mTESR and maintained for two days before initiating the A-BVF mesodermal induction as described above. This vector dose yielded ~50% transduction of hPSC based on FACS analysis of GFP expression in the derived hEMP cells ($CD326^{-}CD56^{+}$).

Transduction of mesoderm progenitors (hEMP)

CD326-CD56+ hEMP were FACS-isolated on day 3.5 of mesoderm induction and resuspended at a concentration of $3x10^5$ cells/mL in EGM-2 complete medium supplemented with: 10 μM Rock inhibitor (Y27632 hydrochloride; R&D Systems), 10 μM SB-431542 (R&D Systems)**,** 5 μg/mL fibronectin (BD Biosciences)**,** 10 μg/mL gentamicin (Gibco, Thermo Fisher Scientific, Waltham, MA) and FUGW virus at $5x10^6$ IU/mL; and transferred to 96 well U bottom low cell binding plate (NUNC, Penfield, New York) at 50 μL/well (equivalent to 15,000 cells/well). The plate was then centrifuged at 1,000 RPM for 5 min at 4°C. After centrifugation, cultures were placed at 37°C for 12 h and then aggregated cell colonies were carefully transferred onto OP9 for differentiation as described above. This vector dose yielded ~80% transduction based on flow cytometry detection of GFP expression in cells after two week of hEMP differentiation. To limit further viral integration as hEMP began differentiation, integrase inhibitor (raltegravir, Merck & Co, White House Station, NJ) was added at 1 μM as soon as OP9 coculture was initiated and was supplemented throughout the 13 days of differentiation.

K562 transduction

K562 (American Type Culture Collection) cells were maintained in RPMI 1640 medium (Mediatech, Manassas, VA) supplemented with 10% heat-inactivated fetal bovine serum

(Gemini Bio-Products, Sacramento, CA) and 1X Penicillin:Streptomycin solution (Gemini Bio-Products). To determine how the timing of integrase inhibitor (raltegravir) addition affects integration number, K562 were transduced with FUGW at MOI 4 and raltegravir was added to the culture at 6 h, 12 h, 24 h, 2 d or 3 d post transduction. DMSO was used as a vehicle control. The transduced cells were cultured for three weeks and their genomic DNA was isolated to determine the vector copy number (VCN) per cell.

Quantification of vector copy number (VCN) by droplet digital PCR

Bio-Rad droplet digital PCR (ddPCR) was used for absolute quantitation of VCN in transduced cells without the requirement of copy number standards as in conventional qPCR. Genomic DNA of transduced cells was isolated as described above and ddPCR was run using primers recognizing the vector packaging signal ψ (forward primer ACCTGAAAGCGAAAGGGAAAC, reverse primer CGCACCCATCTCTCTCCTTCT, and probe FAM-AGCTCTCTC-ZEN-GACGCAGGACTCGGC-Iowa Black FQ). VCN was normalized by the human syndecan 4 gene (SDC4) (forward primer CAGGGTCTGGGAGCCAAGT, reverse primer GCACAGTGCTGGACATTGACA, and probe HEX-CCCACCGAA-ZEN-CCCAAGAAACTAGAGGAGAAT-Iowa Black FQ) or uc378 (forward primer CGCCCCCTCCTCACCATTAT, reverse primer CATCACAACCATCGCTGCCT, and probe HEX-TTACCTTGC-ZEN-TTGTCGGACCAAGGCA-Iowa Black FQ)

ddPCR was carried out as described[2]. Briefly, 1.1µl of template gDNA was added to the ddPCR master mix containing ddPCR Supermix for Probes (Bio-Rad, Hercules, CA) and one unit of DraI enzyme (New England Biolabs, Ipswich, MA). Predigestion was carried out in the PCR reaction mixes for 1–2 h at 37°C before droplet generation with QX100 Droplet Generator and thermal cycling on a T100 Thermal Cycler per manufacturer's instructions (Bio-Rad). The

thermal cycling program conducted was: 95°C 10 m; 55 cycles of 94°C 30 s, 60°C 1 min; 98°C 10 m. After the PCR was completed, the droplets were analyzed using a QX100 Droplet Reader (Bio-Rad). A no template negative control and a positive control genomic DNA with VCN 2 were used to assist gating.

Flow cytometry and cell sorting

Flow cytometry analysis was performed on LSRII or LSRFortessa and cell sorting on a FACSAria II (all from Becton Dickinson). Cultured cells were dissociated into single cell suspension with Accutase (Innovative Cell Technologies, San Diego, CA) and immunofluorescence staining was performed with human- and mouse-specific monoclonal antibodies. Nonspecific binding was blocked with intravenous immunoglobulin (1%) (CSL Behring, King of Prussia, PA) before staining with fluorochrome-conjugated antibodies. The following antibodies were used for staining: CD45-PE-Cy7, CD73-PE, CD73-PE-Cy7, CD144-PerCP-Cy5.5 (all from BD Bioscience) and CD45-BV605, CD31-APC, CD326-PerCP-Cy5.5, CD56-APC and mCD29-APC-Cy7 (all from BioLegend, San Diego, CA). Identification and isolation of lineages was performed using the following gating sequence: mCD29-APC-Cy7 was used to gate out murine cells. CD45-PE-Cy7 was used to gate the hematopoietic population ($CD45^+$). From the non-hematopoietic compartment ($CD45^-$), the CD31-APC and CD73-PE-Cy7 double positive population was first gated, and CD144-PerCP-Cy5.5 positivity was then used to define the endothelial population ($CD45^-CD31^+CD73^+CD144^+$). From the non-hematopoietic, non-endothelial compartment ($CD45^-CD31^-CD144^-$), the mesenchymal population was identified based on CD73-PE-Cy7 positivity ($CD45^-CD31^-CD144^-CD73^+$). Lineage negative cells were $mCD29^-$ cells that could not be assigned to a lineage based on immunophenotype ($mCD29^-CD45^-CD31^-CD73^-$). Unstained cells and fluorescence-minus-one (FMO) were used as

controls for gating. Cell acquisition was performed using FACSDiva (Becton Dickinson) and the analysis was performed using FlowJo (Tree Star, Ashland, OR). Forward and side scatter (FSC/SSC) and 4',6-diamidino-2-phenylindole (DAPI) (Invitrogen) were used to identify live cells.

RNA-Seq and analysis

Total RNA from H1 cultured in mTESR medium (Stem Cell Technologies) and Matrigel (growth factor reduced, no phenol red; BD Biosciences) or mesoderm progenitors (CD326−CD56+ cells) generated in the A-BVF conditions was extracted with Trizol and purified using miRNeasy Mini Kit (Qiagen). 500ng-2ug of total RNA was input to generate cDNA using Nugen Ovation RNA-Seq System v2 and the sequencing libraries were generated using prepX DNA library enzyme kit (IntegenX Inc.) per manufacturer's instructions. Paired-end 100bp sequencing was performed on Illumina HiSeq 2000 with six samples multiplexed per lane. Raw sequence files were obtained using Illumina's proprietary software and are available at NCBI's Gene Expression Omnibus (accession number GSE77879).

RNA-Seq reads were aligned using STAR v2.3.0 [3]. The GRCh37 assembly (hg19) of the human genome and the corresponding junction database from Ensemble's gene annotation were used as reference for STAR. The count matrix for genes in the Ensembl genome annotation was generated with HTSeq-count v0.6.1p2 [4]. DESeq v1.14.0 [5] was used for normalization (using the geometric mean across samples), differential expression (to classify genes as differentially expressed, Benjamini-Hochberg adjusted p-value < 0.01) and to compute moderate expression estimates by means of variance-stabilized data. Heatmaps were built with GENE-E (http://www.broadinstitute.org/cancer/software/GENE-E/) applying relative min/max normalization to moderate expression estimates.

nrLAM-PCR for lentiviral tag sequencing

Genomic DNA was isolated from FACS-sorted cells using the PureLink Genomic DNA Mini kit (Invitrogen) or NucleoSpin Tissue XS kit (Clontech, Mountain View, CA), depending on starting cell number. 5µl of DNA (or a maximum of 75,000 lentiviral tags/reaction) was used as starting input for non-restrictive linear amplification-mediated PCR (nrLAM-PCR) [6]. For repeated sampling, four to ten independent nrLAM-PCR reactions were performed on the genomic DNA from each population. PCR primer sequences and conditions are found in Supplemental Methods Briefly, 50 cycles of linear amplification were performed using AccuPrime Taq (Life Technologies, Carlsbad, CA) with primer HIV3linear (biotin-AGTAGTGTGTGCCCGTCTGT). Linear amplification products were purified using 1.5 volumes of AMPure XP beads (Beckman Genomics, Danvers, MA) and captured onto M-280 streptavidin Dynabeads (Life Technologies). Captured single strand DNA was ligated to read 2 linker (5'Phos-AGATCGGAAGAGCACACGTCTGAACTCCAGTCAC-3C spacer) using CircLigase II (Epicentre, Madison, WI) in a 10-µL reaction at 65° for 1 h. PCR was performed on these beads using primer HIVright PCR (AATGATACGGCGACCACCGAGATCTACACTGATCCCTCAGACCCTTTTAGTC) and an appropriate indexed reverse primer (CAAGCAGAAGACGGCATACGAGAT-index-GTGACTGGAGTTCAGACGTGT).

1. Evseenko D, Zhu Y, Schenke-Layland K et al. Mapping the first stages of mesoderm commitment during differentiation of human embryonic stem cells. **Proceedings of the National Academy of Sciences of the United States of America**. 2010;107:13742-13747.
2. Baldwin K, Urbinati F, Romero Z et al. Enrichment of Human Hematopoietic Stem/Progenitor Cells Facilitates Transduction for Stem Cell Gene Therapy. **Stem Cells**. 2015.
3. Dobin A, Davis CA, Schlesinger F et al. STAR: ultrafast universal RNA-seq aligner. **Bioinformatics**. 2013;29:15-21.

4. Anders S, Pyl PT, Huber W. HTSeq-a Python framework to work with high-throughput sequencing data. **Bioinformatics**. 2015;31:166-169.
5. Anders S, Huber W. Differential expression analysis for sequence count data. **Genome Biol**. 2010;11:R106.
6. Paruzynski A, Arens A, Gabriel R et al. Genome-wide high-throughput integrome analyses by nrLAM-PCR and next-generation sequencing. **Nat Protoc**. 2010;5:1379-1395.

Table 1. Approximation of the degree of undersampling in transduced polyclonal hPSC

Total cell # harvested		% of each lineage from FACS	Cell # yield expected[a]	Cell # yield obtained from FACS	gDNA quant. (uc378/ul)[b]	vol of gDNA eluted (ul)	total cell equivalent after gDNA extraction[c]	vol/ nrLAM PCR (ul)	# nrLAM PCR reactions	total cell equivalent for nrLAM PCR libraries input[d]	% recovery after FACS[e]	% recovery after DNA isolation[f]	% recovery after library prep[g]
3.10E+07	H	22.41%	6.95e6	6.44e6	27600	100	1.38e6	1	10	1.38e5	92.64%	21.44%	2.14%
	E	2.78%	8.60e5	8.57e5	11280	50	2.82e5	1	10	5.64e4	99.67%	32.89%	6.58%
	M	0.98%	3.04e5	2.26e4	2580	50	6.45e4	1.5	10	1.94e4	74.10%	28.60%	8.58%

[a]Cell yield expected is calculated as [Total cell# harvested from OP9 week2] æ [% of each lineage from FACS]
[b]Absolute gDNA quantification is performed by droplet digital PCR using uc378 as a diploid reference target locus
[c]Total cell equivalent after gDNA extraction is calculated as ([gDNA quantification (uc378/ul)] æ [volume of gDNA eluted (ul)])¼2, assuming that each diploid cell contains two copies of uc378
[d]Total cell equivalent for nrLAM PCR libraries input is calculated as ([volume/nrLAM PCR] æ [#nrLAM PCR] æ [total cell equivalent after gDNA extraction])¼[volume of gDNA eluted (ul)]
[e]%recovery after FACS is calculated as ([Cell yield obtain from FACS sorter]¼[Cell yield expected]) æ 100
[f]%recovery after DNA isolation is calculated as ([total cell equivalent after gDNA extraction]¼[Cell yield obtain from FACS sorter]) æ 100
[g]%recovery after library prep is calculated as ([total cell equivalent for nrLAM PCR libraries input]¼[Cell yield obtain from FACS sorter]) æ 100

Table 1: Approximation of the degree of undersampling of DNA from cells differentiated from transduced polyclonal hPSC

The approximate degree of cell/DNA loss was calculated at each experimental stage after transduction of polyclonal hPSC,mesoderm induction and lineage differentiation. After two week of differentiation on OP9, a total of 3.10E+07 cells were harvested from cultures for isolation by flow cytometry. The isolated hematopoietic, endothelial and mesenchyme populations were then used to extract gDNA, from which an aliquot was used to prepare libraries for VISA. Taking these considerations into account, the yield by the end of this process is approximately 2-9% depending on the lineage of interest.

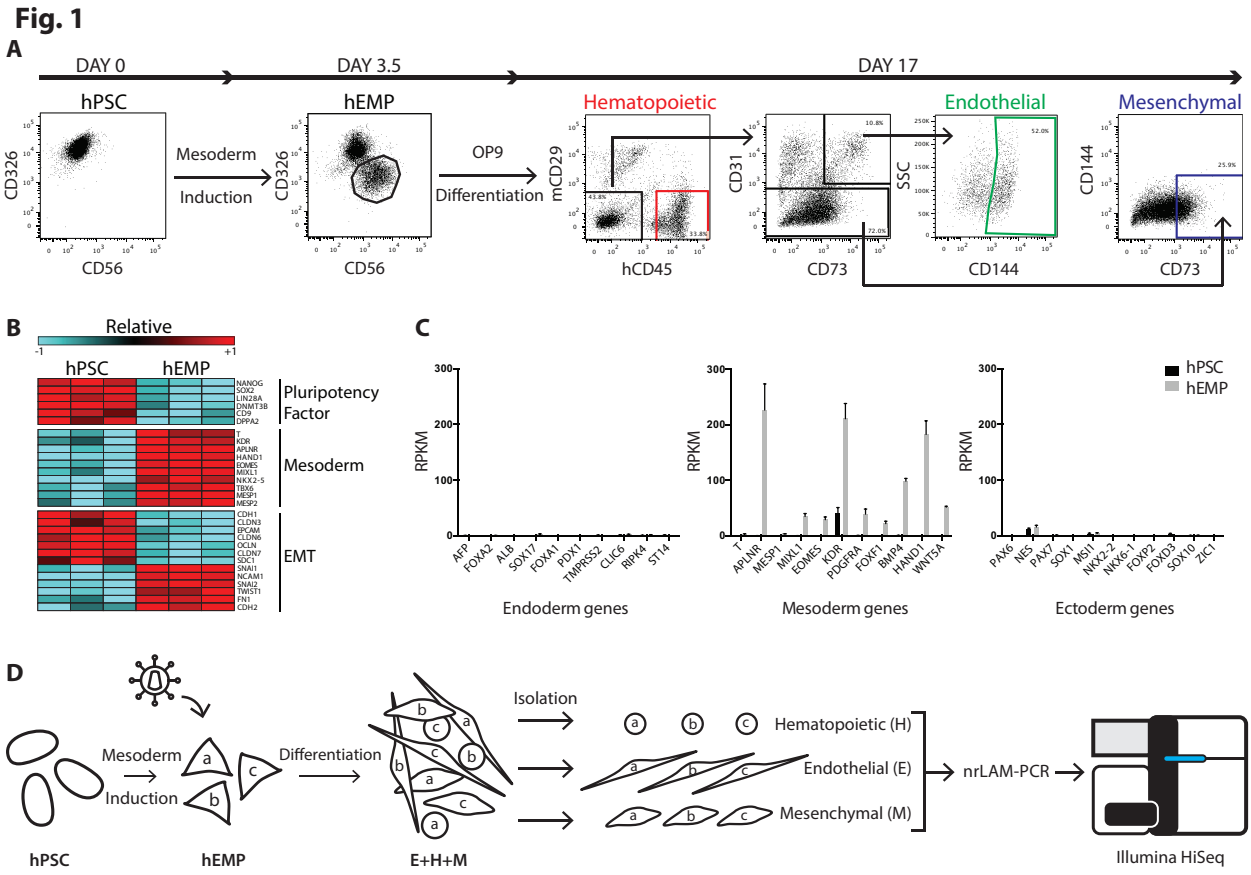


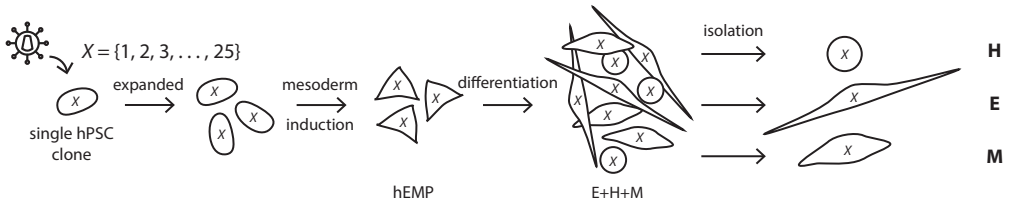
Fig. 1. Generation of a human embryonic mesodermal progenitor population from hPSC with hematopoietic, endothelial and mesenchymal potential

