UNIVERSITY OF CALIFORNIA

SANTA CRUZ

Lineage tracing models reveal novel fate determinants for hematopoietic stem and progenitor cells

A dissertation in partial satisfaction

of the requirements for the degree of

DOCTOR OF PHILOSOPHY

In

MOLECULAR, CELLULAR, AND DEVELOPMENTAL BIOLOGY

By

Jessica Perez-Cunningham

June 2016

Professor E. Camilla Forsberg, Chair

Distinguished Professor Susan Strome

Associate Professor Jeremy Sanford

Assistant Professor, Zhu Wang

Tyrus Miller
Vice Provost and Dean of Graduate Studies

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

ABSTRACT

Lineage tracing models reveal novel fate determinants for hematopoietic stem and progenitor cells

Jessica Perez-Cunningham

Hematopoietic stem cells (HSC) must constantly replenish the blood system by making billions of new cells every day. This high demand for new cells has led to a tightly controlled series of intrinsic and extrinsic signals which enable HSC to balance the need for self-renewal and differentiation. The purpose of this thesis is to explore the myriad signals that regulate the choice between differentiation and self-renewal in HSC and their immediate progeny, multipotent progenitors (MPP). First I describe a novel lineage tracing mouse model that streamlines previously complex flow cytometry staining panels into a single color reporter. This model can now be used for microscopy or flow cytometry analyses to interrogate HSC location and their niche *in situ*. Additionally, I present evidence for the role of the spleen as a microenvironment particularly important for MPP differentiation. Loss of the spleen leads to significantly lowered red blood cell production after transplantation by MPP but not HSC, which indicates that these cell types have different sensitivities to the signals produced there. Finally, I characterized a novel subset of fetal HSC and showed that it is lymphoid biased at a molecular level. This project also generated an RNA sequencing library that can be mined for future characterization of these fetal HSC. Together this work has given us insight into the signals and microenvironments that lead to HSC differentiation *in situ*, after transplantation, and during fetal development.

ACKNOWLEDGEMENTS

I would like to thank my colleagues, mentors, family, and friends for their unending support throughout the pursuit of this degree. There are too many of you to name individually, but your love, kindness, and joy have allowed me to accomplish this goal and I hope we can all continue to lift one another up as every one of us tackles the many adventures in our lives.

CHAPTER 1: Lineage tracing models in the hematopoietic system

INTRODUCTION

For my thesis I worked with several lineage tracing mouse lines and generated a new transgenic lineage tracing system to study hematopoietic stem cells (HSC). Lineage tracing models have been used to interrogate lineage bias and fate decisions in a variety of stem cell populations both in embryonic and adult environments[1]–[3]. Many years of work in this field have led to a wide variety of lineage tracing strategies from pulse-chase experiments to Cre recombinase controlled reporter expression. This chapter focuses on the use of various transgenic mouse lines for lineage tracing. Transgenic mice give researchers the advantage of being able to assess lineage potential *in vivo* which provides a more complex and physiologically relevant environment for study than *in vitro* systems.

We will be discussing transgenic lineage tracing models within the context of the hematopoietic system. Unlike other tissues, the cells of the blood system are in constant motion. As hematopoietic cells mature they leave their tissue of origin (bone marrow, spleen, thymus etc.) and enter the peripheral blood. Immune cells travel through lymph, blood, and tissue to survey for invading pathogens. Red blood cells move in a continuous loop between lungs, heart, and distant tissue to deliver oxygen and remove carbon dioxide. The blood system is an inherently dynamic organ where progeny cells do not necessarily remain near progenitors or stem cells, so there are no spatial landmarks to determine lineage hierarchy. Lineage tracing models provide the solution to this problem, allowing researchers to robustly interrogate hierarchical relationships within the known blood cell types.

An advantage to working with the hematopoietic system is the ability to transplant populations or even single cells and reconstitute the blood system of a recipient. This has given the field extensive insight into many cellular relationships, but does inherently create a

perturbed environment as recipients must be conditioned with either drugs or irradiation. Unlike the traditional experiments that rely on transplantation and conditioning of a recipient mouse, lineage tracing models allow interrogation of unperturbed *in situ* hematopoiesis which can be compared and contrasted to transplantation models and in vitro assays.

Although a lineage tree for the hematopoietic stem cell does already exist there are multiple conflicting versions and the stringency of these relationships is constantly being questioned and revised[4]–[6]. By comparing multiple experimental models of in situ and perturbed hematopoiesis new insights into human diseases such as leukemia and lymphoma can be gained.

In this chapter, I discuss basic considerations when designing a transgenic linage tracing model such as choosing a reporter induction system, vector insertion strategies, and which promoter elements to use. We also discuss how to characterize newly generated transgenic models and particular issues to consider. Finally, we discuss a few transgenic lineage tracing models in the hematopoietic system that already exist, as well as what experimental applications they may be useful for. This review aims to provide a resource for researchers considering using or generating transgenic mouse lines in novel experiments of increasing complexity and utility to the hematopoiesis community.

DESIGNING A TRANSGENIC REPORTER MOUSE LINE

This section will describe the basic considerations one must make to design a transgenic mouse system to perform lineage tracing. To be a lineage tracing system a heritable change must occur in the target cell population that is trackable through multiple cell divisions. This allows conclusions to be drawn regarding hierarchical relationships. The biggest issues to consider when designing a mouse for a lineage trace are reporter specificity

and experimental control of reporter expression. Reporter specificity is influenced by promoter and enhancer elements, vector insertion strategies, and strain differences in gene expression patterns. Experimental control of a reporter system is most influenced by the induction system used. This review will discuss the pros and cons of the most commonly used induction systems and suggest options best suited for use in the hematopoietic system. Ultimately the issue of control over when and where a reporter is expressed allows lineage tracing models to answer important questions about cells' relationships and differentiation potential.

Promoter and Enhancer Elements

Determining which gene elements will drive expression of a construct is arguably the most important decision when generating a transgenic mouse line for lineage tracing. The expression pattern of a construct will be determined primarily through gene promoter and enhancer elements. Thus, it is essential to fully consider known expression patterns for endogenous gene elements in both adulthood and embryonic development. The Stanford Gene Expression Commons [https://gexc.stanford.edu/] is an excellent resource to interrogate known expression patterns of genes within the adult hematopoietic system.

Gene inserts follow a typical format in their construction. Usually cDNA is generated for the target gene, either a component of an induction system or a reporter such as lacZ or a fluorescent protein, and a promoter sequence is fused to the 5' flanking region. A polyadenylation sequence is added to terminate transcription and help stabilize the construct for translation. The polyadenylation sequence is commonly derived from SV40, the growth

hormone (GH) gene, or the β-globin gene [7]. Additional boundary elements can be added to a construct to help mitigate the effect of neighboring genes on construct expression. For example, matrix-attachment regions (MAR) can be added which help anchor a region to the nuclear matrix and promote a euchromatin configuration [8]. Alternatively, insulator elements can be added to help stop the spread of heterochromatin and promote expression of a construct [9]. The completed gene construct can then be cloned into a vector for either targeted insertion or pronuclear injection discussed in the next section.

Vector Insertion Strategies

Another consideration when designing a model is whether to introduce a vector through targeted insertion or pronuclear injection. Targeted insertion gives more control over exactly where in the genome a vector will be knocked-in and how many copies will be present. Alternatively, the pronuclear injection method does not give the researcher as much control over insertion site or copy number, but does allow for greater expression of the inserted construct. This section will discuss methods and specific concerns for each vector insertion strategy.

Targeted insertion is accomplished through homologous recombination in ES cells, producing chimeric mice with transgene germline expression. Historically these mice are generated on the 129 mouse strain background in established embryonic cell lines[11]. Repeated backcrossing of these founder mice and their offspring allows for the vast majority of the genome of the transgenic colony to be on the desired strain background. However, the portion of the genome immediately flanking the targeted insertion gene will remain 129. This means that any genes flanking the transgene during homologous recombination are of 129 background and will be retained as passenger genes even after extensive backcross to other

strains. This is known as the “flanking gene” problem and the likelihood of passenger genes being retained after various numbers of back crossings is summarized well by a table in Lusis et.al. [12]. This can be avoided by using ES cell lines derived from the inbred strain that the final model will be maintained in. The use of ES lines matched to desired mouse background has only recently become an option as more and more ES cell lines are developed from established inbred strains.

Locus choice can greatly affect reporter expression levels and must be considered when targeted insertion is performed. For example, in 1991 the Rosa26 locus was discovered in embryonic stem cells and was found to promote high expression of sequences inserted there with minimum undesired phenotypes [10]. This locus has since become an extremely popular site to knock genes into and many strains are available with various reporter genes inserted there.

Inserting a gene by homologous recombination into a specific locus limits the maximum number of copies to two, whereas randomly inserting a sequence into the genome can boost reporter signal or induction system activity through sheer number of expressed copies. Random insertions are achieved through pronuclear injections and the ease of injection can be greatly improved by choosing specific mouse strains such as FVB mice. Not only does this strain produce oocytes with particularly large and easily visualized pronuclei, but it also has a higher likelihood that injected embryos will survive manipulations and mature to adulthood[13]. This was shown in a study that directly compared embryo survival after pronuclear injection of the same DNA construct into oocytes derived from C57/BL6 mice and FVB mice. Researchers found that 23% of injected FVB embryos yielded newborns compared to only 13% from C57/BL6 embryos. It is interesting to note that although FVB females do not give higher numbers of oocytes when superovulated compared to C57/BL6, 20 embryos and 35 embryos respectively, the health and quality of the oocytes derived from

FVB females is much greater[13], [14]. Due to the relative ease of pronuclear injection in the FVB strain, many Cre recombinase lines are generated on an FVB background. These mice must then be backcrossed onto the desired strain which poses an additional investment of time and resources to produce a useful transgenic tool.

One concern of random insertion is that it offers less control and knowledge about the copy number and location of a construct. This results in a greater likelihood that endogenous genes may be disrupted. It is always a great concern when generating any kind of reporter mice that the reporter construct being used does not alter the normal phenotype of the mouse. This is because one advantage of generating lineage tracing systems for the hematopoietic system is interrogating in situ hematopoiesis and any alteration of the system by gene disruption will necessarily change the conclusions that can be drawn. One way to combat this is to sequence the genomes of mouse lines generated by random insertion to detect which, if any, genes have been disrupted by the inserted construct.

Strain Differences

Genes, and by extension reporter and induction system constructs, can be influenced by the genetic background of the mice they are expressed in. Within the hematopoietic system some commonly used markers have different expression patterns depending on the mouse strain assayed. For example, in C57/Bl6 mice all hematopoietic stem cells (HSC) are Sca-1+ but in BALB/c mice only about 25% of HSC are Sca-1 positive [15], [16]. If promoter and enhancer elements from Sca-1 were used to drive expression of a reporter transgene with the goal of labeling HSC, the genetic background of the transgenic mouse used would hugely influence the results of that experiment.

Additionally, the baseline function and distribution of cells within the immune system can be affected by strain differences. Many of these differences are well reviewed by Sellers et al. [17]. For example, upon infection different subsets of T cells will respond to the same stimulus depending on the genetic background of the mouse strain used. This gives either a TH-1 biased response in C57/Bl6 strain mice or a TH-2 response in BALB/C, DBA/2, A/J etc. strain mice [18]. Thus if a lineage trace was performed on T cell lineages upon response to stimuli it is clear that the strain of the reporter mouse used would directly impact T cell subset expansion.

The issue of strain differences cannot be avoided. The best way to overcome possible difficulties is to properly research known expression patterns and differences between strains. As most hematopoietic models are generated in or backcrossed to the C57/Bl6 strain this option is well characterized, and an excellent choice when considering generating a new reporter system to combine with established models.

Choosing an Induction System

An essential element to lineage tracing is experimental control of changes in reporter expression. This allows for the labeling of a specific subset of cells to be traced through their development and maturation. Control over changes in reporter expression is often accomplished through the use of two very well characterized induction systems: Cre/LoxP and TetOn/Off systems (Figure 1). This section of the review will discuss the basic mechanism of both sections, the advantages of each, and the possibility of combining these systems in novel ways.

The first model we will discuss is the combination of Cre recombinase with its 34 nucleotide recognition sequence, commonly called a loxP site. To design this model loxP

sites are placed both upstream and downstream of a stop sequence which prevents expression of a reporter protein, such as lacZ, or a fluorescent protein, such as GFP. When Cre recombinase is present it excises the stop sequence and allows reporter expression. Because this is a genetic change, it is inherited by all of the target cell's progeny and thus can be used to perform a lineage trace. More sophisticated models have been developed from this basic system such as multi-color switch systems where instead of simply a stop sequence there is an additional color reporter with a stop sequence flanked by loxP sites[19]. With this model, when the Cre recombinase is activated it excises the first color reporter with its stop sequence and allows expression of a second color. Two color systems allow for more elaborate labeling schemes because cells can be labeled before and after Cre recombinase expression[19]. Multiple color reporters also make separation of populations by flow cytometry more definitive and easier to quantify. This is why flow cytometry is a common tool that is paired with lineage tracing models to interrogate the blood system in mice.

An additional layer of control can be added to the Cre recombinase system with the addition of an estrogen receptor domain to the Cre recombinase enzyme[20], [21] (Figure 2). The estrogen receptor added to the Cre recombinase binds heat shock protein 90 (Hsp 90) and sequesters the enzyme in the cytoplasm, preventing Cre recombinase from entering the nucleus. Once the mouse is dosed with tamoxifen, the tamoxifen will bind the estrogen receptor on the Cre recombinase, cause it to be released by Hsp90, and allow Cre recombinase's entry into the nucleus where it will perform its recombinase activity. One appeal of this system is that tamoxifen can be used to activate Cre recombinase and then administration of the drug can be stopped. This allows for a limited window of labeling of cells which can then be followed through many cell divisions. The advantage of this system can readily be seen in developmental studies where subsets of blood cells can be exclusively labeled within certain developmental windows. Overall, the Cre recombinase / loxP system

has been well characterized and reviewed in a number of excellent articles and is an important tool in transgenic mouse lineage tracing[22], [23].

An alternative system that allows for tempo-spatial control of reporter expression is the Tet system[24] (Figure 1). There are two varieties of control in this system: either Tet-On or Tet-Off. Both depend on administration of doxycycline to an animal expressing either the Tet-on element, tTA, or Tet-off element, rtTA. In the Tet-on system, tTA will bind to tetO in the responsive promoter and drive target gene expression unless doxycycline is introduced to perturb this activation. Once doxycycline is removed activation by the tTA element resumes. A similar, but opposite, system is used in the Tet-off model, where rtTA will not bind the tetO elements of the promoter unless it binds doxycycline. In the Tet-off model there is no target expression without doxycycline. An appeal of this system is that the tetO elements are extremely sensitive to any signaling by their transcription factors, so low levels of doxycycline are needed; however, this has resulted in substantial leakiness being reported in these models[25]

Neither of the Tet-on nor Tet-off models offer permanent inheritable changes, but they can be combined with the Cre/loxP system to generate a lineage tracing system. It is easy to envision models with the tTA or rtTA elements downstream of a stop sequence flanked by loxP sites [25]–[27]. In this model, the cells capable of reacting to doxycycline must first have expressed Cre recombinase. Or alternatively expression of Cre can be placed under control of rtTA or tTA elements, so that only cells expressing appropriate doxycycline inducible elements can recombine their DNA . When appropriate regulatory elements are incorporated into all levels of these models a wide array of creative reporting schema could be imagined which would allow for labeling of subsets within subsets of the blood system. A final thought on combinations of models: it is important to think critically regarding how much control is necessary to answer a research question as the more elements there are in a

model the more complicated the breeding schema and colony maintenance can become. Experimental design should be only as complicated as needed to answer a research question; however, the possibilities for exciting novel combinations of drivers and reporters allows researchers to push models to the limits of what they can imagine.

CHARACTERIZING A TRANSGENIC MOUSE LINE

For a lineage tracing model to work efficiently there must be robust reporter expression in a target cell type while maintaining a minimum of off target “background” labeling of other cells types. Important factors to consider when characterizing a newly generated transgenic reporter mouse strain are to determine the level of signal in the target cell population, determine the amount of off target signal, and determine the efficiency of a reporter control method (ex. Tet system or Cre/lox).

When working with a new transgenic model the first step is to determine how robustly a target reporter is expressed. It can be difficult to ascertain exactly which population of cells should express the reporter in a model as it isn’t always clear the exact kinetics of gene expression during dynamic processes such as cell differentiation and maturation. This begs the question, are cells not expressing a reporter construct because the construct isn’t working properly or because they truly aren’t expressing the gene elements driving the reporter or induction system? One strategy to overcome this is to have a two color reporter system. Then cells are labeled whether the recombination event has occurred or not, thus it is possible to determine if there is an error in reporter expression or in the induction system. Alternatively, a reporter can be driven with elements of a gene which is expressed on a cell surface. Comparing reporter expression with endogenous gene levels makes it possible to determine

when cells aren't expressing a construct that they should. The information on expression efficiency allows selection of the best founder mice to breed for a colony and ultimately to generate a robust reporter transgenic model.

Along with determining whether cells are correctly expressing a reporter, it is also important to determine if cells are incorrectly expressing a reporter when they shouldn't be. A systemic analysis of all major organ systems should be performed when characterizing a new transgenic reporter model. When working with the reporter in the hematopoietic system one particular population to check for off target expression is the endothelial system. As endothelial cells and hematopoietic cells share a developmental origin, they express many of the same genes and are most likely to incorrectly express a reporter[28]–[30] .

Another thought to consider when working with lineage tracing systems in the hematopoietic system is the issue of how quickly cells can differentiate and that the reporter may not properly reflect quick changes in cell biology. The average half-life for fluorescent proteins is 3-5 days, which can pose significant problems when considering cell differentiation and lineage potential [31]. A cell may only transiently express the gene elements for a reporter, but the signal can then persist for days after the true gene expression is gone. This can be assessed by comparing endogenous gene expression to reporter expression, which is easier to assay for when the endogenous gene is expressed on the cell surface, but that isn't always feasible for all constructs. An alternative solution is to incorporate a PEST motif into a reporter which shortens the protein half-life to around 3 hours [31]. The first step to handling off target expression of a reporter is to understand which cells are most likely to be affected in a unique transgenic mouse line and whether that will impact its utility.

Besides confirming correct expression patterns, it is paramount to confirm that a reporter construct is not aberrantly expressed when it should be controlled. We have discussed the advantages of using inducible models but the biggest concern with these

models is “leakiness” or inappropriate reporter expression. "Leakiness" can be assessed by checking for reporter expression in naïve animals. First check animals that completely lack the induction system, either the mice that don’t express the Cre recombinase element or Tet inducible elements. This will determine if there is aberrant expression from a reporter construct. Additionally, if a system has a Cre-ERT2 or Tet model, a reporter mouse can be crossed to a mouse carrying the induction system and assayed for reporter expression before dosing with tamoxifen or doxycycline. This determines how tightly exposure to the target drug controls reporter expression and if an induction system is leaky. Depending on the experimental design the transgenic reporter line will be used in, the role of leakiness within a model will have varying effects on experimental outcome.

TRANSGENIC REPORTER MICE CURRENTLY AVAILABLE

As technology continues to improve, more fluorescent reporter lineage tracing models are being generated. Table 1 shows a few selected transgenic mouse lines available for lineage tracing within the hematopoietic system. This table shows a variety of Cre lines with specific activity within the blood system. There are also a few reporter lines shown here with a variety of fluorescent colors available. Finally, there are two lines that are shown to mimic disease states. They combine a disease phenotype with the ability to lineage trace the cells effected. These lines are only a small sample of the many transgenic lineage tracing systems available.

Table 1

Name of Mouse	Model Description	Other notes
CD4CreER^{T2} [32]	Cre-ER^{T2} activity is under control of CD4 gene elements. In mice given tamoxifen, 79% of peripheral blood CD4+ T cells show recombinase activity	No difference in phenotype between homozgyotes and hemizygotes
Pf4-Cre [33]	Cre expression under control of platelet factor 4 elements and recombinase activity is seen in the majority of megakaryocytes	Normal blood counts are observed in these mice
HBB-cre [34]	Cre expression under control of human beta hemoglobin gene elements. Recombination occurs at 50-100% efficiency in erythroid/megakaryocytic cell lineages beginning at onset of hematopoiesis at approximately embryonic day 7.5	Mice hemizygous for the transgenic insert are viable, fertile, normal in size and do not display any gross physical or behavioral abnormalities
Vav-GFP mice [39]	GFP expression is driven in the entire blood system (except RBC) and is lost upon Cre recombination	Mice are fertile and show no gross defects
mTmG[19]	Mice express Tomato fluorescent protein in all cells unless Cre recombinase activity is present. Then Tomato is excised and GFP is expressed	Dual color reporter expressing fluorescence in all tissues
R26-stop-EYFP[35]/ R26-stop-ECFP[35]	Cre recombinase activity excises the stop sequence and allows expression of eYFP or eCFP	Mice appear fertile with no obvious morphological or behavioral defects
Csf1r-HBEGF/mCherry[36]	A loxP flanked stop codon precedes a mCherry fluorescent protein fused to a diphtheria toxin receptor. The entire construct is under control of a *Csf1r* promoter.	After Cre recombinase activity has occurred cells can be treated a simple lineage trace be performed or mice can be treated with diphtheria toxin to ablate target cells.
Bcrtm1(BCR/ABL)Tsr[37]	A loxP flanked stop codon precedes GFP mused human BCR/ABL.	When Cre recombinase activity allows expression of this construct mice develop leukemias that can be lineage traced using the GFP reporter in this model

CONCLUSION

Bone marrow transplantation was the first clinically approved stem cell treatment and has been used continuously for the past 60 years to treat a wide variety of conditions. Understanding normal and perturbed hematopoiesis is essential to continuing to improve this procedure and human health in general. Transgenic lineage tracing systems are uniquely suited to interrogate the hematopoietic system because they allow analysis of hematopoiesis in a variety of settings. First, genetically labeling subsets of the blood allows for study of *in situ* hematopoiesis, answering important questions about normal blood development. In addition, lineage tracing can be coupled to disease models to understand, for example, what cell types and lineages are the cause of different kinds of cancer. Finally, lineage tracing can be coupled with the ability to transplant cells and recapitulate whole or part of a new blood system from a few donor cells . By using transgenic murine systems that allow labeling of donor cells and tracing of the lineage of their progeny questions can be asked about the ability of progenitor populations to give rise to cells with various lineages and function. The ability to transplant cells also allows manipulation of the recipient environment that cells are transplanted into, thus allowing researchers to tease apart the role of the host environment on cell fate decisions. Studying blood development both in situ and under stressed or disease conditions allows for conclusions to be drawn that are extremely relevant to human health.

One potential drawback to lineage tracing is that it often surveys cells at a population level rather than at a single cell level. Populations within the hematopoietic system are thought to be variable at a single cell level [5], [38]. However, it is easy to imagine combining single cells analysis with lineage tracing in a variety of interesting ways. For example, a lineage tracing system could be used to identify a subset of progenitors and then flow

cytometry could be combined with a fluidigm system to analyze them at a single cell level for RNA content. Alternatively, single hematopoietic cells can be sorted and transplanted into conditioned recipients, which allows for analysis of output of an individual cell [5]. Combining existing tools for single cell analysis with transgenic lineage tracing systems allows for interesting questions to be asked at a single cell rather than a population level.

Together the information presented in this review is meant to serve as a guide in moving forward with the ever-growing technology of transgenic mouse construction and lineage tracing within the blood system. Presented here are a few issues to consider when designing and characterizing novel transgenic reporter constructs as well as a topical overview of a few mouse lines already available. As more and more interesting tools are being developed, our ability to manipulate the mouse blood system is becoming extremely advanced and is fast approaching a point where the only limitation is the imagination of the researcher.

Figure 1

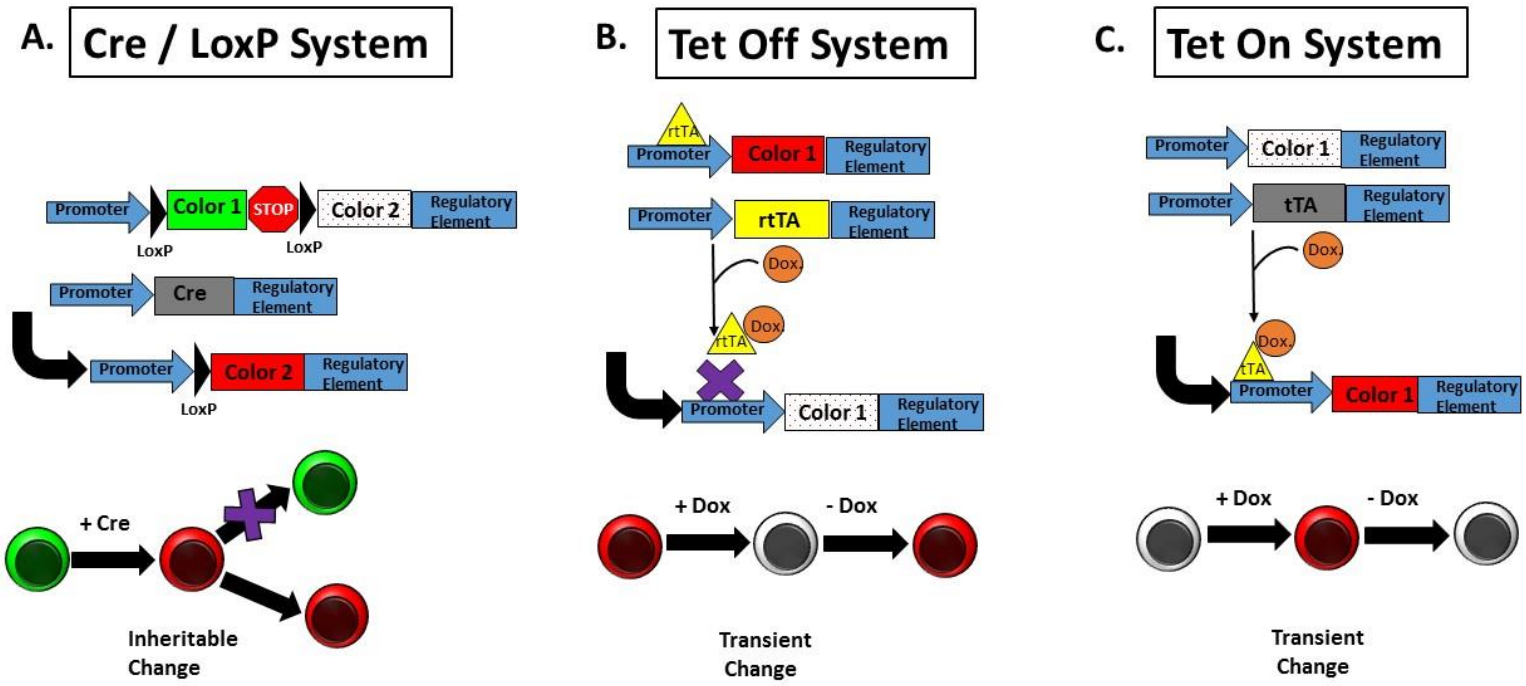

A.
Cre / LoxP System
Promoter
Color 1
STOP
Color 2
Regulatory Element
LoxP
LoxP
Promoter
Cre
Regulatory Element
Promoter
Color 2
Regulatory Element
LoxP
+ Cre
Inheritable Change
B.
Tet Off System
rtTA
Promoter
Color 1
Regulatory Element
Promoter
rtTA
Regulatory Element
Dox
rtTA
Dox
Promoter
Color 1
Regulatory Element
+ Dox
- Dox
Transient Change
C.
Tet On System
Promoter
Color 1
Regulatory Element
Promoter
tTA
Regulatory Element
Dox
Dox
tTA
Promoter
Color 1
Regulatory Element
+ Dox
- Dox
Transient Change

Figure 1. **Schematic of the most commonly used reporter induction systems.** A. When expressed Cre recombinase is capable of removing sections of a construct flanked by loxP sites. This allows for an inheritable genetic change. B. In the Tet off system, rtTA binds to tet elements in a promoter to induce expression of a reporter unless doxycycline is present to inhibit binding of rtTA to promoter elements. C. In the Tet on system, tTA must first bind doxycline before it can bind tet elements in a promoter to allow for transcription of a target gene. Neither the Tet on nor Tet off system produce permanent inheritable changes.

Figure 2

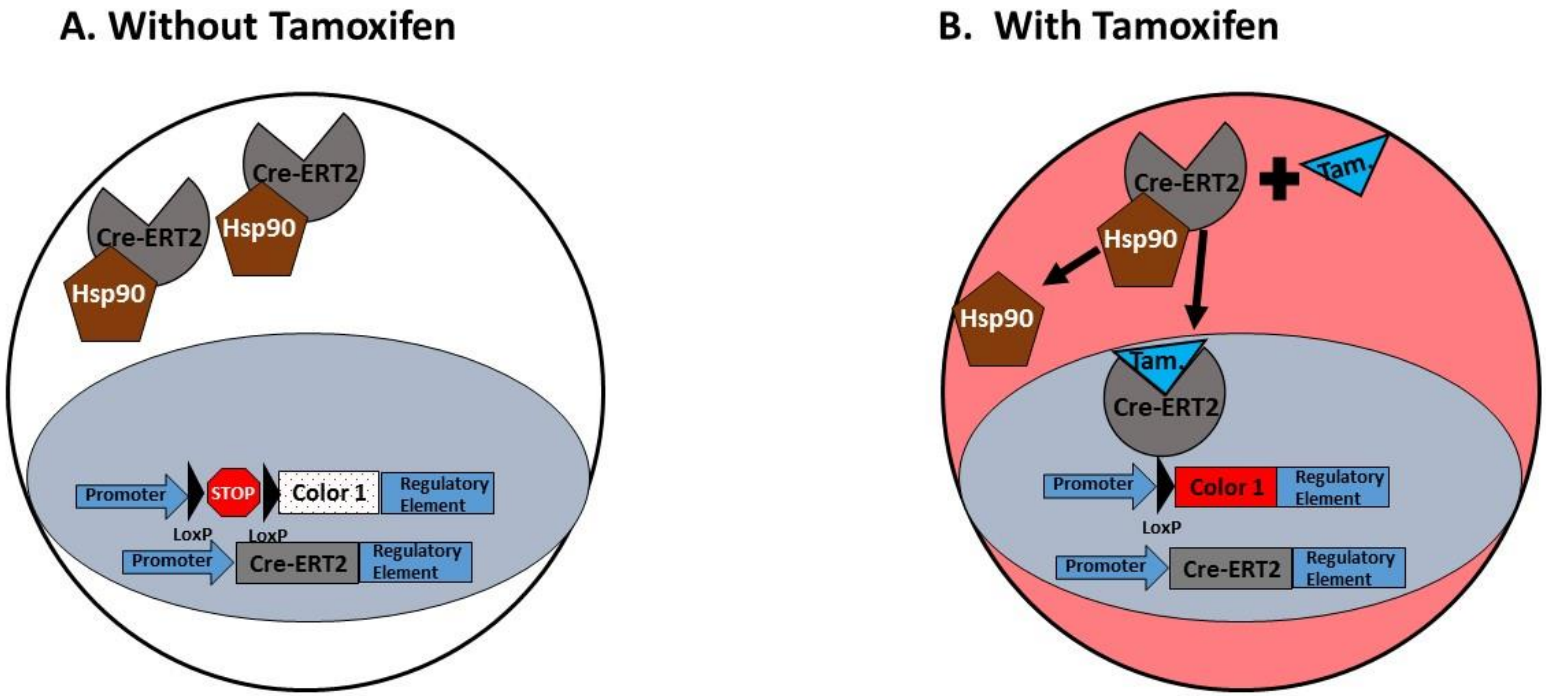

A. Without Tamoxifen
Cre-ERT2
Hsp90
Cre-ERT2
Hsp90
Promoter
STOP
Color 1
Regulatory Element
LoxP
LoxP
Promoter
Cre-ERT2
Regulatory Element
B. With Tamoxifen
Cre-ERT2
Tam.
Hsp90
Hsp90
Tam.
Cre-ERT2
Promoter
Color 1
Regulatory Element
LoxP
Promoter
Cre-ERT2
Regulatory Element

Figure 2. Schematic of the mechanism of action of Cre-ERT2 enzyme with and without tamoxifen. A. Without tamoxifen the Cre-ERT2 enzyme is sequestered in the cytosol of the cell by heat shock protein 90 (Hsp 90) binding to the estrogen receptor and cannot act to recombine loxP sites. B. Upon binding by tamoxifen, Cre-ERT2 is released by Hsp 90 which will allow it to traffic into the nucleus. There is can recombine loxP sites on a reporter construct to allow for expression of a reporter such as a fluorescent protein.

REFERENCES

[1] Y. Kawaguchi, K. Takaori, and S. Uemoto, “Genetic lineage tracing, a powerful tool to investigate the embryonic organogenesis and adult organ maintenance of the pancreas.,” *J. Hepatobiliary. Pancreat. Sci.*, vol. 18, no. 1, pp. 1–5, Jan. 2011.

[2] K. Kretzschmar and F. M. Watt, “Lineage tracing.,” *Cell*, vol. 148, no. 1–2, pp. 33–45, Jan. 2012.

[3] B. D. Humphreys and D. P. Dirocco, “Lineage-tracing methods and the kidney.,” *Kidney Int.*, pp. 1–8, Oct. 2013.

[4] Ema, Hideo, Y. Morita, and T. Suda, “Heterogeneity and hierarchy of hematopoietic stem cells,” *Exp. Hematol.*, vol. 42, no. 2, pp. 74–82, 2014.

[5] R. Yamamoto, Y. Morita, J. Ooehara, S. Hamanaka, M. Onodera, K. L. Rudolph, H. Ema, and H. Nakauchi, “Clonal analysis unveils self-renewing lineage-restricted progenitors generated directly from hematopoietic stem cells.,” *Cell*, vol. 154, no. 5, pp. 1112–26, Aug. 2013.

[6] A. E. Beaudin, S. W. Boyer, and E. C. Forsberg, “Flk2/Flt3 promotes both myeloid and lymphoid development by expanding non–self-renewing multipotent hematopoietic progenitor cells,” *Exp. Hematol.*, vol. 42, no. 3, pp. 218–229.e4, 2014.

[7] C. Liu, “Strategies for Designing Transgenic DNA Constructs,” vol. 1027, 2013.

[8] N. Harraghy, A. Gaussin, and N. Mermod, “Sustained transgene expression using

MAR elements.," *Curr. Gene Ther.*, vol. 8, no. 5, pp. 353–366, 2008.

[9] M. Gaszner and G. Felsenfeld, "Insulators: exploiting transcriptional and epigenetic mechanisms.," *Nat. Rev. Genet.*, vol. 7, no. 9, pp. 703–713, 2006.

[10] G. Friedrich and P. Soriano, "Promoter traps in embryonic stem cells: A genetic screen to identify and mutate developmental genes in mice," *Genes Dev.*, vol. 5, no. 9, pp. 1513–1523, 1991.

[11] B. Conference and G. Background, "Mutant Mice and Neuroscience : Recommendations Concerning," vol. 19, pp. 755–759, 1997.

[12] A. J. Lusis, J. Yu, and S. S. Wang, "The problem of passenger genes in transgenic mice.," *Arterioscler. Thromb. Vasc. Biol.*, vol. 27, no. 10, pp. 2100–3, Oct. 2007.

[13] M. Taketo, a C. Schroeder, L. E. Mobraaten, K. B. Gunning, G. Hanten, R. R. Fox, T. H. Roderick, C. L. Stewart, F. Lilly, and C. T. Hansen, "FVB/N: an inbred mouse strain preferable for transgenic analyses.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 88, no. 6, pp. 2065–9, Mar. 1991.

[14] C. Luo, J. Zuñiga, E. Edison, S. Palla, W. Dong, and J. Parker-Thornburg, "Superovulation strategies for 6 commonly used mouse strains.," *J. Am. Assoc. Lab. Anim. Sci.*, vol. 50, no. 4, pp. 471–8, Jul. 2011.

[15] N. Uchida and I. L. Weissman, "Searching for hematopoietic stem cells: evidence that Thy-1.1lo Lin- Sca-1+ cells are the only stem cells in C57BL/Ka-Thy-1.1 bone marrow.," *J. Exp. Med.*, vol. 175, no. 1, pp. 175–184, 1992.

[16] G. J. Spangrude and D. M. Brooks, "Mouse strain variability in the expression of the hematopoietic stem cell antigen Ly-6A/E by bone marrow cells.," *Blood*, vol. 82, no. 11, pp. 3327–32, 1993.

[17] R. S. Sellers, C. B. Clifford, P. M. Treuting, and C. Brayton, “Immunological Variation Between Inbred Laboratory Mouse Strains: Points to Consider in Phenotyping Genetically Immunomodified Mice,” *Vet. Pathol.*, vol. 49, no. 1, pp. 32–43, 2012.

[18] C. D. Mills, K. Kincaid, J. M. Alt, M. J. Heilman, and A. M. Hill, “M-1/M-2 Macrophages and the Th1/Th2 Paradigm,” *J. Immunol.*, vol. 164, no. 12, pp. 6166–6173, 2000.

[19] M. D. Muzumdar, B. Tasic, K. Miyamichi, L. Li, and L. Luo, “A Global Double-Fluorescent Cre Reporter Mouse,” *Genesis*, vol. 45, pp. 593–605, 2007.

[20] R. Feil, J. Brocard, B. Mascrez, M. LeMeur, D. Metzger, and P. Chambon, “Ligand-activated site-specific recombination in mice.,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 93, no. 20, pp. 10887–90, 1996.

[21] R. Feil, D. Metzger, and P. Chambon, “Regulation of Cre Recombinase Activity by Mutated Estrogen Receptor Ligand-Binding Domains,” vol. 757, no. 237, pp. 752–757, 1997.