(A) Flow cytometry analysis of EPCAM/CD326 and NCAM/CD56 expression in undifferentiated hPSC and human embryonic mesoderm progenitors (hEMP) generated after 3.5 days of sequential morphogen induction with Activin A, BMP4, VEGF and bFGF. hEMP were FACS isolated as the $CD326^{-}CD56^{+}$ population. Further differentiation on the murine stromal line OP9 generated hematopoietic, endothelial and mesenchymal cells after 2 weeks. After exclusion of $mCD29^{+}$ murine cells, the cell surface marker CD45 was used to define human hematopoietic cells. The endothelial lineage was isolated by the markers $CD144^{+}CD31^{+}CD73^{+}CD45^{-}$, a phenotype that defines non-hemogenic endothelium[14]. The mesenchymal lineage was isolated by the markers $CD73^{+}CD31^{-}CD45^{-}CD144^{-}$. **(B)** Transcriptome comparison via RNA-Seq of hPSC and hEMP showed downregulation (blue) of pluripotency factors, upregulation (red) of mesoderm genes and changes in gene expression indicative of EMT. Both FDR < 0.01 and fold change > 2-fold were applied as filters. Color scale in the heatmap shows the relative expression for each gene using its min/max moderate expression estimates as reference. **(C)** Gene expression in hPSC (black) and hEMP (gray) showed marked upregulation of mesodermal genes enrichment in hEMP without ectoderm or endoderm gene expression. **(D)** Schematic of lentiviral tagging approach to interrogate lineage potential of hEMP. After mesodermal induction of hPSC, hEMP were isolated at day 3.5, and transduced with a lentiviral vector. For the purposes of illustration, three distinct lentiviral tags created by transduction are shown as "a," "b" and "c." These transduced hEMP were co-cultured on OP9 stroma with conditions supporting differentiation of hematopoietic, endothelial and

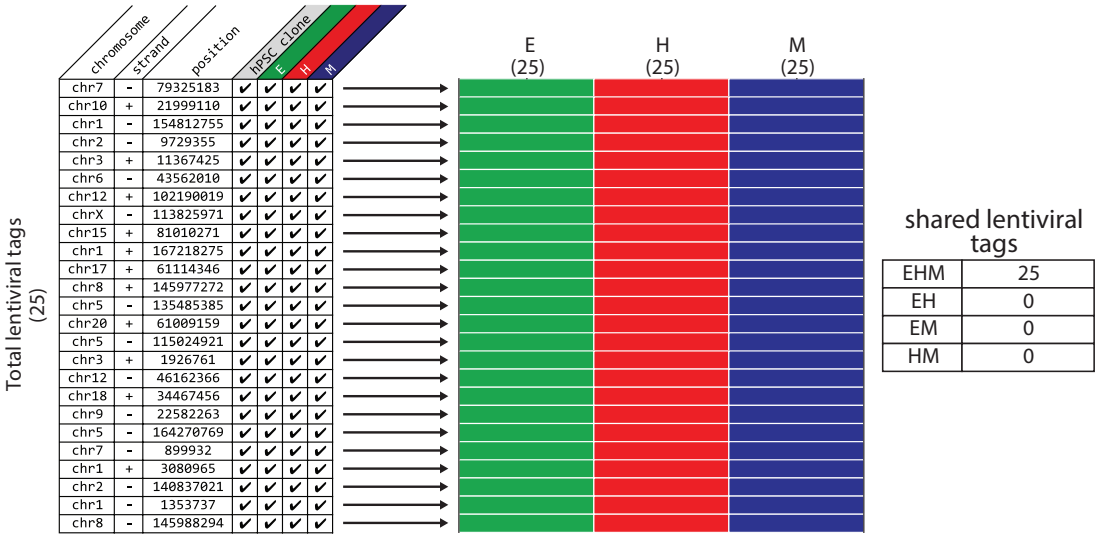
mesenchymal cells; these lineages as well as lin neg cells were FACS isolated after two weeks (as in Fig. 1A). Integration sites in each lineage (E, H and M) were identified by nrLAM-PCR followed by high throughput sequencing on an Illumina HiSeq.

Fig. 2

A



B

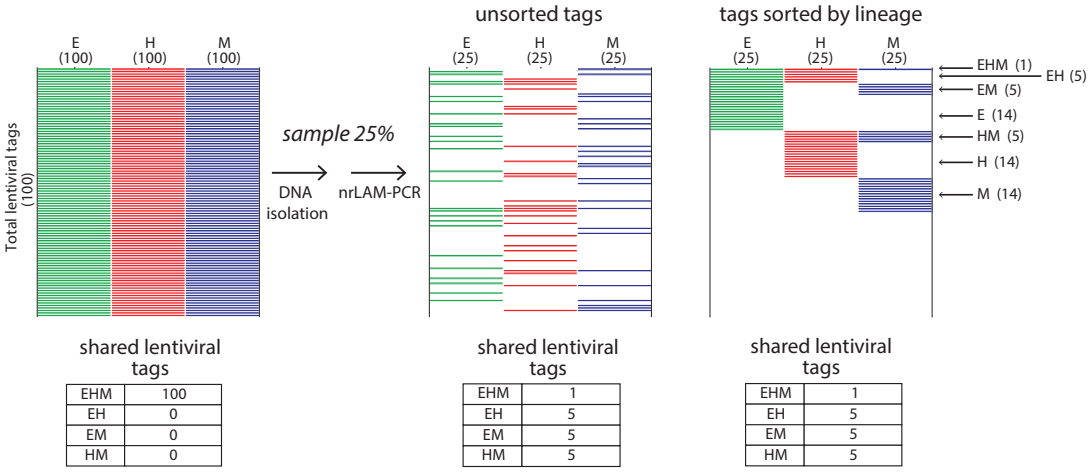


chromosome	strand	position	hPSC clone	E	H	M
chr7	-	79325183	✔	✔	✔	✔
chr10	+	21999110	✔	✔	✔	✔
chr1	-	154812755	✔	✔	✔	✔
chr2	-	9729355	✔	✔	✔	✔
chr3	+	11367425	✔	✔	✔	✔
chr6	-	43562010	✔	✔	✔	✔
chr12	+	102190019	✔	✔	✔	✔
chrX	-	113825971	✔	✔	✔	✔
chr15	+	81010271	✔	✔	✔	✔
chr1	+	167218275	✔	✔	✔	✔
chr17	+	61114346	✔	✔	✔	✔
chr8	+	145977272	✔	✔	✔	✔
chr5	-	135485385	✔	✔	✔	✔
chr20	+	61009159	✔	✔	✔	✔
chr5	-	115024921	✔	✔	✔	✔
chr3	+	1926761	✔	✔	✔	✔
chr12	-	46162366	✔	✔	✔	✔
chr18	+	34467456	✔	✔	✔	✔
chr9	-	22582263	✔	✔	✔	✔
chr5	-	164270769	✔	✔	✔	✔
chr7	-	899932	✔	✔	✔	✔
chr1	+	3080965	✔	✔	✔	✔
chr2	-	140837021	✔	✔	✔	✔
chr1	-	1353737	✔	✔	✔	✔
chr8	-	145988294	✔	✔	✔	✔

Fig. 2. Lentiviral tagging and HTS demonstrates trilineage mesoderm differentiation of monoclonal hPSC.

(A) Schematic of monoclonal hPSC lineage tagging. A single transduced hPSC with 25 lentiviral tags was clonally expanded, induced to form embryonic mesoderm progenitors (hEMP) and differentiated into hematopoietic (H), endothelial (E), and mesenchymal (M) lineages, which were then isolated as separate populations by FACS and subjected to vector integration site analysis (VISA) by HTS. *X* is the set of all 25 lentiviral tags found in the initial hPSC clone. **(B)** Table listing all 25 lentiviral tags detected by HTS in the hPSC clone and the three lineages differentiated from hEMP generated from the hPSC clone, with chromosome, strand and position information. Checkmarks indicate that a particular tag was detected in the lineage indicated at the top of the column. Each row in schematic heatmap to the right represents a distinct lentiviral tag, equivalent to the rows of the table on left. A filled in box indicates the presence of that tag in the population, and the three populations (E, H and M) are annotated as green, red and blue colors respectively for clarity. Numbers in parentheses under each lineage label refer to the total number of tags found in each population, and numbers on the y-axis indicate the total number of tags found among all populations. All 25 lentiviral tags were detected as shared among the three lineages, as indicated in the table to the right of the heatmap.

Fig 3.



$$\hat{N}_{EHM} = 100 \times 0.25^3 \approx 1$$
$$\hat{N}_{EH} = 100 \times 0.25^2 - \hat{N}_{EHM} \approx 5$$
$$\hat{N}_{EM} = 100 \times 0.25^2 - \hat{N}_{EHM} \approx 5$$
$$\hat{N}_{HM} = 100 \times 0.25^2 - \hat{N}_{EHM} \approx 5$$

Fig. 3. Theoretical effect of sampling on shared lentiviral tag detection. Schematic heatmaps are shown for the theoretical scenario of a population in which all cells are multipotent and marked by a total of 100 lentiviral tags. Each multipotent cell goes on to form endothelial (E), hematopoietic (H) and mesenchymal (M) progeny, and all three lineages therefore contain all 100 tags in the culture vessel (left panel). We show how the results of this experiment appear using a model in which only 25% of tags are found from each lineage because of cumulative undersampling during cell and DNA processing. The tags are unsorted in the middle panel to reflect the randomness of the sampling; in the right panel the same tags are sorted according to lineage to aid visualization. With this degree of undersampling (25%), an average experiment will only detect 58 of the 100 lentiviral tags, only 15 will be shared by two lineages, and most importantly, only one tag will be detected as shared between the three lineages.

Fig. 4

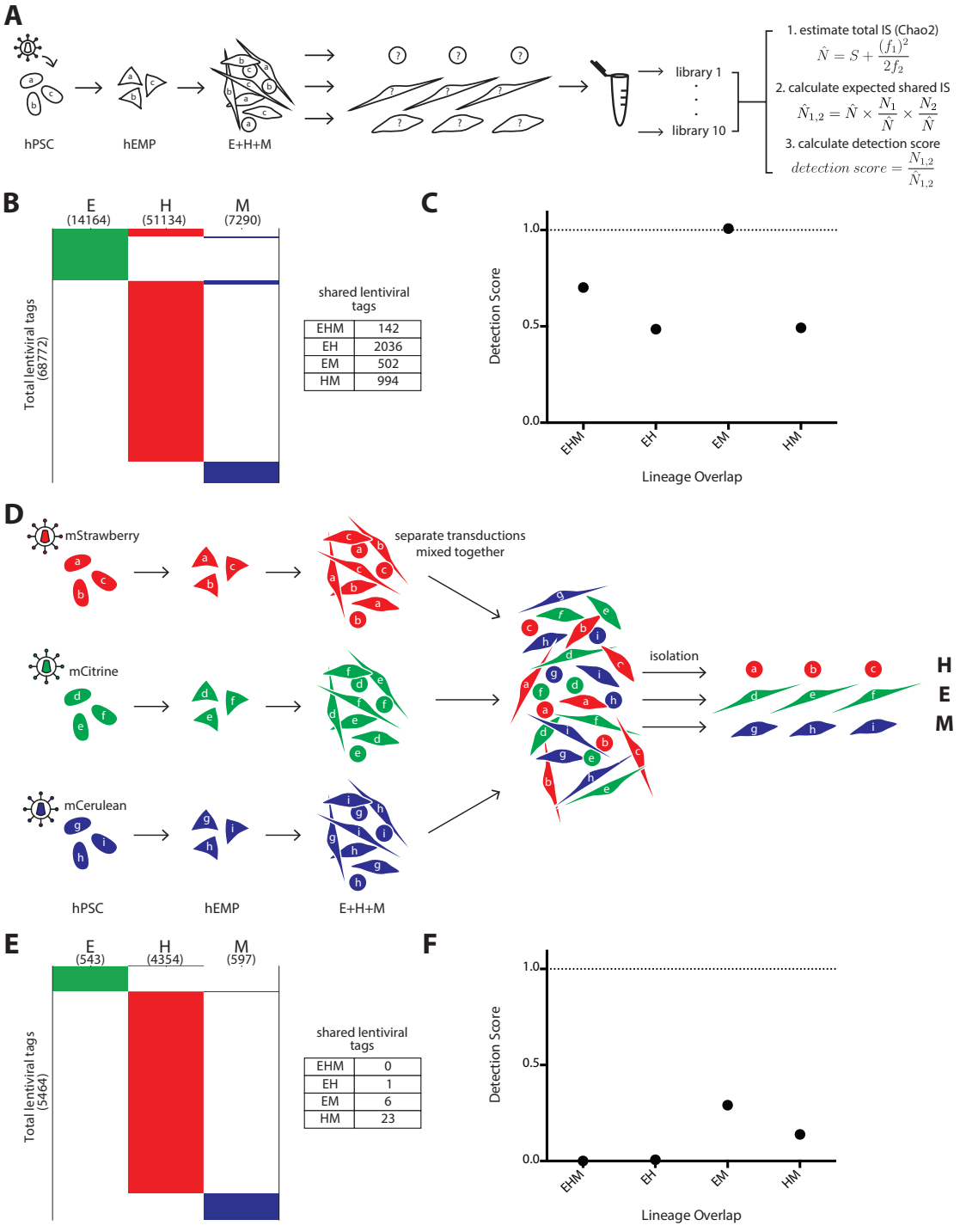


Fig. 4. Positive and negative control experiments to determine the maximal and minimal expectations for lentiviral tag sharing. (A-C) Positive control, and **(D-F)** Negative control for clonal detection limits. **(A)** Schema of polyclonal hPSC tagging experiment performed to define upper boundary of detection. A pool of hPSC were transduced, expanded briefly without selection, and induced to become hEMP, which were then differentiated and analyzed via nrLAM-PCR and HTS. For each lineage, ten separate DNA samplings were taken for nrLAM-PCR and sequencing, with each sampling labeled by a unique barcode. These data were used for

Chao2 estimation of total tag count in each lineage, and this information was used in our model to calculate a shared tag "detection score" ($\widehat{N}$= estimated total number of lentiviral tags, S=observed number of total tags, f_1= number of tags observed only once, f_2 = number of tags observed only twice). $\widehat{N}_{1,2}$ = expected shared tags between population 1 and 2, $N_{1,2}$ = observed shared tags between population 1 and 2 (see also Supplementary Fig. 3, 4 and Methods for Chao2 explanation and model).
(B) Lentiviral tags identified in each lineage derived from a polyclonal pool of transduced hPSC. A total of 68,772 tags were detected in one or more lineage. Each horizontal line in the schematic heatmap represents a specific lentiviral tag and tags are clustered based on their presence in lineages; tags identified in two or more lineages are shown in the same horizontal position. The total number of tags sequenced from each lineage is listed in parentheses below the lineage label. 142 tags were shared among all three lineages. 2036, 502 and 994 shared tags were found between EH, EM and HM, respectively. The majority of tags (65,098) were detected in only one lineage, illustrating the compounded effect of sampling during cell harvest and sequencing preparation. **(C)** The numbers of expected shared tags were compared to the observed shared tags in the form of a detection score (observed/expected). The detection score for EHM, EH, EM, and HM were 0.7, 0.5, 1.0 and 0.5, respectively. Since hPSC are assumed to be pluripotent, the detection score between any set of lineages represents the maximum tag sharing one can detect in this experimental system. **(D)** Three separate pools of hPSCs were transduced with lentiviral vectors expressing distinct fluorescent markers and differentiated in parallel on OP9. The three pools of differentiated cells were combined, and each lineage was then isolated from this pool based on the sets of lineage markers (Fig. 1A), each of which was paired with a distinct fluorescent marker (mStrawberry hematopoietic; mCitrine endothelium; mCerulean mesenchyme). This negative control served to determine the frequency of false clonal overlap by chance and sorting errors. **(E)** Only 1, 6 and 23 shared tags were found between EH, EM and HM, respectively and no tags were shared with all lineages (EHM). **(F)** The detection score for EHM, EH, EM, and HM were 0, 0.007, 0.3 and 0.1, respectively. These scores represent the false positive rate of tag sharing in this experimental system.