[22] C. S. Heffner, C. Herbert Pratt, R. P. Babiuk, Y. Sharma, S. F. Rockwood, L. R. Donahue, J. T. Eppig, and S. a Murray, “Supporting conditional mouse mutagenesis with a comprehensive cre characterization resource.,” *Nat. Commun.*, vol. 3, p. 1218, Jan. 2012.

[23] S. Feil, N. Valtcheva, and R. Feil, *Inducible cre mice*, vol. 530. 2009.

[24] M. Gossen and H. Bujard, “Tight control of gene expression in mammalian cells by tetracycline-responsive promoters.,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 89, no. 12, pp. 5547–51, 1992.

[25] Y. Sun, X. Chen, and D. Xiao, “Tetracycline-inducible Expression Systems: New Strategies and Practices in the Transgenic Mouse Modeling,” *Acta Biochim. Biophys.*

Sin. (Shanghai)., vol. 39, no. 4, pp. 235–246, Apr. 2007.

[26] G. Belteki, J. Haigh, N. Kabacs, K. Haigh, K. Sison, F. Costantini, J. Whitsett, S. E. Quaggin, and A. Nagy, “Conditional and inducible transgene expression in mice through the combinatorial use of Cre-mediated recombination and tetracycline induction,” *Nucleic Acids Res.*, vol. 33, no. 5, pp. 1–10, 2005.

[27] H.-M. I. Yu, B. Liu, S.-Y. Chiu, F. Costantini, and W. Hsu, “Development of a unique system for spatiotemporal and lineage-specific gene expression in mice.,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 102, no. 24, pp. 8615–20, 2005.

[28] E. Dzierzak and N. A. Speck, “Of lineage and legacy: the development of mammalian hematopoietic stem cells.,” *Nat. Immunol.*, vol. 9, no. 2, pp. 129–36, Feb. 2008.

[29] D. Kanz, M. Konantz, E. Alghisi, T. E. North, and C. Lengerke, “Endothelial-to-hematopoietic transition: Notch-ing vessels into blood,” *Ann. N. Y. Acad. Sci.*, p. n/a–n/a, 2016.

[30] E. Gritz and K. K. Hirschi, “Specification and function of hemogenic endothelium during embryogenesis,” *Cell. Mol. Life Sci.*, vol. 73, no. 8, pp. 1547–1567, 2016.

[31] X. Li, X. Zhao, Y. Fang, T. Duong, C. Fan, C. Huang, S. R. Kain, and X. Jiang, “Generation of Destabilized Green Fluorescent Protein as a Transcription Reporter Generation of Destabilized Green Fluorescent Protein as a Transcription Reporter *,” *J. Biol. Chem.*, vol. 273, no. 52, pp. 34970–34975, 1998.

[32] K. Aghajani, S. Keerthivasan, Y. Yu, and F. Gounari, “Generation of CD4CreER(T^2) transgenic mice to study development of peripheral CD4-T-cells.,” *Genesis*, vol. 50, no. 12, pp. 908–13, 2012.

[33] R. Tiedt, T. Schomber, H. Hao-Shen, and R. C. Skoda, “Pf4-Cre transgenic mice allow

the generation of lineage-restricted gene knockouts for studying megakaryocyte and platelet function in vivo," *Blood*, vol. 109, no. 4, pp. 1503–1506, 2007.

[34] K. R. Peterson, H. Fedosyuk, L. Zelenchuk, B. Nakamoto, E. Yannaki, G. Stamatoyannopoulos, S. Ciciotte, L. L. Peters, L. M. Scott, and T. Papayannopoulou, "Transgenic Cre expression mice for generation of erythroid-specific gene alterations," *Genesis*, vol. 39, no. 1, pp. 1–9, 2004.

[35] S. Srinivas, T. Watanabe, C. S. Lin, C. M. William, Y. Tanabe, T. M. Jessell, and F. Costantini, "Cre reporter strains produced by targeted insertion of EYFP and ECFP into the ROSA26 locus.," *BMC Dev. Biol.*, vol. 1, p. 4, Jan. 2001.

[36] H. a Schreiber, J. Loschko, R. a Karssemeijer, A. Escolano, M. M. Meredith, D. Mucida, P. Guermonprez, and M. C. Nussenzweig, "Intestinal monocytes and macrophages are required for T cell polarization in response to Citrobacter rodentium.," *J. Exp. Med.*, vol. 210, no. 10, pp. 2025–2039, 2013.

[37] S. B. Foley, Z. L. Hildenbrand, A. A. Soyombo, J. A. Magee, Y. Wu, K. I. Oravecz-Wilson, and T. S. Ross, "Expression of BCR/ABL p210 from a Knockin Allele Enhances Bone Marrow Engraftment without Inducing Neoplasia," *Cell Rep.*, vol. 5, no. 1, pp. 51–60, 2013.

[38] C. Benz, M. R. Copley, D. G. Kent, S. Wohrer, A. Cortes, N. Aghaeepour, E. Ma, H. Mader, K. Rowe, C. Day, D. Treloar, R. R. Brinkman, and C. J. Eaves, "Hematopoietic stem cell subtypes expand differentially during development and display distinct lymphopoietic programs.," *Cell Stem Cell*, vol. 10, no. 3, pp. 273–83, Mar. 2012.

[39] J. Perez-Cunningham, S. W. Boyer, M. Landon, E. C. Forsberg. "Hematopoietic stem cell-specific GFP-expressing transgenic mice generated by genetic excision of a pan-

hematopoietic reporter gene". *Experimental Hematology*. Accepted for publication

Chapter II. Hematopoietic stem cell-specific GFP-expressing transgenic mice generated by genetic excision of a pan-hematopoietic reporter gene

The text contained within this chapter of the dissertation is a reprint of the following previously published material: [Jessica Perez-Cunningham, Scott W. Boyer, Mark Landon, and E. Camilla Forsberg. Hematopoietic stem cell-specific GFP-expressing transgenic mice generated by genetic excision of a pan-hematopoietic reporter gene. *Experimental Hematology*. Accepted for publication].

ABSTRACT

Selective labeling of specific cell types by expression of green fluorescent protein (GFP) within the hematopoietic system would have great utility in identifying, localizing, and tracking different cell populations in flow cytometry, microscopy, lineage tracing, and transplantation assays. In this report, we describe the generation and characterization of a new transgenic mouse line with specific GFP labeling of all nucleated hematopoietic cells, as well as platelets. This new "Vav-GFP" mouse line labels the vast majority of hematopoietic cells with GFP during both embryonic development and adulthood, with particularly high expression in hematopoietic stem and progenitor cells. With the exception of transient labeling of fetal endothelial cells, GFP expression is highly selective for hematopoietic cells, and persists in donor-derived progeny after transplantation of hematopoietic progenitor cells. Finally, we also demonstrate that the loxP-flanked reporter allows for specific GFP labeling of different hematopoietic cell subsets when crossed to various Cre reporter lines. Notably, by crossing Vav-GFP mice to Flk2-Cre mice, we obtained robust and highly selective GFP expression in HSCs. Together, these data describe a new mouse model capable of directing GFP labeling exclusively of hematopoietic cells or exclusively of HSCs.

INTRODUCTION

There have been many efforts to generate transgenic mice with transgene expression exclusively in the hematopoietic compartment[1]. The *vav1* gene has been the focus of many such studies as it is highly expressed throughout hematopoietic development from the embryonic day 11.5 (e11.5) embryo through adulthood [2] There appears to be very limited expression in other tissues in the adult mouse, with the exception of the developing tooth bud[2]. Vav1 has been shown to activate the Rac/Jun kinase pathway and gene disruption assays have shown it to be essential for signaling through the antigen receptors of lymphocytes [3–5]. Interestingly, even though Vav1 is highly expressed throughout the hematopoietic system, it is not essential for the development of blood cells in general [6].

The unique expression pattern of the *vav1* gene has led to generation of several *vav*-driven cre mouse lines as well as *vav*-driven direct reporter mouse lines [7–12]. These mouse lines have generally had great success in labeling the hematopoietic compartment with minimal off-target expression. Of particular note is Stadtfeld and Graf's model where Cre recombinase is driven with promoter and enhancer elements of the *vav* gene [10]. When crossed to a stop-lox-YFP reporter line, this model accomplished almost 100% labeling in all nucleated bone marrow (BM) cells and platelets in adult mice. They also found that nearly all KLS (ckit+, lin-, sca+) cells were labeled in the e13.5 fetal liver and approximately half of CD45+ (hematopoietic) cells from the e10.5 fetal liver were reporter positive [10]. While this mouse line demonstrated great success in labeling the entire hematopoietic compartment, it does not allow for the resolution of specific cell populations within the hematopoietic lineage needed for experiments such as lineage tracing from hematopoietic stem cells (HSCs) and/or progenitor cells (HSPCs) or localization of HSCs/HSPCs. To enable fluorescent labeling of specific hematopoietic cell populations, we modified Stadtfeld's construct so that the *vav* enhancer/promoter elements drive a fluorescent reporter that can be excised in specific

hematopoietic cell subsets using Cre-mediated recombination. This new mouse line, called Vav-GFP mice, allows for two levels of specificity: firstly, the fluorescent reporter is under control of *vav* promoter elements and, secondly, it can be crossed to a multitude of Cre lines to drive excision of the reporter and thereby restricting fluorescence to a desired population of HSCs or HSPCs.

In this study we characterized the fluorescence of the Vav-GFP mouse line in BM and peripheral blood in both adult and fetal mice. In addition, we showed that the Vav-GFP cells can be distinguished from wild type host cells after transplantation as this is a likely application of the new mouse line. Finally, we also crossed the Vav-GFP mice to a Flk2-driven Cre mouse line to achieve targeted labeling exclusively of HSCs within the BM compartment [13,14]. These data collectively show that the Vav-GFP mouse line generated here represents a novel tool to interrogate HSC differentiation and trafficking by providing hematopoietic-specific expression of a reporter construct under control of Cre mediated recombination.

RESULTS AND DISCUSSION

Characterization of Reporter Expression in Hematopoietic Cell Populations

Our goal was to generate a dual-purpose transgenic mouse line that allows for pan-hematopoietic or, in combination with selected Cre-expressing mouse lines, labeling of a subset of HSCs/HSPCs. To generate Vav-GFP mice, we used the murine regulatory elements of the *vav* gene to drive expression of a dual color reporter. A vector consisting of Vav regulatory elements and Loxp-flanked EGFP was linearized and injected into pronuclei of C57/B6l6 mice (Figure 1A). In this model, GFP is expressed until Cre-mediated recombination causes excision of GFP and a stop codon (Figure 1A and 5A).

To investigate the ability of the reporter construct to fluorescently label hematopoietic cells, HSPCs and mature cell populations were isolated from BM and peripheral blood (PB) of Vav-GFP mice. Flow cytometry analysis revealed reporter expression in all HSPCs and mature cells, including platelets and erythroid progenitors, but not in mature circulating red blood cells (RBCs) (Figure 1B and 1C). All HSCs, multipotent progenitors (MPPs) and myeloid progenitors displayed strong GFP expression that was clearly distinguishable from WT control cells (Figure 1B and 1D). Mature BM cell populations also expressed GFP, although at lower fluorescent intensities compared to HSPCs (Figure 1B, D, E). In the PB, virtually all hematopoietic cells expressed GFP at levels readily distinguishable from WT control cells (Figure 1C, F, G). Notably, GFP expression was robust in platelets, but not detected in circulating RBCs. Also, the reporter expression in circulating GM was significantly higher than in bone marrow resident GM (Supplemental Figure 1). This may be due to differences in Vav expression between these populations. Together, these data show strong reporter expression in all BM and circulating hematopoietic cells, with the exception of RBCs, making the Vav-GFP transgenic mouse a valuable tool for interrogating the blood system.

Reporter expression in adult mice is restricted to hematopoietic cells

The Vav-GFP model was designed to only fluoresce in hematopoietic cells and no other tissues or cell types. To test the reporter specificity, we investigated GFP expression in non-hematopoietic cells of brain, liver, heart, and lungs. Whole organs were isolated, prepared into single cell suspensions, and stained with pan-hematopoietic (CD45) and the erythroid (Ter119) markers to exclude hematopoietic cells. Each organ was then analyzed by flow cytometry and CD45-Ter119- cells were tested for GFP expression. While GFP expression was readily detected in co-isolated CD45+ cells, there was no detectable off-target GFP expression found in non-hematopoietic cells in any of the organs surveyed (Figure 2A,B).

The most likely population of cells to exhibit off-target expression in the Vav-GFP model are endothelial cells (ECs), as other studies using *vav* regulatory elements have reported mixed results of off-target expression in ECs. For example, Georgiades et al showed that all CD31+ cells were labeled with b-galactosidase with their vav-cre line, whereas Ogilvy et al reported no vav-driven hCD4 in non-hematopoietic tissues by immunohistochemical analysis [7,8]. To test endothelial GFP expression in our model, we isolated CD45-Ter119-CD31+Sca1+ ECs from the BM of Vav-GFP mice. In all (n=6) but one mouse surveyed, GFP expression was undetectable in BM ECs (Figure 2C,D). One mouse had approximately 5% of its ECs labeled with GFP, but no other mouse surveyed, including littermates, showed similar expression patterns.

Vav-GFP is selectively expressed in fetal hematopoietic cells

Given that *vav* expression has been detected in the embryo as early as embryonic day 11.5 [2], we wanted to test GFP expression in embryonic hematopoietic cells in Vav-GFP mice. Most, but not all (~85%), of CD45+ckit+ hematopoietic cells isolated from the caudal half of e11.5 Vav-GFP embryos displayed robust GFP expression (Figure 3A). The partial labeling may be due to inadequate accumulation of GFP as the transgene is just beginning to be expressed at this time point, or due to heterogeneity of the cell types included in the CD45+ckit+ phenotypic compartment.

As both progenitor and mature cell populations have been well characterized in fetal livers at e14.5, we investigated the GFP expression of hematopoietic cell subsets at this stage. We observed strong labeling in all hematopoietic cells surveyed, including HSPCs and mature cells (Figure 3B). Compared to the adult counterpart, the level of GFP expression varied in some cases, with fetal HSCs displaying lower GFP intensity than adult HSCs, whereas embryonic B cells appeared brighter than their adult equivalents (Figure 3B).

To test the specificity of GFP expression in fetal Vav-GFP mice, we isolated ECs from e14.5 fetal livers. In contrast to the lack of GFP expression in adult ECs, a large proportion (~50%) of ECs in the e14.5 fetal liver exhibited GFP expression (Figure 3C). Their EC identity was confirmed by costaining with anti-Tie2 and Vcam1 antibodies (data not shown). GFP expression in ECs is clearly transient as it was not detected in the adult and may reflect a brief period of vav expression by early endothelial cells or progenitors. This off-target expression is unlikely to affect adult studies with our Vav-GFP mouse line, but it may require exclusion of ECs to study developmental hematopoiesis in these mice. For example, the ECs can be excluded from flow cytometry assays by either MECA 32 expression or lack of CD45 expression. Conversely, endothelial GFP expression may be used to investigate fetal EC populations, or the relationship between endothelial and hematopoietic development.

Robust reporter expression is detected after transplantation

To test whether GFP expression remains robust during reconstitution following transplantation, we performed transplants with KLS cells isolated from Vav-GFP mice. All mature cell and progenitor lineages surveyed in recipient mice showed clear separation by GFP expression between host and donor populations (Figure 4A,B). In addition, analysis of live, Ter119- cells in chimeric animals showed that all donor derived cells (CD45.2+) were also GFP+. This result verifies that the transgene clearly labels donor cells upon transplantation and supports the utility of the Vav-GFP mice to track transplanted hematopoietic cells in reconstitution assays.

HSC-specific GFP expression accomplished by crossing Vav-GFP and Flk2-Cre lines

The flanking of GFP with loxP sites enables excision of the reporter gene in desired cell populations. To test whether HSC-specific labeling could be accomplished, we crossed the

Vav-GFP mice to Flk2-Cre transgenic mice [15]. We and others have previously shown that Flk2 is expressed by MPPs [16–19] and that Flk2-Cre labeling combined with a color switch model is capable of labeling the entire hematopoietic system downstream of the HSC [13,14]. By crossing Vav-GFP mice to Flk2-Cre line we anticipated that all HSCs would be labeled with GFP, and that all downstream progenitors and mature cells would lack GFP expression due to excision of the reporter gene in Flk2+ progenitor cells[20]. In addition, non-hematopoietic cells would lack GFP expression due to the hematopoietic-specific expression of Vav-driven transgenes (Figures 1 and 2). Indeed, consistent with our previous findings, we did observe strong and highly selective GFP expression in HSCs, whereas the vast majority of hematopoietic progenitors and mature cells were unlabeled (Figure 5). Expression of GFP in a small proportion of downstream populations is a result of incomplete floxing from the specific Cre mouse used. In the case of Flk2-Cre mice, the recombination efficiency varies between individual Flk2-Cre mice [13,14]. Thus, the combination of Vav-GFP and Flk2-Cre enables very bright, highly specific GFP labeling of HSCs.

In conclusion, the Vav-GFP mouse line directs GFP expression exclusively in the hematopoietic system, with the option of differentiation stage specificity by Cre-mediated recombination. We detected strong GFP labeling of all hematopoietic cell types assayed, except for RBCs, with minimal off-target expression in adult tissues. We also detected strong labeling of hematopoietic cells in the developing embryo, with limited off-target GFP expression in ECs. Transplantation experiments demonstrated GFP expression can be used to distinguish donor-derived cells, from host cells in hematopoietic reconstitution assays. We showed that the loxP elements are functional for Cre-mediated recombination by crossing the Vav-Cre mice to a Flk2-Cre line, which lead to highly selective GFP labeling of HSCs in the BM compartment. The Vav-GFP x Flk2-Cre cross provides an excellent example of how the

Vav-GFP line can be used for lineage tracing studies, as well as direct visualization of HSC in adult BM for *in situ* assays, and for tracking HSC migration upon transplantation or mobilization. Our mice can also be crossed to commercially available floxed-stop-reporter mice to achieve labeling of HSC progeny with a second color. For example, a cross to a Rosa26-lox-stop-tomato mouse would result in mice with GFP+ HSCs, Tomato+ hematopoietic progenitor and mature cells, and unlabeled non-hematopoietic cells. Collectively, our data show that the Vav-GFP mouse line, alone, or by breeding to specific Cre lines to obtain selective GFP labeling in hematopoietic subpopulations, represents a novel tool for interrogating the hematopoietic system without fluorescence interference from non-hematopoietic cells. Lastly, our strategy to achieve HSC-specific reporter expression can be utilized for expression of any other transgene specifically in HSCs to assess transgene function exclusively in HSCs.

METHODS

Generation of Vav-GFP transgenic mice

EGFP and HcRed were cloned into the pZ/EG plasmid [21], replacing lacZ and EGPF, respectively, to generate an EGFP-stop-loxP-HcRed-loxP reporter construct. The reporter fragment was then migrated into the vavINS-Cre-IRES-YFP plasmid [10] after SfiI and Not I digestion and 3-piece ligation to replace Cre-IRES-YFP with the EGFP-stop-loxP-HcRed-loxP reporter fragment. The vector was linearized and injected into pronuclei of C57/Bl6 mice at the UCSC transgenic facility. Multiple founders were used to establish a colony, but founder lines were not analyzed separately. Characterization of this line revealed no HcRed fluorescence after floxing. All mice were maintained and investigated in the UCSC vivarium according to IACUC-approved protocols.

Flow Cytometry Analysis

Bone marrow and peripheral blood cells were isolated, processed and analyzed using a four-laser FACSAria or LSRII (BD Biosciences, San Jose, CA) as described previously [19,22,23]. Flowjo Software (Ashland, OR) was used for data analysis and display. Mean fluorescent intensity (MFI) was determined for each cell population by calculating the average intensity for the entire population of both experimental and wild type controls, then subtracting the average wild type intensity from each experimental replicate. Cell populations were defined by the following cell surface phenotypes: HSC ($Lin^{-}Sca1^{+}c\text{-}kit^{+}CD48^{-}Slamf1^{+}Flk2^{-}$), MPP ($Lin^{-}Sca1^{+}$ $c\text{-}kit^{+}CD48^{+}Slamf1^{-}Flk2^{+}$), CMP ($Lin^{-}Flk2^{-}Sca1^{-}$ $c\text{-}kit^{+}FcgR^{mid}CD34^{mid}$), GMP ($Lin^{-}Flk2^{-}Sca1^{-}$ $c\text{-}kit^{+}FcgR^{hi}CD34^{hi}$), MEP ($Lin^{-}Flk2^{-}Sca1^{-}$ $c\text{-}kit^{+}FcgR^{lo}CD34^{lo}$), CLP ($Lin^{-}Sca1^{mid}$ $c\text{-}kit^{mid}Flk2^{+}IL7Ra^{+}$), GM ($Ter119^{-}Mac1^{+}Gr1^{+}CD3^{-}B220^{-}$), Plts ($FSC^{lo}Ter119^{-}CD61^{+}$), B cell ($Ter119^{-}Mac1^{-}Gr1^{-}CD3^{-}B220^{+}$), T cell ($Ter119^{-}Mac1^{-}Gr1^{-}CD3^{+}B220^{-}$), and erythroid progenitors ($B220^{-}$,$CD3^{-}$,$Mac1^{-}$,$Gr1^{-}$, $Ter119^{+}$,$Cd71^{+}$). The lineage cocktail consisted of antibodies recognizing CD3, CD4, CD5, CD8, B220, Gr1, Mac1, and Ter119. Mac1 antibodies were excluded from the lineage cocktail when analyzing fetal progenitors [24].

Transplantation Assays

CD117-enriched bone marrow cells were isolated and double-sorted from Vav-GFP mice using a FACSAriaIII then transplanted into sublethally irradiated (500 rads) WT mice (C57BL6) [13]. 1000 KLS ($ckit^{+}$, lin^{-}, Sca^{+}) cells were transplanted into each recipient and peripheral blood was analyzed from tail bleeds 5 weeks after transplantation.

Non-hematopoietic cell isolation

Whole organs were dissected from Vav-GFP and wild type C57B6 mice and homogenized using a mortar and pestle for all organs. Cell suspensions were passed through 70 micron

filters and stained with anti-CD45 and Ter119 antibodies to exclude hematopoietic cells from analysis. Endothelial cells were isolated as previously described [22]. Briefly, bone marrow cells were isolated from Vav-GFP and wild type mice and were digested in a 3mg/mL collagenase solution for one hour. Similarly, e14.5 fetal livers were isolated from Vav-GFP and wild type embryos and digested with a 1mg/mL collagenase solution. Samples were then filtered and stained with anti-CD45, CD31, Sca1, Tie-2, and Vcam1 antibodies and analyzed using flow cytometry.

ACKNOWLEDGEMENTS

We thank Drs. Thomas Graf and Matthias Stadtfeld for the vavINS-Cre-IRES-YFP plasmid, Corrine Lobe for the Z/EG plasmid, Armen Shamamian at the UCSC vivarium for generation of transgenic mice, Stephanie Smith-Berdan and other Forsberg lab members for technical assistance and helpful discussions. This work was supported by an NIH/NIAID award (R21AI103656), an NIH/NIDDK award (R01DK100917) and UCSC startup funds to ECF; by NIH Training Grant 2T32GM008646 and an HHMI Gilliam Fellow Award to JP-C; CIRM Training grant TG2-01157 to SWB; and by CIRM Shared Stem Cell Facilities (CL1-00506) and CIRM Major Facilities (FA1-00617-1) awards to UCSC. ECF is the recipient of a California Institute for Regenerative Medicine (CIRM) New Faculty Award (RN1-00540) and an American Cancer Society Research Scholar Award (RSG-13-193-01-DDC). Conflicts of interest: The authors have no conflicts to declare.

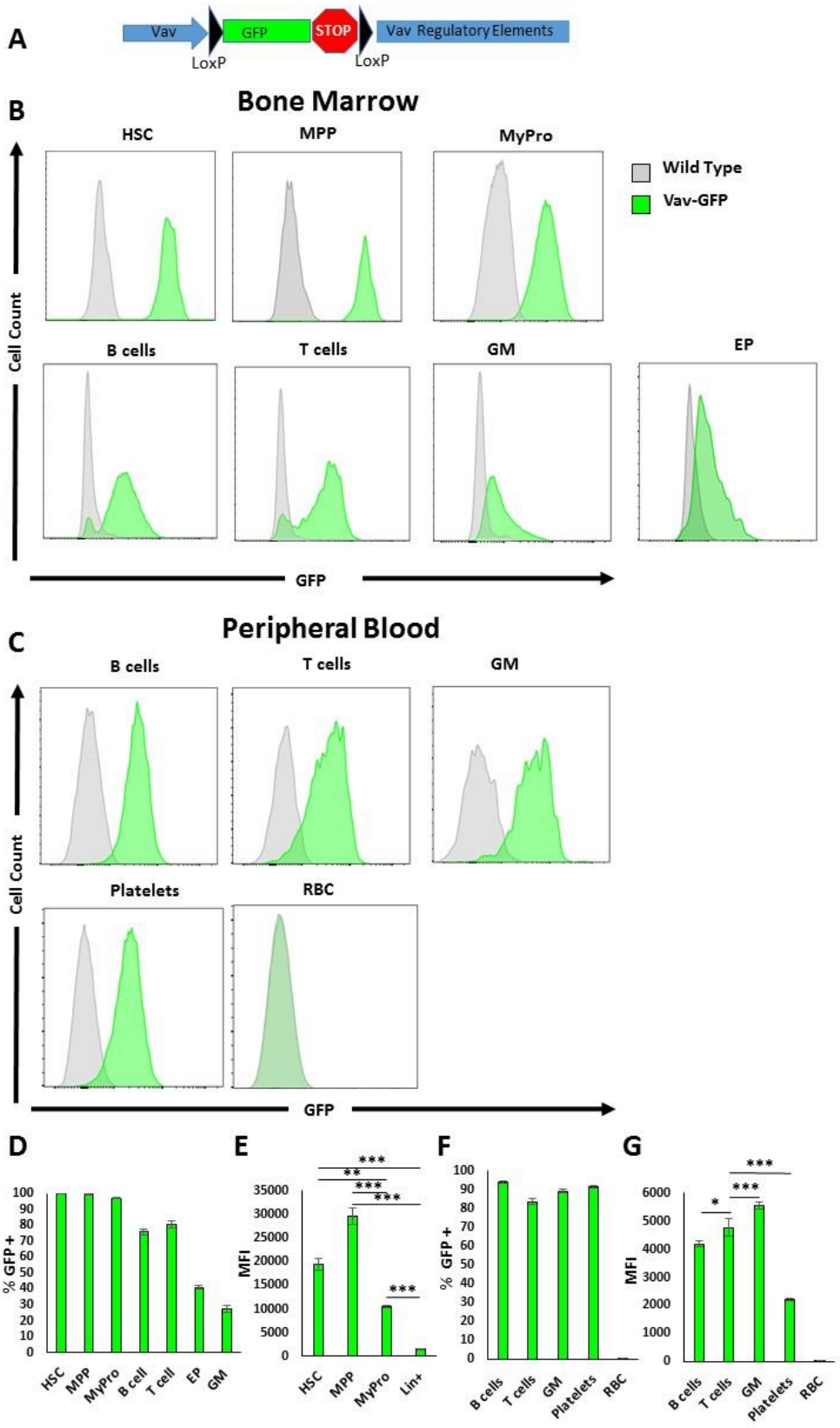

A
Vav
GFP
STOP
Vav Regulatory Elements
LoxP
LoxP
B
Bone Marrow
HSC
MPP
MyPro
Wild Type
Vav-GFP
B cells
T cells
GM
EP
Cell Count
GFP
C
Peripheral Blood
B cells
T cells
GM
Platelets
RBC
Cell Count
GFP
D
% GFP+
HSC
MPP
MyPro
B cell
T cell
EP
GM
E
MFI
HSC
MPP
MyPro
Lin+
F
% GFP+
B cells
T cells
GM
Platelets
RBC
G
MFI
B cells
T cells
GM
Platelets
RBC

Figure 1. Vav-driven GFP expression labels all nucleated hematopoietic cell types as well as platelets in bone marrow and peripheral blood.

(A) Schematic diagram of the Vav-GFP reporter construct.

(B-G) Reporter expression in bone marrow (BM) and peripheral blood (PB) cells in Vav-GFP mice was tested by flow cytometry.

(B) Representative flow cytometer histograms indicating GFP expression levels in the indicated BM cell populations of Vav-GFP (green) and WT (grey) mice.

(C). Representative flow cytometer histograms indicating GFP expression levels in PB cells in Vav-GFP and WT mice.

(D) Proportion of reporter expressing cells in hematopoietic subsets isolated from the BM of Vav-GFP mice is shown as percent of cells positive for GFP.

(E) Mean fluorescence intensity indicates intensity of GFP reporter expression in the indicated BM cell populations.

(F) Proportion of reporter expressing cells in hematopoietic subsets isolated from the PB of Vav-GFP mice is shown as percent of cells positive for GFP.

(G) Mean fluorescence intensity of indicates intensity of reporter expression in the indicated PB cell populations. In all histograms, grey bars represent MFI in wild type samples and green bars represents MFI in Vav-GFP mice.

GFP, green fluorescent protein; WT, wild type; BM, bone marrow; PB, peripheral blood; HSC, hematopoietic stem cell; MPP, multipotent progenitor; MyPro, myeloid progenitor; GM, granulocyte/macrophage; RBC, red blood cell; Lin+, lineage positive (lineage cocktail defined in methods) MFI, mean fluorescent intensity; EP, erythroid progenitor. N = 3-4 mice per experiment. Statistically significant differences were determined by an equal variance T test. * $p<0.05$, ** $p<0.01$, *** $p<0.001$.

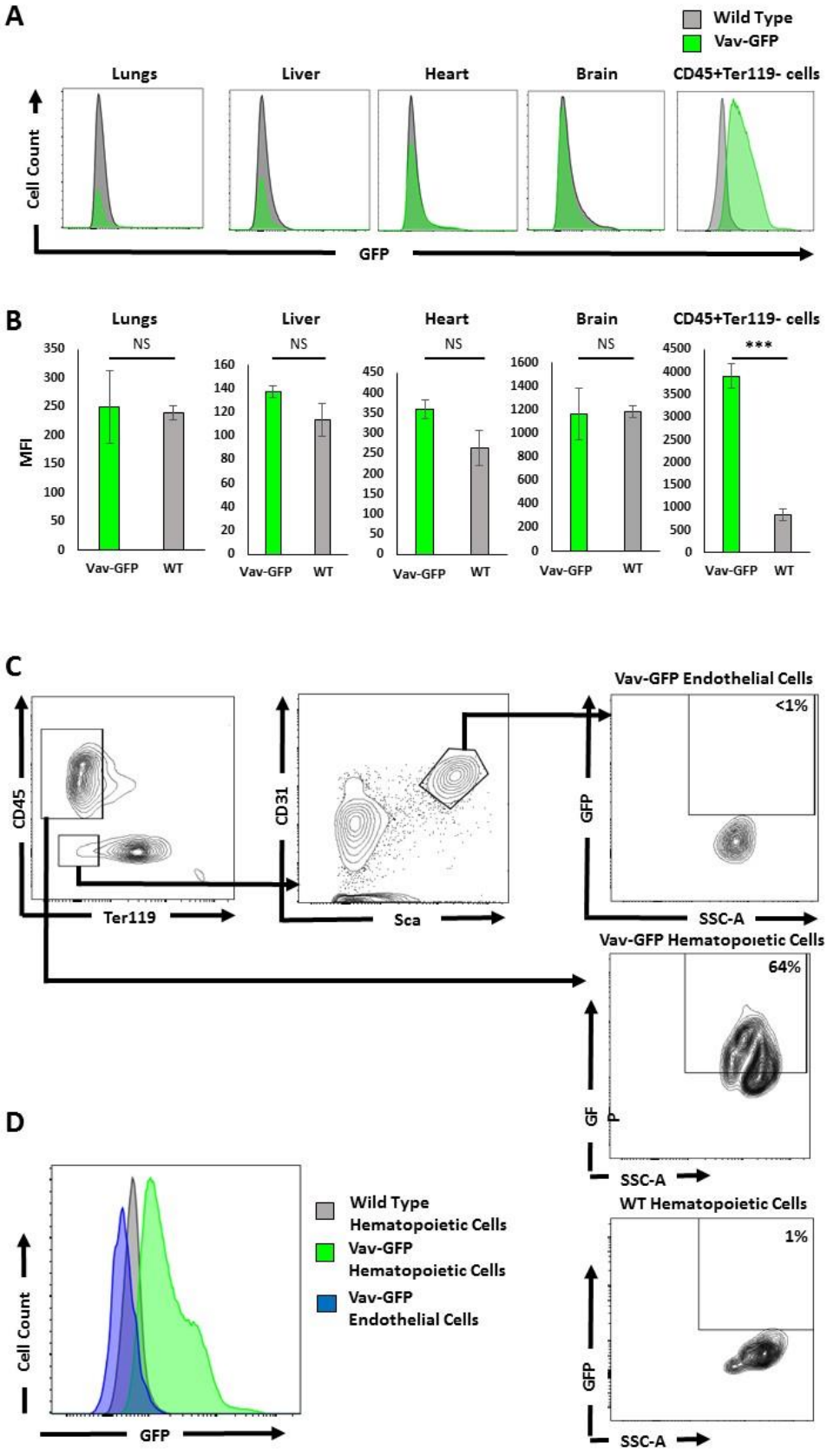

A
Wild Type
Vav-GFP
Lungs
Liver
Heart
Brain
CD45+Ter119- cells
Cell Count
GFP
B
Lungs
Liver
Heart
Brain
CD45+Ter119- cells
NS
NS
NS
NS
***
MFI
Vav-GFP
WT
C
CD45
Ter119
CD31
Sca
Vav-GFP Endothelial Cells
<1%
GFP
SSC-A
Vav-GFP Hematopoietic Cells
64%
D
Cell Count
GFP
Wild Type Hematopoietic Cells
Vav-GFP Hematopoietic Cells
Vav-GFP Endothelial Cells
WT Hematopoietic Cells
1%

Figure 2. GFP expression is largely restricted to hematopoietic cells

(A-B) Vav-GFP reporter expression was observed in hematopoietic (CD45+), but not non-hematopoietic (CD45-), cells in lungs, liver, heart, or brain. Whole organs were isolated from adult Vav-GFP mice and processed into single cell suspensions before antibody staining and analysis by flow cytometry. Representative histograms (A) and mean fluorescent intensity (B) of GFP expression reveal no statistical differences between WT and Vav-GFP mice for non-hematopoietic cells. N = 3 mice. NS, not significant.

(C-D) Endothelial cells (CD45-, Ter119-, Sca1+, CD31+) isolated from the bone marrow of Vav-GFP mice did not express GFP, whereas concurrently isolated CD45+ hematopoietic cells demonstrated robust labeling. (C) Flow cytometric gating strategy used to define endothelial and hematopoietic cells. (D) Histogram plot showing representative GFP expression profiles from hematopoietic and endothelial cells isolated from WT and Vav-GFP mice. Percentages indicate representative frequencies of GFP+ cells for each cell population. N = 6 mice from two independent experiments.

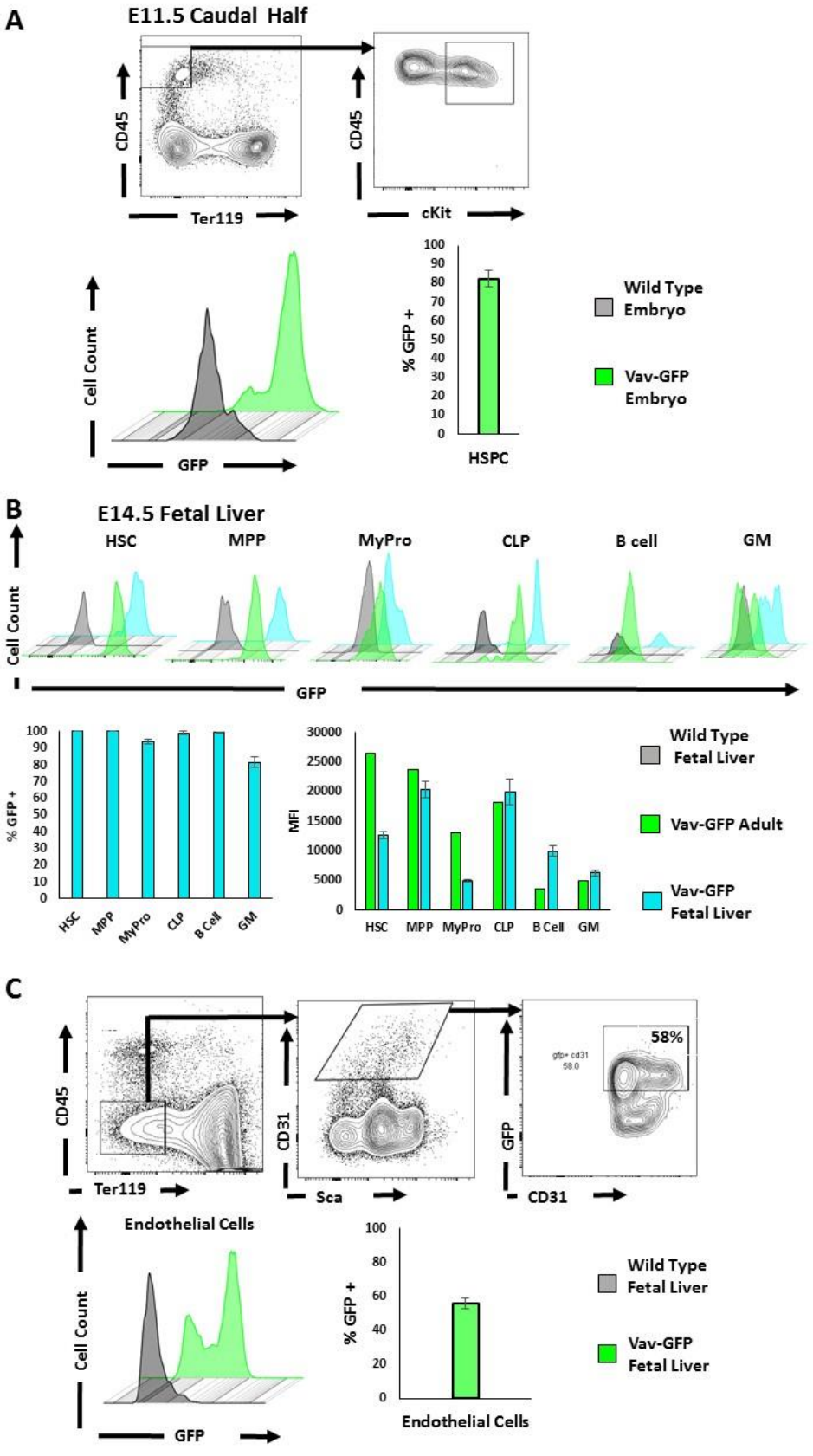

A
E11.5 Caudal Half
CD45
Ter119
CD45
cKit
Cell Count
GFP
% GFP +
HSPC
Wild Type Embryo
Vav-GFP Embryo
B
E14.5 Fetal Liver
HSC
MPP
MyPro
CLP
B cell
GM
Cell Count
GFP
% GFP +
HSC
MPP
MyPro
CLP
B Cell
GM
MFI
Wild Type Fetal Liver
Vav-GFP Adult
Vav-GFP Fetal Liver
C
CD45
Ter119
CD31
Sca
GFP
CD31
58%
Endothelial Cells
Cell Count
GFP
% GFP +
Endothelial Cells
Wild Type Fetal Liver
Vav-GFP Fetal Liver

Figure 3. Vav-GFP reporter expression marks hematopoietic cells across development

(A) Vav-GFP reporter expression was observed as early as E11.5. Representative gating strategy, histograms, and quantification by flow cytometric analysis of reporter activity in HSPCs (CD45+,ckit+, Ter119-) in the caudal half of E11.5 Vav-GFP and WT control embryos. N = 7 embryos representing two independent experiments.

(B) Vav-GFP reporter expression was detected in stem, progenitor, and mature cells isolated from E14.5 fetal liver of Vav-GFP embryos. Representative histograms depict flow cytometric analysis of progenitor and mature cell populations. N = 3 embryos. Grey histograms represent wild type embryos, green represents Vav-GFP adult cells, and blue represents Vav-GFP fetal liver cells.

(C) Reporter expression was detected in a proportion of endothelial cells during development. Gating strategy, representative histogram, and proportion of GFP+ endothelial cells (CD45-, Ter119-, CD31+, $Sca1^{mid}$) isolated from E14.5 fetal liver of Vav-GFP and WT embryos are shown. N = 13 embryos representing two independent experiments.

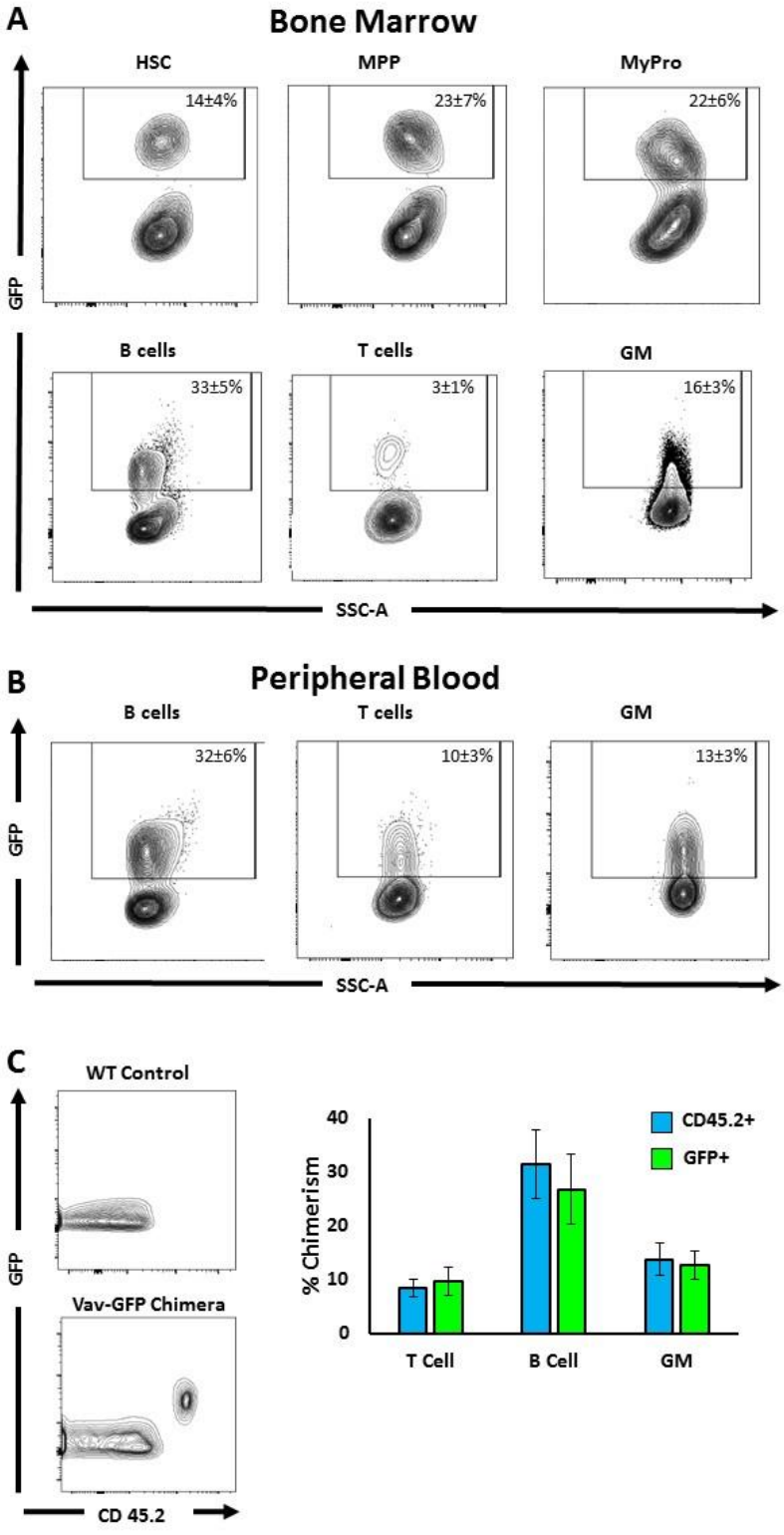

A
Bone Marrow
HSC
MPP
MyPro
14±4%
23±7%
22±6%
GFP
B cells
T cells
GM
33±5%
3±1%
16±3%
SSC-A
B
Peripheral Blood
B cells
T cells
GM
32±6%
10±3%
13±3%
GFP
SSC-A
C
WT Control
GFP
Vav-GFP Chimera
CD 45.2
% Chimerism
40
30
20
10
0
CD45.2+
GFP+
T Cell
B Cell
GM

Figure 4. Robust Vav-GFP expression is maintained in mature cell subsets after transplantation. Adult KLS cells were isolated from Vav-GFP mice and transplanted into sublethally irradiated wild type recipients. Peripheral blood and bone marrow was analyzed 24 weeks after transplantation to assess donor contribution and expression of the Vav-GFP reporter. (A)Representative flow plots depict reporter expression in all mature and progenitor populations analyzed at 24 weeks. (B) Representative flow plots also show reporter expression in mature peripheral blood subsets at 24 weeks. (C) Flow plots showing live, $Ter119^{-}$ cells from an untransplanted recipient (CD45.1) as well as chimeric Vav-GFP (CD45.2). All observed donor cells were double-positive for CD45.2 and GFP. This is quantified on the right for mature peripheral blood cells with CD45.2+ cells shown in blue and GFP+ cells shown in green. Data represent 4 biological replicates.

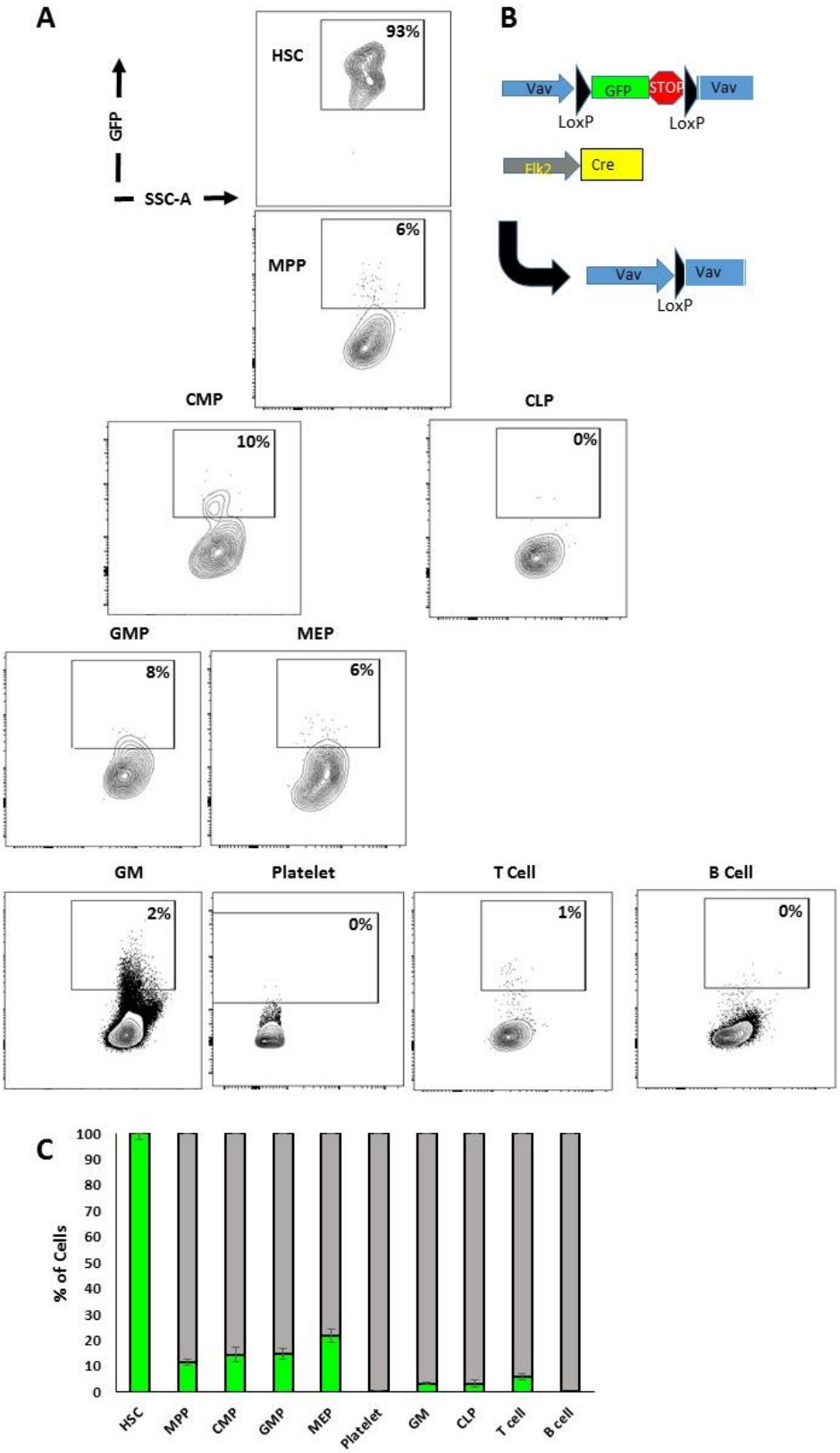

A
B
GFP
SSC-A
HSC
93%
MPP
6%
CMP
10%
CLP
0%
GMP
8%
MEP
6%
GM
2%
Platelet
0%
T Cell
1%
B Cell
0%
Vav
GFP
STOP
Vav
LoxP
LoxP
Elk2
Cre
Vav
Vav
LoxP
C
% of Cells
100
90
80
70
60
50
40
30
20
10
0
HSC
MPP
CMP
GMP
MEP
Platelet
GM
CLP
T cell
B cell

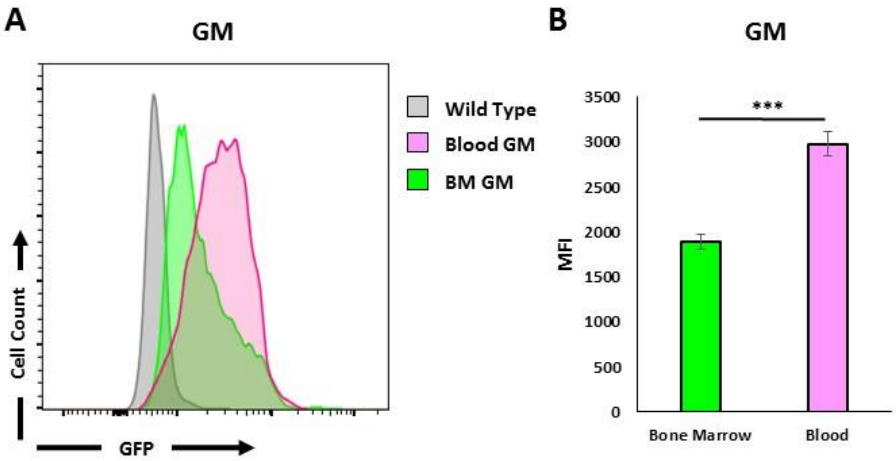
Figure 5. Flk2Cre-mediated excision of Vav-driven GFP selectively labels HSCs.

Vav-GFP mice were crossed to Flk2-Cre mice to assess the ability to specifically label HSCs.

(A) Schematic showing excision of GFP reporter by Cre-mediated recombination.

(B) Representative flow cytometric analysis of reporter expression in hematopoietic subsets isolated from the bone marrow of Vav-GFP/Flk2-Cre double transgenic mice. Numbers indicate the frequency of GFP-expressing cells in each gated cell population of a representative Vav-GFP/Flk2-Cre mouse.

(C) Percentage of cells that expressed GFP in each stem, progenitor, and mature population in the bone marrow. Data represent 11 mice from 3 independent experiments.



Supplementary Figure 1. Circulating granulocyte/monocyte (GM) cells show significantly higher reporter expression than bone marrow resident GM cells. (A) Flow cytometry histograms showing GFP expression levels of wild type circulating GM cells (grey), bone marrow resident Vav-GFP GM cells (green), and circulating Vav-GFP GM cells (magenta). (B) Mean fluorescent intensity of bone marrow resident and circulating GM cells in Vav-GFP mice. *** $p<0.001$.

REFERENCES

1. Duong KL, Das S, Yu S, et al. Identification of hematopoietic-specific regulatory elements from the CD45 gene and use for lentiviral tracking of transplanted cells. *Exp Hematol*. 2014;42(9):761–772.e1–10.

2. Bustelo XR, Rubin SD, Suen KL, Carrasco D, Barbacid M. Developmental expression of the vav protooncogene. *Cell Growth Differ*. 1993;4(4):297–308.

3. Tarakhovsky A, Turner M, Schaal S, et al. Defective antigen receptor-mediated proliferation of B and T cells in the absence of Vav. *Nature*. 1995;374(6521):467–470.

4. Zhan, G R, Alt FW, Davidson L, Orkin SH, Swat W. Defective signalling through the T- and B-cell antigen receptors in lymphoid cells lacking the vav proto-oncogene. *Nature*. 1995;374(6521):470–473. doi:10.1038/374470a0.

5. Fischer KD, Zmuidzinas A, Gardner S, Barbacid M, Bernstein A, Guidos C. Defective T-cell receptor signalling and positive selection of Vav-deficient CD4+ CD8+ thymocytes. *Nature*. 1995;374(6521):474–477. doi:10.1038/374474a0.

6. Zmuidzinas A, Fischer KD, Lira S a, et al. The vav proto-oncogene is required early in embryogenesis but not for hematopoietic development in vitro. *EMBO J*. 1995;14(1):1–11.

7. Georgiades P, Ogilvy S, Duval H, et al. vavCre transgenic mice: A tool for

mutagenesis in hematopoietic and endothelial lineages. *Genesis*. 2002;34(4):251–256. doi:10.1002/gene.10161.

8. Ogilvy S, Metcalf D, Gibson L, Bath ML, Harris a W, Adams JM. Promoter elements of vav drive transgene expression in vivo throughout the hematopoietic compartment. *Blood*. 1999;94(6):1855–1863.

9. de Boer J, Williams A, Skavdis G, et al. Transgenic mice with hematopoietic and lymphoid specific expression of Cre. *Eur J Immunol*. 2003;33(2):314–325. doi:10.1002/immu.200310005.

10. Stadtfeld M, Graf T. Assessing the role of hematopoietic plasticity for endothelial and hepatocyte development by non-invasive lineage tracing. *Development*. 2005;132(1):203–213. doi:10.1242/dev.01558.

11. Kim W-I, Wiesner SM, Largaespada D a. Vav promoter-tTA conditional transgene expression system for hematopoietic cells drives high level expression in developing B and T cells. *Exp Hematol*. 2007;35(8):1231–9. doi:10.1016/j.exphem.2007.04.012.

12. Almarza E, Segovia JC, Guenechea G, Gómez SG, Ramírez Á, Bueren JA. Regulatory elements of the vav gene drive transgene expression in hematopoietic stem cells from adult mice. *Exp Hematol*. 2004;32(4):360–364. doi:10.1016/j.exphem.2004.01.005.

13. Boyer SW, Schroeder A V., Smith-Berdan S, Forsberg EC. All Hematopoietic Cells

Develop from Hematopoietic Stem Cells through Flk2/Flt3-Positive Progenitor Cells. *Cell Stem Cell*. 2011;9(1):64–73. doi:10.1016/j.stem.2011.04.021.

14. Boyer SW, Beaudin AE, Forsberg EC. Mapping differentiation pathways from hematopoietic stem cells using Flk2/Flt3 lineage tracing. *Cell Cycle*. 2012;11(17):3180–3188. doi:10.4161/cc.21279.

15. Benz C, Martins VC, Radtke F, Bleul CC. The stream of precursors that colonizes the thymus proceeds selectively through the early T lineage precursor stage of T cell development. *J Exp Med*. 2008;205(5):1187–1199. doi:10.1084/jem.20072168.

16. Christensen JL, Weissman IL. Flk-2 is a marker in hematopoietic stem cell differentiation: a simple method to isolate long-term stem cells. *Proc Natl Acad Sci U S A*. 2001;98(25):14541–14546. doi:10.1073/pnas.261562798.

17. Adolfsson J, Borge OJ, Bryder D, et al. Upregulation of Flt3 expression within the bone marrow Lin-Sca1+c-kit+ stem cell compartment is accompanied by loss of self-renewal capacity. *Immunity*. 2001;15(4):659–669. doi:10.1016/S1074-7613(01)00220-5.

18. Forsberg EC, Serwold T, Kogan S, Weissman IL, Passegué E. New Evidence Supporting Megakaryocyte-Erythrocyte Potential of Flk2/Flt3+ Multipotent Hematopoietic Progenitors. *Cell*. 2006;126(2):415–426. doi:10.1016/j.cell.2006.06.037.

19. Beaudin AE, Boyer SW, Forsberg EC. Flk2/Flt3 promotes both myeloid and lymphoid development by expanding non–self-renewing multipotent hematopoietic progenitor cells. *Exp Hematol*. 2014;42(3):218–229.e4. doi:10.1016/j.exphem.2013.11.013.

20. Ema, Hideo, Morita Y, Suda T. Heterogeneity and hierarchy of hematopoietic stem cells. *Exp Hematol*. 2014;42(2):74–82. doi:10.1016/j.exphem.2013.11.004.

21. Novak A, Guo C, Yang W, Nagy A, Lobe CG. Z/EG, a double reporter mouse line that expresses enhanced green fluorescent protein upon Cre-mediated excision. *Genesis*. 2000;28(3-4):147–155. doi:10.1002/1526-968X(200011/12)28:3/4<147::AID-GENE90>3.0.CO;2-G.

22. Smith-Berdan S, Nguyen A, Hong MA, Forsberg EC. ROBO4-Mediated Vascular Integrity Regulates the Directionality of Hematopoietic Stem Cell Trafficking. *Stem cell reports*. 2015;4(2):255–268. doi:10.1016/j.stemcr.2014.12.013.

23. Ugarte F, Sousae R, Cinquin B, et al. Progressive Chromatin Condensation and H3K9 Methylation Regulate the Differentiation of Embryonic and Hematopoietic Stem Cells. *Stem Cell Reports*. 2015;5. doi:10.1016/j.stemcr.2015.09.009.

24. Morrison SJ, Hemmati HD, Wandycz M, Weissman IL. The purification and characterization of fetal liver hematopoietic stem cells. *Proc Natl Acad Sci U S A*. 1995;92(22):10302–6. doi:10.1073/pnas.92.22.10302.

CHAPTER III: The spleen represents a significant microenvironment for red blood cell production by multipotent progenitor cells but not hematopoietic stem cells

The text contained within this chapter of the dissertation contains a reprint of the following work submitted for publication: [Scott W. Boyer, Anna E. Beaudin, Jessica Perez-Cunningham, Stephanie Smith-Berdan, Christa Cheung, Herman Tsang, Mark Landon, E. Camilla Forsberg. Clonal and quantitative in vivo assessment of hematopoietic stem and progenitor cell differentiation reveals strong erythroid fate bias of multipotent cells. Cell Stem Cell. In Revision.].

STUDENT'S CONTRIBUTION

Rationale: There has been extensive experimental evidence showing that hematopoietic stem cells (HSC) are multipotent at the single cell and not just population level (Dykstra et al., 2007; Gerrits et al., 2010; Grosselin et al., 2013; Lu et al. 2011; Osawa et al., 1996; Wagers et al., 2002; Yamamoto et al., 2013). However, there have been conflicting reports on whether multipotent progenitors (MPP), the cell type directly downstream of HSC, are multipotent at a single cell level or whether the population consists of several subtypes that have already begun to lose the ability to differentiate into some lineages (Akashi et al., 2000; Beaudin et al., 2014; Boyer et al., 2011, 2012; Forsberg et al., 2006; Kondo et al., 1997; Schlenner et al., 2010). Many of the studies suggesting that MPP are not capable of making various lineages have used *in vitro* assays, so our lab performed a series of *in vivo* transplantion assays to investigate MPP potential at both the single cella and population level.

The spleen is one of the hematopoitic organs that HSCs and MPPs traffic to upon transplantion and provdies an excellent tool to study single cell potential because within the spleen individual transplanted cells form colonies composed of a variety of mature cell types. We found that the cells produced in the highest abundance in these spleen colonies were red blood cells (RBC) regardless of donor cell type, HSC or MPP (Figure 5 of submitted

manuscript). From these data, we hypothesized that the spleen was an important microenvironment for red blood cell (RBC) production by either MPP or HSC.

Method: I surgically removed the spleens from recipient mice and transplanted either HSC or MPP donor cells. I tested the contribution of the transplanted cells to all mature lineages in the peripheral blood. Analyzing the peripheral blood in these mice allows changes to tracked over time in the same mouse which gives us insight into the kinetics of production of various lineages in the same experimental setting. If our hypothesis was correct and the spleen represented an important microenvironment for production of some lineages, then we should see a decrease in contribution of donor cells to those lineages in mice lacking spleens compared to control mice.

Results: MPP donor cells produced significantly fewer RBC in a splenectomized mouse compared to a control mouse. However, they did produce detectable levels of all mature cells types including RBC, platelet, B cells, T cells, and granulocyte/macrophage populations. Transplanted HSC did not produce significantly less of any mature cell type when transplanted into splenectomized recipients compared to control recipients. (Figure 5 and Supplemental Figure 4 of submitted manuscript).

Conclusion: The spleen represents an important microenvironment for RBC production by transplanted MPPs but not HSC. As we know that MPP and HSC express different cell surface receptors, they likely respond differently to the differentiation signals produced by the splenic microenvironment. However, since both transplanted populations produced detectable levels of all mature cells types, the spleen is not a necessary microenvironment for any type of blood cell production.

SUBMITTED MANUSCRIPT

Clonal and quantitative in vivo assessment of hematopoietic stem and progenitor cell differentiation reveals strong erythroid fate bias of multipotent cells

SUMMARY

The lineage potential, heterogeneity, and relationships of hematopoietic stem and progenitor cell populations are currently unclear. To gain new insights on these questions, we performed quantitative analyses of mature cell production from hematopoietic stem cells (HSCs) and multiple hematopoietic progenitor populations. Assessment of the absolute numbers of mature cell types produced by each progenitor cell revealed a striking erythroid bias of all myeloid-competent progenitors assessed, accompanied by strong platelet production that far exceeded generation of myeloid, B and T cells. Clonal analysis by single cell transplantation and spleen colony assays revealed that a significant fraction of HSCs and multipotent progenitors have multilineage potential at the single-cell level. These new insights prompt an erythroid-biased model of hematopoietic differentiation.

INTRODUCTION

Hematopoietic stem cells (HSCs) differentiate via multiple, progressively committed progenitor cell populations to maintain a balanced number of mature blood cells. Despite extensive investigation, the lineage potential, heterogeneity and relationships of hematopoietic stem and progenitor cells (HSPCs) are under intense debate. Data from both single-cell transplantation and barcoding analysis support the existence of long-term, multi-lineage reconstituting clonal

HSCs (Dykstra et al., 2007; Gerrits et al., 2010; Grosselin et al., 2013; Lu et al. 2011; Osawa et al., 1996; Wagers et al., 2002; Yamamoto et al., 2013). However, differential lineage contribution from single cells suggests heterogeneity even within strictly defined HSC compartments (Benz et al., 2012; Yamamoto et al., 2013). Similarly, the heterogeneity and physiological roles of hematopoietic progenitor cells is hotly debated. Evidence from multiple studies indicate that Flk2-positive multipotent progenitors (MPP^F) are utilized as a developmental intermediate, prior to the generation of progenitors restricted to either myeloid (common myeloid progenitors, CMPs) or lymphoid cell (common lymphoid progenitors, CLPs) fates (**Fig S1A**) (Akashi et al., 2000; Beaudin et al., 2014; Boyer et al., 2011, 2012; Forsberg et al., 2006; Kondo et al., 1997; Schlenner et al., 2010). While the existence of multipotent HSCs is widely accepted and recent *in situ* evidence support the existence of multilineage progenitor cells (Boyer et al., 2011, 2012; Busch et al., 2015; Sun et al., 2014), the degree of lineage commitment of hematopoietic populations remains controversial.

Several factors have made it difficult to assess the level of lineage commitment and linage bias within hematopoietic subtypes. Tracking of mature RBC and Plt production from hematopoietic progenitor subsets *in vivo* was developed relatively recently; therefore the full spectrum of mature cell types is rarely simultaneously assessed. Substitute assays, such as hematopoietic differentiation *in vitro,* do not always accurately reflect differentiation *in situ* or upon transplantation *in vivo* (discussed in Boyer et al., 2012; Richie Ehrlich et al, 2011; Schlenner and Rodewald, 2010). In addition, mature cell output from transplanted hematopoietic subtypes is seldom measured quantitatively, precluding accurate comparison of lineage output from specific hematopoietic subsets. Here, we use side-by-side absolute quantification of mature cell production and single cell *in vivo* assays to address the lineage contribution and functional heterogeneity of HSPCs. Our new insights were combined with

previous data into a model of hematopoietic differentiation that reconciles multiple longstanding controversies in hematopoietic stem cell biology.

RESULTS

Lineage potential of hematopoietic cell populations by traditional donor chimerism

To qualitatively and quantitatively assess the differentiation potential of distinct HSPC populations (**Figure S1A and B**), we performed comprehensive analysis of mature cell production upon transplantation into sublethally irradiated mice. Utilization of UBC-GFP mice (Schaefer et al., 2001) allowed for the simultaneous detection of donor-derived red blood cells (RBCs), platelets (Plts), granulocytes/myelomonocytes (GMs), B and T cells (**Figure S1C**). To enable detection of rare and transiently generated cell types, the peripheral blood (PB) of recipient mice was monitored for donor-derived cells at frequent and early time points post-transplantation and a large number of events (≥2,500,000) were collected.

We first displayed reconstitution of each mature cell type as donor chimerism (percentage of donor-derived cells relative to host cells), as is commonly done (**Figure 1A-G**). Aside from a few notable exceptions and the addition of RBC analysis, our results largely agreed with previous reports on transplantation of similarly defined cell populations (Akashi et al., 2000; D'Amico and Wu, 2003; Forsberg et al., 2006; Oguro et al., 2013, Yamamoto et al., 2013). Thus, HSCs gave rise to all five lineages analyzed, without evidence of decline for the duration of the experiment (16 weeks) (**Figure 1A**). MPP^F also gave rise to all five lineages analyzed, with clear declines in chimerism after 21-51 days posttransplantation, depending on the mature cell type (**Figure 1B**). Interestingly, although the Plt contribution from MPP^F was lower than GM, B, or T chimerism, as reported previously (Forsberg et al., 2006; Lai and Kondo, 2006), the RBC chimerism was similar to that of nucleated white blood cells. Both Flk2-negative and Flk2-positive CMPs produced detectable levels of RBCs, Plts and GM, but not B and T, cells in the PB (**Figure 1C and D**). GMPs, MEPs and CLP^F contributed primarily to the GM, RBC, and B cell compartments, respectively (**Figure 1E-G**). Overall, these results agree with the lineage potential previously attributed to the each HSPC populations.

Quantifying absolute numbers of mature cells produced by distinct progenitor populations

Reconstitution displayed as chimerism depends on both donor cell production and on the number of mature host cells present. To compare the effect of radiation conditioning on different types of host cells, we measured mature cell numbers at several time points post-sublethal irradiation. This analysis uncovered a dramatically cell type-specific variation in both the magnitude and kinetics of host cell decrease and recovery, with a rapid, greater than 1000-fold decrease in B cell numbers and only a ~3-fold, slower decrease in RBC numbers (**Figure 1H**). These variations in host cell numbers affect the perceived cell generation from transplanted cells when reconstitution is displayed as donor-to-host chimerism. To remove this variable from the assessment of HSPC reconstitution capacity, we determined the absolute number of each donor-derived mature cell type in the PB after transplantation of different progenitor populations. These data are displayed as the as number of donor-derived cells per microliter of PB (**Figure 1M-S**). To account for differential tissue distribution of each cell type and convey the total number of cells produced per transplanted HSPC in the entire body of the recipient, we also displayed the same data as donor-derived cells generated per mouse per transplanted progenitor cell (**Figure S2A-G**). These assessments of mature cell numbers and tissue distribution between major hematopoietic organs revealed that, as expected, RBCs far outnumbered the other cell types in PB (**Figure 1I**) and that the vast majority of the total RBCs present in a mouse were located in the PB (**Figure 1J**). Plts had a similar distribution pattern, whereas most GM cells were found in the BM. B cells distributed (in order of abundance) between spleen, BM, PB and lymph nodes, and T cells between spleen, thymus, lymph nodes, PB, and BM. Conversely, displaying each tissue based on the abundance of cell types revealed that blood is composed almost entirely of RBCs (87%) and Plts (12%) and that CD3+ T cells make up 88% of the hematopoietic cells in the thymus, whereas other tissues were less

dominated by one cell type (**Figure 1K**). Combining the tissue distribution with PB cell counts provides an estimate of the total numbers of each cell type in a mouse (**Figure 1L**) that are consistent with previous assessments of murine blood cell numbers (Kakumitsu et al., 2005; Nemzek et al., 2001; Seebach et al., 1995; Wilkinson et al., 2001). Making the assumption that donor-derived cells distribute to tissues according to the pattern observed in Figure 1J, these data enabled us to use the PB data (**Figure 1M-S**) to assess the absolute number of each mature cell type generated by each transplanted cell population in each recipient mouse (**Figure S2A-G**).

While the magnitude of the difference between cells that mainly distribute to the blood (RBCs and Plts) and cell types that are dispersed between other tissues (GM, B and T cells) decreased when whole-body distribution was taken into account, the relative order of cell types produced was not altered (compare **Figure 1M-S** with **Figure S2A-G**). However, even though these data were derived from the same transplantation experiments as for **Figure 1A-G**, the absolute quantification of mature cells produced conveyed a different perspective on the ability of different progenitors to reconstitute hematopoiesis (compare **Figure 1A-G** with **Figures 1M-S**). Importantly, the absolute quantification revealed that RBCs were by far the most abundant cell type produced by each progenitor, including GMPs (**Figures 1M-S, S2A-G and Table 1**). The only exception was CLP^F, which stayed true to their reported lymphoid commitment (Forsberg et al 2006; Karsunky et al., 2008; Kondo et al., 1997; Schlenner et al., 2010a, 2010b] by only producing B and T cells (**Figure 1S**). Of note, T cell production by CLP^F was more readily detectable when displayed as absolute numbers (**Figure 1S**) than as chimerism (**Figure 1G**). After RBCs, Plts were the next most abundant mature cell type produced. Despite the low Plt chimerism after transplantation of MPP^F, CMPs, CMP^F, MEPs, and GMPs (**Figure 1B-F**), donor-derived Plts outnumbered the GM, B, and T cells produced from each population (**Figure 1N-R, S2B-F, Table 1**). The observation that GMPs gave rise to RBCs and Plts, and at greater

(RBCs) or ~equal (Plts) numbers than GM cells (**Figure 1Q, S2E, Table 1**) was surprising given the greater GM contribution when data is displayed as chimerism (**Figure 1E**) and their previously reported commitment to the GM lineage (Akashi et al., 2000; Forsberg et al., 2006; Na Nakorn et al., 2002). HSCs, CMPs and MEPs displayed a more expected reconstitution pattern. Notably, CMP^F gave rise to greater numbers of RBCs and Plts than GM cells, (**Figure 1P, S2D, Table 1**). Similarly, though MPP^F displayed complete multipotency, they produced RBCs and Plts in much greater abundance than they produced nucleated mature cells (**Figure 1N, S2B, Table 1**), defying reports that Flk2 expression signifies loss of MegE potential (Adolfsson et al., 2005).

Direct comparison of the mature cell production capacity by different progenitor populations

To assess the relative reconstitution capacity of each progenitor cell type, we then compared the total output of each mature cell type per transplanted HSPC. Each HSC generated more of each mature cell type than any progenitor tested (**Figure 2A-E**). In addition, the time between transplantation and the peak of mature cell production ("time-to-peak") was the longest for HSCs, and mature cell levels persisted without evidence of decline (**Figure 2A-E and 2A'-E'**). These properties are consistent with self-renewal and with HSCs existing at the top of the hematopoietic hierarchy (**Figure S1A**). Per transplanted cell, MPP^F gave rise to more RBCs, Plts, GMs, B-cells, and T-cells than any lineage-restricted progenitor (**Figure 2A-E**). In addition, the time-to-peak for mature cell production from transplanted MPP^F was shorter than for HSCs, but longer compared to lineage-restricted progenitors (**Figure 2A-E, 2A'-E'**). Likewise, the timing and total cell production from both CMP populations were in between the MPP^F and MEP. Relative to CMP^F, CMPs displayed more robust RBC and Plt generation (**Figure 2A-B; A'-B'**), while CMP^F excelled in GM production (**Figure 2C and C'**). While the capacity of CMP^F

to produce RBC and Plt *in vivo* contrasts with *in vitro* data where MegE output from CMP^F was not observed (Nutt et al., 2005), our results are consistent with the relative lineage preferences from previously described *in vitro* data, with CMP^F exhibiting a relative preference for GM production compared to CMP (**Figure 2A'-C' and Table 2**). GMPs produced the same cell types as CMP and CMP^F, but in fewer numbers and with the shortest time-to-peak for any HSPC (**Figure 2A-E, A'-E'**). Overall, these data are consistent with the developmental relationship between HSPCs displayed in **Figure S1A**.

MPP^F give rise to myeloid progenitors in vivo

The lineage potential and absolute number of cells produced can provide insights to the relative hierarchy of cell populations, but not direct mother-daughter relationships. To directly determine if MPP^F can give rise to myeloid-restricted progenitors, we performed phenotypic analysis of donor-derived cells shortly after transplantation (**Figure 3A**). Two days after transplantation of 2,500 HSCs, phenotypic HSCs could be identified in the BM (**Figure 3B**). Similarly, although we were unable to detect either host or donor Flk2+ cells at early timepoints post-transplantation (**Figure S4C**), we detected MPPs defined as KLS, CD150-, CD48+ cells in the BM 2 days after transplantation of MPP^F, but not HSCs, CMPs, or CMP^F (**Figure 3B, Figure S3**). By day 14, transplanted HSCs maintained the self-renewing HSC pool while also repopulating phenotypic MPPs, myeloid progenitors, and mature cells (**Figure 3C**). While transplanted MPP^F did not persist as phenotypic KLS cells for more than a few days, MPP^F repopulated myeloid progenitors and mature cells by day 7 (**Figure 3C**). These data provide direct evidence that MPP^F produce myeloid-restricted progenitor cells.

Transplantation of single MPP^F provides multilineage reconstitution

It is clear from the data presented in **Figures 1M-N and S2A-B** that transplantation of multiple MPP^F result in a similar mature cell production profile as HSCs. However, population data cannot determine if the MPP^F compartment is functionally homogeneous with each cell being multipotent (**Figure 4A**, Multipotency model; Boyer et al., 2012), if MPP^F are a heterogeneous population of committed progenitors that share a common phenotype (**Figure 4A**, Commitment model), or a combination of the two. To differentiate between these possibilities, we implemented two *in vivo* approaches capable of evaluating multilineage readout from single progenitor cells: single-cell transplantation and analysis of single-cell derived CFU-S.