Fig. 5

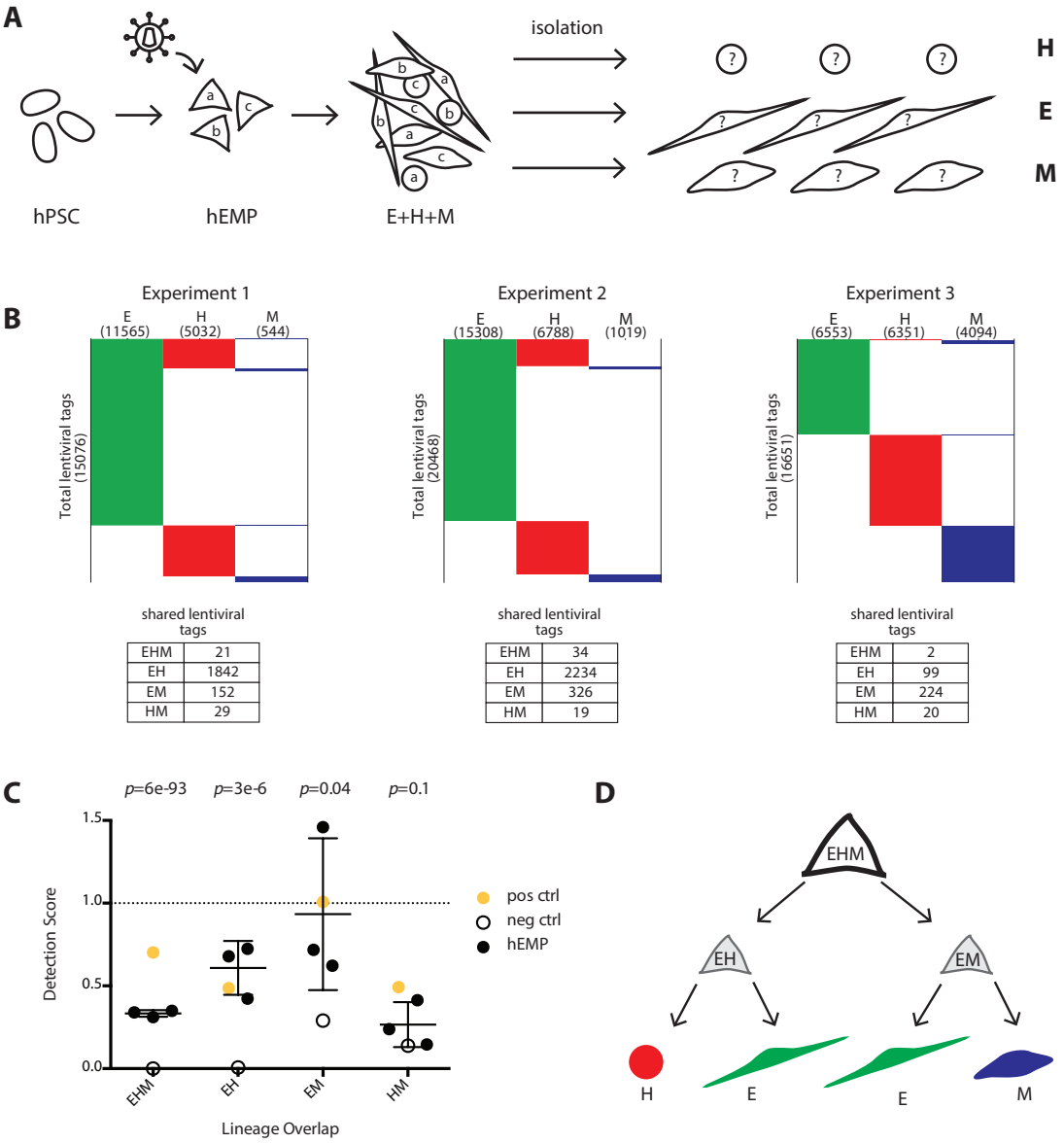
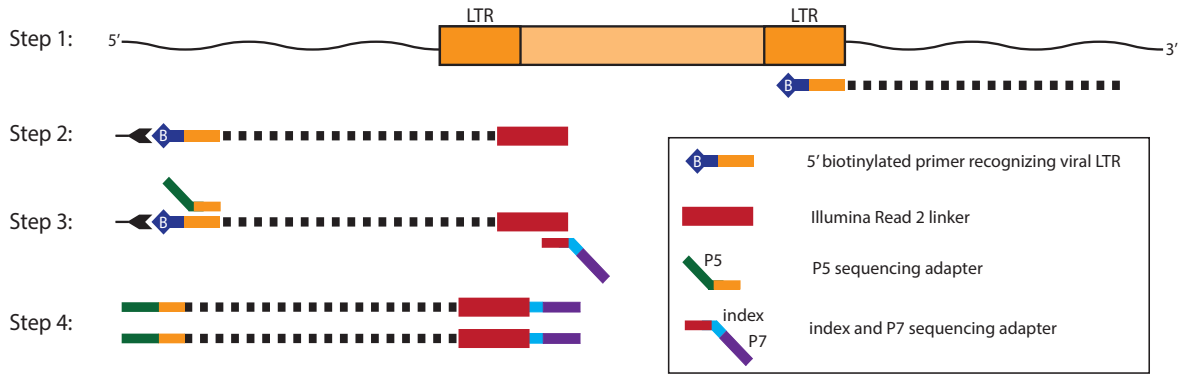


Fig. 5. Lentiviral tagging of human embryonic mesodermal progenitors reveal subpopulations with bipotent and tripotent potential. (A) Experimental design to track clonal output of a pool of tagged mesoderm progenitors (hEMP) differentiated into hematopoietic (H), endothelial (E), and mesenchymal (M) lineages. **(B)** Data from three independent experiments showing number of lentiviral tags identified in each lineage. Experiment 1 was performed with the integrase inhibitor raltegravir added after 12 hour of transduction while experiments 2 and 3 were performed without raltegravir. The addition of raltegravir did not alter the detection score of shared lentiviral tags. **(C)** The detection scores from marking mesoderm progenitors are shown (black dots) as well as the maximum detection expectation (set by positive control, yellow dot) and the false positive rate (set by negative control, empty circle). The negative control detection scores for EHM, EH and EM tag sharing were significantly lower than the detection scores in the mesoderm progenitor tagging experiments (*p*-values $6x10^{-93}$, $3x10^{-6}$ and 0.04 respectively for EHM, EH and EM). The HM detection scores were similar between the

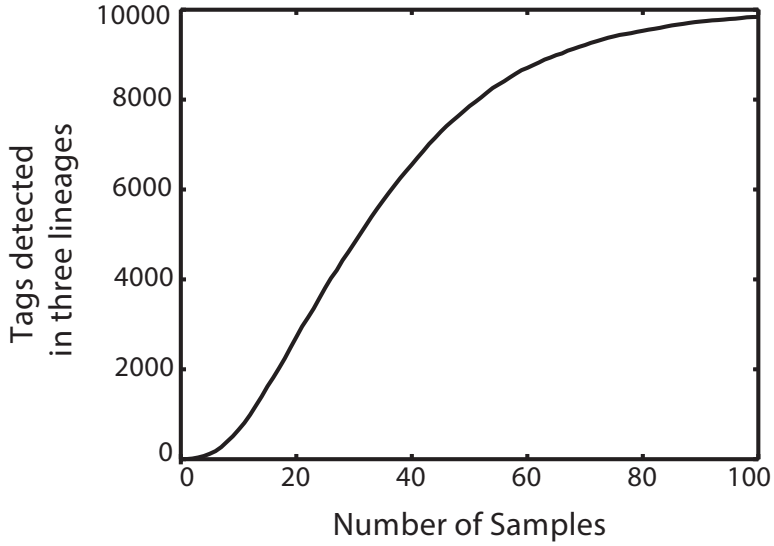
progenitor tagging experiments and the negative control (*p*-value=0.1). *p*-value calculation used an extreme value test performed using the cumulative distribution function of a normal distribution with mean and standard deviation calculated from progenitor experimental data. **(D)** Proposed model of the developmental potential of human mesodermal progenitor defined by unbiased lentiviral marking. During mesodermal differentiation from hPSC, a tripotent progenitor (EHM) with limited self-renewal ability rapidly generates either endothelial/hematopoietic progenitors (EH) or endothelial/mesenchyme progenitor (EM). The lentiviral tagging data do not support the existence of a bipotent HM progenitor.

Supplementary Figure 1



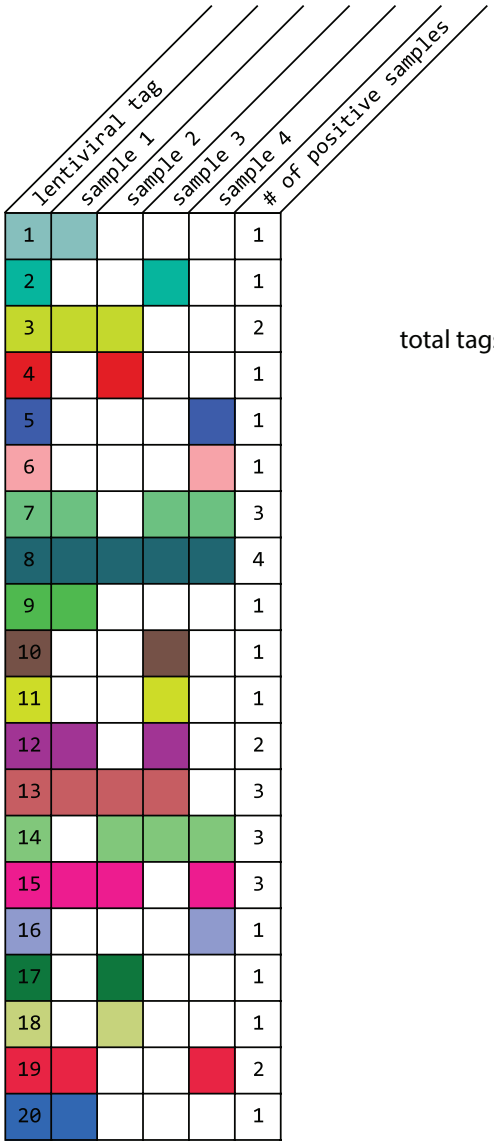
Supplementary Fig. 1. nrLAM-PCR protocol used for HTS library preparation. Step 1: a biotinylated vector-specific primer is used to produce a short ssDNA containing the end of the vector with some flanking genomic sequence. Step 2: after 50 cycles of this linear amplification, biotinylated product is captured onto streptavidin-coated paramagnetic beads, and an ssDNA linker is ligated using an RNA ligase (CircLigase II). Step 3: a nested vector-specific primer is then used with a linker-specific primer containing a distinct index (barcode) for each sample to be sequenced in parallel. Products from multiple samples are then pooled, quantified via digital PCR, and sequenced on an Illumina HiSeq 2000.

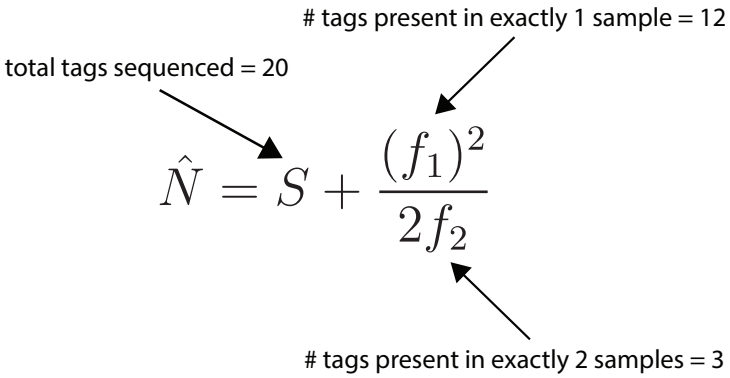
Supplementary Figure 2



Supplementary Fig. 2. Rarefaction curve shows increased detection of shared lentiviral tags as sampling is increased.
The rarefaction curve to detect shared lentiviral tags among 3 populations was modeled *in silico*. Assuming a multipotent progenitor population with 10,000 clones each marked by a single lentiviral tag, if only 5% of the total pool of progeny was sampled each time, approximately 100 samples would be needed to detect nearly all of the lentiviral tags shared by the three lineages.

Supplementary Figure 3





$$\hat{N} = 20 + \frac{(12)^2}{2 \times 3}$$

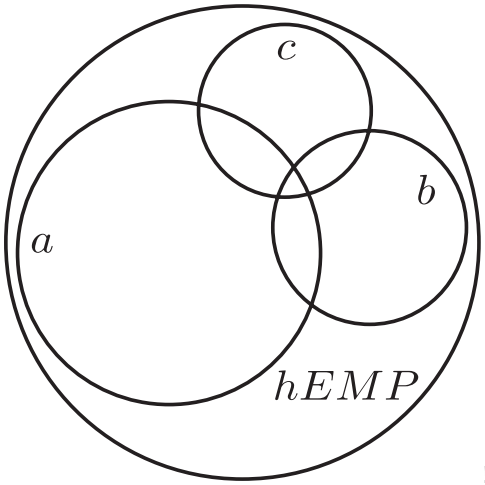
$$\hat{N} = 44$$

Supplementary Fig. 3. Example of how Chao2 formula is applied to estimate the "unseen" lentiviral tags. In a modeling experiment, four samplings (samples 1-4) are taken from a single DNA sample. Twenty lentiviral tags are detected among the four samplings, and tag detection in a particular sample is denoted by a filled-in cell in the table. Most tags are only detected in some samplings (only tag #8 is detected in all four samples). The Chao2 estimator equation only requires the total number of tags detected among all samplings (S=20), the number of tags detected in only one sampling (f_1=12), and the number of tags detected in exactly two samplings (f_2=3). In this example, the Chao2 method estimates that the total number of tags in the set ($\hat{N}$) is at least 44, and therefore only 20 out of 44 or more tags were detected.

Supplementary Figure 4

The model we use for calculating expectations assumes that the sets of lentiviral tags in the lineages are all randomly and independently drawn from the set of tags in the hEMP population before differentiation. This random draw models chance fate decisions, cell sampling during sorting, genomic DNA sampling during isolation and library preparation, and sampling during sequencing. In the graphical example shown below, the size of the circle indicates the size of the set of lentiviral tags. Lineage *a* has more tags than lineage *b*, which has more tags than lineage *c*. All three sets are contained by the total set of tags in the hEMP population.!!

!

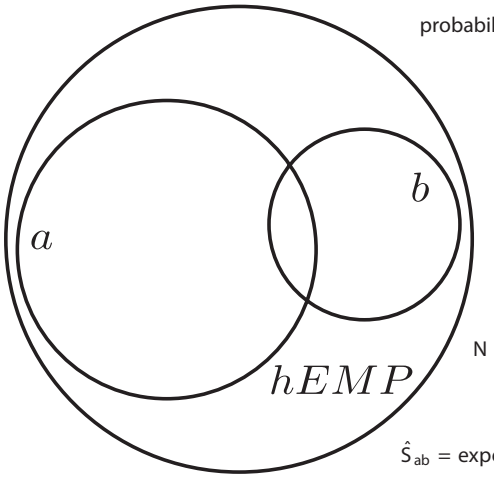


!

We illustrate below how the expected numbers of shared lentiviral tags between lineages *a* and *b* (^S_{ab}) would be calculated under this model using the Chao2 estimates *N* of the number of tags in the original tagged hEMP population:

probability that tag in hEMP is sampled and sequenced in a = $\frac{S_a}{N}$

probability that tag in hEMP is sampled and sequenced in b = $\frac{S_b}{N}$



$$\hat{S}_{ab} = N \times \frac{S_a}{N} \times \frac{S_b}{N}$$

N = Chao2 estimate of # of tags in hEMP population

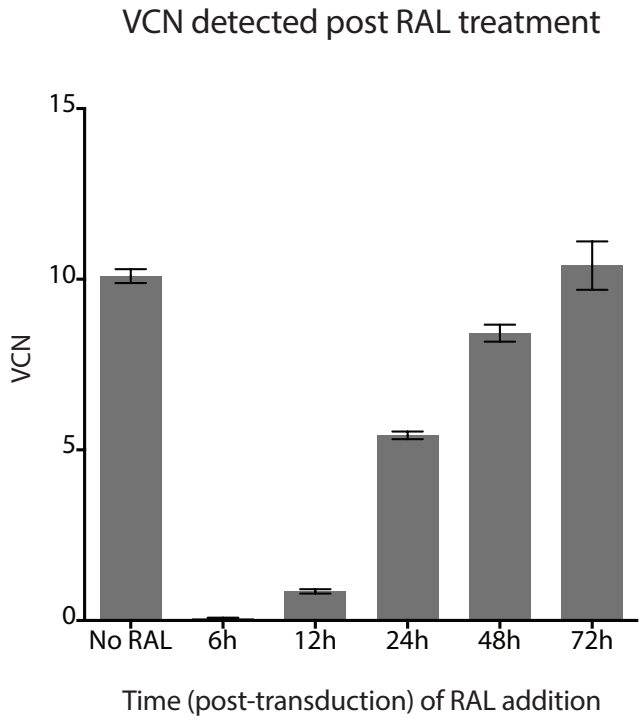
S_a = # of tags sequenced in population a

S_b = # of tags sequenced in population b

$\hat{S}_{ab}$ = expectation for # of IS expected to be detected in both a and b, assuming complete multipotency

Supplementary Fig. 4. Graphical description of model used to calculate expected tag sharing

Supplementary Figure 5



Supplementary Fig. 5. Demarcating the lentiviral tagging window using an integrase inhibitor. The kinetics of lentiviral integration in the presence of raltegravir were tested in K562 cells. Raltegravir was added at 6, 12, 24, 48 or 72h post transduction. Cells were then cultured for a total of 22 days to allow washout of unintegrated vector and then harvested for Vector Copy Number (VCN) analysis by qPCR (see methods). Raltegravir effectively blocked further insertion of viral DNA into the host genome when added from 6-48 hours post transduction. Earlier addition of raltegravir led to progressively fewer vector copies in the transduced cells. 12 hours was chosen as the optimal time point as it created a brief window of lentiviral tagging while still allowing for integration of a significant number of tags.