Single-cell transplantation of one HSC or one MPP^F led to detectable levels of donor-derived cells in 19% and 17% of mice, respectively (**Figure 4B**). 55% of mice reconstituted with a single HSC produced cells of at least one myeloid lineage (RBCs, Plts, and/or GMs) and one lymphoid lineage (B-cells and/or T-cells) and were therefore categorized as multipotent (**Figure 4B and C**). The remaining positive recipients of single HSCs produced only myeloid cells, whereas none produced only lymphoid cells (**Figure 4B and C**). None of the single MPP^F gave rise to all five cell types analyzed. Importantly, however, some single MPP^F gave rise to four different types of mature cells, and a considerable fraction of single MPP^F (43%; 9/21) generated both myeloid and lymphoid cells (**Figure 4B and D**). Increasing the number of transplanted MPP^F to 5 or 25 cells led to a higher frequency of recipients with detectable donor cells (**Figure 4B**), an increased number of cell types detected (**Figure 4D-F**), and an increase in donor chimerism levels (**Figure 4H-J**; average overall reconstitution from 1, 5 and 25 MPP^F was 0.9%, 1.5% and 2.2%, respectively). 25 MPP^F were sufficient for combined myeloid and lymphoid detection in 100% of recipients, whereas mice transplanted with 5 or 1 MPP^F or a single HSC were multilineage reconstituted at lower frequencies (37.5%, 43%, and 55%, respectively; **Figure 4B-F**). Although a single MPP^F did not contribute as robustly to recipient chimerism as a single HSC (**Figure 4G-H**; average overall reconstitution was 0.9% and 8%,

respectively), more RBCs than any other mature cell type were generated from both HSCs and MPP^F (**Figure 4K-N**).

Despite high frequency of combined myeloid and lymphoid reconstitution by MPP^F, we noted that MPP^F led to only GM, only B and/or combined GM/B cell reconstitution of some recipients, whereas HSCs did not. In addition, some myelo/lympho reconstitution occurred in unexpected combinations, such as only RBCs and B cells (mouse #6 for single MPP^F, mouse #9 for 25 MPP^F, and mice #s 8 and 9 for 5 MPP^F). GM and B cells occurred together in several mice, without detection of T cells. Indeed, T-cell readout was the most difficult lineage to detect from both HSC and MPP^F transplants (**Figure 4C-F**). As different mature cell types in the host are depleted to different extents upon sublethal (**Figure 1H**) and lethal (**Figure 4O**) irradiation, and donor-derived progeny are produced with different kinetics from different HSCs and MPP^F (**Figure 2A'-E'**), we hypothesized that the kinetics of host and donor cells differentially affects donor-derived cell detection. Superimposing the decline and recovery of host cells with the production of mature cells by HSCs or MPP^F revealed that RBC production from both HSCs and MPP^F largely coincide with the reduction in host RBCs (**Figure 4P**), facilitating detection of donor-derived RBCs in recipient mice. In contrast, neither HSCs nor MPP^F produce large numbers of T cells until after host T cell numbers have significantly recovered (**Figure 4T**), contributing to poor detection of T cells in mice transplanted with a single cell (**Figure 4C and D**). Similarly, the timing of GM and B cell production from MPP^F occur near the low point of host GM and B cells, whereas the main contribution of HSCs to GM and B cells occur after host cell recovery (**Figure 4R and S**). The fact that B cells is the host cell type most affected by irradiation (**Figure 1H and 4O**) and that MPP^F produce B cells prior to host recovery whereas HSCs do not (**Figure 4S**) likely contribute to the relatively high frequency of B cell detection in mice transplanted by a single MPP^F, as well as the more apparent contribution of CLP^F to B cells than T cells (**Figure S4**).

Multilineage reconstitution from single HSCs and MPP^F in spleen colony assays

To test the multipotency of clonal HSCs and MPP^F by an independent method, we analyzed donor-derived colony-forming units of the spleen (CFU-S). Like single cell transplants, this assay measures clonal lineage potential, as each colony consists of progeny from a single cell (Becker et al., 1963; Weber et al., 2011). In support of the clonal origin of CFU-S, we observed only single-color colonies when a mixture of Tomato+ cells and GFP+ cells were transplanted into the same recipient (**Figure 5A**). At early timepoints, most of the clonally-derived progeny are localized within the CFU-S. We reasoned that the resulting high concentration of donor-derived cells in a confined location would reduce the impact of the variables affecting the detection limits in the blood stream. Thus, we utilized the CFU-S assay to test for multilineage potential of single cells by examining individually dissected CFU-S for erythroid progenitors (EPs), megakaryocytes (Megs), GMs, and B-cells (**Figure 5B**). It is highly improbable that CFU-S colonies were contaminated with circulating, donor-derived mature cells, as donor contribution within individual CFU-S was substantially higher than the donor cell chimerism in the peripheral blood (compare **Figure 5B** to **Figure 5C**). T-cell output was not assessed as T-cells require extended development in the thymus, and are not produced in the spleen or in the timeframe of this assay.

Of 13 HSC-derived CFU-S, 6 (46%) contained all four lineages that can be detected in this assay (EPs, Megs, GMs, and B-cells; **Figure 5D, red bar; Figure 5E**). The remaining colonies lacked B cells and were thus a mixture of myeloid lineages. Similarly, 6 of the 18 (33%) single-MPP^F-derived colonies contained all four lineages (**Figure 5D, E**). Other MPP^F-derived CFU-S contained various combinations of myeloid only and/or myeloid plus B cells. All colonies visible to the eye contained erythroid cells. In fact, the vast majority of total cells produced were erythroid cells (**Figure 5F**), and CFU-S derived from CMP, CMP^F, and MEP contained only

EPs (**Figure 5D**). The absence of additional lineages from myeloid progenitors is likely due to their reduced total cell production compared to HSCs and MPP^F (**Figure 2A-C**), limiting our ability to detect mature cells that are not produced in as high abundance as erythrocytes. The CFU-S data, like the single cell transplantations, demonstrate that a substantial fraction of both HSCs and MPP^F possess multilineage reconstitution at the single cell level.

The spleen is an important, but not essential, site of megakaryocyte/erythroid potential

The observation that the vast majority (>90%) of total cells produced by MPP^F in CFU-S are RBCs stands in stark contrast to *in vitro* assays where MPP^F have been reported to lack significant MegE potential. Although the predominant *in vivo* RBC production is also evident in the PB (**Figure 1N**) and in whole-body analysis (**Figure S2B**), those RBCs could be generated in the spleen and then dispersed into the circulation. To test whether the spleen is necessary for MegE production, we transplanted HSCs or MPP^F into splenectomized mice and compared the absolute numbers of donor-derived mature cells in the PB to transplanted control mice. All five investigated cell types were detected from both HSCs and MPP^F in both control and splenectomized recipients (**Figures 5G and S5**), demonstrating that the spleen is not essential for multilineage differentiation. However, while the numbers of GM, B and T cells were comparable in control and splenectomized mice (**Figure S5**), splenectomy resulted in a dramatic reduction in MegE production. RBC generation from transplanted HSCs and MPP^F was 20 and 30 times lower, respectively, in splenectomized mice (**Figure 5G**). Of note, we also detected up to an 18-fold decrease in MPP^F-derived platelet numbers in splenectomized compared to control mice, and only a 2-fold, transient decrease in platelets by HSCs, indicating that the splenic environment is differentially important for MPP^F-mediated megakaryopoiesis. However, both HSCs and MPP^F are also capable of multilineage differentiation outside the spleen.

DISCUSSION

A quantitative perspective of progenitor repopulation capacity

By simply quantifying the absolute number of mature cells generated per transplanted stem and progenitor cell, we provide a new perspective of the regenerative capacity of HSCs and several progenitor populations. Our strategy revealed that all progenitor cells with myeloid potential produce far more RBCs and Plts than any other cell type (**Table 1 and Figures 1M-S, 2A-E, 2A'-E', 4K-N and 5E**). As illustrated by the concordance with previous findings when the same results were displayed as traditional donor chimerism (**Figure 1A-G**), our results are not due to differences in gating strategies or transplantation methods. Rather, the new insights were reached by eliminating the drastically variable fluctuations in host cell disappearance and recovery (**Figures 1H and 4O**) and by taking into account the absolute numbers of each mature cell type produced by each stem and progenitor cell (**Figures 2A-E, 2A'-E', S3**). As transplanted donor cells differ not only in the types of cells generated, but also in the number, timing and duration of mature cell production, each progenitor cell has a distinct reconstitution pattern. In addition, the half-lives of mature cell types vary considerably. Collectively, these dynamics affect the ability to detect and quantify the contribution of different donor populations to each mature lineage, as exemplified in **Figure 4P-T**.

Unexpected lineage potential of hematopoietic progenitor populations

The lineage potential that we uncovered was in some cases unexpected. For example, GMPs have been reported to lack MegE potential (Akashi et al., 2000; Forsberg et al., 2006) and CMP^F produce few megakaryocyte/erythroid colonies *in vitro* (Nutt et al., 2005), whereas our experiments revealed that both CMP^F and GMPs produce more RBCs and Plts than GM cells (**Figure 2A-C, A'-C', S3, Table 1**). Factors that contribute to these contradictory findings include differences between the in vitro and in vivo assay conditions (discussed below), limits

of detection, and the relatively recent development of mice that makes it possible to track RBCs and Plts *in vivo*. The cell production capacity varies drastically between different progenitors: our estimates revealed that each MPP^F produces, on average, ~800,000 cells, whereas one GMP gives rise to ~45 progeny (43 RBCs, 1 Plt and 1 GM cell; **Figure 2A'-E'**). The low burst size of GMPs precludes detection of progeny in single cell transplants and in CFU-S. In contrast, the numbers of RBCs produced by CMP and CMP^F (averaging ~140,000 and 34,000, respectively) was sufficient for detection of EPs in CFU-S, whereas detection of Plts from CMP, CMP^F, GMPs and MEPs required transplantation of higher numbers of progenitors. While absolute purity of bulk populations cannot be guaranteed, contamination cannot explain the differences between the lineage potential apparent in **Figure 1A-G** and **Figure 1M-S** as these data are derived from the exact same experiments.

CLP^F appear restricted to B and T cell generation

The only population in our study that did not produce RBCs and Plts was CLP^F. Like MEPs, CLP^F also lacked detectable GM potential and therefore appear committed to the lymphoid fate, in agreement with their initial designation as CLPs (Kondo et al., 1997) and with *in vivo* lineage tracing experiments (Schlenner et al., 2010a). Interestingly, while B cell production was readily detectable, T cell capacity was more apparent by absolute quantification than by donor chimerism data (compare **Figure 1G** to **1S**). The relative difficulty in detecting CLP^F-derived T cells can be attributed, in part, to host T cells being less affected by irradiation compared to host B cells (~20-fold versus ~1,000 fold, respectively; **Figure 1H**) and because host T cell numbers recover before CLP^F-derived T cells start accumulating (~day 19; **Figure 1S and S4**). In contrast, CLP^F-derived B cells are detected as early as Day 9 after irradiation/transplantation, closely coinciding with the sharp reduction in host B cell numbers (**Figure S4**). Their more limited lineage potential and lower expansion capacity is consistent

with CLP^F residing hierarchically downstream of MPP^F. However, we noted a difference in the B-to-T cell ratio of MPP^F and CLP^F. Our calculations estimated that one MPP^F gives rise to equal numbers of B cells (~810) and T cells (~880), whereas one CLP^F gives rise to 10-fold more B cells than T cells (140 versus 14) (**Figure 2D'-E'**). This finding appears inconsistent with a direct and exclusive MPP^F to CLP^F transition, but rather evokes additional intermediate populations and/or flexibility in differentiation pathways from upstream progenitors.

Multilineage differentiation potential of single HSCs and MPP^F

It is clear from both the chimerism (**Figure 1B**) and absolute cell number data (**Figures 1N and 2A-E, S3**), as well as previous publications (Forsberg et al., 2006; Miyawaki et al., 2014; Yamamoto et al., 2013), that MPP^F are multipotent at the population level. Here, we also show that a fraction of MPP^F are multipotent at the single cell level. In CFU-S assays, MPP^F produced both myeloid and lymphoid cells in 67% of the colonies, with the remaining colonies consisting of myeloid cells only (**Figure 5D-E**). While the CFU-S frequency from MPP^F is lower than that from HSCs (~1/78 versus 1/33; Beaudin et al., 2014; Forsberg et al., 2006), CFU-S capability is clearly an underestimate of the MegE potential of a cell as far more than 1/33 HSCs are multipotent in other assays. Indeed, the CFU-S frequency increases ~10-fold when HSCs or MPP^F are injected directly into the spleen as opposed to IV (Beaudin et al., 2014). In addition, our splenectomy experiments demonstrate that HSCs and MPP^F also have multilineage potential outside the spleen (**Figure 5G and S5**). Collectively, these data demonstrate that the ability of MPP^F to generate MegE cells is substantially greater than estimated by CFU-S alone.

A similar proportion of HSCs and MPP^F displayed combined myelo/lympho potential at the clonal level upon single cell transplantation (55% of HSCs and 43% of MPP^F; **Figure 4B-C**). Our numbers are similar to those of Yamamoto et al, who found that 56% (9/16 mice) of "LMPPs" (in their report defined as KLS CD34+Flk2+, a population significantly overlapping

with our MPP^F) displayed combined myelo/lymphoid readout (Yamamoto et al., 2013). *In vivo* evidence for the existence of clonal MPP was also provided by in situ barcoding studies (Sun et al., 2014). The exact proportion of uncommitted cells within the HSC and MPP^F compartment is difficult to estimate, as variable lineage outputs could be a result of detection limits (progeny were produced but not in sufficient numbers to detect), stochasticity (the transplanted cell happened to encounter a particular combination of cytokines), population heterogeneity (a proportion of the cells are multipotent, whereas some are not), or a combination of the three. Few studies have tested the full lineage potential of MPP^F, but the heterogeneity of HSCs has been tested and debated extensively (Ema et al., 2014; Hock, 2010; Schroeder, 2010). The HSC results are highly relevant for the heterogeneity of MPP^F: a lineage restricted cell (a committed HSC) cannot give rise to a cell with greater lineage potential (a multipotent MPP). Thus, the proportion of multipotent cells within the MPP^F fraction must be equal to or smaller than the fraction of HSCs that are multipotent (**Figure 4A**; Boyer et al., 2012). Conversely, if some MPP^F are uncommitted, an equal or greater proportion of HSCs and other upstream populations should also be uncommitted. In our experiments, the frequency of single cells with combined myelo/lympho potential is similar for HSCs and MPP^F (55% and 43% for singly transplanted HSCs and MPP^F, respectively, and 46% and 67% for HSCs and MPP^F, respectively, in CFU-S). Given that it is harder to detect progeny from MPP^F than HSCs, the extent of lineage restriction upon the transition from HSCs to MPP^F appear quite low.

RBC production as the default hematopoietic fate

While the overwhelming dominance of RBC production should not be unexpected, models of hematopoietic differentiation rarely take the vastly different numbers of mature cells into account. We propose a model where RBC production is the default pathway of HSCs (**Figure**

6). In this model, MegE potential is gained upon specification of HSCs from a hematovascular precursor. Differentiation into alternative fates is accomplished by the combination of two mechanisms: increased expression of proteins promoting GM, B, and/or T cell differentiation, and a concurrent decrease in expression of genes driving RBC and Plt differentiation. In vivo, continuous expression of MegE-promoting genes does not appear necessary, conceivably because MegE differentiation has already been initiated. Unless exposed to sufficient concentrations of a combination of factors to alter this path, most hematopoietic progenitors will produce primarily RBCs and Plts. Conversely, downregulation of MegE-promoting genes may be necessary for deviation from this default pathway to allow for the relatively rare production of GM, B and T cells. When the default pathway is interrupted (removal of cells from their natural environment), progenitors that have downregulated MegE-specifying receptors (MPP^F, CMP^F and GMPs) are unable to *reinitiate* MegE production. Thus, they perform relatively poorly in MegE assays *in vitro* (Adolfsson et al., 2005; Nutt et al., 2005). As these cell types have acquired expression of receptors that promote GM or lymphoid fates, they are reactive to the corresponding cytokines and consequently readily differentiate *in vitro* into cell types that normally (*in vivo*) represent alternative fates. The proposed model provides an explanation for the discordant lineage potential by *in vivo* and *in vitro* strategies. Collectively, MPP^F may better viewed as cells that have *gained* GM, B and T potential than as cells that have *lost* MegE potential. This view is supported by our previous demonstration that all hematopoietic lineages are derived via a Flk2+ stage (Boyer et al., 2011, 2012), by Flk2 promoting expansion of all mature blood cell types (Beaudin et al., 2014), and by recent reports pointing to MPPs, rather than HSCs, as the major source of mature hematopoietic cells during steady-state hematopoiesis *in situ* (Busch et al., 2014; Sun et al., 2014).

MATERIALS AND METHODS

Transplantation assays

Hematopoietic cells were isolated from BM isolated from murine femurs and tibias from wild-type (C57Bl6) or UBC-GFP mice (Schaefer et al., 2007; The Jackson Laboratory, Stock # 004353) in accordance with UCSC guidelines, as described (Beaudin et al., 2014; Smith-Berdan et al., 2011, 2015; Ugarte et al., 2015). CD117-enriched bone marrow cells were double-sorted using a FACSAriaIII then transplanted into sublethally (500 rads) or lethally (~1000 rads) recipients. HSC (Lineage-, cKit+, Sca1+, CD150+, Flk2-), MPP (Lineage-, cKit+, Sca1+, CD150-, Flk2+), CMP^{-} (Lineage-, $cKit^{+}$, $Sca1^{-}$, $FcγRα^{mid}$, $CD34^{mid}$, $Flk2^{-}$), CMP^{+} (Lineage-, $cKit^{+}$, $Sca1^{-}$, $FcγRα^{mid}$, $CD34^{mid}$, $Flk2^{+}$), MEP (Lineage-, $cKit^{+}$, $Sca1^{-}$, $FcγRα^{lo}$, $CD34^{lo}$), GMP (Lineage-, $cKit^{+}$, $Sca1^{-}$, $FcγRα^{hi}$, $CD34^{hi}$), CLP (Lineage-, $cKit^{mid}$, $Sca1^{mid}$, $IL7Rα^{+}$, $Flk2^{+}$). The lineage cocktail was comprised of CD3, CD4, CD5, CD8, Ter119, Mac1, Gr1, and B220.

Mature cell quantification

A known volume of peripheral blood was mixed with an antibody solution (Ter119, CD61, Mac1, Gr1, B220, CD3) containing a known quantity of Calibrite-APC beads (Catalog no. 340487) prior to flow cytometry analysis. For tissues, a known quantity of Calibrite-APC beads was added to each tissue preparation prior to antibody staining and analysis. The number of beads counted by flow cytometry for blood and tissue samples was used to calculate the number of mature cells per microliter of blood or within each tissue. RBC (FSC^{lo-mid}, Ter119+, CD61-, Mac1-, Gr1-, B220-, CD3-), Platelets (SSC^{lo}, Ter119-, CD61+, Mac1-, Gr1-, B220-, CD3-), GM (FSC^{mid-hi}, Ter119-, CD61-, Mac1+, Gr1+, B220-, CD3-), B-cell (FSC^{mid}, Ter119-, CD61-, Mac1-, Gr1-, B220+, CD3-), T-cell (FSC^{mid}, Ter119-, CD61-, Mac1-, Gr1-, B220-, CD3+). The distribution of mature hematopoietic cells in a mouse was measured in the blood obtained by perfusion; in bone marrow by analysis of two femurs and tibias; spleen; thymus; and lymph nodes (inguinal, axillary, and superficial cervical).

HSPC analysis post-transplantation

2.5K HSC, 20K MPP^F, 50K CMP and 50K CMP^F were FACS purified from UBC-GFP mice and transplanted into sublethally irradiated (500 rads) WT recipients (C57BL6). Between 2-14 days post-transplant, BM was isolated from femurs and tibias were analyzed by flow cytometry for donor-derived cells. KLS (GFP+, Lineage-, cKit+, Sca1+), MyPro (GFP+, Lineage-, cKit+, Sca1-), HSC (GFP+, Lineage-, cKit+, Sca1+, CD48-, CD150+), MPP^F (GFP+, Lineage-, cKit+, Sca1+, CD48+, CD150-), Mature (Lineage+, cKit-).

Single-cell transplants

Individual HSCs and MPP^F were double-sorted into separate wells on Terasaki plates using a FACSAriaIII from lineage-depleted bone marrow cells from UBC-GFP mice. Fluorescence microscopy was used to verify that only one cell occupied each well. Individual cells were loaded into a 0.5 mL syringe pre-loaded with 200,000 WT BM cells. One syringe was used per lethally irradiated (1,000 rads) WT recipient to inject one single HSCs or one single MPP^F retroorbitally per recipient. Donor contribution to mature cells was assessed in the peripheral blood over time.

CFU-S analysis

Lethally irradiated (1,000 rads) WT mice were transplanted with an equal mixture of double-sorted cells isolated from mT/mG (Muzumdar et al., 2007) and UBC-GFP mice. On day 8.5 (MEP), 9.5 (CMP and CMP^F), 11.5 (MPP^F), and 13.5 (HSC) post-transplantation, mice were sacrificed and perfused to remove peripheral blood. Individual CFU-S were removed with a scalpel under a fluorescent dissecting scope. Single-cell suspensions of dissected colonies were labeled with the following antibodies: Ter119, CD41, Mac1, Gr1, and B220. Cell types were defined as follows: Erythroid Progenitor (EP; FSC^{mid-hi}, Ter119+, CD41-, Mac1-, Gr1-,

B220-); Megakaryocyte (Meg; FSC^{mid-hi}, Ter119-, CD41+, Mac1-, Gr1-, B220-); GM (FSC^{mid-hi}, Ter119-, CD41-, Mac1+, Gr1+, B220-); B-cell (FSC^{mid-h},Ter119-, CD41-, Mac1-, Gr1-, B220+).

Splenectomies

Spleens were surgically removed from WT mice (C57BL6) and mice were allowed to recover for four weeks before transplantation. Mice were then sublethally irradiated (500 rads) and transplanted with either 200 HSCs (Lineage-, cKit+, Sca1+, CD150+, Flk2-) or 1000 MPP^F (Lineage-, cKit+, Sca1+, CD150-, Flk2+) isolated from UBC-GFP mice using a FACS ARIA III. Mature blood cell readout was analyzed as described above in the “Mature cell quantification” section of the methods.

TABLES

Table 1. Approximate proportion of the absolute number of mature cells generated by each transplanted progenitor cell type. nd, not detected.

	RBC	Plt	GM	B	T
HSC	94.7	3.48	0.406	0.91	0.46
MPP^F	99.4	0.38	0.060	0.10	0.11
CMP	99.8	0.17	0.001	nd	nd
CMP^F	99.8	0.20	0.041	nd	nd
GMP	98.3	0.57	1.143	nd	nd
MEP	99.9	0.08	nd	nd	nd
CLP	nd	nd	nd	90.9	9.1

Figure 1

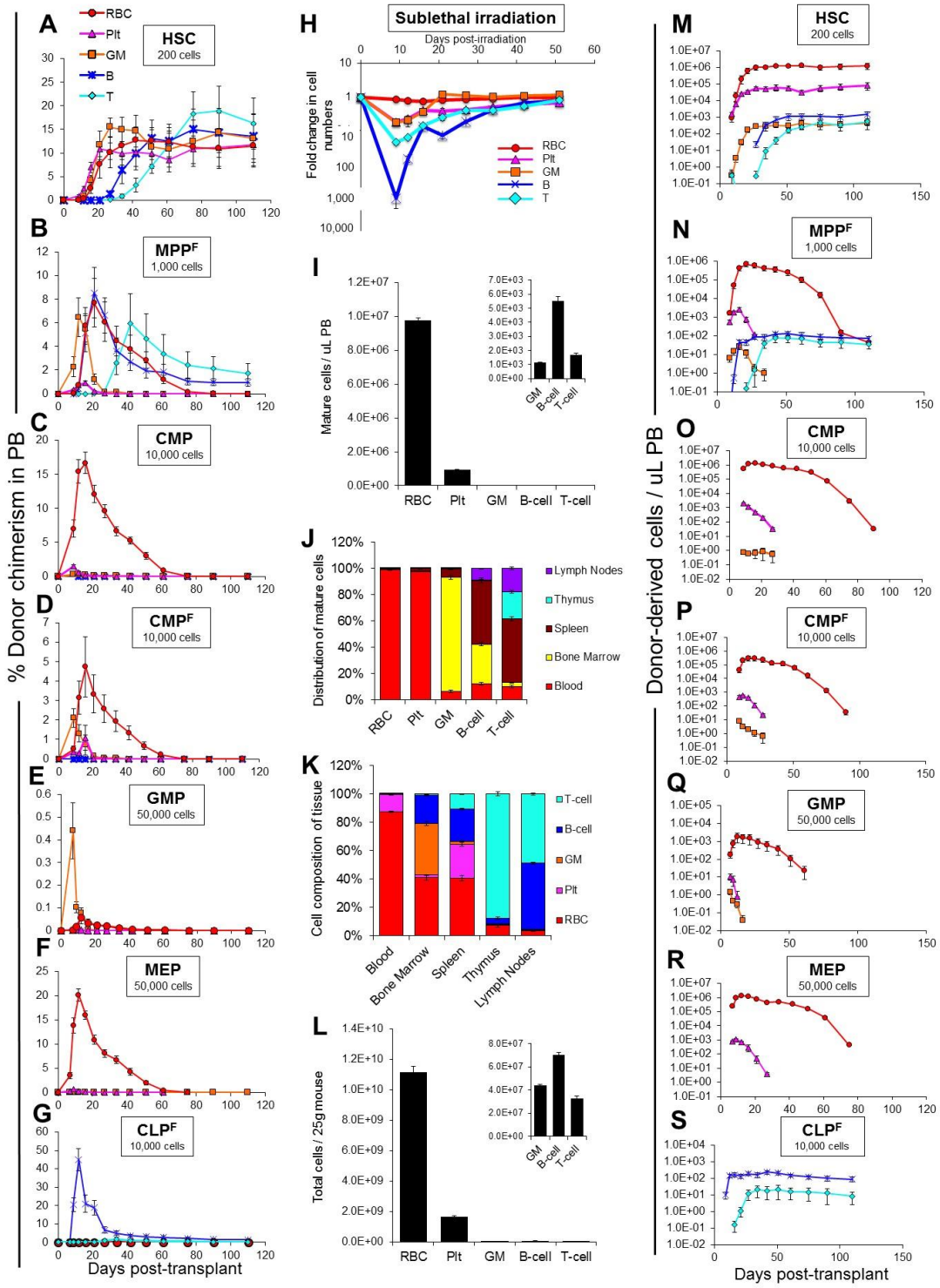

A
HSC
200 cells
RBC
Plt
GM
B
T
B
MPP^F
1,000 cells
C
CMP
10,000 cells
D
CMP^F
10,000 cells
E
GMP
50,000 cells
F
MEP
50,000 cells
G
CLP^F
10,000 cells
% Donor chimerism in PB
Days post-transplant
H
Sublethal irradiation
Days post-irradiation
Fold change in cell numbers
I
Mature cells / uL PB
RBC
Plt
GM
B-cell
T-cell
J
Distribution of mature cells
Lymph Nodes
Thymus
Spleen
Bone Marrow
Blood
K
Cell composition of tissue
T-cell
B-cell
GM
Plt
RBC
Blood
Bone Marrow
Spleen
Thymus
Lymph Nodes
L
Total cells / 25g mouse
M
HSC
200 cells
N
MPP^F
1,000 cells
O
CMP
10,000 cells
P
CMP^F
10,000 cells
Q
GMP
50,000 cells
R
MEP
50,000 cells
S
CLP^F
10,000 cells
Donor-derived cells / uL PB
Days post-transplant

Figure 1. Reconstitution potential of transplanted hematopoietic stem and progenitor cell populations.

(**A-G**) Percent donor chimerism over 110 days from HSCs (**A**), MPP^F (**B**), CMPs (**C**), CMP^F (**D**), GMPs (**F**), MEPs (**E**), or CLP^F (**G**) upon transplantation into sublethally irradiated mice.

(**H**) B cell numbers display a rapid and more drastic decline (1,000-fold) after sublethal irradiation than other mature cell types (3-, 6-, 6-, and 150-fold for RBCs, Plts, GM and T cells, respectively). Data displayed are fold changes in mature cell numbers in the PB of sublethally irradiated mice over time. n=7.

(**I**) The number of mature hematopoietic cells in a microliter of peripheral blood (PB) at steady-state. n=10.

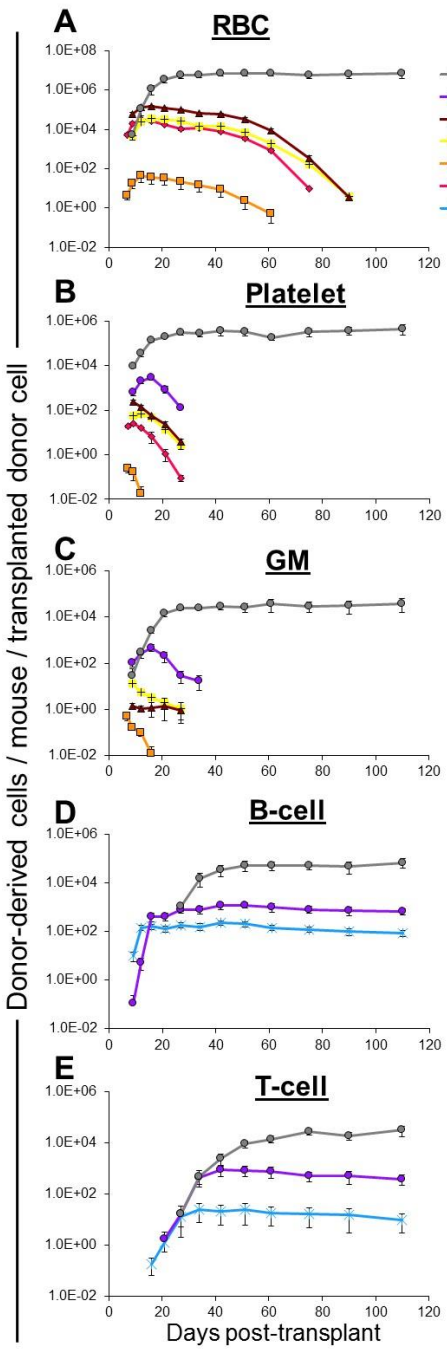
(**J**) The distribution of mature hematopoietic cells between blood, bone marrow, spleen, thymus, and lymph nodes of a mouse. n=10.

(**K**) The composition of mouse blood, bone marrow, spleen, thymus, and lymph nodes displayed as a percentage of total mature hematopoietic cells. n=10.

(**L**) The number of mature hematopoietic cells in a 25 gram mouse at steady-state. n=10.

(**M-S**) Reconstitution data from A-G replotted as the absolute number of donor-derived cells per microliter PB. Transplantation data in A-G and M-S are representative averages from at least 7 recipient mice per cell type. All data are means ± SEM. HSC – Hematopoietic Stem Cell; MPP^F – Multipotent Progenitor; CMP – Common Myeloid Progenitor; CLP^F – Common Lymphoid Progenitor; GMP – Granulocyte/Myelomonocyte Progenitor; MEP – Megakaryocyte/Erythrocyte Progenitor; RBC - Red Blood Cell; Plt - Platelet; GM - Granulocyte/Myelomonocyte.

Figure 2



A'

	Day of peak	No. cells at peak
HSC	27 +	5.6E6 (±1.8E6)
MPP^F	21	7.9E5 (±2.4E5)
CMP	12	1.4E5 (±1.8E4)
CMP^F	16	3.4E4 (±1.2E4)
GMP	12	4.3E1 (±2.3E1)
MEP	12	3.2E4 (±3.2E3)
CLP^F	N/A	N/A

CMP > CMP^F p=0.0002

B'

	Day of peak	No. cells at peak
HSC	27 +	3.1E5 (±7.3E4)
MPP^F	16	3.0E3 (±6.8E2)
CMP	9	2.4E2 (±3.5E1)
CMP^F	12	6.9E1 (±1.1E1)
GMP	7	2.5E-1 (±9.7E-2)
MEP	9	2.6E1 (±1.9E0)
CLP^F	N/A	N/A

CMP > CMP^F p=0.0006

C'

	Day of peak	No. cells at peak
HSC	27 +	2.4E4 (±3.4E3)
MPP^F	16	4.8E2 (±1.6E2)
CMP	9	1.4E0 (±4.1E-1)
CMP^F	9	1.4E1 (±3.1E0)
GMP	7	5.0E-1 (±1.6E-1)
MEP	N/A	N/A
CLP^F	N/A	N/A

CMP^F > CMP p=0.003

D'

	Day of peak	No. cells at peak
HSC	51 +	5.4E4 (±2.4E4)
MPP^F	27	8.1E2 (±2.3E2)
CMP	N/A	N/A
CMP^F	N/A	N/A
GMP	N/A	N/A
MEP	N/A	N/A
CLP^F	12	1.4E2 (±3.4E1)

E'

	Day of peak	No. cells at peak
HSC	75 +	2.7E4 (±8.5E3)
MPP^F	42	8.8E2 (3.9E2)
CMP	N/A	N/A
CMP^F	N/A	N/A
GMP	N/A	N/A
MEP	N/A	N/A
CLP^F	27	1.4E1 (±8.3E0)

Figure 2. Comparative mature cell production capacity and kinetics by transplanted hematopoietic stem and progenitor cell populations.

(**A-E**) The total number of RBCs (**A**), Plts (**B**), GMs (**C**), B-cells (**D**), and T-cells (**E**) generated per mouse per transplanted HSPC.

(**A'-E'**) The approximate timepoint and cell number for peak donor-derived mature cells in the PB from A-E. Data were generated from the transplantation and cell distribution experiments of Figure 1. The mature cell production capacity was always in the same order: HSCs > MPP^F > CLP^F (for B and T cells) and HSCs > MPP^F > CMP > CMP^F > MEP > GMP (for RBCs, Plts, and GM cells). The only exception was that CMP^F produced significantly more GM cells than did CMP (**C** and **C'**; $p<0.003$ by Student's two-tailed t-test). All other comparisons were also significant.

Figure 3

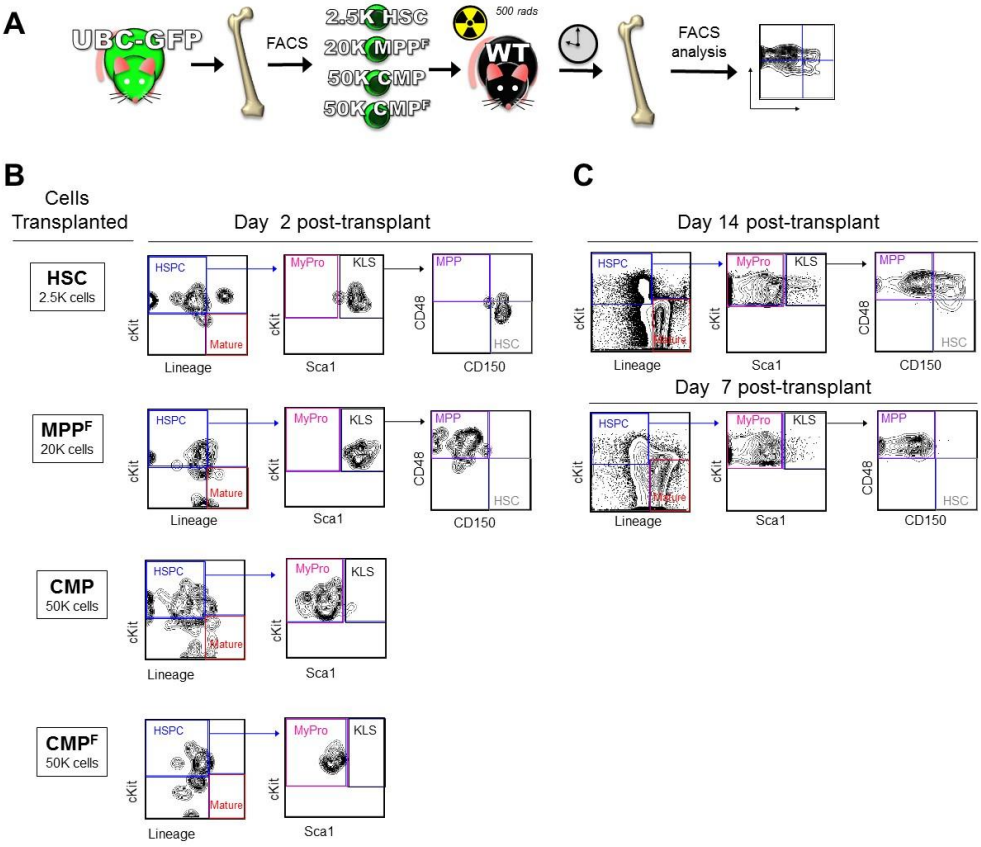
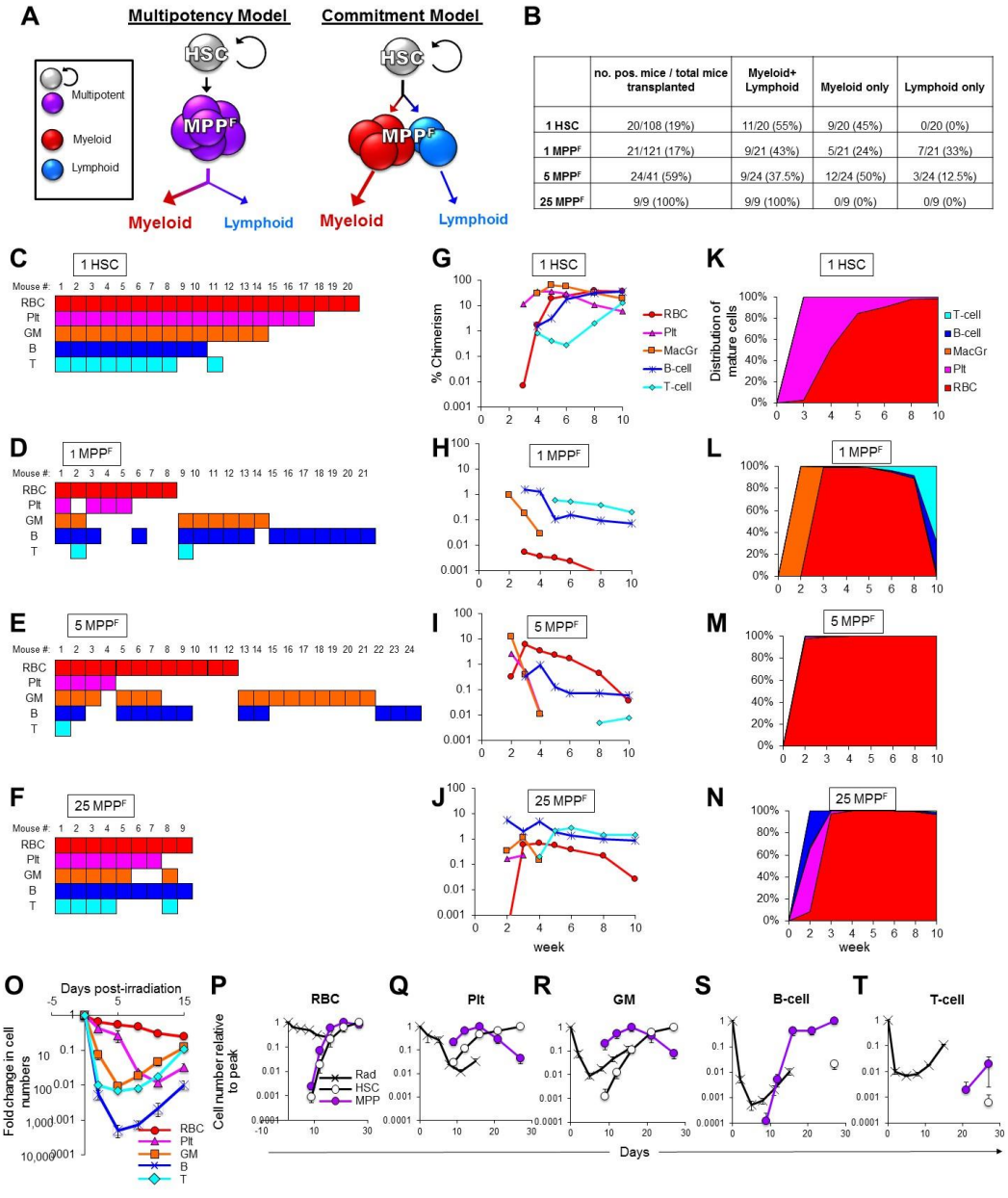

A
UBC-GFP
FACS
2.5K HSC
20K MPP^F
50K CMP
50K CMP^F
500 rads
WT
FACS analysis
B
Cells Transplanted
Day 2 post-transplant
HSC
2.5K cells
MPP^F
20K cells
CMP
50K cells
CMP^F
50K cells
HSPC
Mature
MyPro
KLS
MPP
HSC
cKit
Lineage
Sca1
CD48
CD150
C
Day 14 post-transplant
Day 7 post-transplant

Figure 3. Both HSCs and MPP^F generate myeloid progenitors after transplantation.