Supplementary Figure 6

hPSC experiment sample	sequence reads
Lin neg replicate 1	1189184
Lin neg replicate 2	2574531
Lin neg replicate 3	2796234
Lin neg replicate 4	1112701
Lin neg replicate 5	2192998
Lin neg replicate 6	1946064
Lin neg replicate 8	2201297
Lin neg replicate 9	1838262
Lin neg replicate 10	7025150
Endo replicate 1	2272245
Endo replicate 2	2417017
Endo replicate 3	2242602
Endo replicate 5	1180544
Endo replicate 6	722642
Endo replicate 7	1360803
Endo replicate 8	1477416
Hem replicate 1	784341
Hem replicate 2	333257
Hem replicate 3	1007918
Hem replicate 4	1057099
Hem replicate 5	1462369
Hem replicate 6	7233616
Hem replicate 7	2868065
Hem replicate 8	3494205
Hem replicate 9	1993129
Hem replicate 10	1789315
Mes replicate 1	910508
Mes replicate 2	877073
Mes replicate 3	1078692
Mes replicate 4	495860
Mes replicate 6	1067230
Mes replicate 7	1787316
Mes replicate 8	978519
Mes replicate 9	4076401
Mes replicate 10	1727158

neg ctrl experiment sample	sequence reads
Endo mCit replicate 1	2617422
Endo mCit replicate 2	1581080
Endo mCit replicate 3	1679044
Endo mCit replicate 4	213235
Hem mStr replicate 1	1659138
Hem mStr replicate 2	2269970
Hem mStr replicate 3	1888961
Mes mCer replicate 1	6280301
Mes mCer replicate 2	413124
Mes mCer replicate 3	1570069
Mes mCer replicate 4	804127

hEMP experiment 1 sample	sequence reads
Endo replicate 1	1246803
Endo replicate 2	840652
Endo replicate 3	953043
Hem replicate 1	1160294
Hem replicate 2	859364
Hem replicate 3	789210
Mes replicate 1	97545
Mes replicate 2	1545205
Mes replicate 3	941272
Lin neg replicate 1	1081870
Lin neg replicate 2	998882
Lin neg replicate 3	1404188

hEMP experiment 2 sample	sequence reads
Endo ral replicate 1	1249441
Endo ral replicate 2	1410173
Endo ral replicate 3	1248145
Hem ral replicate 1	1327326
Hem ral replicate 2	766698
Hem ral replicate 3	1004382
Mes ral replicate 1	675396
Mes ral replicate 2	638379
Mes ral replicate 3	234359
Lin neg ral replicate 1	1431616
Lin neg ral replicate 2	1564562
Lin neg ral replicate 3	1310101

hEMP experiment 3 sample	sequence reads
Endo replicate 1	1918009
Endo replicate 2	2382874
Endo replicate 3	1868462
Endo replicate 4	1942129
Hem replicate 1	1651724
Hem replicate 2	2282962
Hem replicate 3	1868711
Hem replicate 4	1812654
Mes replicate 1	1651083
Mes replicate 2	2329124
Mes replicate 3	1712285
Mes replicate 4	2264889
Lin neg replicate 1	1788908
Lin neg replicate 2	1850487
Lin neg replicate 3	1820167
Lin neg replicate 4	1505202

Supplementary Fig. 6. The numbers of sequence reads obtained for samples analyzed in the manuscript. For each sample, three or more independent nrLAM-PCRs were generated and sequenced by Illumina technologies to obtain $>1\times10^6$ sequence reads. The sequence reads were then mapped to unique chromosomal locations within the human genome to profile the lentiviral tags as detailed in Methods. (Endo: Endothelial; Hem: Hematopoietic, Mes: Mesenchymal; Lin neg: lineage negative cells that could not be assigned to Endo, Hem or Mes based on immunophenotype)

Chapter 4

Pericyte-like cells generated from human pluripotent stem cells support hematopoietic stem and progenitors *ex vivo*

Chee Jia Chin[1], Mirko Corselli[2], David Casero[1], Yuhua Zhu[1], Chong Bin He[1], Reef Hardy[3,4], Bruno Péault[3,4], Gay M. Crooks[1,4,5,6]

[1]Department of Pathology & Laboratory Medicine, David Geffen School of Medicine (DGSOM), University of California, Los Angeles (UCLA)

[2]Becton Dickinson, San Diego, CA

[3]Department of Orthopedic, DGSOM. Orthopedic Hospital Research Center, UCLA

[4]Broad Stem Cell Research Center, UCLA

[5]Department of Pediatrics, DGSOM, UCLA

[6]Jonsson Comprehensive Cancer Center, UCLA

CJC: Conception and design, collection and assembly of data, data analysis and interpretation, manuscript writing; MC: Conception and design; collection of data; DC: RNA-Seq data analysis and interpretation; YZ and CBH: Collection of data; RH and BP: RNA-Seq data collection; GMC: Conception and design, data analysis and interpretation, manuscript writing, financial support, final approval of manuscript.

Abstract

Various mesenchymal cell types have been identified as critical components of the hematopoietic stem/progenitor cell (HSPC) niche. Although several groups have described the generation of mesenchyme from human pluripotent stem cells (hPSC), the capacity of such cells to support hematopoiesis has not been reported. Here we demonstrated that distinct mesenchymal

subpopulations co-emerge from mesoderm during hPSC differentiation. Despite co-expression of common mesenchymal markers (CD73, CD105, CD90, PDGFRβ), a subset of cells defined as $CD146^{++}CD140a^{low}$ supported HSPC *ex vivo* while a separate subset expressing $CD146^{-}CD140a^{+}$ drove HSPC differentiation. The $CD146^{++}$ subset expressed genes associated with the HSPC niche and supported hematopoiesis in part through contact-dependent Notch signaling and modulation of Wnt signaling. Molecular profiling revealed remarkable similarity in transcriptional regulation between hPSC-derived $CD146^{++}$ and primary human CD146++ perivascular cells. The derivation of diverse pools of mesenchymal populations from hPSC opens potential avenues to model their developmental and functional differences and to improve cell-based therapeutics from hPSC.

Introduction

Maintenance of bona fide, self-renewing hematopoietic stem and progenitor cell (HSPC) *ex vivo* remains challenging in part because of our limited ability to recapitulate the human HSPC niche in culture. Intensive research efforts have begun to uncover the cellular and molecular constituents of the niche that regulate self-renewal and differentiation of HSPC. Through the use of knockout and transgenic mice several cell populations have been described in terms of their spatial relationship to the bone and vascular components of the bone marrow, and their differential expression of various markers and bioactive molecules (Ding et al., 2012; Kunisaki et al., 2013).

We have recently shown that high expression of melanoma-associated cell adhesion molecule (MCAM/CD146) identifies human pericytes, a cell type that ensheaths arterioles and sinusoids (Crisan et al., 2008; Kunisaki et al., 2013) and which can establish a heterotopic HSC niche

when transplanted into immunodeficient mice (Sacchetti et al., 2007). Unlike CD146- mesenchyme, monolayers of CD146++ cells from primary tissue (adipose tissue and fetal bone marrow) can support human HSPC *ex vivo* for at least 2 weeks in the absence of exogenous cytokines (Corselli et al., 2013).

We and others have shown that mesenchymal cells can be differentiated from human pluripotent stem cells (hPSC)-derived mesoderm (Chin et al., 2016; Evseenko et al., 2010; Vodyanik et al., 2010; Wanjare et al., 2014). These previous studies identified mesenchyme as a single population defined mostly by expression of CD73 or CD105 and absence of hematopoietic and endothelial markers. We now report that the mesenchyme generated from hPSC is functionally and transcriptionally heterogeneous. Our studies identified a distinct subpopulation of hPSC-derived mesenchyme with high CD146 and low CD140a expression, capable of supporting clonogenic, engraftable and self-renewing human HSPCs *ex vivo* without exogenous cytokines. In contrast CD146lo/- CD140a+ mesenchyme showed little ability to support HSPC. Transcriptome analysis revealed that the CD146++ cells expressed significantly higher levels of perivascular markers (*NES* and *CSPG4*) and niche factors known to have critical roles in HSC maintenance (*JAG1, SCF, IGFBP2*). HSPC survival was dependent on cell-cell interactions and Notch signaling as well as modulation of Wnt signaling. Closer transcriptional analysis, combining data from both hPSC and human primary tissue, revealed that dominant pathways shared by the CD146++ populations that support HSPC were those related to vascular development, cell adhesion and motility. Our data suggest that hPSC-derived mesoderm can generate mesenchymal cells phenotypically, functionally and molecularly similar to previously identified primary pericytes that comprise the human HSPC niche.

Results

Heterogeneity of embryonic mesoderm-derived mesenchymal cells

We have previously characterized a human embryonic mesoderm progenitor (hEMP) population derived from hPSC that marks the onset of mesoderm commitment and has the potential to generate broad mesodermal derivatives (Chin et al., 2016; Evseenko et al., 2010). The isolation of hEMP offers a platform to systematically examine downstream mesoderm derivatives for their ability to support HSPC.

Culture of isolated hEMP using conditions that favor mesenchymal differentiation, generated at least two distinct mesenchymal populations which could be discriminated based on expression of CD146 and PDGFRA (CD140a, a marker commonly used to define murine mesenchyme (Houlihan et al., 2012; Morikawa et al., 2009)) **(Fig 1A)**. High co-expression of CD146 and CD73 identified a largely CD140a- population, whereas CD140a+ cells expressed low levels of CD73 and CD146. This reciprocal expression pattern between CD146 and CD140a was consistent with primary human lipoaspirate-derived mesenchyme (**Supplementary Fig 1**). Both hPSC-derived mesenchymal subsets expressed classic mesenchymal markers CD90, CD105, CD44, PDGFRB (**Supplementary Fig 2**). BrdU pulsing of hEMP-derived mesenchymal cultures showed that a significantly higher frequency of $CD146^{++}$ cells were cycling relative to $CD146^{-}$ cells (**Fig 1B**) ($p \leq 0.05$). Mesenchymal differentiation from two other hPSC lines, UCLA3 and UCLA6, yielded similar mesenchymal populations to the H1 line; although the relative frequency of CD146++ and CD146- cells varied, the inverse relationship of CD146 and CD140a expression in the mesenchyme was seen across all three hPSC lines (**Fig 1A, Supplementary Fig 3A**).

To further analyze the mesenchymal subpopulations, we isolated cells that represented the extremes of these phenotypes: a $CD146^{++}$ $CD73^{++}$ $CD140a^{Dim/-}$ population ("CD146++") and a $CD146^{-}$ $CD73^{dim}$ $CD140a^{+}$ population ("CD146-")(**Fig 1B**).

RNA-Seq profiling of the isolated populations demonstrated 916 genes were differentially expressed (558 up- and 358 down-regulated in CD146++ relative to CD146- cells (**Fig 1C)**. $CD146^{++}$ mesenchyme expressed high levels of *CSPG4* (which encodes NG2), *NES* (NESTIN) and *LEPR* (leptin receptor) **(Fig 1C and Supplementary Fig 2B Supplementary Fig 3B**), characteristic markers of pericytes (Crisan et al., 2008; Kunisaki et al., 2013). Consistent with the FACS analysis, *PDGFRA* was expressed by the $CD146^{-}$ mesenchyme at significantly higher levels and CD73 at lower levels than the CD146++ population (**Fig 1B**). HSPC regulatory genes such as *KITLG*, *IGFBP2* and *TNC*, were also expressed at significantly higher levels in CD146++ mesenchyme than CD146- mesenchyme.

Co-culture with $CD146^{++}$ mesenchyme supports HSPCs *ex vivo*

The ability of hEMP-derived $CD146^{++}$ and $CD146^{-}$ cells to support HSPCs *ex vivo* was assessed by co-culturing cord blood (CB)-derived CD34+ cells on monolayers generated from each mesenchymal population. To determine the specific effect of each stromal subset, short term (2 week) cultures were performed in basal medium with a low concentration of serum (5%) and without addition of hematopoietic cytokines.

Although total hematopoietic (CD45+ cell) output was similar with both mesenchymal populations (**Fig 2A**), co-cultures with $CD146^{++}$ cells retained a significantly higher frequency and number of CD34+ HSPC than co-cultures with $CD146^{-}$ mesenchymal cells (**Fig 2B, C**)

($p<0.01$ n=16). Immature HSPC (defined as CD34+lin- cells based on absence of the myeloid markers CD33 and CD14 and lymphoid markers CD10 and CD19) were also maintained at significantly higher numbers on CD146++ monolayers ($p<0.001$, n=16, **Fig 2D)**. Myeloid output was slightly higher in CD146++ co-cultures whereas more B lymphoid cells were generated on $CD146^-$ mesenchyme (**Fig 2E, F**). Consistent with CD34+ cell data, significantly more clonogenic cells (CFU-C) were maintained in $CD146^{++}$ co-cultures than CD146- co-cultures ($p<0.001$, n=16, **Fig 2G**). Of note, multiple other hPSC-derived mesenchyme phenotypic subsets were screened based on their CD146 and CD140a expression. Irrespective of CD140a expression, those subsets that expressed high levels of CD146 best supported CD34+Lin- cells (**Supplementary Fig 4**).

$CD146^{++}$ mesenchyme maintains human HSCs with repopulating ability and self-renewal potential

To assess support of functional HSPC *in vivo*, equal numbers of CD34+ cells were co-cultured with CD146++ or CD146- cells in the absence of exogenous cytokines. After 2 weeks, equal hematopoietic (CD45+) output from the co-cultures was transplanted into sublethally irradiated NOD/SCID/IL-2 receptor γ -/- (NSG) mice. Strikingly, mice transplanted with hematopoietic cells co-cultured with $CD146^{++}$ cells engrafted at significantly higher levels than those from $CD146^-$ cocultures ($P<0.001$ n=22 mice over 5 independent experiments, **Fig 2H**). Human $CD34^+$ progenitors, $CD19^+$ lymphoid cells and $CD14^+$ myeloid cells were detected in the bone marrow, spleen and peripheral blood demonstrating multilineage human reconstitution in the transplanted mice (**Fig 2I**). Self-renewal potential of HSPCs from $CD146^{++}$ co-cultures was demonstrated by multilineage engraftment in secondary NSG recipients (**Fig 2J**).

Together, these data indicate that at least two functionally distinct mesenchymal populations are generated from hPSC-derived mesoderm; high CD146 and CD73 expression marks a population of CD140a-/lo hPSC-derived mesenchyme capable of maintaining HSPC *ex vivo*, while CD146-CD140a+ mesenchyme have limited supportive capacity.

Hematopoietic support from hPSC-derived mesenchyme is sustained through inter-cellular interactions, mediated in part by Notch signaling

Prevention of direct contact between $CD146^{++}$ cells and CB CD34+ cells using a transwell culture system significantly reduced the number of $CD34+Lin^{-}$ cells maintained in culture ($P<0.05$, n=3, **Fig 3A**). Transwell culture with CD146- cells also reduced output of CD34+lin- cells although to a lesser extent. Thus intercellular mechanisms provide at least part of the HSPC supportive capacity of $CD146^{++}$ mesenchyme.

Notch signaling is a well-established cell contact-dependent paracrine factor and deletion of Jagged1 (encoded by *Jag1*) has been shown to result in premature exhaustion of the murine HSC pool (Houlihan et al., 2012; Morikawa et al., 2009; Poulos et al., 2013). RNA-Seq data showed both mesenchymal subsets expressed the Notch ligand *JAG1,* with higher levels in $CD146^{++}$ mesenchyme (**Fig 3B**); *DLL4* and *DLL1* were not expressed in either population (not shown). Addition of a neutralizing antibody to human JAG1 caused a profound (>300-fold) decrease in number of CD34+Lin- cells in CD146++ cocultures (**Fig 3C,** $P \leq 0.05$, n=3) and 250 fold decrease in CD146- co-cultures. Of note, cell output was not significantly reduced when human Jagged 1 blocking antibody was added to cocultures on OP9, a mouse stromal cell line, thus excluding non-specific cytotoxicity from the blocking antibody (**Fig 3D**). Thus HSPC survival in hPSC-mesenchymal cocultures is dependent on JAG1-mediated activation of Notch signaling.

Although Wnt inhibitors are highly expressed in CD146^{++} mesenchyme, impact on hematopoietic support is unclear

Across the entire RNA-Seq data set, the most differentially expressed gene between CD146^{++} and CD146^{-} mesenchyme was the Wnt antagonist secreted Frizzled-Related Protein 1 (*SFRP1)* (33 fold difference, FDR< 0.01) (**Fig 3E**). Other Wnt inhibitors were also significantly upregulated in CD146^{++} mesenchyme including *SFRP2* (FDR< 0.001) and Dickkopf 1 (*DKK1*) (FDR< 0.001) (**Fig 3E**) . The cysteine rich domain of sFRPs contains high homology to the Wnt-binding domain of the Frizzled receptors (Fzd), therefore they compete with the Fzd receptors for Wnt molecules and inhibit Wnt activation (Poulos et al., 2013; Schreck et al., 2014).

Despite addition of soluble SFRP1 had no effect on HSPC in either CD146++ or CD146- co-cultures suggesting that the differential expression of this wnt inhibitor is not directly related to hematopoietic support (Fig 3F). However addition of a Wnt agonist, the GSK-3 inhibitor CHIR99021 (CHIR), promoted rapid differentiation into both myeloid and lymphoid lineages and loss of CD34+Lin- cells in both CD146++ and CD146- co-cultures (**Fig 3G**). This observation is consistent with previous findings that canonical Wnt signaling regulates hematopoiesis in a dosage-dependent fashion, with high Wnt activation detrimental to HSC self-renewal leading to depletion of the LT-HSC pool (Luis et al., 2011). Overall, the high expression of Wnt inhibitors in CD146^{++} mesenchyme suggests that they could be involved in suppressing the level of Wnt signaling to promote HSC maintenance. Additional studies will be required to define the role of individual Wnt inhibitors during stroma coculture.

hPSC-derived CD146++ mesenchyme shares a coordinated gene expression profile that resembles primary human pericytes

Given the complexity of signals produced in the HSC niche, it seems likely that no single factor or mechanism mediates the differential support of hematopoiesis from different mesenchymal subpopulations. As molecular characterization and functional studies of hPSC-derived $CD146^{++}$ and CD146- mesenchyme showed strong parallels to populations from primary human tissue (fetal bone marrow and adult adipose tissue), we performed a comparative analysis of the transcriptomes of CD146++ and CD146- derived from hPSC-or primary adult tissue (adipose, either from fresh tissue isolation or after culture and expansion).

Gene set enrichment analysis (GSEA, Reference) was applied as follows. The most up- and down-regulated genes in hPSC-CD146++ cells (compared to hPSC-$CD146^{-}$ cells) were defined as *positive* and *negative signatures*, respectively, using FDR <0.01 and 2-fold difference as cutoff. These signatures of hPSC-derived CD146++ cells were compared to the expression profiles of primary CD146+ pericytes, ranked by differential expression to primary CD146- cells. Significant enrichment was found between the positive signatures of hPSC-derived CD146++ cells with genes upregulated in the adult pericytes (**Fig 4A**, enrichment score of 1.36 and 1.78 in fresh and cultured pericytes respectively). Similarly, the negative signatures of hPSC-CD146++ were significantly enriched in genes highly down-regulated in primary pericytes (**Fig 4A**, enrichment score of -1.77 and -1.39). Thus hPSC- and adipose-derived CD146++ cells exhibited similar transcriptional signatures.

The core transcriptional signature obtained from both hPSC- and adipose-derived CD146++ cells was imported into the STRING database to identify known and predicted protein interactions.

This set of genes showed significantly more interactions than would be expected for a random set of proteins of similar size drawn from the genome (**Fig 4B**). Of note, the dominant functional categories identified in the STRING analysis were those enriched in biological processes characteristic of pericytes (i.e. cell adhesion, vasculature development, regulation of cell motion and wound healing) (**Fig 4B**).

Discussion

Mesenchymal stroma cells, either derived from primary tissues or differentiated from hPSC (Corselli et al., 2013; Dar et al., 2012), are conventionally isolated and expanded as a heterogeneous population of cells loosely defined by CD73, CD105 and CD90 expression (Corselli et al., 2013; Murray et al., 2014). Our data strongly argues that phenotypically, molecularly and functionally distinct populations coexist within these conventional definitions of mesenchymal stroma. Previously, CD146 was shown to define mesenchymal precursors residing on the vasculature of multiple human organs including bone marrow and adipose tissues (Corselli et al., 2013; Crisan et al., 2009). Additionally, we showed that purified human CD146+ perivascular cells from these primary tissues, compared with unfractionated mesenchymal stroma and CD146− cells, were able to uniquely sustain human HSPCs in coculture in the absence of exogenous cytokines (Corselli et al., 2013). We have now expanded our findings, demonstrating that $CD146^{++}CD73^{++}CD140a^{-}$ serves as distinguishing cell surface phenotypes to isolate subsets from hPSC-derived mesenchyme that support human HSPC ex vivo.

hPSC-derived $CD146^{++}$ mesenchyme showed high expression in *NES, LEPR* and *CSPG4* markers of peri-sinusoidal and peri-arteriolar cells revealed by seminal papers characterizing the murine HSC niche (Dar et al., 2012; Ding et al., 2012; Kunisaki et al., 2013; Méndez-Ferrer et

al., 2010; Zhou et al., 2014). The relevance of these cellular constituents has just begun to be addressed in the context of human hematopoiesis (Corselli et al., 2013; Murray et al., 2014; Pinho et al., 2013). In the present study, we have used hPSC-derived mesenchyme as a platform to dissect the mechanisms underlying human HSC maintenance. Both $CD146^{++}$ and CD146- mesenchyme expressed known HSC niche genes such as *SLIT2* an *ANGPT1*. However other niche factors such as such as *KITL, JAG1* and *IGFBP2* were specifically upregulated in CD146++ cells. As is likely for the components of the hematopoietic niche in vivo, we conclude the CD146++ mesenchymal subset supports HSPC through a combination of mechanisms that involve cell-to-cell contact (e.g. Notch signaling and adhesion molecules) and secreted factors (e.g. Wnt signaling and cytokines). On the other hand, the $CD146^{-}$ subset of hPSC-derived mesenchyme drove differentiation and favored lymphoid commitment, consistent with our prior studies with primary human mesenchyme (Corselli et al).

In summary, we demonstrate that $CD146^{++}$ population is a molecularly and functionally distinct subset of hPSC-derived mesenchyme able to directly support the ex vivo maintenance of human HSPCs. The ability to prospectively isolate a human stroma with HSC-supporting capacity is particularly relevant given clinical interest in using stroma coculture as platform for HSC expansion (Crisan et al., 2009). CD146+ population has also been reported to resolve other functional heterogeneity within primary human mesenchymal populations for better therapeutic potential in immune modulation (Corselli et al., 2013; Wu et al., 2016) and tissue regeneration (Russell et al., 2010).

While this study focused on defining cellular subsets within hPSC-derived mesenchyme that would support a well-characterized HSPC population, the existence of distinct mesenchymal

subsets during hPSC differentiation has broader implication for the directed differentiation of hPSC towards definitive, functional HSC. Further study will aim to define optimal hPSC-derived mesenchyme that would promote HSC specification. Clarifying and defining heterogeneities among mesenchyme could lead to improved cell-based tissue regeneration strategies and a better understanding of human developmental processes.

Materials And Methods

Mesoderm differentiation from hPSC to generate hEMP

The hPSC line H1 (WiCell, Madison, WI) and UCLA 3 (UCLA hESC Core Bank) were maintained and expanded on irradiated primary mouse embryonic fibroblasts (EMD Millipore, Billerica, MA). Mesoderm commitment was induced as previously described (see supplemental methods). $CD326^{-}CD56^{+}$ embryonic mesoderm progenitors (hEMP) were isolated by flow cytometry at day 3.5 (**Fig. 1A**) and seeded onto matrigel for endothelial and mesenchymal differentiation over the next 14 days.

Mesenchymal differentiation from hEMP

To induce mesenchymal differentiation, hEMP generated were seeded at 50,000 cells/well into 6-well tissue culture plates, pre-coated with matrigel for two hours. For the first 7 days of differentiation, MesenCult Proliferation Kit, Human (STEMCELL Technologies) supplemented with ALK 4/5/7 blocker SB-431542 (10urs. &D Systems) was used. During day 1 to 3, 10μM Rock inhibitor (Y27632 hydrochloride; R&D Systems) and 10 μg/mL gentamicin (Gibco, Thermo Fisher Scientific, Waltham, MA) were also added. From day 7-14, the media was switched to EGM-2 complete medium (Lonza) supplemented with SB-431542 (102 (10&D Systems). Half medium change was performed every other day. At day 14-18, cells were

dissociated into single cell suspension with Accutase (Innovative Cell Technologies, San Diego, CA), stained with monoclonal antibodies and analyzed or isolated by FACS (see supplemental methods).

Coculture of stromal cells and CB CD34$^+$ cells

Freshly sorted hEMP-derived mesenchyme subset (CD146^{++} and CD146$^-$) were plated on 96 well tissue culture-treated plates at 6.5×10^3 cells per well in 100 μ l of EGM-2. Three days later, CB CD34$^+$ cells (1×10^5 per well) were plated on top of the stromal layer (80% confluent) in 200 μ l of coculture media. The CB CD34$^+$ cells were FACS to ensure >80% CD34$^+$ purity prior to initiating coculture. Coculture media consists of RPMI 1640 with L-glutamine, 5% fetal bovine serum, 1× penicillin/streptavidin. No supplemental cytokines were ever added. Half media change was performed every 4 days. Cells were harvested after 2 weeks and were subjected to FACS analysis, CFU assay and in vivo repopulating assay (see Supplemental methods). Description of transwell, Notch and Wnt pathways perturbations can be found in Supplemental methods.

Graphical and Statistical Analysis

Graphs were generated and statistics analyzed using GraphPad Prism software. Unpaired parametric two-tailed t-tests were used to calculate *p*-values. $p < 0.05$ was considered statistically significant.

References

Chin, C.J., Cooper, A.R., Lill, G.R., Evseenko, D., Zhu, Y., He, C.B., Casero, D., Pellegrini, M., Kohn, D.B., and Crooks, G.M. (2016). Genetic Tagging During Human Mesoderm Differentiation Reveals Tripotent Lateral Plate Mesodermal Progenitors. Stem Cells *34*, 1239-1250.

Corselli, M., Chin, C.J., Parekh, C., Sahaghian, A., Wang, W., Ge, S., Evseenko, D., Wang, X., Montelatici, E., Lazzari, L.*, et al.* (2013). Perivascular support of human hematopoietic stem/progenitor cells. Blood *121*, 2891-2901.
Crisan, M., Chen, C.W., Corselli, M., Andriolo, G., Lazzari, L., and Péault, B. (2009). Perivascular multipotent progenitor cells in human organs. Ann N Y Acad Sci *1176*, 118-123.
Crisan, M., Yap, S., Casteilla, L., Chen, C.W., Corselli, M., Park, T.S., Andriolo, G., Sun, B., Zheng, B., Zhang, L.*, et al.* (2008). A perivascular origin for mesenchymal stem cells in multiple human organs. Cell Stem Cell *3*, 301-313.
Dar, A., Domev, H., Ben-Yosef, O., Tzukerman, M., Zeevi-Levin, N., Novak, A., Germanguz, I., Amit, M., and Itskovitz-Eldor, J. (2012). Multipotent vasculogenic pericytes from human pluripotent stem cells promote recovery of murine ischemic limb. Circulation *125*, 87-99.
Ding, L., Saunders, T.L., Enikolopov, G., and Morrison, S.J. (2012). Endothelial and perivascular cells maintain haematopoietic stem cells. Nature *481*, 457-462.
Evseenko, D., Zhu, Y., Schenke-Layland, K., Kuo, J., Latour, B., Ge, S., Scholes, J., Dravid, G., Li, X., MacLellan, W.*, et al.* (2010). Mapping the first stages of mesoderm commitment during differentiation of human embryonic stem cells. Proceedings of the National Academy of Sciences of the United States of America *107*, 13742-13747.
Houlihan, D.D., Mabuchi, Y., Morikawa, S., Niibe, K., Araki, D., Suzuki, S., Okano, H., and Matsuzaki, Y. (2012). Isolation of mouse mesenchymal stem cells on the basis of expression of Sca-1 and PDGFR- α . Nat Protoc *7*, 2103-2111.
Kunisaki, Y., Bruns, I., Scheiermann, C., Ahmed, J., Pinho, S., Zhang, D., Mizoguchi, T., Wei, Q., Lucas, D., Ito, K.*, et al.* (2013). Arteriolar niches maintain haematopoietic stem cell quiescence. Nature *502*, 637-643.
Luis, T.C., Naber, B.A., Roozen, P.P., Brugman, M.H., de Haas, E.F., Ghazvini, M., Fibbe, W.E., van Dongen, J.J., Fodde, R., and Staal, F.J. (2011). Canonical wnt signaling regulates hematopoiesis in a dosage-dependent fashion. Cell Stem Cell *9*, 345-356.
Morikawa, S., Mabuchi, Y., Kubota, Y., Nagai, Y., Niibe, K., Hiratsu, E., Suzuki, S., Miyauchi-Hara, C., Nagoshi, N., Sunabori, T.*, et al.* (2009). Prospective identification, isolation, and systemic transplantation of multipotent mesenchymal stem cells in murine bone marrow. J Exp Med *206*, 2483-2496.
Murray, I.R., West, C.C., Hardy, W.R., James, A.W., Park, T.S., Nguyen, A., Tawonsawatruk, T., Lazzari, L., Soo, C., and Péault, B. (2014). Natural history of mesenchymal stem cells, from vessel walls to culture vessels. Cell Mol Life Sci *71*, 1353-1374.
Méndez-Ferrer, S., Michurina, T.V., Ferraro, F., Mazloom, A.R., Macarthur, B.D., Lira, S.A., Scadden, D.T., Ma'ayan, A., Enikolopov, G.N., and Frenette, P.S. (2010). Mesenchymal and haematopoietic stem cells form a unique bone marrow niche. Nature *466*, 829-834.
Pinho, S., Lacombe, J., Hanoun, M., Mizoguchi, T., Bruns, I., Kunisaki, Y., and Frenette, P.S. (2013). PDGFR α and CD51 mark human nestin+ sphere-forming mesenchymal stem cells capable of hematopoietic progenitor cell expansion. J Exp Med *210*, 1351-1367.
Poulos, M.G., Guo, P., Kofler, N.M., Pinho, S., Gutkin, M.C., Tikhonova, A., Aifantis, I., Frenette, P.S., Kitajewski, J., Rafii, S.*, et al.* (2013). Endothelial Jagged-1 is necessary for homeostatic and regenerative hematopoiesis. Cell Rep *4*, 1022-1034.
Russell, K.C., Phinney, D.G., Lacey, M.R., Barrilleaux, B.L., Meyertholen, K.E., and O'Connor, K.C. (2010). In vitro high-capacity assay to quantify the clonal heterogeneity in trilineage potential of mesenchymal stem cells reveals a complex hierarchy of lineage commitment. Stem Cells *28*, 788-798.

Sacchetti, B., Funari, A., Michienzi, S., Di Cesare, S., Piersanti, S., Saggio, I., Tagliafico, E., Ferrari, S., Robey, P.G., Riminucci, M., *et al.* (2007). Self-renewing osteoprogenitors in bone marrow sinusoids can organize a hematopoietic microenvironment. Cell *131*, 324-336.
Schreck, C., Bock, F., Grziwok, S., Oostendorp, R.A., and Istvánffy, R. (2014). Regulation of hematopoiesis by activators and inhibitors of Wnt signaling from the niche. Ann N Y Acad Sci *1310*, 32-43.
Vodyanik, M.A., Yu, J., Zhang, X., Tian, S., Stewart, R., Thomson, J.A., and Slukvin, I.I. (2010). A mesoderm-derived precursor for mesenchymal stem and endothelial cells. Cell Stem Cell *7*, 718-729.
Wanjare, M., Kusuma, S., and Gerecht, S. (2014). Defining differences among perivascular cells derived from human pluripotent stem cells. Stem Cell Reports *2*, 561-575.
Wu, C.C., Liu, F.L., Sytwu, H.K., Tsai, C.Y., and Chang, D.M. (2016). CD146(+) mesenchymal stem cells display greater therapeutic potential than CD146(-) cells for treating collagen-induced arthritis in mice. Stem Cell Res Ther *7*, 23.
Zhou, B.O., Yue, R., Murphy, M.M., Peyer, J.G., and Morrison, S.J. (2014). Leptin-receptor-expressing mesenchymal stromal cells represent the main source of bone formed by adult bone marrow. Cell Stem Cell *15*, 154-168.

SUPPLEMENTAL METHODS

hESC culture and mesoderm induction conditions

The hESC line H1 (WiCell, Madison, WI) was maintained and expanded on irradiated primary mouse embryonic fibroblasts (EMD Millipore, Billerica, MA). To induce mesoderm differentiation, hESC colonies were cut into uniform-sized pieces using the StemProEZPassage tool (Invitrogen, Thermo Fisher Scientific, Waltham, MA), transferred into 6-well plates pre-coated for 1 hour with Matrigel (growth factor reduced, no phenol red; BD Biosciences, San Jose, CA), and cultured initially in mTESR medium (Stem Cell Technologies, Vancouver, BC) until 50–60% confluent (typically 3 days). To initiate differentiation, mTESR medium was replaced with basal induction medium X-Vivo 15 (Lonza, Walkersville, MD). Basal induction medium was supplemented with human growth factors activin A, BMP4, VEGF and bFGF for 3.5 days (all at 10 ng/mL; R&D Systems, Minneapolis, MN), with the inclusion of activin A only on day 1 ("A-BVF" condition [1]). After mesoderm induction, $CD326^{-}CD56^{+}$ cells were isolated by fluorescence activated cell sorting (FACS) at day 3.5 (typically 5-10% of total population,

Figure 1A) and placed onto matrigel coated 6 well plates for further differentiation over the next 14 days.