(**A**) Schematic of HSPC transplantation from UBC-GFP mice and short-term analysis of donor-derived progenitor cells.

(**B**) Analysis of donor-derived cells shortly after transplantation of HSCs, MPP^F, CMP, and CMP^F into sublethally irradiated recipients. Cells were pre-gated on GFP+ to only display donor-derived cells. At Day 2 post-transplantation, the phenotype of GFP+ cells was predominantly the same as the phenotype of the input cell type. (**C**) At Day 7 (MPP^F) and 14 (HSCs), both cells of the input phenotype and cells with a myeloid progenitor phenotype were present in recipients transplanted with either HSCs or MPP^F. KLS (Lineage-, cKit+, Sca1+); MyPro (Lineage-, cKit+, Sca1-); HSC (Lineage-, cKit+, Sca1+, CD150+, CD48-); MPP (Lineage-, cKit+, Sca1+, CD150-, CD48+). n=4 recipients in 2 independent experiments (HSC and MPP^F); n=4 recipients in 1 experiment (CMP and CMP^F).

Figure 4



	no. pos. mice / total mice transplanted	Myeloid+ Lymphoid	Myeloid only	Lymphoid only
1 HSC	20/108 (19%)	11/20 (55%)	9/20 (45%)	0/20 (0%)
1 MPP^F	21/121 (17%)	9/21 (43%)	5/21 (24%)	7/21 (33%)
5 MPP^F	24/41 (59%)	9/24 (37.5%)	12/24 (50%)	3/24 (12.5%)
25 MPP^F	9/9 (100%)	9/9 (100%)	0/9 (0%)	0/9 (0%)

Figure 4. Single HSCs and MPP^F reconstituted both myeloid and lymphoid lineages *in vivo*.

(**A**) Alternative models of MPP^F multipotency. The multilineage reconstitution by MPP^F transplanted in bulk (Figures 1B and 1N) may be derived from clonally multipotent MPP^F (left panel) or from the combined contribution of lineage-committed cells with a shared MPP^F surface phenotype (right panel).

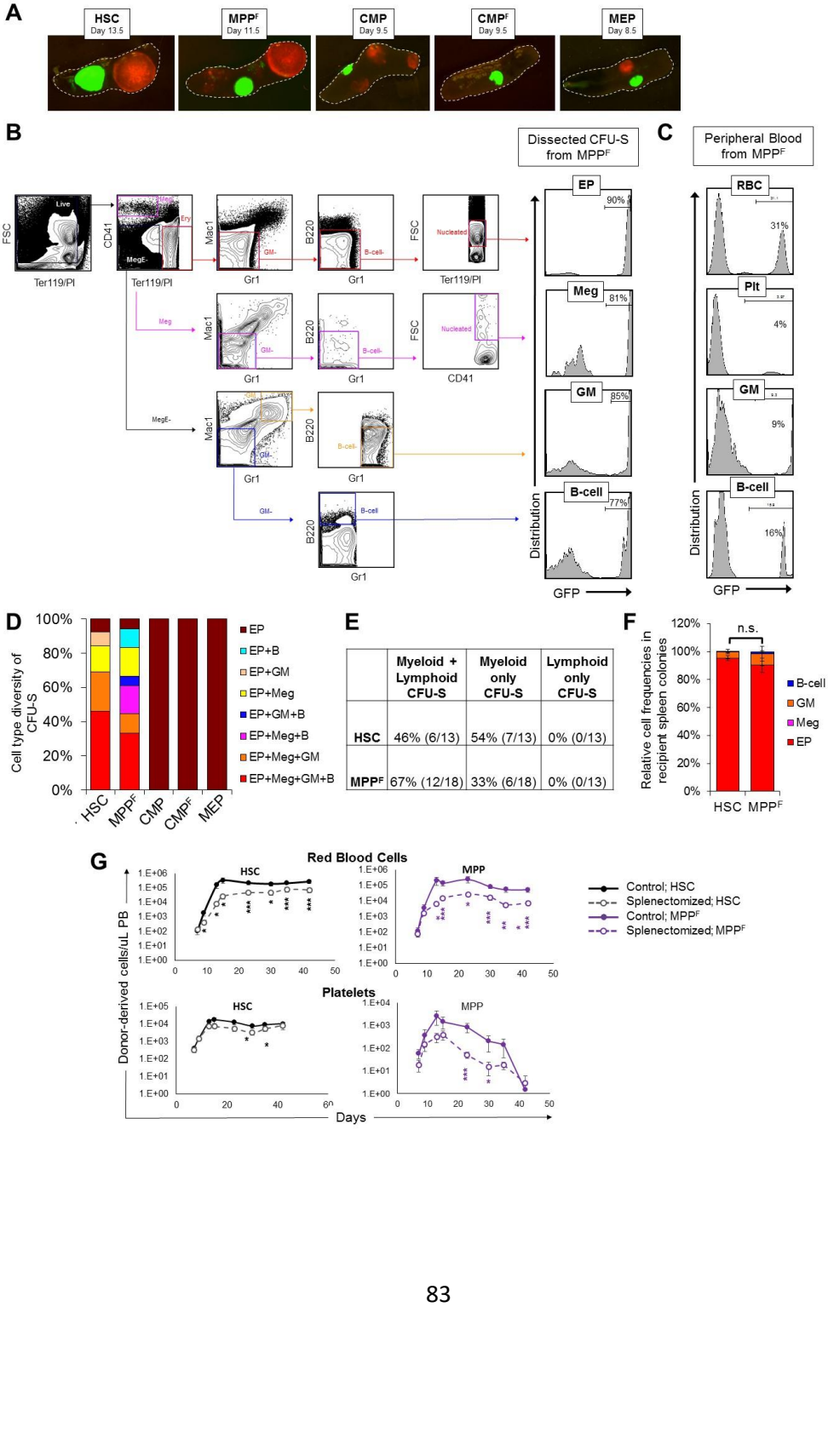
(**B**) Summary of the results from single transplanted HSCs and from single and multiple MPP^F.

(**C-F**) Lineages detected in individual recipients receiving either single HSCs or 1, 5 or 25 MPP^F. (**G-J**) Percent donor chimerism and (**K-N**) distribution of mature cells from single HSCs and 25, 5 or single MPP^F. Representative recipients that displayed multilineage reconstitution are shown.

(**O**) B cell numbers display a rapid and more drastic decline (2,000-fold) after lethal irradiation than other mature cell types (4- to 150-fold for RBCs and T cells, respectively). Data displayed are fold change in mature cell numbers in the PB of lethally irradiated mice over time. n=4.

(**P-T**) The magnitude and timing of host cell depletion and recovery affects the ability to detect reconstitution from transplanted progenitor cells. Numbers from panel (**O**) superimposed with the kinetics of mature cell generation by HSCs and MPP^F for each lineage, displayed relative to the peak reconstitution (set at 1.00) for each population.

Figure 5



	Myeloid + Lymphoid CFU-S	Myeloid only CFU-S	Lymphoid only CFU-S
HSC	46% (6/13)	54% (7/13)	0% (0/13)
MPP^F	67% (12/18)	33% (6/18)	0% (0/13)

Figure 5. Multilineage reconstitution by single HSCs and MPP^F in *in vivo* spleen colony assays.

(**A**) In support of the clonal origin of spleen colonies, CFU-S were comprised of only a single color, either red or green, after transplantation of a mixture of Tomato+ and GFP+ cells of the same cell type. Representative fluorescent microscopy images are shown.

(**B-F**) Spleen colonies derived from single HSCs or MPP^F contain erythroid, megakaryocytic, GM, and B-cell lineages. Donor cells from UBC-GFP mice were transplanted into lethally irradiated mice. CFU-S were dissected and analyzed by flow cytometry at the timepoint post-transplantation of largest colony size; HSCs on day 13.5 (n=13), MPP^F on day 11.5 (n=18), CMPs on day 9.5 (n=16), CMP^F on day 9.5 (n=8), and MEPs on day 8.5 (n=9).

(**B**) Gating strategy and analysis of a representative MPP^F-derived CFU-S. The percentages of donor-derived (GFP+) cells for each cell type within representative colonies are shown in the histograms on the right, with gating strategies shown on the left.

(**C**) Substantially lower donor contribution was observed in the PB (**C**) than within CFU-S (**B**), indicating that the detection of multiple lineages within a colony is not due to contamination of circulating cells. Of note, there were no detectable EPs or Megs in the PB. Representative flow cytometry plots are from a recipient of MPP^F.

(**D**) Proportion of individual CFU-S containing detectable erythroid progenitors (EPs), megakaryocytes (Meg), GM, and/or B cells.

(**E**) Summary of the CFU-S data shown in panel D.

(**F**) The frequency of EPs, Meg, GM and B cells in dissected CFU-S colonies from (D). Percentages are shown as donor-derived (GFP+) cells only.

(**G**) The spleen represents an important, but not essential, site for RBC and Plts production. The number of RBCs and Plts derived from HSCs or MPP^F transplanted into control and splenectomized mice. Recipients received either 200 HSC or 1000 MPP^F after conditioning by

sublethal irradiation. Splenectomy led to decreased production of RBCs and Plts, but not GM, B or T cells (**Figure S6**). Data represent 4 independent experiments with a total of 12-17 mice per group. *p<0.05, **p<0.01, ***p<0.001 by unpaired t-test.

Figure 6

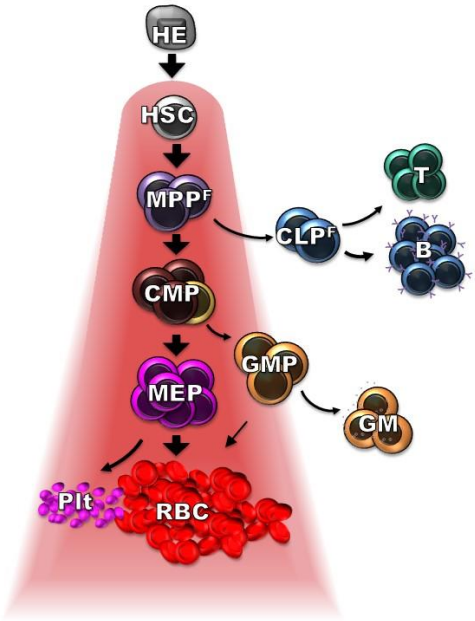
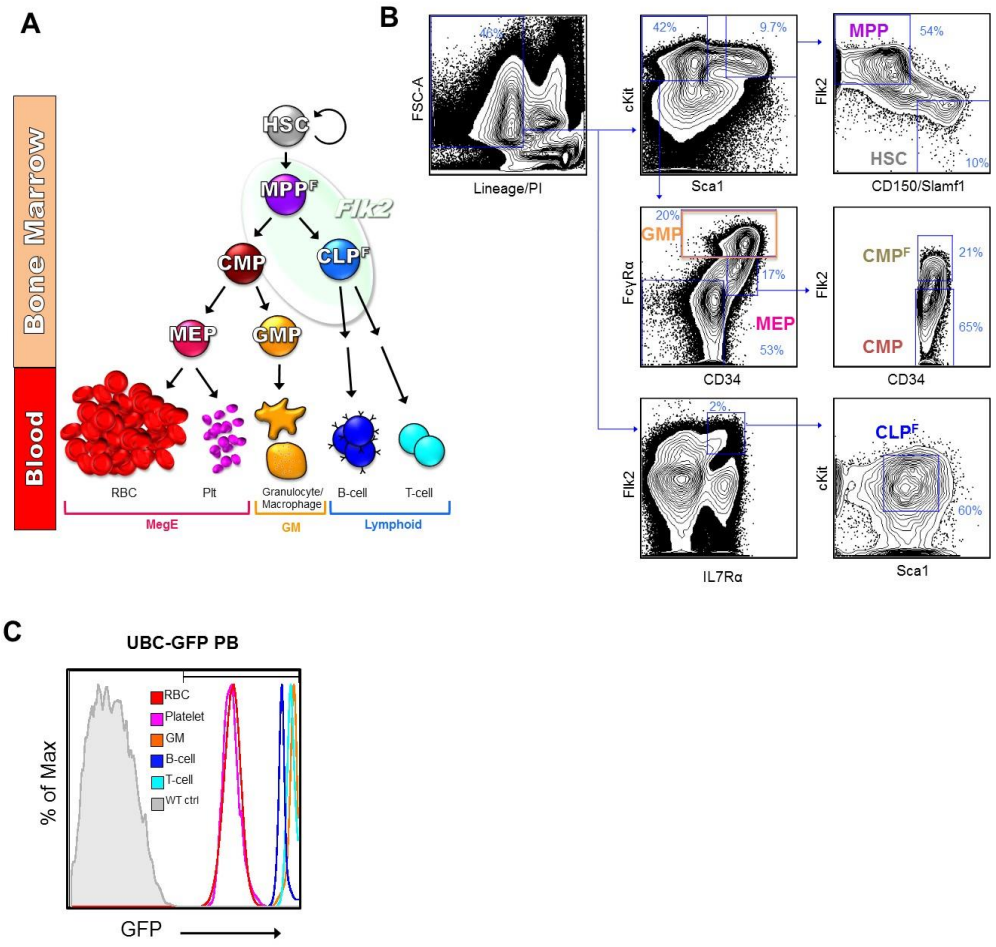


Figure 6. Hematopoietic differentiation model where erythroid production represents the default fate of HSCs. Both functional experiments and gene expression data indicate that the capacity to generate RBCs and Plts is acquired upon specification of HSCs from a hematovascular precursor. Despite downregulation of genes that drive MegE development, RBC and Plt production remain the predominant fates of MPP^F and other non-lymphoid committed progenitors. Relatively rare production of GM, B and T cells occurs through combined downregulation of MegE drivers and a gain of genes promoting the alternative fates. Initiation of non-MegE fates in MPP^F may thus be viewed as gain of GM and lymphoid potential, rather than loss of capacity to generate RBCs and/or Plts.

Supplemental Figure 1



Supplemental Figure 1. Transplantation strategy for evaluating lineage potentials from various HPSCs double-sorted from UBC-GFP mice.

(**A**) Schematic of hematopoietic differentiation to illustrate the terminology used for cell types and lineages in the text.

(**B**) FACS-sorting strategy for isolating hematopoietic subtypes. Cells were pre-gated for singlets only (FSC-Wlow). Bone marrow was cKit-enriched prior to FACS-sorting of transplanted

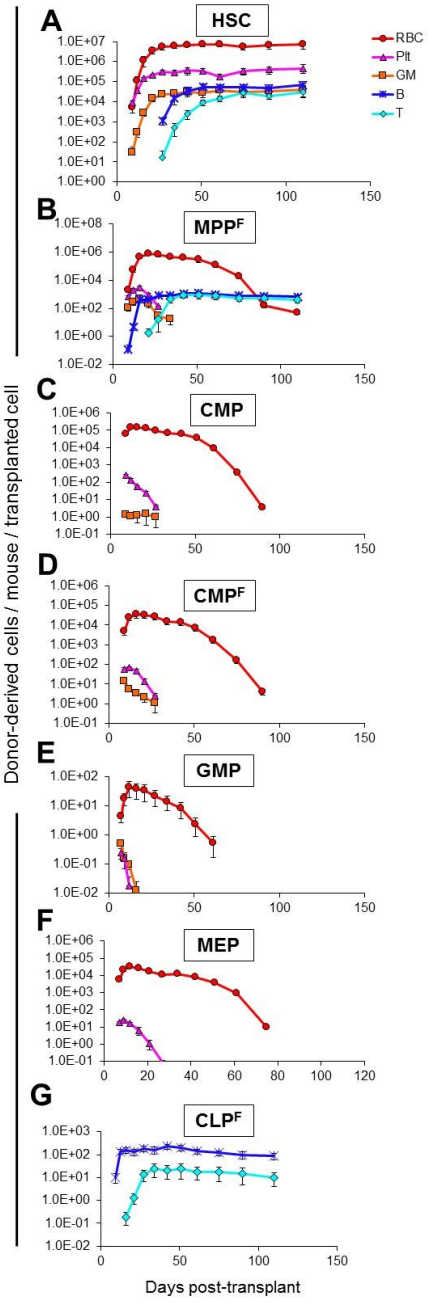
cell types, with the exception of CLPs where lineage-depletion was used instead of cKit enrichment.

(**C**) Flow cytometry analysis of peripheral blood (PB) cells from UBC-GFP mice showing high levels of GFP expression in both nucleated and enucleated hematopoietic cell types. HSC – Hematopoietic Stem Cell; MPPF – Multipotent Progenitor; CMP – Common Myeloid Progenitor;

CMPF – Flk2+ Common Myeloid Progneitor; CLPF – Common Lymphoid Progenitor; GMP – Granulocyte/Myelomonocyte Progenitor; MEP – Megakaryocyte/Erythrocyte Progenitor; RBC -

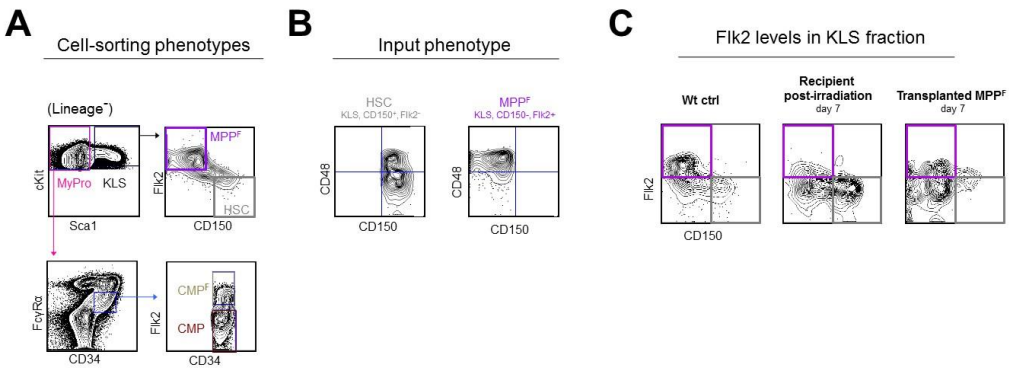
Red Blood Cell; Plt - Platelet; GM - Granulocyte/Myelomonocyte.

Supplemental Figure 2



Supplemental Figure 2. Total numbers of mature cells generated per transplanted cell. (**A-G**) Absolute number of donor-derived mature cells present in a mouse over time post transplantation per transplanted donor cell. Enumeration of donor-derived mature cells in Figure 1M-S was used in combination with total mature cell number in the PB (**Figure 1I**) and mature
cell distribution (**Figure 1J**) to estimate the total number of mature cells generated in the entire
recipient per HSPC after transplantation.

Supplemental Figure 3



Supplemental Figure 3. Reconstitution of mature cells by KLS cells separated by CD150

and CD48 expression.

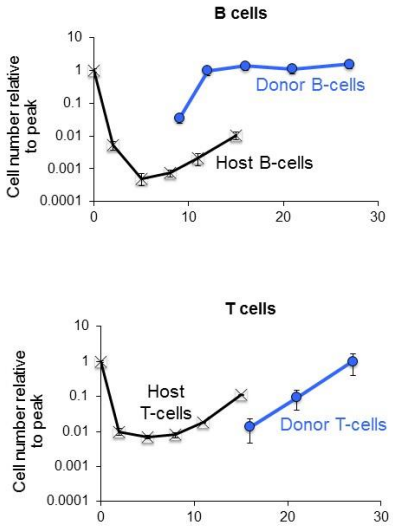
(**A**) Gating strategy used for double-FACS purification of HSCs, MPPF, CMP, and CMPF in Figure 3. Cells were pregated for singlets and Lineage- cells.

(**B**) FACS analysis of CD48 distribution in HSCs and MPPF isolated as indicated in panel (**A**). Note that the majority of MPPF are CD48-positive.

(**C**) FACS analysis revealed that Flk2 cell surface levels decrease relative to unirradiated controls (left) in both host cells (middle) and donor cells (right) 7 days after a sublethal dose of

irradiation (500 rads).

Supplemental Figure 4

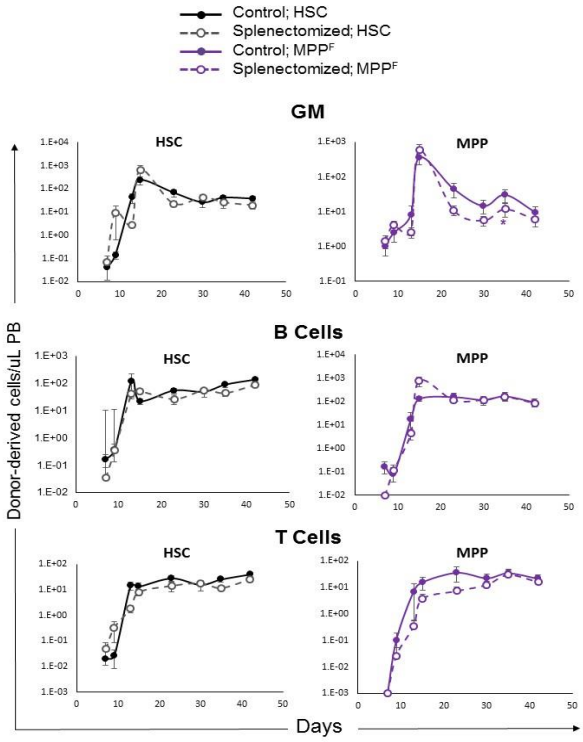


Supplemental Figure 4. CLPF-derived B cells accumulate near the low point of host B cell

decline, whereas host T cells recover prior to CLP-derived T cells are accumulating. Black

lines depict the decline and recovery of host B cells (top) and T cells (bottom) after lethal irradiation. Blue lines indicate donor-derived B cells (top) and T cells (bottom) after transplantation of CLPF.

Supplemental Figure 5



Supplemental Figure 5. Splenectomy of recipient mice does not alter the numbers of GM,

B or T cells generated after transplantations of HSCs or MPPF. The number of GM, B and T

cells derived from HSCs or MPPF transplanted into control and splenectomized mice. Recipients

received either 200 HSC or 1000 MPPF after conditioning by sublethal irradiation. The same recipient cohort displayed significantly fewer donor-derived RBCs and Plts as a result of splenectomy (**Figure 5G**). Data represent 4 independent experiments with a total of 12-17 mice

per group.

REFERENCES

Adolfsson, J., Mansson, R., Buza-Vidas, N., Hultquist, A., Liuba, K., Jensen, C.T., Bryder, D., Yang, L., Borge, O.J., Thoren, L.A., et al. (2005). Identification of Flt3+ lympho-myeloid stem cells lacking erythro-megakaryocytic potential a revised road map for adult blood lineage commitment. Cell 121, 295–306.

Akashi, K., Traver, D., Miyamoto, T., and Weissman, I.L. (2000). A clonogenic common myeloid progenitor that gives rise to all myeloid lineages. Nature 404, 193–197.

Beaudin AE, Boyer SW, Forsberg EC. (2014). Flk2/Flt3 promotes both myeloid and lymphoid development by expanding non-self-renewing multipotent hematopoietic progenitor cells. Exp Hematol. 42(3):218-229.e4. doi: 10.1016/j.exphem.2013.11.013.

Becker AJ, McCulloch EA, Till JE. (1963). Cytological demonstration of the clonal nature of spleen colonies derived from transplanted mouse marrow cells. Nature. 197:452-4.

Benz, C., Copley, M.R., Kent, D.G., Wohrer, S., Cortes, A., Aghaeepour, N., Ma, E., Mader, H., Rowe, K., Day, C., Treloar, D., Brinkman, R.R., Eaves, C.J. (2012). Hematopoietic stem cell subtypes expand differentially during development and display distinct lymphopoietic programs. Cell Stem Cell 10, 273-83.

Boyer SW, Beaudin AE, Forsberg EC. (2012). Mapping differentiation pathways from hematopoietic stem cells using Flk2/Flt3 lineage tracing. Cell Cycle. 11(17):3180-8.

Boyer SW, Schroeder AV, Smith-Berdan S, Forsberg EC. (2011). All hematopoietic cells develop from hematopoietic stem cells through Flk2/Flt3-positive progenitor cells. Cell Stem Cell. 9(1):64-73.

Busch, K., Klapproth, K., Barile, M., Flossdorf, M., Holland-Letz, T., Schlenner, S.M., Reth, M., Höfer, T., Rodewald, H.R. (2015). Fundamental properties of unperturbed haematopoiesis from stem cells in vivo. Nature 518, 542-6.

D'Amico, A., and Wu, L. (2003). The early progenitors of mouse dendritic cells and plasmacytoid predendritic cells are within the bone marrow hemopoietic precursors expressing Flt3. J. Exp. Med. 198, 293–303.

Dykstra B, Kent D, Bowie M, McCaffrey L, Hamilton M, Lyons K, Lee SJ, Brinkman R, Eaves C. (2007). Long-term propagation of distinct hematopoietic differentiation programs in vivo. Cell Stem Cell. 1:218-29.

Ema H, Morita Y, Suda T. (2014). Heterogeneity and hierarchy of hematopoietic stem cells. Exp Hematol. 42, 74-82.

Forsberg, E.C., Serwold, T., Kogan, S., Weissman, I.L., and Passegue´, E. (2006). New evidence supporting megakaryocyte-erythrocyte potential of flk2/flt3+ multipotent hematopoietic progenitors. Cell 126, 415–426.

Gerrits A, Dykstra B, Kalmykowa OJ, Klauke K, Verovskaya E, Broekhuis MJ, de Haan G, Bystrykh LV. (2010). Cellular barcoding tool for clonal analysis in the hematopoietic system. Blood 115(13):2610-8.

Grosselin J, Sii-Felice K, Payen E, Chretien S, Roux DT, Leboulch P. (2013). Arrayed lentiviral barcoding for quantification analysis of hematopoietic dynamics. Stem Cells (10):2162-71.

Hock, H. (2010). Some hematopoietic stem cells are more equal than others. J. Exp. Med. 207, 1127–1130.

Kakumitsu H, Kamezaki K, Shimoda K, Karube K, Haro T, Numata A, Shide K, Matsuda T, Oshima K, Harada M. (2005). Transgenic mice overexpressing murine thrombopoietin develop myelofibrosis and osteosclerosis. Leuk Res (7):761-9. Epub 2005 Mar 2.

Karsunky, H., Inlay, M.A., Serwold, T., Bhattacharya, D., and Weissman, I.L. (2008). Flk2+ common lymphoid progenitors possess equivalent differentiation potential for the B and T lineages. Blood 111, 5562–5570.

Kondo, M., Weissman, I.L., and Akashi, K. (1997). Identification of clonogenic common lymphoid progenitors in mouse bone marrow. Cell 91, 661–672.

Lai, A.Y., and Kondo, M. (2006). Asymmetrical lymphoid and myeloid lineage commitment in multipotent hematopoietic progenitors. J. Exp. Med. 203, 1867–1873.

Lu R, Neff NF, Quake SR, Weissman IL. (2011). Tracking single hematopoietic stem cells in vivo using high-throughput sequencing in conjunction with viral genetic barcoding. Nat Biotechnol (10):928-33.

Miyawaki, K., Arinobu, Y., Iwasaki, H., Kohno, K., Tsuzuki, H., Iino, T., Shima, T., Kikushige, Y., Takenaka, K., Miyamoto, T., Akashi, K. (2015) CD41 marks the initial myelo-erythroid lineage specification in adult mouse hematopoiesis: redefinition of murine common myeloid progenitor. Stem Cells 33, 976-87.

Muzumdar, M.D., Tasic, B., Miyamichi, K., Li, L., and Luo, L. (2007). A global double-fluorescent Cre reporter mouse. Genesis 45, 593–605.

Na Nakorn, T., Traver, D., Weissman, I.L., Akashi, K. (2002). Myeloerythroid-restricted progenitors are sufficient to confer radioprotection and provide the majority of day 8 CFU-S. J Clin Invest 109, 1579-85.

Nemzek JA, Bolgos GL, Williams BA, Remick DG. (2001). Differences in normal values for murine white blood cell counts and other hematological parameters based on sampling site. Inflamm Res 50(10):523-7.

Nutt, S.L., Metcalf, D., D'Amico, A., Polli, M., and Wu, L. (2005). Dynamic regulation of PU.1 expression in multipotent hematopoietic progenitors. J. Exp. Med. 201, 221–231.

Oguro H, Ding L, Morrison SJ. (2013) SLAM family markers resolve functionally distinct subpopulations of hematopoietic stem cells and multipotent progenitors. Cell Stem Cell. 2013 Jul 3;13(1):102-16. doi: 10.1016/j.stem.2013.05.014.

Osawa, M., Hanada, K., Hamada, H., and Nakauchi, H. (1996). Long-term lymphohematopoietic reconstitution by a single CD34-low/negative hematopoietic stem cell. Science 273, 242–245.

Richie Ehrlich LI, Serwold T, Weissman IL. (2011). In vitro assays misrepresent in vivo lineage potentials of murine lymphoid progenitors. Blood 117:2618-24.

Schaefer BC, Schaefer ML, Kappler JW, Marrack P, Kedl RM. (2001). Observation of antigen-dependent CD8+ T-cell/ dendritic cell interactions in vivo. Cell Immunol 214(2):110-22.

Schlenner SM, Madan V, Busch K, Tietz A, Läufle C, Costa C, Blum C, Fehling HJ, Rodewald HR. (2010). Fate mapping reveals separate origins of T cells and myeloid lineages in the thymus. *Immunity*. 32(3):426-436.

Schlenner SM, Rodewald HR. (2010). Early T cell development and the pitfalls of potential. Trends Immunol. 31:303-10.

Schroeder, T. (2010). Hematopoietic stem cell heterogeneity: subtypes, not unpredictable behavior. Cell Stem Cell 6, 203–207.

Seebach J, Bartholdi D, Frei K, Spanaus KS, Ferrero E, Widmer U, Isenmann S, Strieter RM, Schwab M, Pfister H, Fontana A. (1995). Experimental Listeria meningoencephalitis. Macrophage inflammatory protein-1 alpha and -2 are produced intrathecally and mediate chemotactic activity in cerebrospinal fluid of infected mice. J Immunol 155(9):4367-75.

Smith-Berdan, S., Nguyen, A., Hassanein, D., Zimmer, M., Ugarte, F., Ciriza, J., Li, D., Garcı´a-Ojeda, M.E., Hinck, L., and Forsberg, E.C. (2011). Robo4 cooperates with CXCR4 to specify hematopoietic stem cell localization to bone marrow niches. Cell Stem Cell 8, 72–83.

Smith-Berdan, S., Nguyen, A., Hong, M.A., Forsberg, E.C. (2015). ROBO4-mediated vascular integrity regulates the directionality of hematopoietic stem cell trafficking. Stem Cell Reports 4, 255-68.

Sun, J., Ramos, A., Chapman, B., Johnnidis, J.B., Le, L., Ho, Y.J., Klein, A., Hofmann, O., Camargo, F.D. (2014) Clonal dynamics of native haematopoiesis. Nature 514, 322-7.

Ugarte F., Sousae, R., Cinquin, B., Martin, E.W., Krietsch, J., Sanchez, G., Inman, M., Tsang, H., Warr, M., Passegué, E., Larabell, C.A., Forsberg, E.C. (2015) Progressive Chromatin Condensation and H3K9 Methylation Regulate the Differentiation of Embryonic and Hematopoietic Stem Cells. Stem Cell Reports 10,728-40.

Wagers, A.J., Sherwood, R.I., Christensen, J.L., and Weissman, I.L. (2002). Little evidence for developmental plasticity of adult hematopoietic stem cells. Science 297, 2256–2259.

Weber K, Thomaschewski M, Warlich M, Volz T, Cornils K, Niebuhr B, Täger M, Lütgehetmann M, Pollok JM, Stocking C, Dandri M, Benten D, Fehse B. (2011). RGB marking facilitates multicolor clonal cell
tracking. Nat Med 17(4):504-9.

Chapter IV. Developmentally restricted fetal hematopoietic stem cells' gene signature correlates with observed lineage bias and developmental constraint

The data contained within this chapter of the dissertation contains a reprint of a portion of the following work submitted for publication: [Anna E. Beaudin, Scott W. Boyer, **Jessica Perez-Cunningham**, Gloria Hernandez, Eric Aaserude, Chethan Jujuvarapu, E. Camilla Forsberg. Flk2/Flt3 lineage tracing identifies a developmentally-restricted hematopoietic stem cell that gives rise to innate-like lymphocytes. *Cell Stem Cell*. In revision.]

STUDENT'S CONTRIBUTION

SUMMARY

Our lab recently discovered a novel, developmentally restricted HSC population with a lymphoid bias and the ability to robustly generate innate-like lymphocyte subsets. My goal with this work has been to discover a molecular basis for these *in vivo* observations using a variety of molecular techniques including RNA sequencing, qPCR, and cell cycle analysis using flow cytometry. With this work I have shown that our novel HSC subset has a unique gene signature from other fetal HSC and a number of differentially regulated genes which could explain the transplantation data we have generated. Additionally, I have revealed unique differences in cell cycle genes between our fetal HSC subsets and other fetal HSC and validated this with DNA content staining. Together these data have helped us gain new insight into a molecular basis for our observed *in vivo* differences and have provided the lab with a valuable resource as we continue to learn more about this novel HSC population.

INTRODUCTION

Unlike the relatively stable hematopoiesis we see in adult animals, fetal hematopoiesis is believed to occur in waves marked by distinct HSC location and production of different cell types[1]–[3]. During the early stages of development, HSC arise from hemogenic endothelium and can be found in the yolk sac, the aorta-gonad-mesonephros, and eventually seed the fetal liver[4]. After birth these HSC migrate and seed the marrow of the long bones, which is where they will reside throughout the life of the adult animal. As these HSC move to different locations during development, and eventually adulthood, their cell surface phenotype, cell cycle status, and proliferation rates change[5]–[7]. As this process occurs, a marked difference in the subsets of cells produced has been observed. This has led to the “layered immune hypothesis” which states that the immune system is generated in discreet waves of hematopoiesis during development [8]. Thus, the changing phenotype observed between fetal and adult HSC likely correlates with physiological differences in their ability to generate various subsets of the immune system.

The immune system is composed of many cell types, some believed to be of uniquely fetal origin. For example, there is a subset of B cells, called B1 cells, that exhibit innate-like properties even though they appear to be derived via lineages of the adaptive immune system[9]. These B1 cells are thought to be laid down early in development and then constitute a self-replenishing subset of the immune system[10]. This is supported by evidence that adult HSC produce B1 cells very inefficiently compared with fetal HSC[11]. There is also a subset of T cells, termed $\gamma\delta$ T cells, that are believed to be similarly fetally derived[12]. The fact that these subsets appear to be generated during a short window of development and then not produced in adult animals supports the hypothesis of a layered immune system. However, there has been no direct evidence of distinct subsets within the

fetal HSC compartment that could be responsible for laying down these fetally derived lineages.