Flow cytometry and cell sorting

Flow cytometry analysis was performed on LSRII or LSRFortessa and cell sorting on a FACSAria II (all from Becton Dickinson). Cultured cells were dissociated into single cell suspension with Accutase (Innovative Cell Technologies, San Diego, CA) and immunofluorescence staining was performed with human-specific monoclonal antibodies. Nonspecific binding was blocked with intravenous immunoglobulin (1%) (CSL Behring, King of Prussia, PA) before staining with fluorochrome-conjugated antibodies. Unstained cells and fluorescence-minus-one (FMO) were used as controls for gating. Cell acquisition was performed using FACSDiva (Becton Dickinson) and the analysis was performed using FlowJo (Tree Star, Ashland, OR). Forward and side scatter (FSC/SSC) and 4',6-diamidino-2-phenylindole (DAPI) (Invitrogen) were used to identify live cells.

Immunophenotype analysis of hEMP-derived mesenchyme

hEMP-derived mesenchyme at d14 of differentiation were harvested and analyzed on an LSR II flow cytometer (Becton Dickinson). Cells were stained with monoclonal antibodies: CD146-FITC (AbD Serotec), CD73-PE Cy7 (BD Biosciences), CD45-BV711 (Biolegend), CD31-APC (Biolegend), CD144-PE (BD Biosciences), CD34-PE Cy7 (BD Biosciences), CD105-PE (Biolegend), CD90-PE (Biolegend), CD44-PE (Biolegend), CD140a-PE (BD Biosciences), CD140b-PE (Biolegend), CSPG4-PE (BD Biosciences). Unstained cells and fluorescence-minus-one (FMO) were used as controls for gating.

To isolate the mesenchymal population by FACS, the following gating sequence was used: CD31-APC (Biolegend) identified the endothelial population. From the non-endothelial compartment ($CD31^-$), the mesenchymal population was identified based on CD73-PE-Cy7 (BD Biosciences) positivity ($CD31^-CD73^+$). The CD146-FITC (AbD Serotec) bright cells were gated as $CD146^{++}$ mesenchyme, and CD146-FITC negative cells were gated as $CD146^-$ mesenchyme.

Flow cytometric analysis of cultured CB $CD34^+$ cells

After 2 weeks of coculture, cells were harvested and stained with the following antibodies: CD45-BV711 (Biolegend), CD34-APC Cy7 (Biolegend), CD14-BV605 (Biolegend), CD10-PE Cy7 (Biolegend), CD33-PE (BD Biosciences), CD19-BV510 (Biolegend).

Colony forming unit assay

After 2 weeks of coculture cells were harvested and 2.5×10^3cells were plated in methylcellulose (Methocult 4434; Stem Cell Technologies). Colonies, here reported as the sum of the progeny of colony forming unit (CFU) granulo-macrophage, burst-forming unit erythroid, and CFU mixed, were scored after 14 days.

In vivo repopulation assay

CB $CD34^+$ cells were cocultured with $CD146^{++}$ or $CD146^-$ cells for 2 weeks. After FACS analysis, an equal number of $CD45^+$ cells (1×10^5) was obtained from the cocultures and was intratibially injected in sublethally irradiated (250 cGy), 6- to 8-week-old NSG mice (The Jackson Laboratory). Mice were sacrificed 6 weeks posttransplantation. Engraftment of human hematopoietic cells was evaluated by FACS analysis after staining with anti-human specific monoclonal antibodies: CD45-APC Cy7, HLA (A/B/C)-PE, CD34-PE Cy7, CD10-FITC, CD19-

FITC, CD14-APC (all from BD Biosciences). For secondary transplantation, bone marrows of left and right tibia and femur from engrafted mice were pooled and intratibially injected into a secondary host. Engraftment was evaluated 6 weeks posttransplantation.

For the animal studies, an animal care and use committee protocol (ARC no. 2008-175-11) was approved for the injection of human cells into immunodeficient mice and for the analysis of engraftment of transplanted cells.

Coculture of stromal cells and CB CD34$^+$ cells

Cocultures in the absence of cell-to-cell contact were performed in HTS Transwell 96 Permeable Support Culture Plate System, 0.4μm Polycarbonate Membrane (Corning). Contact= mesenchyme and CB CD34+ both plated on the bottom well. No Contact= mesenchyme plated on the bottom well, CB CD34+ cells plated on the top well.

For the inhibition of Notch, 10μM γ-secretase inhibitor (GSI), L-685, 458 (R&D Systems) or 1 μg/ml of anti-Jag1 N-17 (sc-34473, Santa Cruz Biotechnology) were added to each well every 48 hours. An equal volume of dimethylsulfoxide (DMSO; Sigma) or an equal concentration of irrelevant Goat IgG (Sigma I9140) was added to wells as negative controls for GSI and anti-Jagged1 antibody, respectively.

For analysis of sFRP-1 and activation of Wnt signaling, 2 μg/ml recombinant human sFRP-1 protein (R&D Systems) or 3uM CHIR99021 (Tocris Bioscience) were added to each well every 48 hours. An equal volume of dimethylsulfoxide (DMSO; Sigma) was added to wells as negative control for CHIR99021 treatment.

Cell cycle analysis

At day 14 of mesenchymal differentiation of hEMP, 1 mM BrdU (diluted in EGM-2 supplemented with SB-431542) was added to the total culture for 40 minutes before harvested for imunophenotype analysis. BrdU incorporation was measured by flow cytometry using the FITC BrdU Flow Kit (BD Biosciences).

The cell-cycle phases were defined as apoptotic ($BrdU^{-}$ and DNA<2n), G0/G1 ($BrdU^{-}$ and 2n DNA), S (BrdU+ and 2n≤DNA≤ 4n) and G2/M ($BrdU^{-}$ and 4n DNA).

Quantitative RT-PCR

Freshly sorted hEMP-derived mesenchyme subset ($CD146^{++}$ and $CD146^{-}$) or those cultured 3-4 days in EGM-2 were processed for RNA extraction using a Qiagen micro kit. An Omniscript reverse transcriptase (RT) kit was used to make complementary DNA, which was subjected to quantitative polymerase chain reaction (qPCR) using Taqman probe-based gene expression assay (Applied Biosystems) and β2 microglobulin (β2M) as housekeeping gene. Best coverage primer/probe sets were selected for all target genes (*MCAM, IGFBP2, KITLG, LEPR, NES* and *CSPG4*). A 7500 real-time PCR system was used (ABI). Data were analyzed using the comparative ΔΔCT method.

Isolation of human primary pericytes from lipoaspirate and RNA-Seq

Human pericytes were derived from human lipoaspirate specimens obtained as discarded anonymous waste tissue and thus deemed exempt from institutional review board approval. One hundred milliliters of lipoaspirate were incubated at 37°C for 30 minutes with digestion solution composed by RPMI 1640 (Cellgro), 3.5% bovine serum albumin (Sigma), and 1 mg/mL collagenase type II (Sigma). Adipocytes were discarded after centrifugation while the pellet was

resuspended and incubated in red blood cell lysis (eBioscience) to obtain the stromal vascular fraction (SVF). An aliquot of SVF was plated in tissue culture-treated flask for the expansion of conventional Adipose-derived Stem Cells (ASCs). Another aliquot of SVF was processed for fluorescence-activated cell sorting (FACS).

Cells were incubated with the following antibodies: CD45-APC Cy7 (BD Biosciences), CD34-APC (BD Biosciences), and CD146-FITC (AbD Serotec). The viability dye 4,6 diamidino-2-phenylindole (DAPI; Sigma) was added before sorting on a FACS Aria III (BD Biosciences). $DAPI^{-}CD45^{-}CD34^{-}CD146^{+}$ were gated as pericytes.

For RNA-Seq, lipoaspirate-derived pericytes were sorted from two donors and cultured for 3 passages in DMEM containing 20% FBS. Following three passages in culture the pericytes were trypsinized, and those cells that retained CD146 expression were FACS collected for RNA extraction. Fifty nanograms of total RNA was extracted from cultured-pericytes obtained from each of two donors and combined to yield a pooled sample of 100ng of total RNA. ASCs similarly cultured were collected for RNA seq as well. The samples were sequenced on an Illumina HiSeq 2000 using paired-end 100bp reads to a depth of coverage of ~ 10-15 million reads.

Isolation of human $CD34^{+}$ cells from CB

Umbilical cord blood (CB) was collected from normal deliveries without individually identifiable information, therefore no institutional review board approval was required. MNCs were isolated by density gradient centrifugation using Ficoll-Paque (GE Healthcare). Enrichment of $CD34^{+}$ cells was then performed using the magnetic-activated cell sorting system (Miltenyi Biotec) as per the manufacturer's instructions.

RNA-Seq and analysis

hEMP ($CD326^{-}CD56^{+}$ cells) and freshly sorted day 14 hEMP-derived mesenchyme subset ($CD146^{++}$ and $CD146^{-}$) were extracted with Trizol and purified using miRNeasy Mini Kit (Qiagen). Three biological replicates were obtained for each population. 500ng-2ug of total RNA was input to generate cDNA using Nugen Ovation RNA-Seq System v2 and the sequencing libraries were generated using prepX DNA library enzyme kit (IntegenX Inc.) per manufacturer's instructions. Paired-end 100bp sequencing was performed on Illumina HiSeq 2000 with six samples multiplexed per lane. Raw sequence files were obtained using Illumina's proprietary software and are available at NCBI's Gene Expression Omnibus (accession number GSEXXXXX).

RNA-Seq reads were aligned using STAR v2.3.0 [2]. The GRCh37 assembly (hg19) of the human genome and the corresponding junction database from Ensemble's gene annotation were used as reference for STAR. The count matrix for genes in the Ensembl genome annotation was generated with HTSeq-count v0.6.1p2 [3]. DESeq v1.14.0 [4] was used for normalization (using the geometric mean across samples), differential expression (to classify genes as differentially expressed, Benjamini-Hochberg adjusted p-value < 0.01) and to compute moderate expression estimates by means of variance-stabilized data. Heatmaps were built with GENE-E (http://www.broadinstitute.org/cancer/software/GENE-E/) applying relative min/max normalization to moderate expression estimates.

Gene Set Enrichment Analysis and STRING protein interaction analysis

Gene Set Enrichment Analysis (GSEA) (Subramanian et al., 2005) was used to perform the comparative analysis between the RNA-Seq expression signatures of hEMP-derived and lipoaspirate pericytes. The differentially expressed genes in hPSC-derived $CD146^{++}$ vs $CD146^{-}$

mesenchyme (fold change>2, FDR<0.01) defined the gene set of interest (*positive signature*: upregulated genes (n=540); *negative signature*: downregulated genes (n=351)). In both cases, the background for enrichment consisted of a ranked gene list where top ranking genes were more expressed in lipoaspirate-derived CD146+ pericytes as compared to CD146- cells. GSEA was run independently using both freshly sorted and cultured cells as background. For the positive (negative) signature, significant enrichment was only found for genes up- (down-) regulated in both freshly sorted and cultured lipoaspirate-derived CD146+ pericytes, The core gene sets provided by GSEA (highlighted in grey in Fig 4A) included those genes that are consistently up or downregulated in $CD146^{++}$ from all sources. This final $CD146^{++}$ gene expression signature was then imported into the STRING database (Jensen et al., 2009). For protein interaction analyses, STRING (version 10) was used with default parameters. GO enrichment analysis was performed to identify the biological processes involved in the predicted STRING protein network.

1. Evseenko D, Zhu Y, Schenke-Layland K et al. Mapping the first stages of mesoderm commitment during differentiation of human embryonic stem cells. **Proceedings of the National Academy of Sciences of the United States of America**. 2010;107:13742-13747.
2. Dobin A, Davis CA, Schlesinger F et al. STAR: ultrafast universal RNA-seq aligner. **Bioinformatics**. 2013;29:15-21.
3. Anders S, Pyl PT, Huber W. HTSeq-a Python framework to work with high-throughput sequencing data. **Bioinformatics**. 2015;31:166-169.
4. Anders S, Huber W. Differential expression analysis for sequence count data. **Genome Biol**. 2010;11:R106.
5. Subramanian A, Tamayo P, Mootha VK et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. **Proc Natl Acad Sci U S A**. 2005;102:15545-15550.
6. Jensen LJ, Kuhn M, Stark M et al. STRING 8--a global view on proteins and their functional interactions in 630 organisms. **Nucleic Acids Res**. 2009;37:D412-416.

Fig. 1

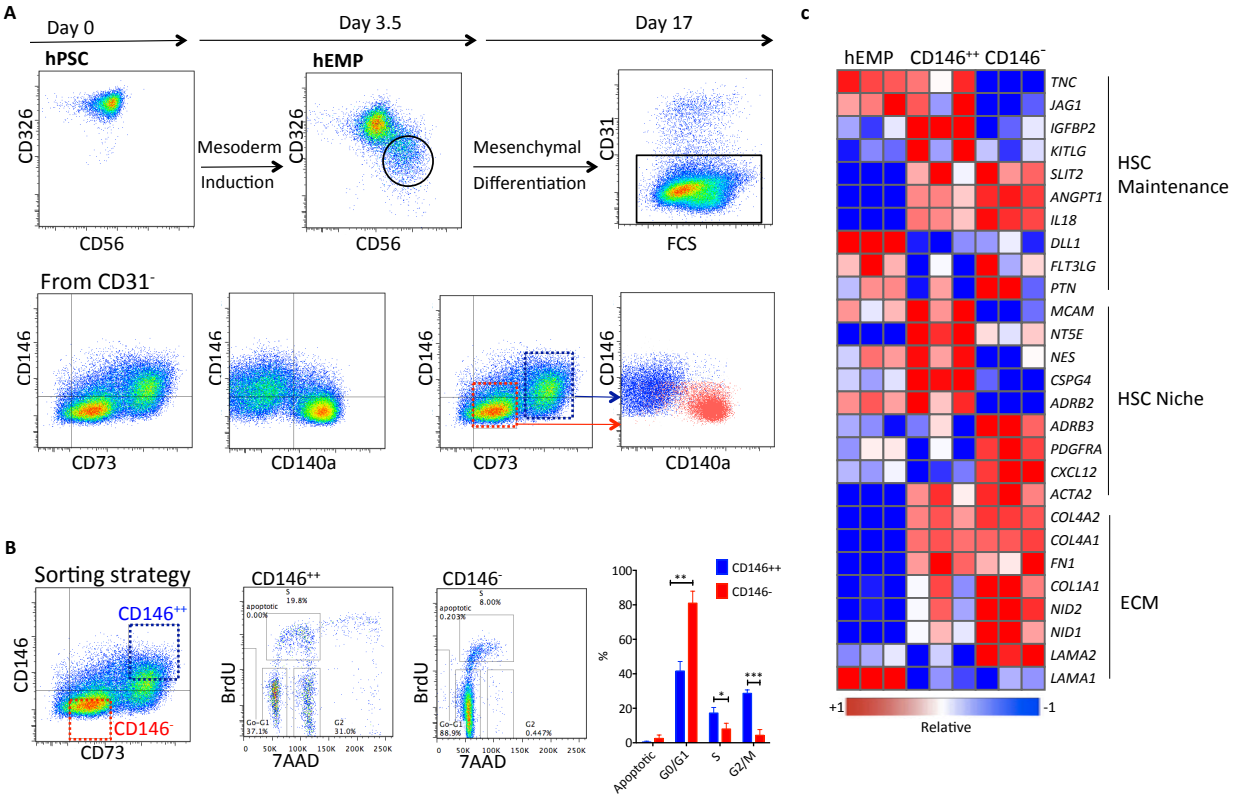


Fig 1. Immunophenotypic fractionation of hPSC-derived mesenchyme reveals transcriptionally distinct populations. (A) Embryonic Mesoderm Progenitors (hEMP) were generated from hPSC (H1 cells) and isolated by FACS for further differentiation in mesenchymal conditions. After two weeks, two distinct populations were identified within the CD31- (non-endothelial) cells;. CD146-CD73^{dim} cells (red box) were CD140a+, and CD146+CD73++ cells (blue box) were CD140a- (representative FACS plot; n=16 independent experiments). (B) Cell cycle analysis of the hEMP-derived mesenchymal differentiation cultures at week 2, pulsed with BrdU for 40 minutes, and gated on CD31- mesenchyme as either "CD146^{++}" or "CD146^{-}.". Representative FACS, n=3 * p<0.05, ** p<0.01, *** p< 0.001. (C) Transcriptional profiling of hEMP, and CD146^{++} and CD146^{-} mesenchyme using RNA-Seq. Heatmaps for selected genes are shown. Color scale shows relative expression for each gene using its min/max moderate expression estimates as reference. n=3 biological replicates