Using a lineage tracing model discussed in depth in the results section, our lab has identified a developmentally restricted HSC subset that we hypothesize is responsible for generating some waves of the “layered immune system”. If correct, this finding would provide direct evidence of the layered immune hypothesis. Our lab has performed extensive characterization of this novel HSC subset through transplantation assays and found significant differences in cell lineage and reconstitution levels. We hypothesized that these differences in lineage bias would also be evident at the molecular level so I performed RNA sequencing on different fetal liver HSC subsets. My work has shown that different fetal HSC subsets have unique molecular signatures as well as physiologically relevant differences in gene expression. I have validated my sequencing findings with cell cycle analysis and qPCR and have generated a rich data set that can now be mined for a multitude of future projects and continue to inform us on the biology of our novel HSC subset for years to come.

RESULTS

Flk 2 expression marks a subset of hematopoietic stem cells that are developmentally restricted

Our lab has established a dual color lineage tracing mouse line to label different subsets of the blood system. In our model, all cells express the red fluorescent protein Tomato, until Cre recombinase is expressed. Cre expression is driven by the regulatory elements of the Flk2 gene, a receptor tyrosine kinase. When Cre is expressed it results in an irreversible switch of the reporter expression to green fluorescent protein (GFP). This means

that GFP+ cells can never give rise to Tom+ cells as they have lost the Tomato gene (Figure 1A)[13], [14].

This lineage tracing model now allows us to ask questions about the differentiation history of various cell populations. For example, there have been contrary findings regarding the ability of Flk2+ MPP cells to generate platelets which likely has arisen because MPP do not display robust platelet generation *in vitro* but they do in *in vivo* transplantation assays[15], [16]. Our lab hypothesized that all mature cells, including platelets, must pass through a Flk2+ MPP stage during *in situ* hematopoiesis, and our dual color mouse reporter line allowed us to analyze this as it eliminated the need for transplantation assays to label cell subsets. When this model was used to interrogate the hematopoietic system of adult mice it was found that the only cells that expressed Tomato protein within the blood system were hematopoietic stem cells and that all other lineages expressed GFP [13], [14] (Figure 1B). This meant that HSC must express Flk2 as they differentiate, thus causing all downstream cells to express Cre recombinase which causes the switch in reporter expression to GFP. This finding clearly shows that the Flk2+ MPP stage is required for differentiation into all cell lineages.

A second conclusion that can be drawn is that adult HSC must never have expressed Flk2 during their development otherwise they would have previously undergone the color switch to GFP at a fetal or neonate time point. However, since Flk2 is known to be an important player in self-renewal, proliferation, and survival the lab hypothesized that it may play an important role in fetal hematopoiesis as well[15], [17], [18]. To test this, HSC in the fetal compartment were analyzed and distinct subsets were found to express either GFP or Tomato proteins (Figure 1C) (Appendix 1). As previously stated, once a cell has expressed GFP it cannot go back to expressing Tomato, so we can conclude that GFP+ fetal HSC do not contribute to the adult HSC compartment which exclusively express Tomato. GFP+ HSC

are developmentally restricted and have been detected as early as embryonic day 11.5 and as late as postnatal day 14. This means two HSC populations coexist during development: one that is Tom+ and likely gives rise to adult HSCs, and one that is GFP+ and cannot be a precursor of adult HSCs.

To determine the authenticity of these GFP+ fetal HSC our lab performed transplantation assays to test if GFP+ HSC are capable of both self-renewal and multilineage reconstitution, the gold standard tests for determining if a cell is a bona fide HSC. Upon transplantation, GFP+ fetal HSC are capable of long term multilineage reconstitution in primary (Figure 1D) and secondary transplants (Appendix I). These data indicate that GFP+ fetal HSC are bona fide HSC, and are thus the first developmentally restricted HSC population that has been identified.

Gene expression profiles reveal unique signature for developmentally restricted hematopoietic stem cells

From the transplantation data in Figure 1 it is clear that GFP+ HSC are true HSC as they contribute to every lineage of mature cells, but they do not have the same reconstitution pattern as Tom+ HSC. We hypothesized that these functional differences arose from a molecular difference at an RNA level. To investigate the possible molecular differences between GFP+ and Tom+ fetal HSC, I performed RNA sequencing on KLS, Slam+ fractions isolated from e14.5 fetal livers. Fetal cells were isolated using flow cytometry and mRNA libraries were prepared for sequencing. To analyze these datasets we performed a hierarchical cluster analysis on fetal GFP+ HSC, fetal Tom+ HSC, and adult HSC. Both fetal subsets cluster together more closely than either fetal group clusters with the adult HSC

samples (Figure 2 A). This indicates that the Tom+ and GFP+ HSC are more similar to each other than either is to the adult HSC.

We wanted to test the hypothesis that the GFP+ HSC have a unique gene signature so we performed a principal component analysis. This analysis revealed three distinct clusters of samples which are directly related to the cell types that were sequenced (Figure 2 B). We see that there are distinct groups for the fetal GFP+ HSC, fetal Tom+ HSC, and adult HSC indicating that there are enough differentially regulated genes to generate unique gene signatures for each group. We also see relatively tight clustering of replicates, indicating that this signature is consistent across all biological samples. From this we can conclude that the GFP+ fetal HSC are distinct from Tom+ fetal HSC at an RNA level. Next, genes were filtered based on differential expression within the Tom+ and GFP+ fetal subsets and we generated a heat map (Figure 2C). This heat map allows us to visualize the overall gene signature and we can clearly see that the GFP+ HSC samples have a unique expression pattern from Tom+ HSC and adult HSC. This supports the conclusion from the principal component analysis that GFP+ HSC have a unique gene signature from other HSC types.

As we had established that GFP+ fetal HSC have a unique signature, we hypothesized that the genes differentially regulated in this cell type would correlate with functional differences we had previously seen *in vivo*. I performed GO term analysis on genes differentially expressed within the two fetal populations and visualized the results with a Revigo gene ontology treemap (Figure 3A). The size of the box in this diagram indicates how strongly genes were enriched for that go term. The most obvious categories that are enriched are immune system genes, homeostasis and cell proliferation, and developmental genes such as those associated with blood vessel morphogenesis. The developmental genes related to blood vessel morphogenesis seem to clearly fit the developmental origin of fetal HSC. It is thought that early HSC arise from primitive endothelium so it is not surprising to

find common genes expressed in fetal HSC and blood vessel formation[4], [19], [20]. It is not clear however why these genes would be differentially regulated between the two fetal populations. Exploration of these genes may reveal interesting insights into the exact origin of our GFP+ HSC. There are two possible scenarios for their origin: either GFP+ HSC arise from Tom+ HSC, or GFP+ HSC develop in a parallel but independent manner to Tom+ HSC. Future studies to understand why this gene set is differentially expressed may add strength to one of these possibilities.

The other two GO terms that are enriched from our analysis are immune regulation genes and cell cycle genes, both of which correlated with our *in vivo* observations. These results will be discussed in later sections.

From our RNA sequencing data, we are able to conclude that GFP+ fetal HSC have a unique gene signature from other HSC populations. When forced to cluster, we see Tom+ HSC are more similar to their fetal counterparts than to adult HSC, denoting a close relationship between both fetal populations. This led us to investigate what differences we see between the two fetal populations. Together, these data provides an excellent resource to mine for physiologically relevant gene expression differences between our GFP+ fetal HSC and other HSC subsets, some of which we have identified and will be explored in later sections of this chapter.

Proliferation differences in fetal HSC subsets may contribute to their developmental restriction

It is interesting to see cell proliferation and homeostasis genes differentially regulated between Tom+ and GFP+ HSC because both of these cells are derived from fetal livers and in a fetal environment one would expect all cells to be rapidly dividing. However, even with

this highly proliferative environment, we see differences between the GFP+ and Tom+ fetal HSC which led us to hypothesize that GFP+ HSC have different cell cycle kinetics than Tom+ fetal HSC. We performed qPCR on two genes: CCD1, a cyclin, and MLL1, which plays an important role in development and early hematopoiesis (Figure 3C). Both genes are significantly upregulated at the RNA level in the GFP+ HSC compared to the Tom+ indicating that there may be physiological differences in the cell cycle rate of these two cell types.

To test our hypothesis that these cells cycle at different rates *in vivo* I performed assays with EdU and propidium iodide (PI). A pregnant mouse was sacrificed and fetal livers were collected from the embryos. The cells were stained with PI to assess their DNA content and analyzed using flow cytometry. From the PI staining we saw significantly more cells were in the G2/S/M phase of their cell cycle in the GFP+ fraction than the Tom+ fraction of HSC (Figure 4 A,B). This was an intriguing result, but due to limitations of the PI assay, it was difficult to separate the individual phases of the cell cycle. We decided to perform assays with EdU, which is a nucleotide analog that is incorporated into actively dividing cells. EdU staining combined with with staining with Hoechst dye for DNA content allowed for two-dimensional separation of cell cycle progression by flow cytometry. For the EdU experiments, pregnant mice we injected with EdU which was allowed to incorporate into any actively dividing cells for one hour. Then the mice were sacrificed and the HSC from the fetal liver were stained for EdU and Hoechst. Samples were analyzed using flow cytometry. From this data we were able to conclude that within the GFP+ HSC subset there were significantly less cells in the G0/G1 phase and more cells in the S phase (Figure 4 C, D). Together this data indicates that the GFP+ HSC are cycling more rapidly than either their fetal Tom+ counterparts or from adult HSC.

Lymphoid bias is supported by gene expression and transplantation profiles

Upon transplantation, GFP+ HSC make significantly more lymphoid cells; specifically, they are predisposed to make several unique subsets of lymphoid cells such as B1 cells and γδ T cells (Appendix I). These subsets of innate-like lymphoid cells are thought to have a fetal origin which would coincide with the incidence of our GFP+ HSC *in situ*. Given our knowledge of these unique functional differences we hypothesized that genes associated with the lymphoid lineage would be upregulated in the GFP+ HSC. As previously stated there was an enrichment of differentially regulated genes related to immune system development and maintenance. We performed qPCR on sorted Tom and GFP fetal HSC populations to confirm these findings from RNA sequencing data (Figure 3B). There are a number of immune genes significantly upregulated in the GFP+ fetal HSC compartment. Among these are Il7rα, Rag1, and TCR-β; all of which are essential to production of the lymphoid arm of the immune system. This data indicates there may be a unique lineage bias at the RNA level within the GFP+ fetal HSC compartment. This RNA bias along with the functional data from transplantation assays supports the hypothesis that GFP+ HSC are involved in a wave of immune development in which several innate-like lymphoid cells are produced. This constitutes the first direct experimental evidence for the origin of layered immunity.

DISCUSSION

The lab has previously shown the existence of GFP+ fetal HSC and that their unique biology leads to functional differences in their lineage bias and ability to reconstitute hosts after irradiation. The data presented here show that these differences can be traced to an RNA level. First we see that GFP+ HSC have a distinct gene signature from adult and other fetal liver HSC populations. Second, we see that observed differences in cell cycle genes can

be validated with qPCR and functional assays. Finally, GFP+ HSC lymphoid bias is seen not only at a functional level but also at an RNA level. Together, these data supports the hypothesis that fetal GFP+ HSC play a unique and defined role in laying down certain innate-like immune cells during development.

It is not currently well understood why GFP+ fetal HSC do not persist into adulthood. Findings from the RNA sequencing data suggest several possible explanations. First we see that GFP+ HSC cycle more rapidly during development. It may be that this population has a distinct role in laying down the fetal immune system and then becomes exhausted so it does not persist. Perhaps HSC destined to seed adult niches cycle less during development to preserved them from replicative stress. This scenario would define GFP+ HSC as having their unique role during development to buffer those HSC destined for adulthood. An alternative explanation may simply be that GFP+ HSC are incapable of seeding the adult bone marrow. One GO-term category we saw highly differentially regulated was cell adhesion and locomotion genes. It may be that GFP+ HSC simply lack the capacity to find appropriate niches in the adult bone marrow to sustain them in the long term, despite them actually being found in the BM at p14. These two theories are not mutually exclusive. Perhaps the benefit of having a unique fetal HSC population is to protect HSC destined for adult niches and the mechanism for this is through down regulation of adhesion molecules on GFP+ HSC.

It is clear that this RNA sequencing library will provide an excellent data source to be mined to help answer these biological questions. This library was prepared from e14.5 fetal liver cells, a time point when we see a very robust population of GFP HSC that are clearly very active in the fetus. It would be of interest to prepare a second RNA sequencing library from GFP+ and Tom+ HSC from p14 mice. This is the latest time point GFP+ HSC have been reliably identified. I believe the key to understanding why GFP+ HSC do not persist into adulthood can be found from these cells about to be lost. Interesting questions could be

asked by comparing my RNA sequencing library with one prepared from older cells. Are the observed gene differences we see in e14.5 populations recapitulated in older cells? Would we see more genes that negatively regulate the cell cycle actively transcribed in the GFP+ HSC from older mice, forcing them to stop reproducing and die off?

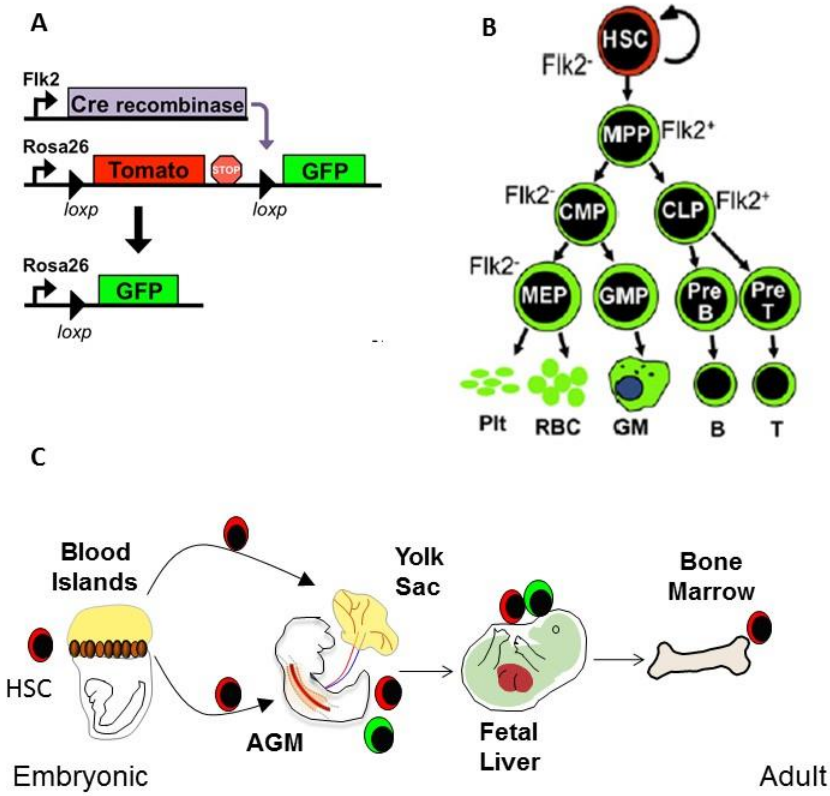
There are still many questions to be answered in this system but I feel confident that the work I have completed allows us to conclude that we have identified a population of HSC that is distinct from all other known populations not only in its functional potential but, as my findings here shows also its gene expression signature.

METHODS

RNA Sequencing: RNA-seq libraries from in vitro cultured cells were prepared using NEBNext® Ultra kit (New England Biolabs) following manufacturer instructions. Libraries were sequenced using HiSeq2000 platform (Illumina) at the genomic sequencing laboratory, UC Berkeley, and analyzed with DESEQ software (Bioconductor.org). GO term analysis was performed with Revgo (http://revigo.irb.hr/).

Further detailed methods are provided in Appendix I for these studies.

FIGURES


A
Flk2
Cre recombinase
Rosa26
Tomato
STOP
GFP
loxp
loxp
Rosa26
GFP
loxp
B
HSC
Flk2-
MPP
Flk2+
Flk2-
CMP
CLP
Flk2+
Flk2-
MEP
GMP
Pre B
Pre T
Plt
RBC
GM
B
T
C
Blood Islands
Yolk Sac
Bone Marrow
HSC
AGM
Fetal Liver
Embryonic
Adult

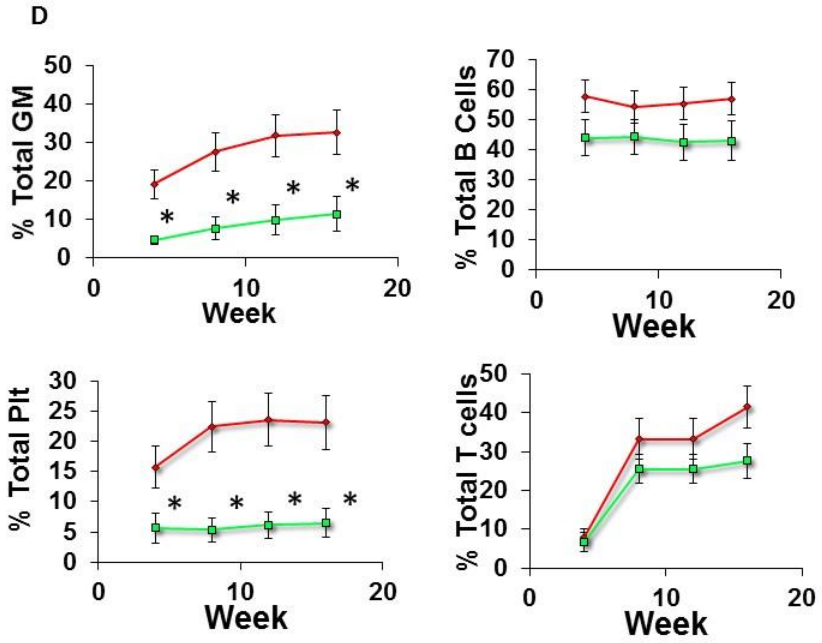
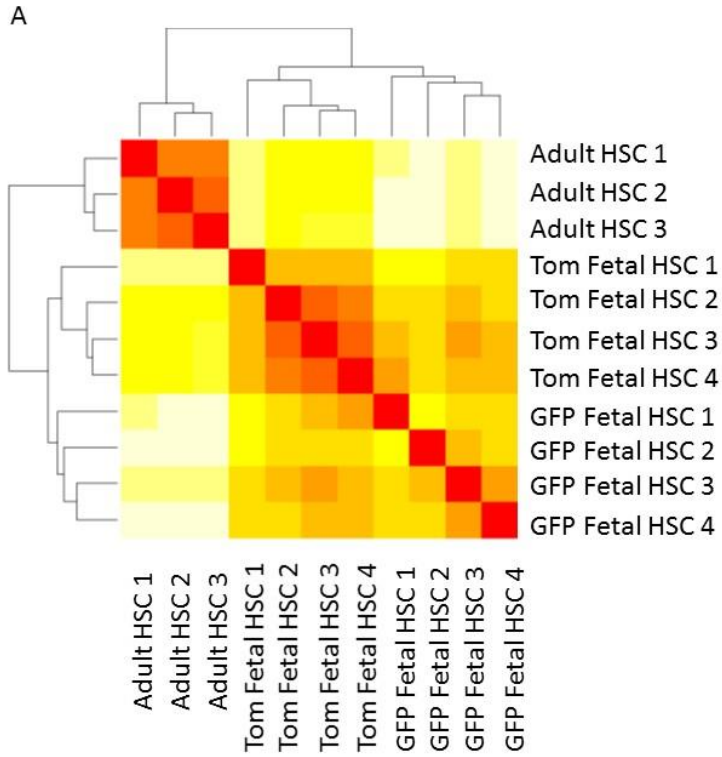
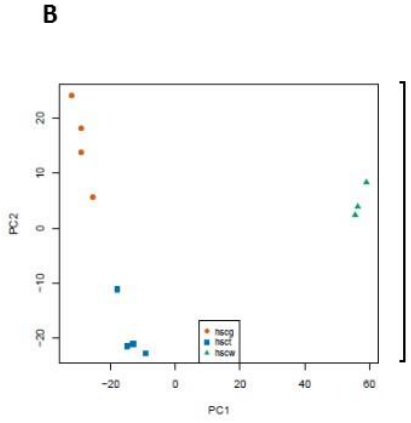

D
% Total GM
Week
% Total B Cells
Week
% Total Plt
Week
% Total T cells
Week

Figure 1. Lineage tracing reveals a developmentally restricted fetal HSC population. A. Schematic of lineage tracing model before and after Cre recombinase expression. B. Lineage tree showing Flk2 expression pattern and reporter expression in adult hematopoietic system [13]. Only adult HSC express Tomato protein while all other lineages express GFP. C. Schematic showing reporter expression during development in lineage tracing mouse line with a transient population of HSC labeled with GFP. D. Primary transplantation data from sorted GFP+ fetal HSC showing percent chimerism over weeks from transplantation indicating long term multi-lineage potential in this population.(Appendix I). HSC: hematopoietic stem cell, MPP: multipotent progenitor, CMP: common myeloid progenitor, CLP: common lymphoid progenitor, MEP: megakaryocyte/erythroid progenitor, GMP: granulocyte-macrophage progenitor, RBC: red blood cell, GM: granulocyte/macrophage, AGM: aorta-gonad-mesonephros.


A
Adult HSC 1
Adult HSC 2
Adult HSC 3
Tom Fetal HSC 1
Tom Fetal HSC 2
Tom Fetal HSC 3
Tom Fetal HSC 4
GFP Fetal HSC 1
GFP Fetal HSC 2
GFP Fetal HSC 3
GFP Fetal HSC 4


B
PC1
PC2
hscg
hsct
hscw

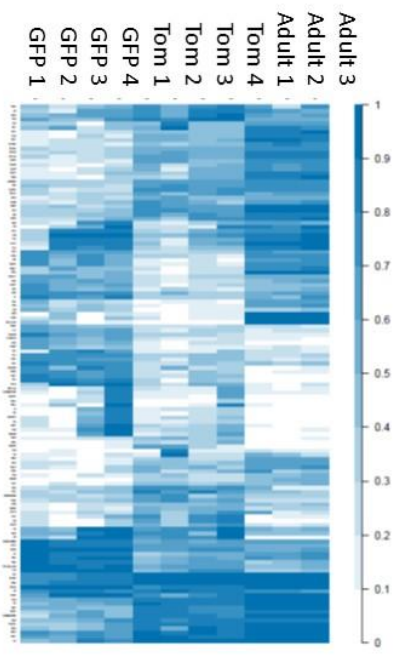

C
GFP 1
GFP 2
GFP 3
GFP 4
Tom 1
Tom 2
Tom 3
Tom 4
Adult 1
Adult 2
Adult 3

Figure 2. RNA Sequencing reveals a unique genetic signature of GFP+ HSC. A. Clustering analysis of Tomato and GFP+ fetal HSC with adult HSC populations shows all populations cluster separately and that the fetal populations cluster more tightly with each other than with the adult HSC . B. Principal component analysis of all fetal and adult HSC populations shows discreet clustering of each population. C. Heat map analysis of all fetal and adult HSC populations and shows differential gene expression in each population. Four independent libraries were prepared from pooled litter-mate fetal livers and three biological replicates were used for the adult libraries.

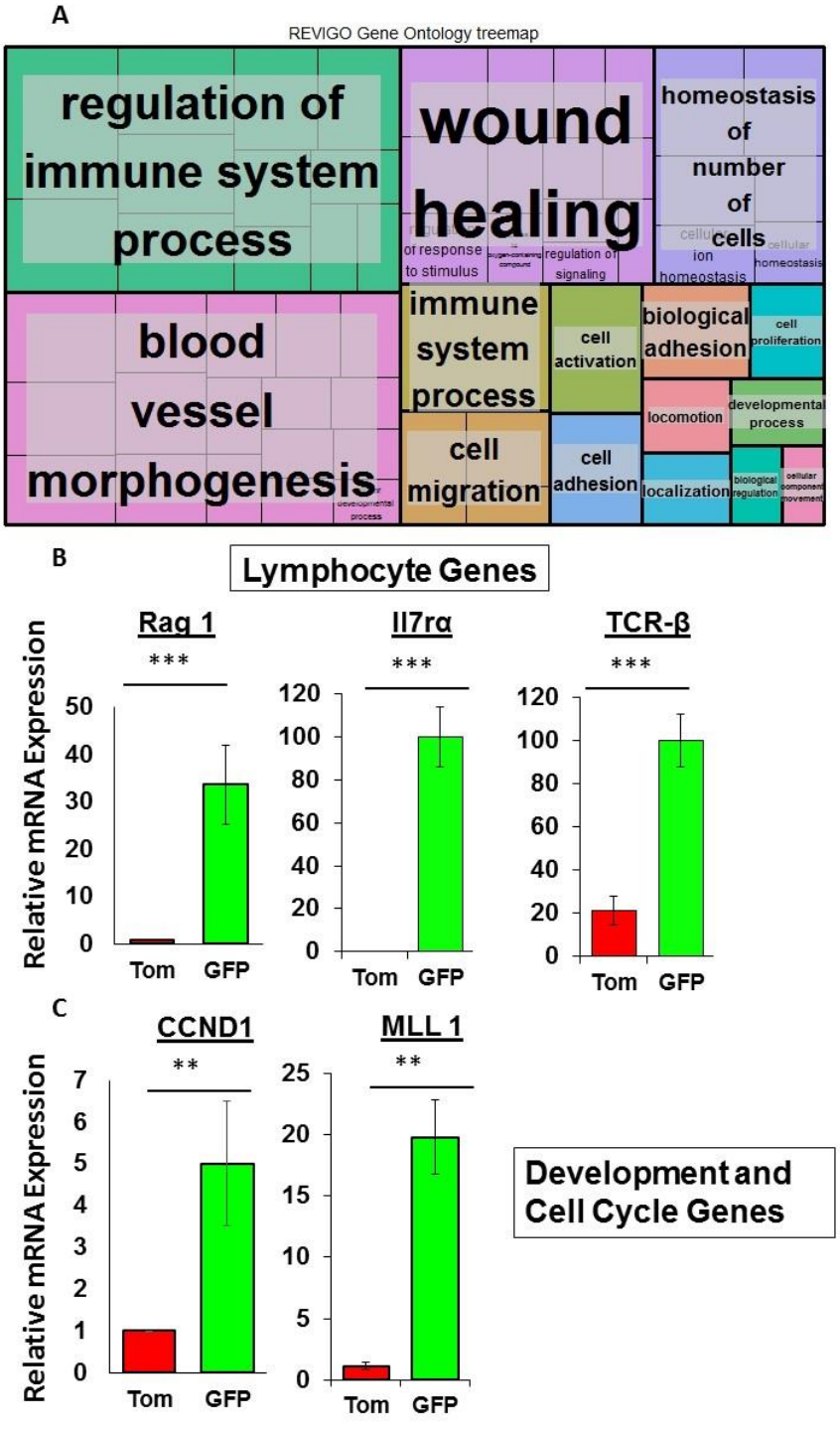

A
REVIGO Gene Ontology treemap
regulation of immune system process
wound healing
homeostasis of number of cells
blood vessel morphogenesis
immune system process
cell activation
biological adhesion
cell proliferation
locomotion
developmental process
cell migration
cell adhesion
localization
B
Lymphocyte Genes
Rag 1
Il7rα
TCR-β
Relative mRNA Expression
***
Tom
GFP
C
CCND1
MLL 1
**
Development and Cell Cycle Genes

Figure 3. Gene ontology analysis of GFP+ HSC gene signature reveals lymphoid bias and cell cycle differences. A. Ontology map of go term analysis generated by Revigo of genes differentially regulated between the GFP+ and Tomato fetal HSC subsets. B and C. qPCR analysis of selected genes from RNA sequencing results shows significant differences in relative gene expression. At least 3 biological replicates were used for the qPCR analysis of each population.

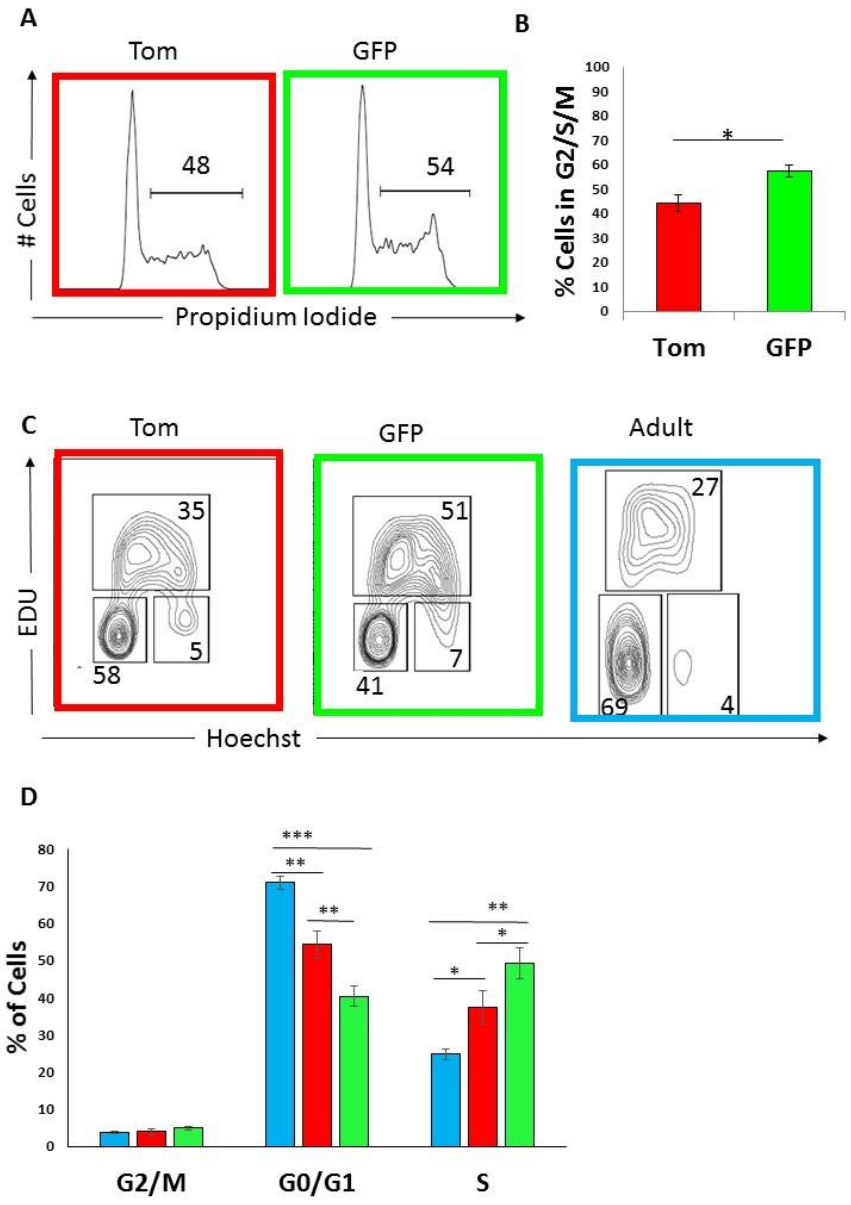

A
Tom
GFP
48
54
Cells
Propidium Iodide
B
% Cells in G2/S/M
100
90
80
70
60
50
40
30
20
10
0
*
Tom
GFP
C
Tom
GFP
Adult
35
5
58
51
7
41
27
69
4
EDU
Hoechst
D
% of Cells
80
70
60
50
40
30
20
10
0
***
**
**
**
*
*
G2/M
G0/G1
S

Figure 4. GFP+ HSC are more actively cycling than other fetal or adult HSC subsets. A. Representative flow cytometry plots of propidium iodide staining. Designated gate shows S/G2/M population. B. Quantification of the propidium iodide shown in A. C. Representation flow cytometry plots of the EdU and Hoechst staining. The lower left hand gate shows the G0/G1 gate, the upper gate is the S gate, and the lower right gate is the G2/M gate. D. Quantification of EdU and Hoechst staining from C. Red represents Tomato+ fetal liver HSC; green represents GFP+ fetal liver HSC; blue represents adult HSC. Data represents 8 biological replicates over three independent experiments. Error bars represent SE. * indicates $p<0.05$.

REFERENCES

[1] E. Dzierzak and N. A. Speck, “Of lineage and legacy: the development of mammalian hematopoietic stem cells.,” *Nat. Immunol.*, vol. 9, no. 2, pp. 129–36, Feb. 2008.

[2] A. Medvinsky, S. Rybtsov, and S. Taoudi, “Embryonic origin of the adult hematopoietic system: advances and questions.,” *Development*, vol. 138, no. 6, pp. 1017–31, 2011.

[3] K. E. McGrath and J. Palis, “Hematopoiesis in the yolk sac: More than meets the eye,” *Exp. Hematol.*, vol. 33, no. 9 SPEC. ISS., pp. 1021–1028, 2005.

[4] A. Ciau-Uitz, O, R. Monteiro, A. Kirmizitas, and R. Patient, “Developmental hematopoiesis: Ontogeny, genetic programming and conservation,” *Exp. Hematol.*, vol. 42, no. 8, pp. 669–683, 2014.

[5] S. Matsuoka, Y. Ebihara, M. J. Xu, T. Ishii, D. Sugiyama, H. Yoshino, T. Ueda, A. Manabe, R. Tanaka, Y. Ikeda, T. Nakahata, and K. Tsuji, “CD34 expression on long-term repopulating hematopoietic stem cells changes during developmental stages,” *Blood*, vol. 97, no. 2, pp. 419–425, 2001.

[6] M. B. Bowie, K. D. Mcknight, D. G. Kent, L. Mccaffrey, P. A. Hoodless, and C. J. Eaves, “Hematopoietic stem cells proliferate until after birth and show a reversible phase-specific engraftment defect,” vol. 116, no. 10, 2006.

[7] V. I. Rebel, C. L. Miller, C. J. Eaves, and P. M. Lansdorp, “The repopulation of fetal liver hematopoietic stem cells in mice exceeds that of their adult bone marrow counterparts,” *Blood*, vol. 87, no. 8, pp. 3500–3507, 1996.

[8] L. A. Herzenberg and L. A. Herzenberg, “Toward a layered immune system,” *Cell*, vol. 59, no. 6, pp. 953–954, 1989.

[9] R. R. Hardy and K. Hayakawa, “A developmental switch in B lymphopoiesis.,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 88, no. December, pp. 11550–11554, 1991.

[10] A. B. Kantor and L. A. Herzenberg, “LINEAGES,” 1993.

[11] K. Hayakawa, R. R. Hardy, L. A. Herzenberg, and L. A. Herzenberg, “Progenitors for Ly-1 B cells are distinct from progenitors for other B cells,” *J Exp Med*, vol. 161, no. 6, pp. 1554–1568, 1985.

[12] W. L. Havran and J. P. Allison, “Developmentally ordered appearance of thymocytes expressing different T-cell antigen receptors.,” *Nature*, vol. 335. pp. 443–445, 1988.

[13] S. W. Boyer, A. V. Schroeder, S. Smith-Berdan, and E. C. Forsberg, “All Hematopoietic Cells Develop from Hematopoietic Stem Cells through Flk2/Flt3-Positive Progenitor Cells,” *Cell Stem Cell*, vol. 9, no. 1, pp. 64–73, 2011.

[14] S. W. Boyer, A. E. Beaudin, and E. C. Forsberg, “Mapping differentiation pathways from hematopoietic stem cells using Flk2/Flt3 lineage tracing,” *Cell Cycle*, vol. 11, no. 17, pp. 3180–3188, 2012.

[15] J. Adolfsson, R. Månsson, N. Buza-Vidas, A. Hultquist, K. Liuba, C. T. Jensen, D. Bryder, L. Yang, O. J. Borge, L. A. M. Thoren, K. Anderson, E. Sitnicka, Y. Sasaki, M. Sigvardsson, and S. E. W. Jacobsen, “Identification of Flt3+ lympho-myeloid stem cells lacking erythro-megakaryocytic potential: A revised road map for adult blood lineage commitment,” *Cell*, vol. 121, no. 2, pp. 295–306, 2005.

[16] R. Yamamoto, Y. Morita, J. Ooehara, S. Hamanaka, M. Onodera, K. L. Rudolph, H. Ema, and H. Nakauchi, “Clonal analysis unveils self-renewing lineage-restricted progenitors generated directly from hematopoietic stem cells.,” *Cell*, vol. 154, no. 5, pp. 1112–26, Aug. 2013.

[17] J. Adolfsson, O. J. Borge, D. Bryder, K. Theilgaard-Mönch, I. Åstrand-Grundström, E. Sitnicka, Y. Sasaki, and S. E. W. Jacobsen, “Upregulation of Flt3 expression within the bone marrow Lin-Sca1+c-kit+ stem cell compartment is accompanied by loss of self-renewal capacity,” *Immunity*, vol. 15, no. 4, pp. 659–669, 2001.

[18] B. Hunte, S. Hudak, and D. Campbell, “flk2/flt3 ligand is a potent cofactor for the growth of primitive B cell progenitors,” *J. Immunol.*, vol. 156, pp. 489–96, 1995.

[19] D. Kanz, M. Konantz, E. Alghisi, T. E. North, and C. Lengerke, “Endothelial-to-hematopoietic transition: Notch-ing vessels into blood,” *Ann. N. Y. Acad. Sci.*, p. n/a–n/a, 2016.

[20] E. Gritz and K. K. Hirschi, “Specification and function of hemogenic endothelium during embryogenesis,” *Cell. Mol. Life Sci.*, vol. 73, no. 8, pp. 1547–1567, 2016.