Fig. 2

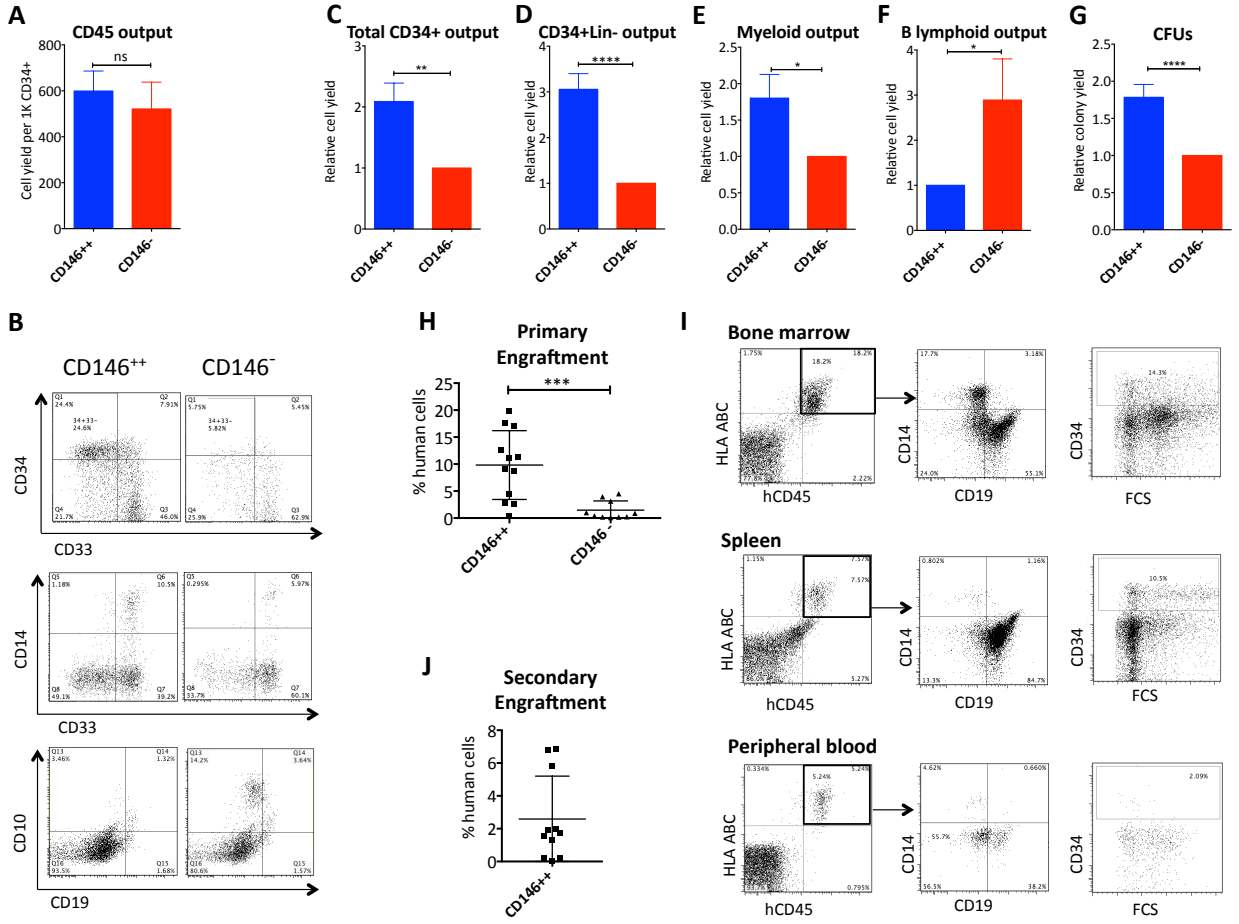


Fig. 2. CD146^{++} mesenchyme supports ex vivo maintenance of clonogenic, engraftable and self-renewing HSCs. CD146++ and CD146- mesenchymal cells generated from H1-derived hEMP, were isolated and co-cultured as monolayers with cord blood CD34+ cells in 5% serum without added cytokines. After 2 weeks , hematopoietic cells were assayed in vitro and in vivo. (A) Total number of DAPI-CD45+ cells recovered from cocultures, shown as cell yield normalized to 1,000 input CD34+ cells seeded Day 0. (B) FACS analysis of viable hematopoietic cells (gated on DAPI-CD45+) from co-cultures. (C-G) Cell yield of CD34^{+}, CD34+Lin−, CD14+ myeloid, CD10+/CD19+ B lymphoid cells and CFU (after re-plating at 2 weeks) recovered from CD146^{++} normalized to yield from CD146^{-} cocultures. n = 16 independent experiments, each experiment performed in duplicate. All data presented as mean ± SEM. * $p<0.05$, ** $p<0.01$, *** $p< 0.001$, **** $p< 0.0001$; unpaired two-tailed Student's t test. (H) Primary engraftment (%human cells defined as HLA-Class1+ hCD45+) of mice 6 weeks after transplant with equal number of CD45+ cells (1x10^{5}) obtained from CD146^{++} mesenchyme coculture or CD146^{-} cocultures (n = 5 independent experiments, 10-12 mice per group; ***P < .0001). (I) Representative flow cytometry for the detection of human cells and multilineage reconstitution in primary transplantation. (J) Secondary transplantation of bone marrow from primary chimeric mice transplanted with CD146^{++} cell coculture (6 weeks post transplantation, n = 5 independent experiments).

Fig. 3

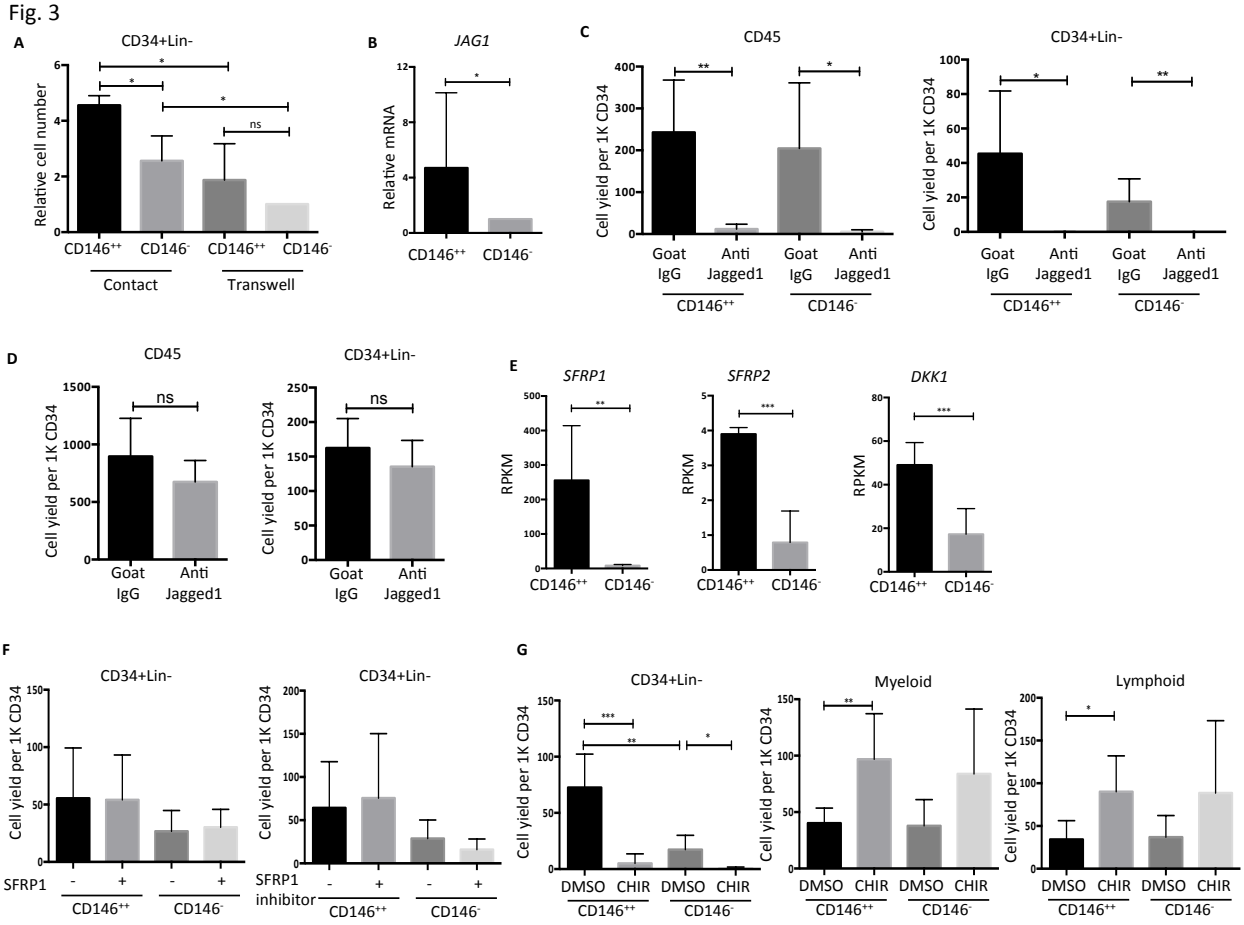


Fig 3. Notch inhibition and Wnt signaling activation in PSC-mesenchyme cocultures leads to loss of HSPC. (A) Cell-cell contact is required for ex vivo HPSC support on PSC-mesenchyme. Shown is relative number of CD34+Lin- cells in contact and transwell coculture recovered from $CD146^{++}$ normalized to yield from $CD146^{-}$ cocultures (n=3 independent experiments). All data are presented as mean ± SEM. * $p<0.05$ (B) Quantitative RT-PCR of *JAG1* expression on $CD146^{++}$ and $CD146^{-}$ mesenchyme (n=14 biological replicates, * $p<0.05$). (C,D) Total and CD34+Lin− cell number (relative to goat IgG control) with addition of human JAG1 blocking antibody to (C) PSC-mesenchyme coculture (*$p<0.05$, n=3 independent experiments) or (D) OP9 coculture (n=4 independent experiments). (F) *SFRP1*, *SFRP2* and *DKK1* expression on $CD146^{++}$ and $CD146^{-}$ mesenchyme (n=3 biological replicates, RNA-Seq ** FDR<0.01, ***FDR<0.001). (F) CD34+Lin− cell numbers after addition of soluble SFRP1 or SFPR1 inhibitor (n=3 independent experiments). (G) CD34+Lin−, myeloid and lymphoid cell numbers after activation of Wnt signaling by addition of 3 μ M CHIR99021 (* $p<0.05$, n=3 independent experiments).

Fig. 4

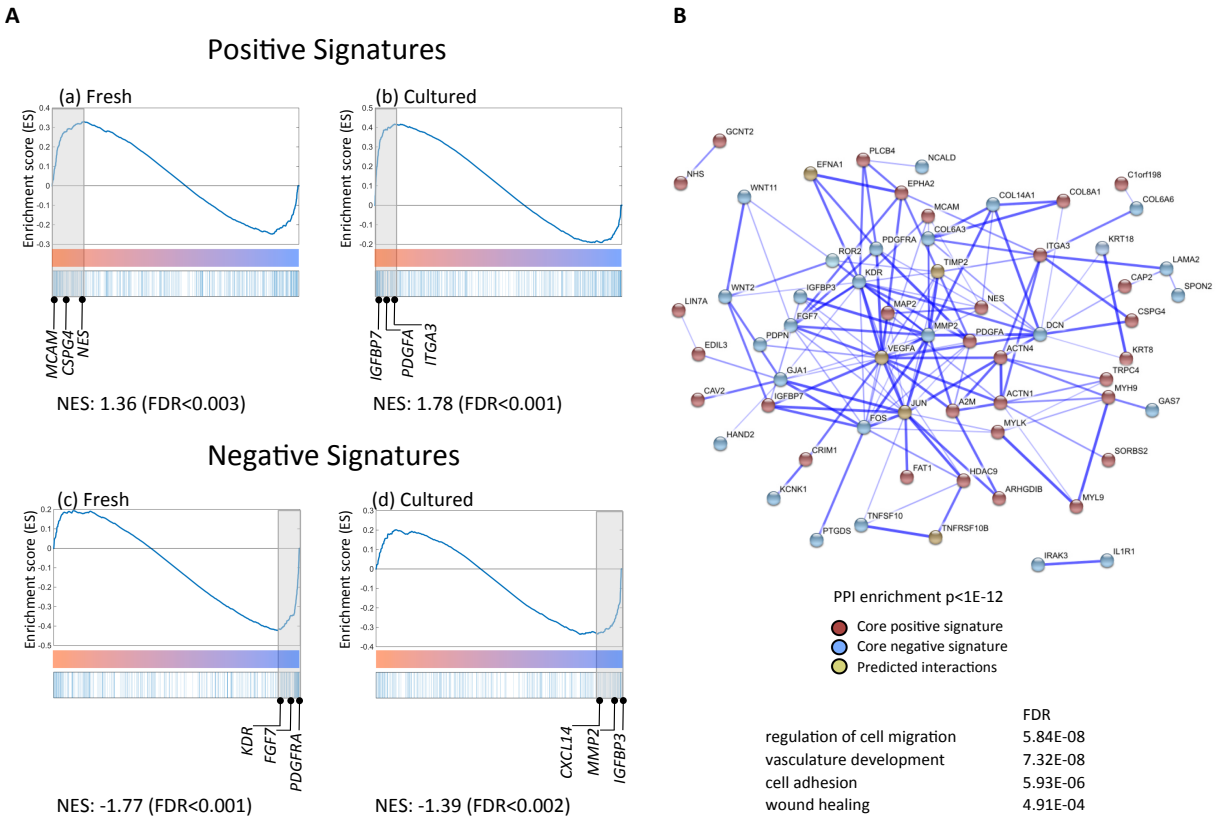
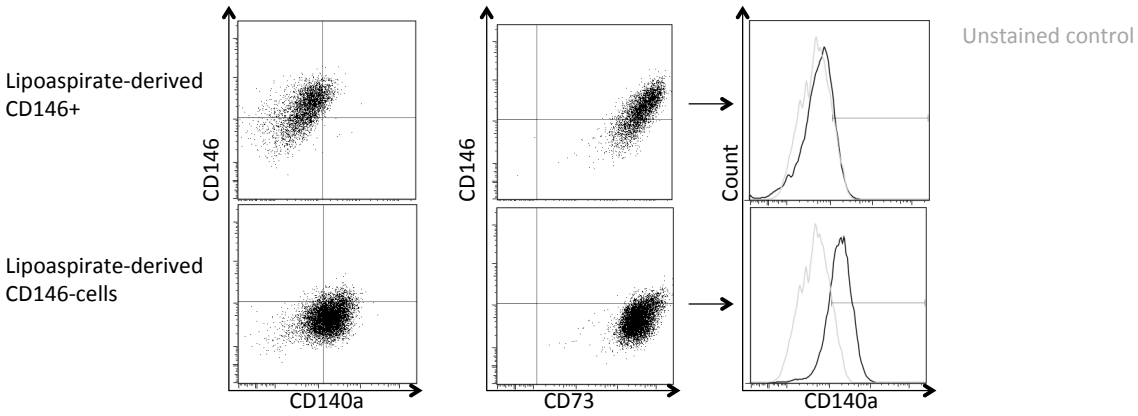


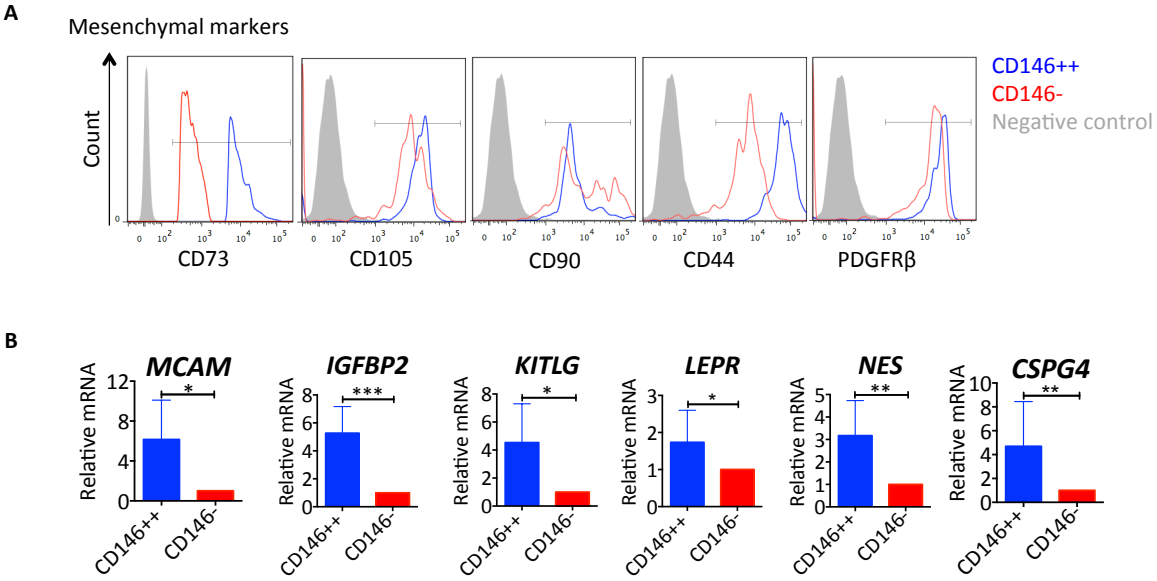
Fig 4. hPSC-derived CD146++ mesenchyme shares a molecular signature with primary human adult pericytes. (A) Gene set enrichment analysis (GSEA) of hPSC-derived mesenchyme and lipoaspirate-derived populations (both freshly sorted and culture passaged). Positive and negative signatures (vertical lines) were defined as up- and down-regulated genes in hPSC-derived CD146^{++} compared to hPSC-derived CD146^{-}. X-axis represents genes ranked from positive to negative fold change in lipoaspirate-CD146+ compared to CD146- cells. Genes in the core enrichment set are highlighted (shaded grey boxes). (B) Up and downregulated genes common to the core sets from both fresh and cultured cells were imported into the STRING database, which identified a network of known protein-protein interactions (PPI enrichment p-value = 0), along with a number of high-confidence (score>0.9999) predicted interactions. Line thickness is proportional to the confidence level of each interaction.

Supplementary Fig 1



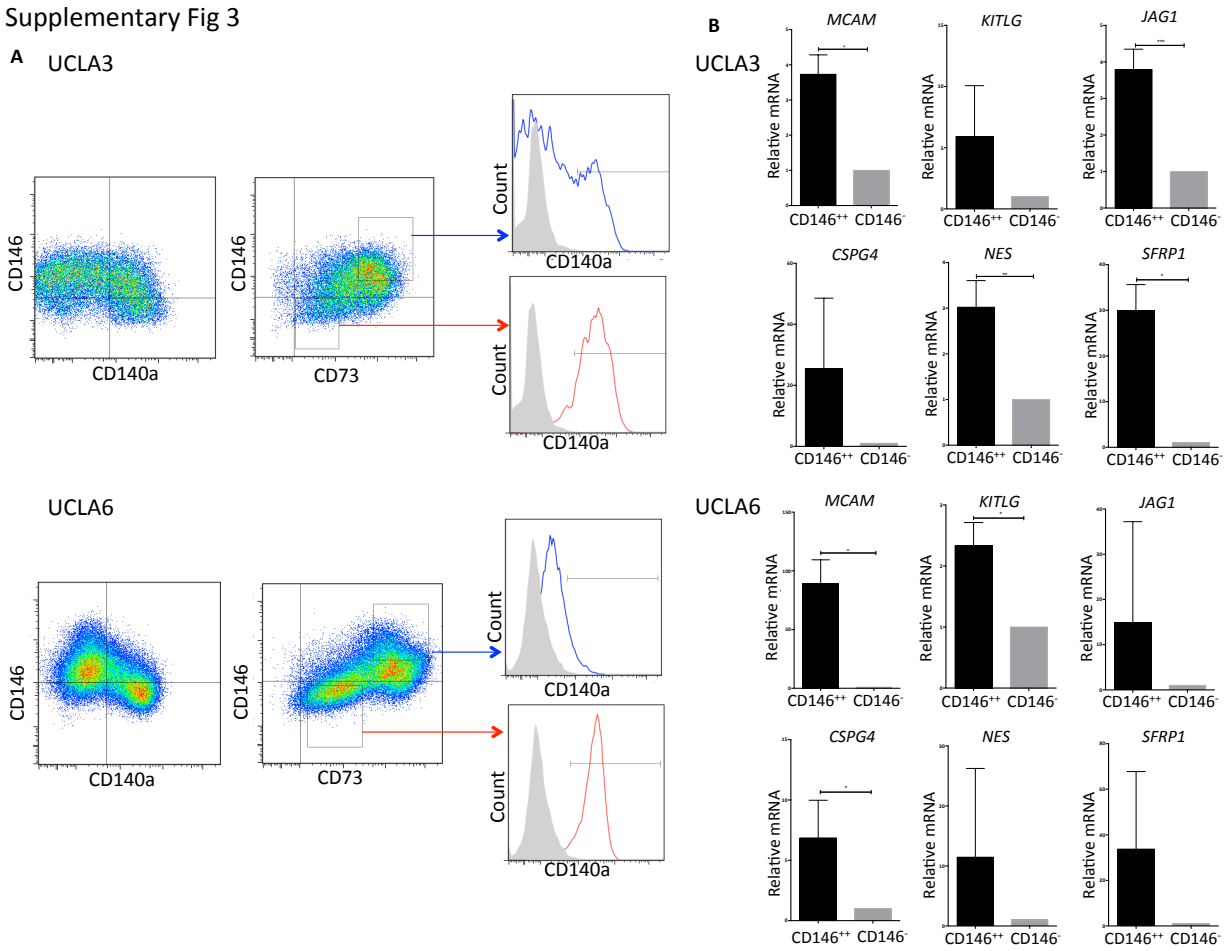
Supplementary Fig 1. Immunophenotypes of mesenchymal cells from primary human tissue are similar to hPSC derived mesenchyme. FACS analysis of lipoaspirate-derived CD146+ and CD146- cells. purified and cultured. Representative data from n= 3 biological replicates, passage 5-13.

Supplementary Fig 2



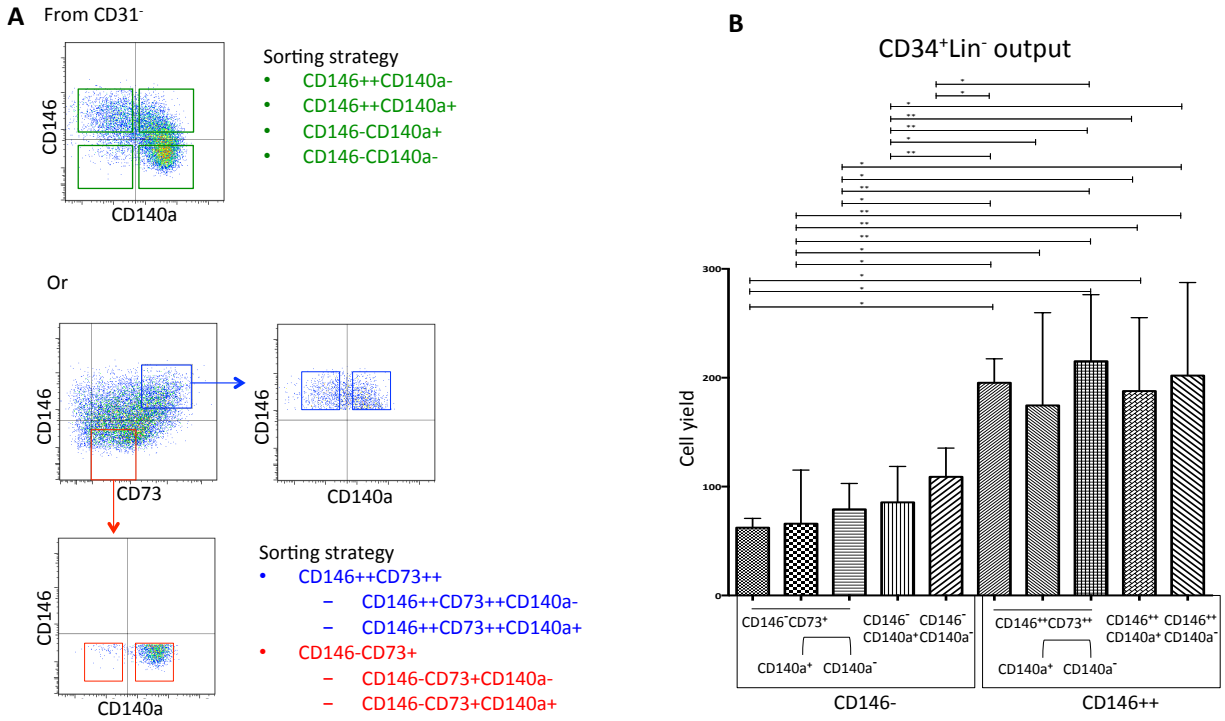
Supplementary Fig 2. CD146^{++} and CD146^{-} cells express conventional mesenchymal markers; HSC niche gene expression enriched in CD146^{++} cells

hEMP derived mesenchymal cultures were analyzed at week 2 (A) by FACS (marker expression after gating on CD31- and either CD146^{++} (blue) or CD146^{-} (red). Data representative of 4 independent analyses with comparable results. (B) Quantitative RT-PCR validation of representative genes from the RNA-Seq analysis. n= 4-6 independent experiments. * $p<0.05$, ** $p<0.01$, *** $p< 0.001$; unpaired two-tailed Student's t test.



Supplementary Fig 3. Mesenchymal populations generated from two additional hESC lines demonstrate similar immunophenotypes to H1. (A) Mesenchymal differentiation culture was initiated from hEMP generated from hESC lines UCLA3 and UCLA6 and analyzed by FACS after two weeks (n=2 biological replicates). (B) Quantitative RT-PCR of representative genes on either $CD146^{++}$ or $CD146^{-}$ derived from UCLA3 and UCLA6. n= 2-3 independent experiments. * $p<0.05$, ** $p<0.01$, *** $p< 0.001$; unpaired two-tailed Student's t test.

Supplementary Fig 4



Supplementary Fig 4. CD146 expression of hPSC-derived mesenchyme defines HSPC supportive capacity. (A) Multiple populations (defined as shown from CD31$^-$) were purified after two-week mesenchymal differentiation cultures from hEMP and tested for their ability to support HSPC ex vivo. (B) High CD146 expression best served as distinguishing cell surface marker to identify HSPC-supporting stroma (n=3 biological replicates, * $p<0.05$, ** $p<0.01$; unpaired two-tailed Student's t test).

Chapter 5

Conclusions

This thesis examined the clonal relationships of hematopoietic, endothelial and mesenchymal lineages during embryonic mesodermal commitment and characterized the mesenchymal subsets of both primary human tissues and hPSC origin to elucidate the regulatory signals in the maintenance of definitive HSC.

Lentiviral tagging and high throughput sequencing have allowed examination, at a clonal level, the lineage potential of human embryonic mesoderm progenitors (hEMP) as they differentiated from human pluripotent stem cells. The results demonstrated that hEMP, which represent the *in vitro* counterparts of the primitive streak, contained not only the previously reported bipotent (hemato-endothelial and mesenchymal-endothelial) progenitors but also multipotent progenitors with differentiation potential of all three lineages. An underappreciated challenge in the application of high-throughput sequencing to clonal tracking was also addressed. Estimating the unobserved population of marked clones and accounting for sampling rate was critical to accurately infer multipotency.

Study of the perivascular cells from both human adipose and fetal bone marrow revealed significant heterogeneity in the composition of mesenchymal populations. CD146+ expression identified a subset that expressed well-known markers of HSC-supportive pericytes in the mouse bone marrow such as Nestin, CXCL12 and LepR. Purified CD146+ cells were capable of supporting the long-term maintenance of transplantable human hematopoietic stem and progenitor cells (HSPC) in co-culture, whereas unfractionated mesenchyme or CD146- cells induce differentiation and compromise ex vivo maintenance of HSPCs.

Characterization of the hEMP-derived mesenchyme revealed that distinct populations also coexisted within conventional definitions of mesenchymal stroma cells. High CD146 and low/absent of CD140a expression defined a hEMP-derived mesenchyme capable of maintaining clonogenic, engraftable and self-renewing HSCs. The CD146++ mesenchyme shared similar transcriptional regulation as primary perivascular cells from human adipose tissues, displaying dominant functional networks in vascular development, wound healing and cell migration.

Future Direction

Transcriptome profiling of the hEMP population opens up opportunity to study cell fate control during early mesoderm development. Both hematopoietic and mesenchymal specifications are known to transition through endothelial intermediate stage that upregulates angiogenic gene expression profiles [1, 2]. According to STRING analysis, the molecular signatures of endothelial cell development such as *CDH5* and *NRP1* were detected in hEMP, suggesting potential endothelial lineage priming. Future direction will aim to investigate whether there is opposing transcriptional machineries that drive fate solidification into EH vs EM bipotent progenitor. Recent study described transcription factor competition as an intriguing cell fate control mechanism. *SCL* mediate fate restriction of early mesodermal cells towards a hematopoietic identity by blocking transcription factor of alternative cardiac lineage, *HAND1*, to assess to cardiac enhancers [3].

RNA sequencing revealed that transcription factors known to directly bind and activate the promoter of *CDH5* (encodes VE-cadherin) such as *SOX7, ETV2, FOXC2* were already present at the hEMP stage [4, 5]. It would be interesting to determine whether these transcription factors

bind to alternative enhancer elements on *CDH5* and drive alternative downstream fate specification.

The ability to prospectively isolate a HSC-supporting subpopulation using a cell surface marker is particularly relevant given clinical interest in using stroma coculture as platform for cord blood HSC expansion[6]. It will be interesting to further examine the use of CD146 as a distinguishing cell surface marker to enrich for HSC supporting stroma from other adult and neonatal tissues such as mesenchymal stem cells (MSC) derived from bone marrow, placenta and umbilical cord.

Besides differentiating into osteoblast/adipocyte/chondrocyte to replace damaged cells, MSC have been shown to mediate immunosuppressive function, to secrete growth factors that promote angiogenesis and wound healing, and to stimulate tissue-resident progenitor cells to participate in repair[7]. Given the wide clinical application of MSCs (390+ registered clinical trials to date)[7], it will be beneficial to examine the potential use of CD146 as marker on MSC that could better resolve functional heterogeneity and provide better therapeutic potential. This is supported by recent findings that CD146+ MSC display greater therapeutic potential than CD146- cells for treating collagen-induced arthritis in mice due to better suppressive effect on T helper 17 cells and better chondrogenesis[8]. In addition, CD146 has also been shown to be a marker predictive of trilineage mutipotency from bone marrow MSC preparations[9].

The goal of the field is to generate functional HSC from hPSC but no studies to date are able to induce correct HSC specification. The hematopoietic cells developed from hPSC do not share the same spectrum of differentiation potential as HSCs and they lack the ability to engraft and self-renew after transplantation that is seen with definitive HSCs[10, 11]. All studies used a stepwise

mesoderm differentiation protocol and data from Chapter 4 here showed the existence of distinct mesenchyme during hPSC mesodermal differentiation. This has broader implication because the mesenchymal populations in the milieu may affect proper HSC differentiation. Further study will aim to define optimal hPSC-derived mesenchyme that would promote HSC specification. The ability to generate HSC from hPSC would enable unlimited, patient-matched cells for transplantation and a development model to study how hematological diseases arise.

1. Vodyanik MA, Yu J, Zhang X et al. A mesoderm-derived precursor for mesenchymal stem and endothelial cells. **Cell Stem Cell**. 2010;7:718-729.
2. Choi KD, Vodyanik MA, Togarrati PP et al. Identification of the hemogenic endothelial progenitor and its direct precursor in human pluripotent stem cell differentiation cultures. **Cell Rep**. 2012;2:553-567.
3. Org T, Duan D, Ferrari R et al. Scl binds to primed enhancers in mesoderm to regulate hematopoietic and cardiac fate divergence. **EMBO J**. 2015;34:759-777.
4. Marcelo KL, Goldie LC, Hirschi KK. Regulation of endothelial cell differentiation and specification. **Circ Res**. 2013;112:1272-1287.
5. Park C, Kim TM, Malik AB. Transcriptional regulation of endothelial cell and vascular development. **Circ Res**. 2013;112:1380-1400.
6. de Lima M, McNiece I, Robinson S et al. Cord-Blood Engraftment with Ex Vivo Mesenchymal-Cell Coculture. **New England Journal of Medicine**. 2012;367:2305-2315.
7. Baldwin K, Urbinati F, Romero Z et al. Enrichment of Human Hematopoietic Stem/Progenitor Cells Facilitates Transduction for Stem Cell Gene Therapy. **Stem Cells**. 2015.
8. Wu CC, Liu FL, Sytwu HK et al. CD146(+) mesenchymal stem cells display greater therapeutic potential than CD146(-) cells for treating collagen-induced arthritis in mice. **Stem Cell Res Ther**. 2016;7:23.
9. Russell KC, Phinney DG, Lacey MR et al. In vitro high-capacity assay to quantify the clonal heterogeneity in trilineage potential of mesenchymal stem cells reveals a complex hierarchy of lineage commitment. **Stem Cells**. 2010;28:788-798.
10. Dravid G, Zhu Y, Scholes J et al. Dysregulated Gene Expression During Hematopoietic Differentiation From Human Embryonic Stem Cells. **Molecular Therapy**. 2011;19:768-781.
11. Dravid GG, Crooks GM. The challenges and promises of blood engineered from human pluripotent stem cells. **Adv Drug Deliv Rev**. 2011;63:331-341.