CHAPTER V: Conclusion

In this dissertation I have explored the role of cell intrinsic and extrinsic signals on hematopoietic stem cells (HSC) fate decisions. Additionally, a literature review has been provided to summarize the generation and use of transgenic mouse lines for lineage tracing. Finally, our lab has generated and characterized a new mouse line that has particular utility when studying HSC fate decisions. Here I will discuss the major conclusions that can be drawn from this work as well as the impact on moving the field forward with understanding HSC fate decisions.

Lineage tracing using transgenic mouse lines

Lineage tracing has proven to be a valuable tool to the hematopoietic community and has been reviewed in the first chapter of this thesis. As more mouse lines are constantly being generated this review serves as a starting place for new researchers to explore the basic decisions involved in transgenic mouse design.

Additionally, in the second chapter, our generation and characterization of a new transgenic mouse line that allows the specific and controlled labeling of different stem and progenitor subsets within the blood system is described. All hematopoietic stem and progenitor cell populations show robust labeling by flow cytometry with less robust, but detectable, labeling seen in nearly all mature populations. The new line also shows no off target tissue labeling in adult mice and a transient level of endothelial cell labeling during fetal development. The reporter in this mouse line takes advantage of the Cre/LoxP system and allows for label expression to be controlled by breeding to a variety of Cre lines. This was

validated here by use of Flk2-Cre to show specific labeling of only HSC. Ultimately, this level of specific and controlled labeling has not been shown before within this stem cell compartment and will serve the field well in a variety of applications. This tool allows for single color labeling of complex stem and progenitor subsets and can now be used to gain insights into HSC-specific questions such as niche microenvironment location and signaling.

Cell extrinsic regulation

To explore the role of cell extrinsic signals on differentiation, HSC and MPP were transplanted into mice whose spleens had been removed surgically. This served to explore the role of the spleen as a microenvironment for differentiation. When MPP were transplanted into mice lacking a spleen they showed a significant decrease in their ability to produce RBC. HSC did not show a similar decrease. This indicates that there are distinct signals within the spleen microenvironment that more greatly effect MPP than HSC differentiation, particularly in their ability to generate RBC. Future studies are necessary to ascertain which signaling pathways are being activated within this microenvironment, and why different stem and progenitor cell types respond to them differently. However, I have shown for the first time that the effects of the spleen microenvironment vary with the input cells.

Cell intrinsic regulation

To explore cell intrinsic signals that affect HSC fate, I performed mRNA sequencing on a developmentally restricted HSC subset that the lab previously identified with a novel lineage tracing system. This led to a number of interesting findings, some which validated

our previous *in vivo* work and some that lead to novel conclusions about the biology of this novel HSC subset. First, I have shown that this subset of HSC is biased toward lymphoid fate decisions at a transcriptional level which is consistent with our *in vivo* findings. Additionally, a novel finding from this gene expression data was that developmentally restricted HSC upregulate a number of cell cycle genes compared to conventional HSC subsets. I validated this finding with cell cycle analysis of freshly isolated fetal liver cells and found that there are significantly more cells actively cycling in our developmentally restricted subset compared with control fetal HSC. Ultimately, this project has identified a novel HSC subset using lineage tracing and my work has provided a rich resource that can be mined for years to come as we continue to explore the role of this new subset in blood system development

In summary, the work presented here has demonstrated the utility of transgenic lineage tracing models in studying HSC fate decisions and the need for continued study of these systems to understand hematopoiesis and further improve human health.

Appendix I: Discovery of a developmentally restricted hematopoietic stem cell that gives rise to innate-like B and T cells

The text contained within this chapter of the dissertation contains a reprint of the following work submitted for publication: [Anna E. Beaudin, Scott W. Boyer, **Jessica Perez-Cunningham**, Gloria Hernandez, Eric Aaserude, Chethan Jujuvarapu, E. Camilla Forsberg. Flk2/Flt3 lineage tracing identifies a developmentally-restricted hematopoietic stem cell that gives rise to innate-like lymphocytes. *Cell Stem Cell*. In revision.].

SUMMARY

The generation of distinct hematopoietic cell types distinguishes fetal from adult hematopoiesis, but the mechanisms responsible for differential cell production during prenatal hematopoietic development remain to be established. Using an irreversible lineage tracing model, we have identified a novel, developmentally restricted hematopoietic stem cell (HSC) that supports long-term multilineage reconstitution upon transplantation into adult recipients, but does not normally persist into adulthood. Despite its multilineage potential, this novel hematopoietic progenitor displays both higher lymphoid cell production and greater capacity to generate innate-like B and T lymphocytes as compared to coexisting fetal HSCs and adult HSCs. Our lineage tracing identifies a developmentally restricted HSC that contributes to the formation of a layered immune system and reveals the mechanism underlying developmentally regulated hematopoietic waves. As early lymphoid cells play essential roles in establishing self-recognition and tolerance, defining their origin and generation has critical implications for understanding the development of autoimmune disease, allergy, and tolerance induction upon organ transplantation.

Main text:

Developmental hematopoiesis occurs in waves, characterized by the generation of distinct cell types [1-3]. The developmental origin and hierarchical relationship between the progenitors responsible for the temporal production of specialized hematopoietic cell types is the subject of intense investigation [4,5]. Elegant fate-mapping and adoptive transfer experiments have confirmed that at least some fetal definitive HSCs, defined by their capacity for long-term multilineage reconstitution of adult recipients, persist into adulthood [6-11]. However, fetal and early neonatal definitive HSCs exhibit markedly different properties as compared to adult HSCs, including differences in surface marker expression [12,13], proliferative state [14], and repopulation capability [15,16]. Additionally, fetal HSCs possess distinct potential for innate-like lymphoid subsets as compared to adult HSCs [17-21]. It remains unclear to what extent these differences reflect disparities in the maturation state of HSCs in a single lineage, the effects of differential niche environments, or the existence of distinct, parallel lineages of HSCs across ontogeny. Recent investigation of human immune development supports the latter by positing the existence of HSCs that gives rise to specialized immune cells promoting tolerance in early life [22,23], but the identification of a developmentally restricted HSC that specifies innate-like lymphoid subsets has remained elusive.

We recently established a lineage tracing model that tracks cells with a history of expression of the tyrosine kinase receptor Flk2 (Flt3) by crossing mice expressing Cre recombinase under control of Flk2 regulatory elements to mice expressing a dual color reporter from the Rosa26 locus (Fig. 1a). Importantly, we have shown that all adult HSCs in this “FlkSwitch” mouse model express Tomato (Tom) and that only Tom+ cells within the adult bone marrow (BM) possess long-term reconstituting potential [24-26] (Extended data Fig.

1a and b). Because the switch from Tom to GFP expression is due to the irreversible excision of the Tom gene, GFP+ cells cannot give rise to Tom+ cells (Extended data Fig. 1c). Thus, adult HSCs cannot be derived from cells with a history of Flk2 expression. The FlkSwitch model thereby provides a unique tool with which to interrogate the developmental relationships between fetal and adult HSCs. Here, using the FlkSwitch model, we identify a novel, definitive HSC that is restricted to early life and efficiently generates innate-like B and T cells.

GFP+ HSCs exist during fetal development

As we have demonstrated previously that all adult HSCs in the FlkSwitch lineage tracing model are Tom+ (Fig. 1j, Extended data Fig. 1b) [24,25], all developmental precursors of adult HSCs must be derived via a Flk2-negative lineage. To gain further insight into HSC development, reporter expression was investigated in phenotypic stem and progenitor cell compartments of FlkSwitch embryos. Neither Flk2 surface protein nor GFP expression were observed prior to embryonic day (E)10 (Fig. 1c; Extended data Fig. 2a; Extended data Fig. 2f). By E10.5, coincident with emergence of definitive hematopoiesis and Flk2 expression, GFP+ cells were also found. Unexpectedly, both Tom+ and GFP+ cells were consistently observed within the phenotypic compartments enriched for the most primitive stem/progenitor cells in the yolk sac (YS), aorta-gonad-mesonephros region (AGM), placenta, and fetal liver (FL) between E10.5 and E14.5 during fetal development (Fig. 1b,d-h), as well as in the neonate BM at postnatal day 14 (P14) (Fig. 1i), regardless of phenotypic markers used (Extended data Fig. 2). Cells appearing both Tom+ and GFP+ had undergone excision of the Tom gene and had thus recently switched from Tom to GFP expression (Extended data Fig. 3a). As shown previously [24,25], GFP+ cells were not present in the phenotypic (Fig. 1j) or functional (Extended data Fig. 1b) adult HSC compartment, whereas the majority of adult Flk2+ multipotent progenitors (MPPs) expressed GFP (Fig. 1j). The presence of both Tom+

and GFP+ cells within the phenotypic stem cell compartments across fetal development opened the possibility of coexistence of two HSC populations during fetal and neonatal hematopoiesis.

GFP+ FL cells are capable of serial, multilineage reconstitution

To determine which population contained functional HSCs, we transplanted Tom+ or GFP+ cells sorted from the cKit+ Lin-Sca1+ (KLS) compartment of E14.5 FL of FlkSwitch embryos into sublethally (750 rad) irradiated adult recipients (Figure 2a). We focused on the E14.5 FL stage, as developmental HSCs are relatively abundant and well characterized at this time point. We transplanted more GFP+ cells (1000) as compared to Tom+ cells (500) because we reasoned that the Tom+ KLS fraction would include a greater proportion of HSCs with long-term multilineage reconstitution (LTMR) capability, based on a greater proportion of CD150+ cells in the Tom+ KLS fractions (Extended data Fig. 3b) [27,28]. Surprisingly, and in direct contrast to transplantation of the equivalent populations from the BM of adult FlkSwitch mice (Extended data Fig. 1b and references [24,25,29]), both Tom+ and GFP+ FL populations were capable of long-term multilineage reconstitution (LTMR) (Fig. 2b-d,Extended data Table 1). Although LTMR for individual mice was defined as reconstitution of all mature lineages in the peripheral blood (PB) at levels > 0.1% for a minimum of 16 weeks, the average reconstitution from both Tom+ and GFP+ cells was remarkably robust (GM, 7%; Plt, 6%). Using the same LTMR criteria as for sublethal recipients, GFP+ FL KLS cells similarly reconstituted lethally irradiated adult recipients in competition with cotransplanted adult cells (Extended data Fig. 4a,b). Although Tom+ FL KLS cells gave rise to both Tom+ and GFP+ cells upon transplantation, as previously described [25], GFP+ FL KLS cells never gave rise to Tom+ cells (Extended data Fig. 3c). The absence of Tom+ cells in mice transplanted with GFP+ cells confirmed that the switch to GFP expression within the HSC compartment is irreversible, and that Tom+ cells did not contaminate transplanted GFP+ cells. We confirmed

a previous report of heterogeneity within the fetal HSC compartment [13] by demonstrating LTMR capability of both Flk2+ and Flk2- fetal KLS cells (Extended data Figure 5). These data corroborate our model, as cell surface expression of Flk2 alone is capable of distinguishing a phenotypically distinct fetal HSC. Thus, both Tom+ and GFP+ KLS cells isolated from the FL were capable of LTMR upon transplantation. As all adult HSCs in the FlkSwitch model are Tom+, these data **suggested** that GFP+ HSCs exist during a limited developmental window.

Examination of BM chimerism in recipients of Tom+ or GFP+ E14.5 FL KLS cells 18 weeks post-transplantation revealed donor-derived cells of all phenotypes examined, including HSCs (Fig. 2e). Limited dilution analysis further confirmed the presence of HSCs within both the Tom+ and GFP+ FL KLS populations, with GFP+ HSCs present at lower frequency than Tom+ HSCs (Fig. 2i). However, as the 'gold standard' for defining functional HSCs is their ability to support LTMR upon serial transplantation, we determined whether both populations of fetal HSCs displayed the self-renewal capability necessary to support serial reconstitution (Fig. 2a). WBM cells from individual recipients demonstrating LTMR after primary transplantation (Figure 2a) were transplanted into one or more secondary recipients and donor-derived contribution was monitored over an additional 16 weeks. Remarkably, the majority of secondary recipients of either Tom+ or GFP+ FL KLS cells displayed LTMR (Fig. 2f-h, Extended data Table 2). As in primary recipients, reconstitution was robust for all mature lineages (averaging 8% and above) and sustained over time, with remarkably similar reconstitution profiles from Tom+ and GFP+ cells. Serial reconstitution was also observed when fewer cells were transferred from primary to secondary recipients, and persisted for longer than 6 months (Extended Data Fig.4c-d). As efficiency of Flk2-cre mediated recombination in the FlkSwitch model is lower during fetal development (Extended data Fig. 6a), the frequency of the GFP+ HSC in our model is likely underestimated (Extended data

Fig. 6b-d). These data are consistent with self-renewal capability of both Tom+ and GFP+ FL KLS cells and thus identify two coexisting populations of fetal HSCs: a Tom+ HSC that likely contributes to the adult HSC compartment, and a developmentally limited GFP+ HSC that can be distinguished from co-existing Tom+ HSCs and from adult HSCs by development via a Flk2-expressing pathway. GFP+ HSCs therefore represent a distinct HSC lineage that exists during a limited developmental window, yet has sustained self-renewal and multilineage engraftment potential upon transplantation.

Cell-extrinsic and cell-intrinsic mechanisms regulate the lifespan of the GFP+ HSC

A fetal HSC capable of LTMR upon transplantation that does not persist into adulthood in situ has never before been identified. The capability of GFP+ HSC to engraft in the adult BM upon transplantation despite their apparent inability to persist into adulthood in situ is reminiscint of the phenotype of CXCR4-/- HSCs [30-33]. To understand the mechanisms limiting the persistence of the GFP+ HSC, we first determined whether GFP+ HSC were capable of seeding the BM. GFP+ and Tom+ FL HSC expressed similar levels of CXCR4 (data not shown) and showed equivalent capacity to migrate towards an SDF1 gradient in vitro (Fig. 3a). Consistent with this ability, GFP+ FL HSCs were also detected within the KLS fraction of the neonate (P14) BM by phenotypic (Figure 1i) and transplantation analysis (Figure 3b, c). Transplantation of 2000 GFP+ or 500 Tom+ KLS cells from P14 BM led to long-term reconstitution of all myeloid and lymphoid lineages, in a pattern similar to that observed for FL cells (Fig 2d). GFP+ HSC therefore arise as early as E10.5 (Figure 1d), are capable of homing to the BM and persist for a limited time. However, they disappear from the BM between 2 weeks and 8 weeks of age, coinciding with a previously described “switch” in hematopoiesis that occurs after 3 weeks of age in mice [14,34,35].

Despite their inability to persist into adulthood *in situ*, GFP+ FL HSCs were able to persist when transplanted into an adult environment. To determine the impact of the microenvironment on the persistence of GFP+ FL HSC, we determined the ability of the GFP+ FL HSC to engraft when transplanted in utero into a fetal environment. We performed in utero transplantation (IUT) of Tom+ or GFP+ FL KLS cells or adult KLS cells into the fetal liver of WT E14.5 embryos (Fig. 3d-f). Adult BM cells performed poorly in IUT assays as compared to FL cells, as previously described for unfractionated adult BM cells[36] (Fig. 3d-f). Amazingly, almost 40% of fetal recipients of 2500-5000 Tom+ FL KLS cells demonstrated LTMR, defined as sustained reconstitution of myeloid and lymphoid lineages at > 0.1% (Fig 3d; Extended data Table 3). In contrast, LTMR was only observed in 10% of fetal recipients of GFP+ FL KLS cells and in none of the fetal recipients of adult KLS cells. The majority of fetal recipients reconstituted by GFP+ FL KLS cells demonstrated short-term multilineage reconstitution (STMR) or reconstitution in the lymphoid lineages only. Nonetheless, chimerism was still observed in the BM, spleen, and thymus of reconstituted recipients (Fig. 3d-f, Fig. 6h-i, and data not shown), albeit at low levels, confirming sustained self-renewal potential. The inefficiency of GFP+ FL HSCs to sustain long-term engraftment when transplanted into an unirradiated, fetal environment is reflective of their inability to persist in this same setting in situ. These data suggest that cell-extrinsic factors within the [transplanted] adult environment are necessary to confer long-term self-renewal of the GFP+ FL HSC.

Fetal HSCs are intrinsically less quiescent than adult HSCs [35,37]. To determine whether differences in quiescence may underlie the inability of the GFP+ HSC to persist in situ, we determined cell cycle parameters of Tom+ and GFP+ FL HSCs in comparison to adult HSCs. CD150 was used to enhance the purity of the HSCs, as we confirmed that FL HSCs were enriched in the CD150+ fraction for both Tom+ and GFP+ FL KLS cells

(Extended data Figure 7)[28]. Propidium iodide staining revealed a greater proportion of GFP+ CD150+ KLS cells in S/G2/M phase as compared to Tom+ CD150+ KLS cells (Figure 3G,H). Additionally, costaining for Hoechst and EdU incorporation revealed an increased proportion of GFP+ CD150+ KLS cells in S phase with a reciprocal reduction of cells in the G0/G1 phase as compared to Tom+CD150+ KLS cells (Figure 3i,j). Our data support previously established differences in the proliferation state between fetal HSCs and adult HSCs (Figure 3j) [14,35], and further show that within the fetal HSC fraction, GFP+ HSCs are more proliferative than Tom+ HSCs. Consistent with increased proliferation, GFP+ FL HSCs also expressed higher levels of the cell cycle regulator cyclind1 as compared to Tom+ FL HSCs (Figure 3k). The less quiescent state of the GFP+ HSC within the fetal liver raises the possibility that they may exhaust more rapidly during development as compared to Tom+ FL HSC, as recently suggested by studies tracing unperturbed fetal HSCs in vivo [38].

GFP+ FL HSCs are functionally distinct from Tom+ HSCs

Quantitative assessment of the cellular output after primary transplantation revealed functional differences between the Tom+ and GFP+ HSCs (Figure 4). Despite robust and persistent generation of myeloid cells, the percent contribution to the myeloid lineages was significantly lower in primary recipients of GFP+ FL KLS cells as compared to recipients of Tom+FL KLS cells (Figures 4a). These differences were also reflected in a significantly lower ratio of donor-derived myeloid to lymphoid cells in individual recipients of GFP+ KLS cells as compared to recipients of Tom+ KLS cells 16 weeks after transplantation (Fig. 4b). Differences in lineage output could not be attributed simply to persistence of long-lasting lymphoid cells from GFP+ HSCs, as we also observed a lower ratio of myeloid to lymphoid progenitors in the BM of recipients of GFP+ HSCs as compared to recipients of Tom+ HSCs (Fig. 4c). Comparison of the levels of BM HSC chimerism in individual recipients further suggested that differences in lineage output were due to higher output of lymphoid cells, as

opposed to lower myeloid output, from GFP+ FL HSCs (Fig. 4d). Differences in lymphoid readout between Tom+ and GFP+ cells were further exaggerated when Flk2 was additionally included as a marker to account for inefficient floxing (Extended data Fig. 5b-d). Evidence for a lymphoid bias was also found in the readout from GFP+ HSC in the P14 neonate BM (Fig. 4e,f), as well as from GFP+ FL HSCs transplanted in utero (Fig. 4g,h). Enhanced lymphoid cell production from GFP+ FL HSCs was further maintained upon serial transplantation, even when secondary transplantations were performed from primary donors with similar levels of GM chimerism (Fig. 4i,j). Lastly, GFP+ FL HSCs exhibited transcriptional lymphoid priming, as evidenced by upregulated expression of Rag1, IL7ra, and TCR-β as compared to Tom+ FL HSCs (Figure 4k), suggesting that the lymphoid bias was intrinsic to the GFP+ HSCs themselves, and not mediated by co-transplanted progenitor cells. Thus, the developmentally restricted GFP+ FL HSCs can be functionally distinguished from their Tom+ counterparts as they displayed transcriptional lymphoid priming and consistently exhibited greater lymphoid cell production across ontogeny and transplantation settings.

GFP+ FL HSCs have enhanced capacity to generate innate-like T-cells within a fetal thymic microenvironment

As the existence of GFP+ FL HSCs during a limited developmental window coincides with production of specialized hematopoietic cell types, we next examined their capacity to generate developmentally regulated mature cell subsets. We detected only definitive, adult-type globin gene expression in erythroid progenitors derived from transplanted GFP+ or Tom+ FL HSCs or adult HSCs (Extended data Fig. 8), suggesting that GFP+ FL HSCs do not contribute to primitive erythropoiesis. These data are consistent with the switch from primitive to definitive erythropoiesis occurring prior to E14.5 [39]; thus, we did not further pursue the developmental erythroid potential of GFP+ and Tom+ FL HSCs. Instead, we compared the ability of GFP+ and Tom+ FL HSCs to generate fetal-specified lymphoid cell subsets.

Specialized B1-type B-cells and certain γδ T-cells are efficiently generated from fetal progenitor cells, whereas adult BM cells have mostly lost this capacity [17-20,40]. These "innate-like" lymphoid subsets differ from conventional B- and T-cells in that they reside primarily within specific tissues and express a limited set of receptors that respond primarily to self-antigens [41].

We investigated the generation of innate-like lymphocytes in mice that had been serially transplanted with Tom+ or GFP+ FL KLS cells (Fig. 3a) and comparable transplants of adult HSCs. Serially transplanted mice were examined to ensure that the donor-derived cells originated from HSCs and not from co-transplanted progenitor cells within the KLS fraction. Secondary recipients of Tom+ or GFP+ FL KLS cells and primary recipients of adult KLS cells with similar levels of T-cell chimerism (Fig. 5a) displayed no differences in the frequency of donor-derived thymic T-cell subsets (Fig. 5b,c). The vast majority of donor-derived CD3+ T-cells in the thymus expressed the adult-type T-cell receptor (TCR)-β, whereas < 1% of donor-derived thymic CD3+ T-cells expressed TCR-γδ (Fig. 5b,c), regardless of developmental origin of the transplanted cells. In mice, T-cells expressing the developmentally limited TCR-Vγ3 are generated only from FL, but not adult, HSCs and comprise the dendritic epidermal T-cell (DETC) population [20]. Consistent with a previous report that the adult thymus does not support the generation of developmentally limited TCR-Vγ3+ cells [17], we observed negligible (< 0.1%) generation of TCR-Vγ3+ cells in the thymi of transplanted adult recipients (Fig. 5c). Examination of donor-derived TCR-Vγ3+ DETC cells within the epidermis in secondary recipients revealed minimal levels of DETC chimerism (Extended data Table 4), with even lower chimerism observed upon in utero transplantation (data not shown). While these results show that FL HSCs do retain the capability to generate DETCs in adult recipients, the requirement for very high levels of overall T-cell chimerism and

the absence of consistent chimerism among transplanted individuals prevented conclusive quantitative comparisons of the DETC capability of the different HSC populations.

As a previous study reported that a fetal thymic microenvironment is necessary for the generation of TCR-Vγ3+ cells [17], we examined differences in developmental T-cell potential using fetal thymic organ cultures (FTOCs). While FL CD150+ KLS cells efficiently differentiated into T cells, adult CD150+ KLS cells consistently failed to contribute robustly to CD3+ T-cells in the FTOC assay (Fig. 5d). Tom+ or GFP+ CD150+ FL KLS cells yielded similar profiles of T-cell differentiation (Fig. 5e,f), with the notable exception of more efficient generation of TCR-Vγ3+ T-cells by GFP+ CD150+ FL KLS as compared to their Tom+ counterparts (Fig. 5g-i). Whereas at least 1/3 of Tom+ cultures failed to yield any TCR-Vγ3+ T-cells, all GFP+ cultures contained TCR-Vγ3+ T-cells (Fig. 5g), and each GFP+ Vγ3-containing culture had significantly higher levels of Vγ3+ T-cells compared to Tom+ FTOCs (Fig. 5h,i). These findings show that both the progenitor cell and thymic environment must be of fetal origin to allow generation of TCR-Vγ3+ T-cells. In addition, these data suggest that GFP+ FL HSCs are a more critical source of TCR-Vγ3+ T-cells during development.

GFP+ FL HSCs efficiently generate innate-like B-cells in vivo

We next investigated the ability of Tom+ and GFP+ FL HSCs to give rise to innate-like B-cell subsets. All analyses were conducted in secondary recipients to ensure that donor-derived cells originated from stem cells and to exclude the possibility of cell production from contaminating lymphoid progenitors present within the originally transplanted fractions. Supporting a fetal origin of peritoneal cavity (PerC) B-cells [19,21], we observed strikingly higher donor-derived contribution to peritoneal IgM+ B-cells in secondary recipients of fetal compared to adult HSCs (Fig. 6a,b, Extended data Fig. 9a). GFP+ FL HSCs gave rise to significantly higher peritoneal B-cell chimerism than Tom+ FL HSCs, even in cohorts of mice

with similar B-cell contribution in the PB (Fig. 6b). To determine whether the low contribution to PerC B-cells by adult HSCs could be attributed to their overall lower lymphoid cell generation, we quantified peritoneal B cell output in recipient mice of adult HSCs or Tom+ or GFP+ FL HSCs that displayed comparable B-cell chimerism in the PB (Extended data Fig. 10). This analysis reinforced the relative myeloid bias of adult versus fetal HSCs, and attenuated the differences in peritoneal B cell contribution between adult HSCs and Tom+ FL HSCs. Strikingly, the superior ability of GFP+ FL HSCs to generate peritoneal B-cells was still apparent (Extended data Fig. 9). Furthermore, not only was the contribution of GFP+ FL HSCs was higher to all peritoneal B-cell subtypes (Fig. 6b and Extended data Figs. 9a, 10), they also gave rise to a significantly higher proportion of B1 type B-cells compared to both adult HSCs and Tom+ FL HSCs (Fig. 6c,d). The few donor-derived peritoneal B-cells found in recipients of adult HSCs were predominantly B2 cells (IgM+ CD5-CD11b-), with very few B1a (IgM+CD5+CD11b+) or B1b (IgM+CD5-CD11b+) cells (Fig. 6c,d). While Tom+ FL HSCs gave rise to a significantly higher proportion of B1a and B1b cells compared to adult HSCs, GFP+ FL HSCs were even more efficient in B1 cell generation (Fig. 6c,d). Despite limited long-term reconstitution upon in utero transplantation (Fig 3d-g, Fig. 6h), GFP+ FL HSC still retained a relatively robust capability to seed the peritoneal cavity and particularly the B1a cell compartment (Figure 6i) upon IUT. Collectively, our data reveal the unique capacity of the developmentally restricted GFP+ HSCs to generate innate-like B cells in the peritoneum and in particular fetal-restricted B1 B-cells.

Like peritoneal B1 B-cells, marginal zone (MZ) B-cells within the spleen exhibit innate-like immune responses and have a suggested fetal origin [40,42]. We examined the contribution of Tom+ and GFP+ FL HSCs and adult HSCs to MZ B-cells in the spleen by strategies analogous to those used for PerC contribution described above. In mice with comparable PB B-cell chimerism, adult HSCs yielded lower overall chimerism within the

spleen as compared to fetal HSCs, as well as significantly reduced chimerism of MZ B-cells (Fig. 6e-g, Extended data Fig. 9b). The lower contribution by adult HSCs to MZ B-cells supports their proposed fetal origin. Tom+ and GFP+ FL secondary recipients did not exhibit significant differences in percent chimerism of splenic B-cells subsets, for either total B-cells (IgM+B220+), follicular (FO, $CD21^{-}CD23^{hi}$) or MZ ($CD21^{hi}CD23^{lo}$) B-cells (Fig. 6e, Extended data Fig. 8b). However, within the donor-derived fraction of splenic B-cells, GFP+ FL HSCs gave rise to a significantly higher ratio of MZ B-cells as compared to both Tom+ FL HSCs and adult HSCs (Fig. 6f,g). Similarly, in utero transplantation revealed the ability of GFP+ FL HSCs to preferentially seed the MZ of the spleen despite lower overall chimerism (Figure 6j, k). Collectively, these data reveal that the developmentally limited GFP+ HSCs are endowed with a unique capability to generate fetal-restricted innate-like B-cells. Retention of this ability upon transplantation into both fetal and adult recipients also shows that this capability is mainly regulated cell-intrinsically and does not require a fetal environment.

In summary, our fate mapping of Flk2 expression identified a novel, definitive HSC that is capable of LTMR upon transplantation into adult recipients, but does not persist into adulthood and is not a precursor to adult HSCs. Our discovery of a 'disappearing' stem cell with unique lineage potential provides a refined framework for understanding the specification and hierarchy of early HSCs, the regulation of self-renewal and engraftment, and the mechanisms underlying developmental waves of hematopoietic cell output. In conjunction with previous fate mapping studies that have detected the persistence into adulthood of cells labeled during early development [6,7,9], our discovery of a transient HSC population supports the existence of at least two distinct fetal definitive HSC lineages: HSCs that contribute to the adult HSC compartment (Tom+), and HSCs that do not (GFP+). Though the demonstration of transient HSCs challenges the fundamental concept that stem cells contribute to tissue regeneration across the lifespan, our findings fit with a recent lineage tracing report

suggesting that fetal hematopoiesis is supported by distinct, rapidly extinguishing HSCs [38]. Collectively, our data support a model of parallel definitive hematopoietic development in which partially overlapping hematopoietic waves arise from distinct progenitor subsets.

Comparison of the lineage output of Tom+ and GFP+ FL HSCs and adult HSCs suggests that the distinct properties manifested by fetal and adult HSCs may be attributed mainly to the population labeled by GFP, as opposed to the entire fetal HSC compartment. The differences observed between Tom+ and GFP+ FL HSCs also explain the basis for previous reports on functional [34] and phenotypic [13] heterogeneity within the fetal HSC compartment. Spatial and temporal analysis of HSCs during mouse development, complemented by genetic fate-mapping approaches, have contributed to a model in which definitive HSCs generated within the AGM of the mid-gestation embryo subsequently seed and expand within the FL prior to relocation and lifelong endurance in the BM [7-9]. The LTMR capability of prenatal cells has played a critical role in their designation as precursors to adult HSCs [8,10,11,43,44]. Our data show that not all developmental HSCs with LTMR capacity contribute to the adult HSC compartment, and therefore that adult repopulating capability cannot be used as the sole criterium to define developmental populations that give rise to adult HSCs. Moreover, we show that cells that do not normally persist into adulthood can display long-term engraftment upon transplantation. This uncoupling of self-renewal capacity *in situ* with that observed upon transplantation suggests that cues associated with transplantation can induce persistence of otherwise transient cells. These findings provide insights for maintaining engraftable HSCs ex vivo and for deriving cells with repopulation capability from pluripotent cells.

By prospectively isolating and functionally testing a transient HSC, we provide direct evidence that discrete HSC populations are responsible for layered immune development, as proposed several decades ago by the Herzenberg Lab [22]. Evidence for this mechanism in

humans was recently supported by studies of human T-cell development [23]. The coincidence in existence of the GFP+ HSCs with the perinatal burst in lymphoid development suggests that these HSCs play important roles in specifying innate-like B- and T-cells, while also contributing to robust establishment of adaptive immunity. Intriguingly, we found that development of distinct B-cell subsets appears to be regulated primarily cell intrinsically, as fetal HSCs generate more fetal-like B-cells than adult HSCs even within an adult environment. In contrast, T-cell development appears to be regulated both cell intrinsically and extrinsically, as both the cell-of-origin and the thymic microenvironment affected the generation of fetal-restricted T-cells. Together, these data indicate that both fetal-specific HSCs and the fetal niche regulate the establishment of innate-like immune cells. These findings shed new light on immune system development and have important implications for understanding the mechanisms regulating maternal-fetal tolerance during development, the development of autoimmune disorders, and determining optimal vaccination efficiency [45,46]. Furthermore, the demonstration of unique mature cell output from a developmentally limited HSC suggests that potential to generate tolerogenic immune cells should be considered when choosing the optimal source of cells for transplantation, particularly for treatment of autoimmune disease.

AUTHOR CONTRIBUTIONS

A.E.B. and E.C.F conceived the study, designed the experiments, and co-wrote the paper. AEB performed experiments and analyzed the data. S.W.B. designed and performed experiments. T.M. designed experiments. S.C.D, G.E.H, C.J, E.A, and J,P-C. performed experiments. All authors read the manuscript.

ACKNOWLEDGEMENTS

The authors thank the vivarium staff at the UCSC vivarium as well as Bari Nazario and the Institute for the Biology of Stem cells/CIRM flow cytometry facility for their support. The authors would also like to acknowledge Chukwuemeka Ajaelo, Elizabeth Lopez, and Shahar Zimmerman for their technical contribution.

METHODS

Mice

All mice were maintained in the UCSC vivarium according to IACUC-approved protocols. FlkSwitch mice were bred and maintained as previously described [24,25,47,48]. WT recipients on a C57Bl/6 background were originally obtained from Jackson Laboratory and maintained in our colony. All adult recipients were 8-12 weeks of age at time of transplant.

Cell isolation and analysis

BM and PB cells were isolated and processed as previously described [49,50] using a four-laser FACSAria or LSRII (BD Biosciences, San Jose, CA). Embryo tissues were isolated by dissection and cells were processed by incubation with 1mg/mL collagenase for 15-20 minutes at 37°C. Analysis and display of flow cytometry data was accomplished using FlowJo analysis software (Ashland, OR). Cell populations in the PB and BM were defined as follows: KLS (Lin- Sca1+ c-kit+), HSC (Lin- Sca1+ c-kit+ Flk2- CD150+ CD48-), MPP (Lin- Sca1+ c-kit+ Flk2+ CD150- CD48+), Myeloid Progenitor (Lin- Sca1- c-kit+), CMP (Lin- Sca1- c-kit+ FcgR^{mid} CD34^{mid}), GMP (Lin- Sca1– c-kit+ FcgR^{hi} CD34^{hi}), MEP (Lin- Sca1- c-kit+ FcgR^{lo} CD34^{lo}), CLP (Lin- Sca1^{mid} c-kit^{mid} IL7Ra+, Flk2+), GM (Ter119- Mac1+ Gr1+ CD3- B220-), EP (Mac1- Gr1- CD3- B220- Ter119+ CD71+), Plts (FSC^{lo} Ter119- CD61+), B-cell (Ter119- Mac1- Gr1- CD3- B220+), and T-cell (Ter119- Mac1- Gr1- CD3+ B220-). The

lineage cocktail consisted of antibodies recognizing CD3, CD4, CD5, CD8, B220, Gr1, Mac1 and Ter119. Mac1 was excluded from the lineage cocktail for staining of fetal KLS populations as FL HSCs have been previously shown to be Mac1+ [16].

Transplantation assays

Uteri were removed from pregnant female mice at E14.5, and fetal livers were isolated from individual embryos and prepared as a single cell suspension. The fluorescence profile of each liver was assessed a priori to determine approximate floxing efficiency of individual embryos; subsequently, only embryos with higher floxing efficiencies were used for transplantation assays. Tom+ and GFP+ KLS fractions were sorted from fetal liver cell preparations enriched for cKit+ fraction using MACS beads (Miltenyi). The same protocol was used to isolate Tom+ and GFP+ KLS cells from the P14 neonate BM. 500 Tom+ and 1000 GFP+ FL or 2000 GFP+ P14 BM KLS cells were transplanted into sublethally irradiated (770 rads) WT adult recipients. For competitive transplantation assays, different cells doses were transplanted into lethally irradiated (1000 rads) WT recipients along with 250,000 unlabeled adult WBM cells. Chimerism in the PB was monitored by analyzing Tom- and GFP-fluorescent cells by flow cytometry over 16 weeks following transplantation. Long-term multilineage reconstitution was defined as reconstitution in all four mature lineages (GM, Plt, B-cell, and T-cell) at levels > 0.1%. Primary recipients were killed after 16 weeks and the frequencies of fluorescent cell populations in the BM were analyzed. For serial transplantation assays, 2 X 10^7 WBM cells from some of the primary recipients that demonstrated engraftment in all lineages and at least > 1% GM engraftment 16 weeks post-transplantation were transplanted into lethally irradiated (1000 rads) WT recipients.

To assess donor cell chimerism in additional organs, spleen and thymus were harvested from secondary recipients and prepared as a single cell suspension. Dendritic epidermal T-cells

were isolated from the ears by incubation at 37°C with 2.5 mg/mL dispase for 2 hours followed by 1mg/mL collagenase type I for 1-1.5 hours. A cell suspension was prepared by forcing the treated tissue through a 19-gage needle. Cells were harvested from the peritoneal cavity by lavage of 5mL PBS with 2% serum and 5mM EDTA.

Analysis of innate-like lymphocyte chimerism was normalized to the reconstitution from Tom+ FL KLS cells in each experiment due to inherent inter-experiment variability among secondary recipients. Analyses were also performed on subsets of secondary recipients that were closely matched for PB B-cell, T-cell or GM chimerism, as appropriate, to avoid unnecessary bias associated with baseline differences in chimerism. Comparisons to adult HSCs were made in either secondary recipients of adult HSCs (Lin-, cKit+, Sca1+, Flk2-), performed as described above for serial transplantation of FL primary recipients, or in lethally irradiated (1000 rad) primary recipients of 500 adult KLS cells (Lin-, ckit+, Sca1+). Comparison was based on comparable levels of PB B-cell, T-cell or GM chimerism. Reconstitution from adult HSCs was normalized to the mean of all recipients of Tom+ FL cells.

In utero transplantation assays

WT C57Bl/6 fetal mice were injected with Tom+ or GFP+ KLS cells or adult KLS cells into isolated as described above. Fetal intrahepatic injections were performed as previously described[51]. Briefly, under isofluorane-induced general anesthesia, a midline laparotomy was performed on pregnant dams, the uterus was exposed, and 5 µL of cell suspension was injected into the fetal liver of each pup using pulled glass micropipettes. The abdominal incision was closed in multiple layers. Chimerism in the PB was monitored beginning at weaning for 12-16 weeks as described above.

qRT-PCR

Quantitative RT-PCR was performed as previously described[50,52], except that reactions were performed on a Viia7 cycler (Applied Biosystems) using Quantace Sensimix Plus SYBR. cDNA amounts were normalized to expression of β-actin between samples.

PCR analysis of Tom and GFP expression

Genomic DNA **(**gDNA) was isolated from FACS-purified FL KLS cells separated based on Tom or GFP expression. gDNA was isolated using the QIAamp DNA Micro kit (Qiagen) . Multiplex PCR was used to simultaneously determine the presence of the Tom allele, Tom excision, and the mT/mG genomic insertion in one PCR reaction.

Cell cycle analysis

Staining for propidium iodide (PI) was performed as described previously [53] on sorted Tom+ or GFP+ fractions of CD150+ FL KLS cells. To perform cell cycle analysis, pregnant FlkSwitch mice were given an intraperitoneal injection of 100 uL of EdU (1mg/mL) and sacrificed 1 hour later. Fetal livers were isolated from E14.5 embryos and Tom+ or GFP+ fractions of CD150+ FL KLS cells were isolated by flow cytometry. Prior to sorting, cells were fixed in 2% formaldehyde. After sorting, cells were permeabilized in 0.15% Triton X, and EdU staining was visualized with the Click-it EdU kit from Life Technologies (C10337). Co-staining for Hoechst was simultaneously performed.

In vitro migration assays

In vitro migration assays were performed as previous described[49]. Briefly, FL cells were lineage depleted by magnetic selection, and preincubated for 1 hr at 37°C. One-fourth FL equivalent was placed in the upper chamber of a transwell insert (5uM pore size) in which the bottom well contained 100 ng/mL SDF1. Cells underwent migration for 2 hrs at 37°C before harvesting and analysis by flow cytometry. HSC migration was determined by the percent migration of CD150+ KLS cells.

Fetal thymic organ culture

Thymic lobes were removed from E14.5 embryos and irradiated (2000 γ) prior to culture. Individual thymic lobes were cultured in a hanging drop in Terasaki plates in 25 uL of culture media with 500 sorted Tom+ or GFP+ CD150+ KLS cells for 72 hours. Thymi were then transferred to Nucleopore filters (Millipore) set atop gelfoam sponges (Upjohn) in 6-well plate and cultured for 18 days before analysis. Cultures were maintained in a 37°C incubator with 5% CO2 in 2 mL of RPMI 1640 media containing 15% FCS, 0.1% Primocin, 10mM HEPES, and 50μM β-mercaptoethanol.

Analysis of globin gene expression

Donor-derived, nucleated erythroid progenitors were isolated from CFU-S colonies 13 days after transplantation of sorted Tom+ or GFP+ CD150+ FL KLS cells into lethally irradiated WT hosts. Total RNA was isolated by Trizol extraction and RT-PCR was performed using the High Capacity cDNA Reverse Transcription kit (Life Technologies). qPCR was performed on a ViiA 7 Real Time system (Applied Biosystems) using Quantace SensiMixPlus SYBR with the following mouse primers: βmajor/minor-For: 5'-TCTGCTATCATGGGTAATGCCAAA, βmajor/minor-Rev: 5'-GAAGGCAGCCTGTGCAGCG, βhl For – 5'-ATCATGGGAAACCCCCGGA, βhl Rev – 5'- GGGTGAATTCCTTGGCAAAATGAGT, εγ- For: 5'-CAAGCTACATGT GGATCCTGAGAA, εγ- Rev: 5'- TGCCGAAGTGACTAGCCAAA, CD71-

For: 5’-TCCGCTCGTGGAGACTACTT, CD71-Rev: 5’-ACATAGGGCGACAGGAAGTG. Expression of CD71 was used to normalize cDNA amounts between samples.

Statistical analysis

Statistically significant differences between groups were assessed by unpaired, two-tailed Student’s t-test in most cases. For analysis of qPCR and in vitro migration, statistical differences were assessed by paired, two-tailed Student’s t-test.

Figure 1

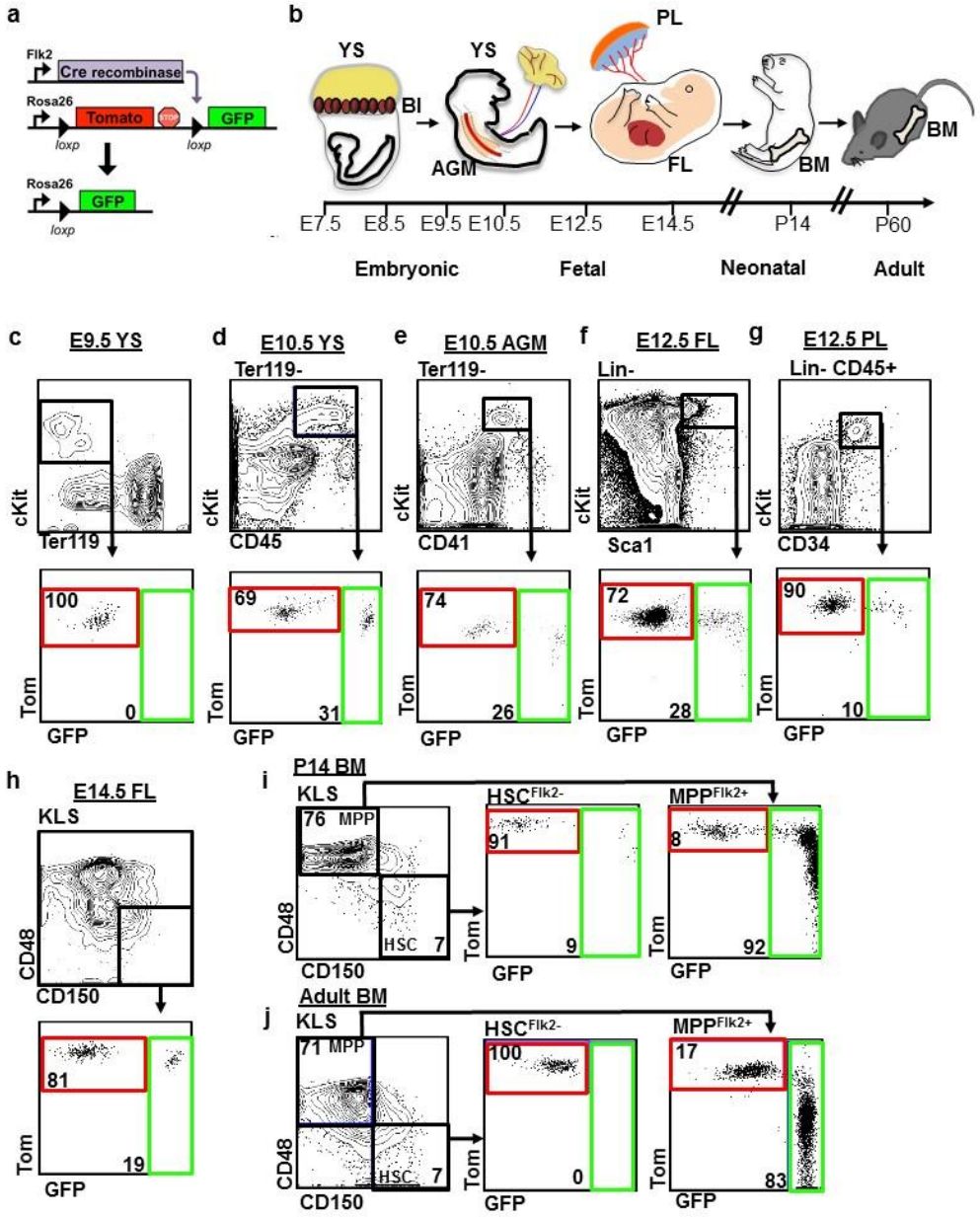

a
Flk2
Cre recombinase
Rosa26
Tomato
STOP
GFP
loxp
loxp
Rosa26
GFP
loxp
b
YS
YS
PL
Bl
AGM
FL
BM
BM
E7.5
E8.5
E9.5
E10.5
E12.5
E14.5
P14
P60
Embryonic
Fetal
Neonatal
Adult
c
E9.5 YS
cKit
Ter119
100
0
Tom
GFP
d
E10.5 YS
Ter119-
cKit
CD45
69
31
Tom
GFP
e
E10.5 AGM
Ter119-
cKit
CD41
74
26
Tom
GFP
f
E12.5 FL
Lin-
cKit
Sca1
72
28
Tom
GFP
g
E12.5 PL
Lin- CD45+
cKit
CD34
90
10
Tom
GFP
h
E14.5 FL
KLS
CD48
CD150
81
19
Tom
GFP
i
P14 BM
KLS
76 MPP
HSC 7
CD48
CD150
HSCFlk2-
91
9
Tom
GFP
MPPFlk2+
8
92
Tom
GFP
j
Adult BM
KLS
71 MPP
HSC 7
CD48
CD150
HSCFlk2-
100
0
Tom
GFP
MPPFlk2+
17
83
Tom
GFP

Figure 1. GFP+ cells coexist with Tom+ cells in fetal stem and progenitor compartments in FlkSwitch mice.

a, Strategy for generation of the FlkSwitch lineage tracing mouse model. The Rosa26 locus drives reporter expression in all cells. Cells derived via a Flk2-negative pathway express Tomato, whereas Flk2-driven expression of Cre recombinase leads to deletion of the Tomato coding region and an irreversible switch to GFP expression in Flk2-expressing cells and their progeny.

b, Chronological depiction of developmental hematopoiesis. Primitive hematopoiesis initiates in the blood islands (BI) of the early embryonic yolk sac (YS). The main sites of definitive hematopoiesis in the mid-gestation embryo include the aorta-gonad-mesonephros region (AGM), the placenta (PL), and the fetal liver (FL). Late fetal hematopoiesis occurs primarily in the FL, with the bone marrow (BM) becoming the main site of HSC residence around birth, where they persist throughout life.

c-h, Phenotypic stem cell- and progenitor-enriched populations from E10.5-14.5 contain both Tom+ and GFP+ cells. Plots depict flow cytometric analysis of reporter activity within the hematopoietic stem cell- and progenitor-enriched compartments of the YS at E9.5 (c); the YS (d) and AGM (e) at E10.5; the FL (f) and PL (g) at E12.5; and the FL at E14.5 (h) in FlkSwitch embryos. Top row depicts gating strategy for stem and progenitor cell compartments. Bottom row depicts reporter gene fluorescence profiles of cells within gated regions shown above.

i and j, A fraction of neonatal HSCs express GFP, whereas all adult HSCs express Tom. Flow cytometric analysis of reporter activity in the P14 (i) and adult (j) BM. Plots (left) indicate the gating strategies used to define HSCs (c-kit+Lineage-Sca1+ (KLS) Flk2-CD150+ CD48- BM cells) and multipotent progenitors (MPPs; KLS Flk2+CD150- CD48+ BM cells).

Numbers indicate representative frequencies of Tom- and GFP-expressing cells from each gated cell population.

Figure 2

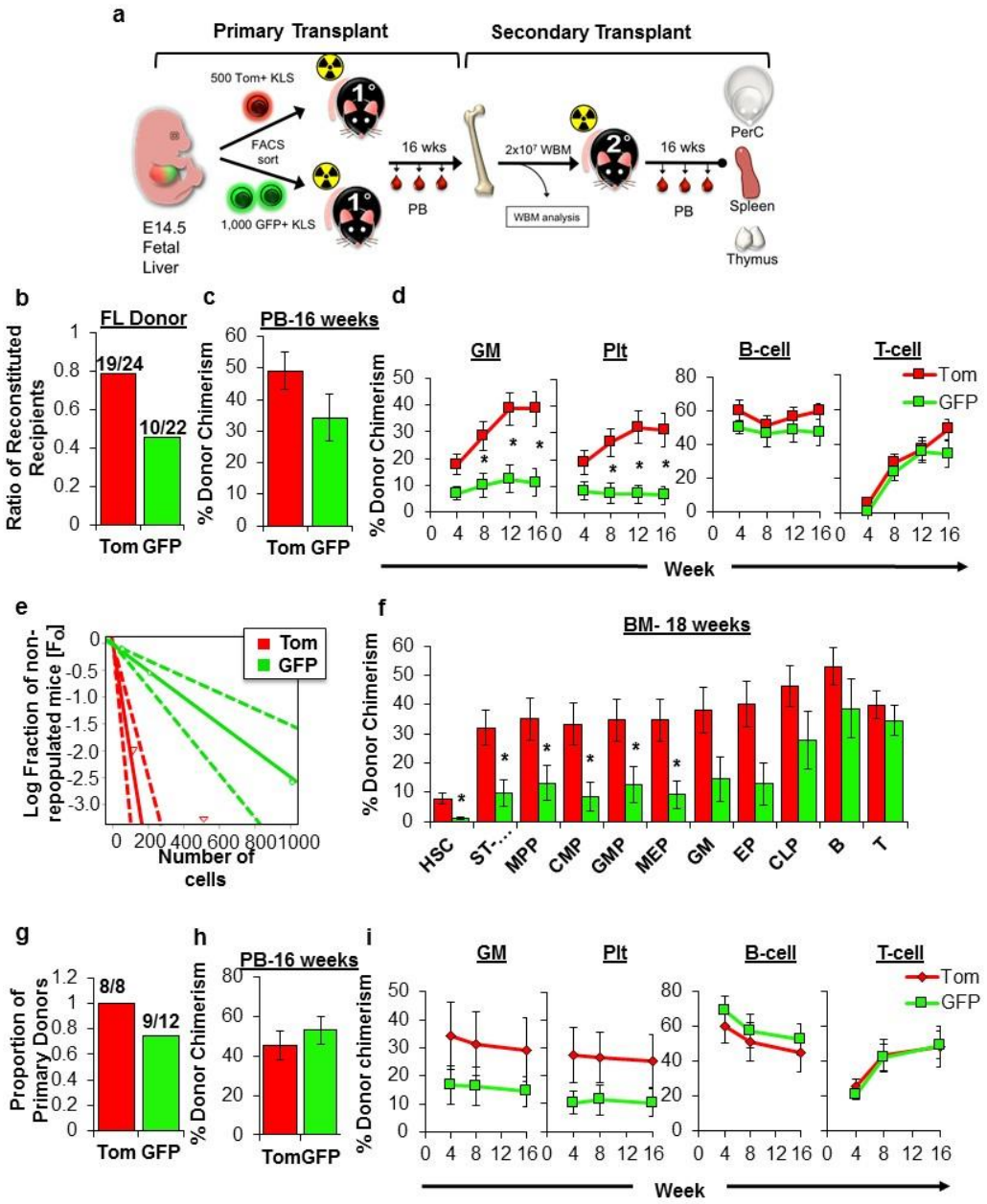

a
Primary Transplant
Secondary Transplant
500 Tom+ KLS
FACS sort
1,000 GFP+ KLS
E14.5 Fetal Liver
16 wks
PB
2x10^7 WBM
WBM analysis
16 wks
PB
PerC
Spleen
Thymus
b
FL Donor
Ratio of Reconstituted Recipients
19/24
10/22
Tom GFP
c
PB-16 weeks
% Donor Chimerism
Tom GFP
d
GM
Plt
B-cell
T-cell
Tom
GFP
% Donor Chimerism
Week
e
Log Fraction of non-repopulated mice [Fo]
Tom
GFP
Number of cells
f
BM- 18 weeks
% Donor Chimerism
HSC
ST-...
MPP
CMP
GMP
MEP
GM
EP
CLP
B
T
g
Proportion of Primary Donors
8/8
9/12
Tom GFP
h
PB-16 weeks
% Donor Chimerism
TomGFP
i
GM
Plt
B-cell
T-cell
Tom
GFP
% Donor chimerism
Week

Figure 2. Both Tom+ and GFP+ fetal HSCs possess serial reconstitution potential

a, Schematic of the experimental approach for primary and secondary transplantation and analyses. Sublethally-irradiated primary recipients were transplanted with Tom+ (500 cells) or GFP+ (1000 cells) FL KLS cells isolated from FlkSwitch fetal mice. Donor-derived chimerism was monitored in the peripheral blood (PB) over 16 weeks. After 18 weeks 2 X 10^7 WBM cells from individual primary donors were transplanted into one or more lethally irradiated secondary recipients. Chimerism was determined in the blood, peritoneal cavity (PerC), spleen, and thymus after 16 weeks in secondary recipients.

b, Proportion of primary recipients exhibiting long-term multilineage reconstitution (LTMR) following transplantation of either 500 Tom+ or 1000 GFP+ FL KLS cells. LTMR was defined as reconstitution > 0.1% in all four lineages over 16 weeks. Additional details can be found in Extended data Table 1.

c, Total white blood cell (WBC) contribution to the PB 16 weeks post-transplantation in in primary recipients exhibiting LTMR over 16 weeks following transplantation. N= 10-19 recipient mice representing 4 independent experiments. Data are shown as mean ± SEM.

d, Peripheral blood (PB) contribution by Tom+ or GFP+ FL KLS cells to the granulocyte/monocyte (GM), platelet (Plt), B-cell and T-cell lineages in the same mice from (c). Data are shown as mean ± SEM. *, $P<0.05$.

e, Donor chimerism of stem, progenitor and mature cells in the bone marrow (BM) 18 weeks post-transplantation in the mice from (c). Cell populations are defined in the online methods.

f, Limited dilution analysis was performed by competitive transplantation of three doses of Tom+ or GFP+ FL KLS cells into lethally-irradiated hosts. Data are shown as the log fraction of non-engrafted (non-repopulated) mice plotted on the y axis versus the transplanted cell

dose on the x axis. ELDA software (http://bioinf.wehi.edu.au/software/elda/) was used to determine HSC frequency (color-coded in bold) and assess statistical significance.

g, Proportion of primary donors that gave rise to LTMR in at least one secondary recipient after secondary transplantation. Additional details can be found in Extended data Table 2.

h, Total WBC donor chimerism in the PB in mice exhibiting LTMR over 16 weeks after secondary transplantation. N = 21-22 recipients per cell type representing 4 independent experiments. Data are shown as mean ± SEM.

i, PB contribution of donor-derived cells to GM, Plt, B-cell, and T-cell lineages over 16 weeks following transplantation in mice from (g). *, $P<0.05$.

Figure 3

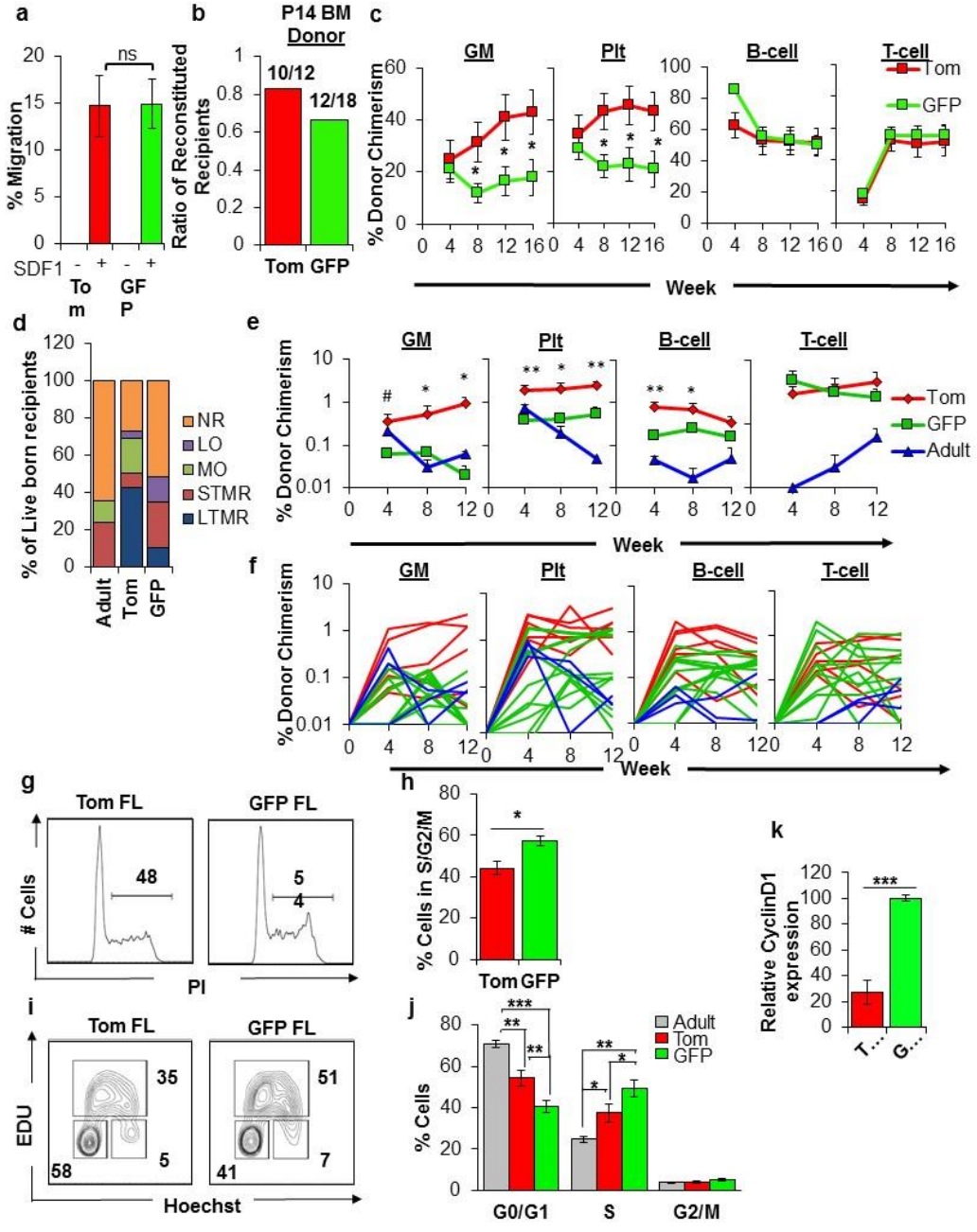

a
% Migration
ns
SDF1 - + - +
Tom
GFP
b
P14 BM
Donor
Ratio of Reconstituted Recipients
10/12
12/18
Tom GFP
c
GM
Plt
B-cell
T-cell
% Donor Chimerism
Tom
GFP
Week
d
% of Live born recipients
NR
LO
MO
STMR
LTMR
Adult
Tom
GFP
e
GM
Plt
B-cell
T-cell
% Donor Chimerism
Tom
GFP
Adult
Week
f
GM
Plt
B-cell
T-cell
% Donor Chimerism
Week
g
Tom FL
GFP FL
Cells
48
5
4
PI
h
% Cells in S/G2/M
Tom GFP
k
Relative CyclinD1 expression
i
Tom FL
GFP FL
EDU
35
5
58
51
7
41
Hoechst
j
% Cells
Adult
Tom
GFP
G0/G1
S
G2/M

Figure 3. Cell-extrinsic and cell-intrinsic mechanisms limit the developmental window of the GFP+ HSC

a-c, GFP+ fetal HSC migrate to and seed the neonate BM.

a, The percentage of Tom+ or GFP+ CD150+ FL KLS cells that migrated towards an SDF1 gradient in vitro. Data are shown as mean ± SEM from four independent experiments performed in triplicate. ns, not significant.

b, Proportion of mice exhibiting LTMR following transplantation of either 500 Tom+ or 2000 GFP+ neonate KLS cells. Cells were isolated from the P14 BM of FlkSwitch mice and transplanted into sublethally-irradiated WT recipients.

c, Peripheral blood (PB) contribution by Tom+ or GFP+ P14 BM KLS cells to the GM, Plt, B-cell and T-cell lineages in mice exhibiting LTMR over 16 weeks following transplantation. N= 10-12 recipient mice representing 3 independent experiments. Data are shown as mean ± SEM. *, $P<0.05$.

d-f, GFP+ fetal HSCs display limited long-term engraftment following *in utero* transplantation.

d, The percentage of live-born recipients of Tom+ or GFP+ FL KLS cells or adult KLS cells transplanted in utero into the FL of WT embryos at E14.5. Live-born recipients were classified on the basis of donor-derived chimerism within the GM, Plt, B-cell and T-cell lineages over 12 weeks post-birth as non-reconstituted (NR), or demonstrating myeloid only (MO), lymphoid only (LO), short-term multilineage reconstitution (STMR) or long-term multilineage reconstitution (LTMR). Additional information on cell doses and total number of live embryos per experiment can be found in Extended data Table 3.

e, Peripheral blood (PB) contribution by Tom+ or GFP+ FL KLS cells or adult KLS cells to GM, Plt, B-cell, and T-cell lineages over 12 weeks post-birth following in utero transplantation

in recipient mice exhibiting STMR or LTMR as described in (d). Data are shown as mean ± SEM. #, P<0.1 *, P<0.05; **, P<0.01.

f, PB chimerism in individual recipients of Tom+ or GFP+ FL KLS cells or adult KLS cells to GM, Plt, B-cell, and T-cell lineages over 12 weeks following in utero transplantation in mice demonstrating STMR or LTMR as described in (d). This panel displays the reconstitution of individual mice that were shown as mean reconstitution in panel (e).

g-k, GFP+ FL HSC are less quiescent than Tom+ FL and adult HSC.

g-i, Analysis of cell cycle status of Tom+ and GFP+ FL and adult CD150+ KLS cells. (**g-h**) Representative flow cytometry plots (**g**) and accompanying quantification (**h**) of propidium iodide (PI) staining for DNA content. Brackets and values indicate the representative percentage of cells in G2/S/M. (**i-j**) Representative flow cytometry plots (**i**) and accompanying quantification (**j**) of cell cycle status as determined by measurement of EdU incorporation and Hoechst DNA staining. Values indicate representative frequencies of gated populations. N = 8 representing three independent experiments. Data are shown as mean ± SEM. *, P < 0.05. **, P < 0.01. *** P < 0.001.

k) Expression of Cyclin D1 is higher in GFP+ CD150+ KLS cells as compared to their Tom+ counterparts. Data are shown as mean ± SEM. N= 4 independent experiments performed in triplicate. *** P< 0.001.

Figure 4

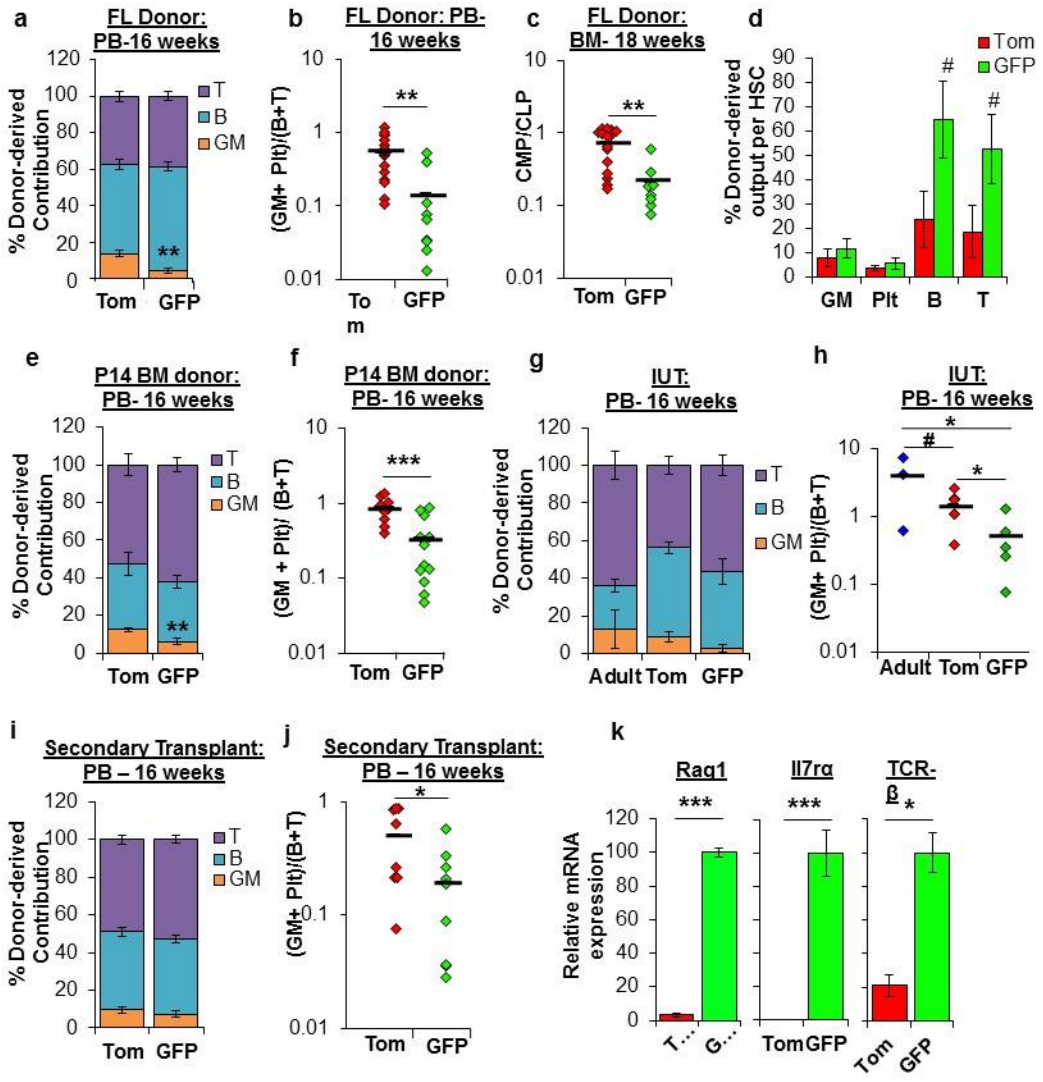

a
FL Donor: PB-16 weeks
% Donor-derived Contribution
T
B
GM
Tom
GFP
b
FL Donor: PB-16 weeks
(GM+ Plt)/(B+T)
c
FL Donor: BM- 18 weeks
CMP/CLP
d
% Donor-derived output per HSC
Tom
GFP
GM
Plt
B
T
e
P14 BM donor: PB- 16 weeks
f
P14 BM donor: PB- 16 weeks
(GM + Plt)/ (B+T)
g
IUT: PB- 16 weeks
Adult
h
IUT: PB- 16 weeks
i
Secondary Transplant: PB – 16 weeks
j
Secondary Transplant: PB – 16 weeks
k
Rag1
Il7rα
TCR-β
Relative mRNA expression

Figure 4. Developmentally-restricted GFP+ HSC are lymphoid-biased

a, Distribution of donor-derived contribution to WBC lineages in the peripheral blood (PB) 16 weeks post-transplantation in primary recipients of FL KLS cells from Figure 2d.

b, The ratio of myeloid (Plt+GM) to lymphoid (B+T) chimerism 16 weeks post-transplantation in the PB of individual primary recipients of Tom+ or GFP+ FL KLS cells from Figure 2d. **, $P < 0.01$.

c, The ratio of myeloid progenitors (CMP) to lymphoid progenitors (CLP) in the BM of individual primary recipients of Tom+ or GFP+ FL KLS cells from Figure 2e. **, $P < 0.01$.

d, Donor-derived contribution to each mature lineage calculated on a per HSC basis, based on HSC chimerism as determined in the BM of individual primary recipients of FL KLS cells (Fig. 2e). See online methods for calculations.

e, Distribution of donor-derived contribution to WBC lineages in the PB 16 weeks post-transplantation in recipients of P14 BM KLS cells from Figure 3c. **, $P< 0.01$.

f, The ratio of myeloid (Plt+GM) to lymphoid (B+T) chimerism 16 weeks post-transplantation in the PB of individual recipients of Tom+ or GFP+ P14 BM KLS cells from Figure 3c. ***, $P < 0.001$.

g, Distribution of donor-derived contribution to WBC lineages in the PB 16 weeks post- in utero transplantation of Tom+ or GFP+ FL KLS or adult KLS cells in recipients from Figure 3e.

h, The ratio of myeloid (Plt+GM) to lymphoid (B+T) chimerism in the PB of individual recipients transplanted in utero with Tom+ or GFP+ FL KLS or adult KLS cells from Figure 3e. #, $P < 0.01$. *, $P< 0.05$.

i, Distribution of donor-derived contribution to WBC lineages in the PB 16 weeks post-transplantation in secondary recipients of FL KLS cells from Fig. 2i.

f, The ratio of myeloid (Plt+GM) to lymphoid (B+T) chimerism in the PB 16 weeks post-transplantation in individual secondary recipients of Tom+ or GFP+ FL KLS cells from Fig. 2i. *, $P < 0.05$.

k, Relative levels of the indicated transcripts in Tom+ and GFP+ CD150+ KLS cells isolated from the E14.5 FlkSwitch FL as quantified by qRT-PCR. N= 2-4 independent experiments performed in triplicate. *, $P < 0.05$. ***, $P < 0.001$.

Figure 5

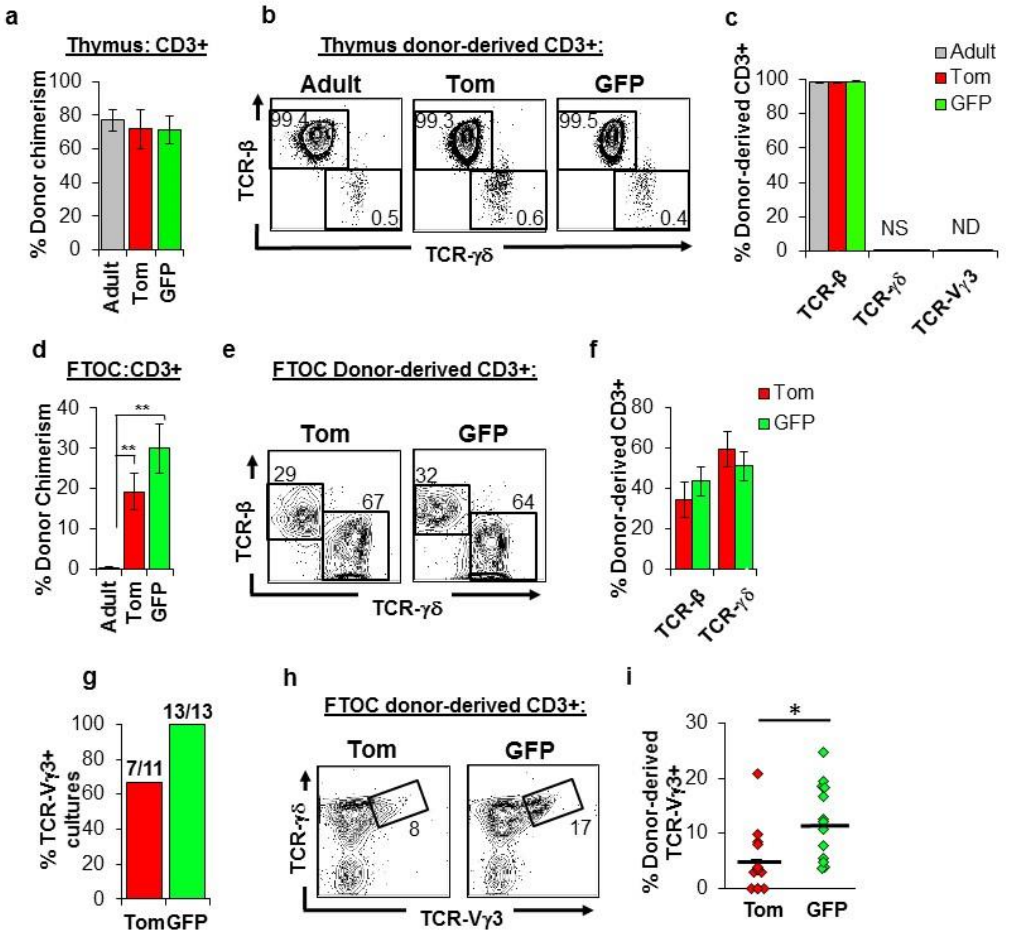

a
Thymus: CD3+
% Donor chimerism
Adult
Tom
GFP
b
Thymus donor-derived CD3+:
Adult
Tom
GFP
TCR-β
TCR-γδ
99.4
0.5
99.3
0.6
99.5
0.4
c
% Donor-derived CD3+
Adult
Tom
GFP
NS
ND
TCR-β
TCR-γδ
TCR-Vγ3
d
FTOC:CD3+
% Donor Chimerism
e
FTOC Donor-derived CD3+:
Tom
GFP
29
67
32
64
f
% Donor-derived CD3+
g
% TCR-Vγ3+ cultures
7/11
13/13
Tom
GFP
h
FTOC donor-derived CD3+:
TCR-γδ
TCR-Vγ3
8
17
i
% Donor-derived TCR-Vγ3+

Figure 5. GFP+ fetal HSCs display superior ability to give rise to developmentally-restricted T-cells in a fetal thymic microenvironment.

a-c, Neither fetal nor adult HSCs give rise to substantial numbers of TCR□□-expressing T cells upon transplantation into adult recipients.

a, Donor contribution to CD3+ thymocytes within the thymus of adult, secondary recipients of Tom+ and GFP+ fetal KLS cells or comparable transplants of adult KLS cells 18-weeks post-transplantation. N = 10-12 representing three independent experiments. Data are shown as mean ± SEM.

b-c, Representative FACS plots (b) and accompanying quantification (c) of donor-derived CD3+ thymocyte subsets for the mice described in (a). Numbers indicate representative frequencies of gated cell subsets. Data are shown as mean ± SEM.

d-i, Fetal, but not adult, HSCs efficiently produce T cells in a fetal thymic microenvironment.

d, Donor contribution to CD3+ thymocytes within fetal thymic organ cultures (FTOCs) seeded with Tom+ or GFP+ fractions of CD150+ FL KLS cells or adult CD150+ KLS cells. N = 12-13 representing three independent experiments. Data are shown as mean ± SEM. *, $P<0.05$. **, $P< 0.01$.

e-f, Representative flow cytometry plots (e) and accompanying quantification (f) of donor-derived contribution to TCR□- and TCR□□□expressing subsets in FTOC assays described from (d) are shown. Numbers indicate representative frequencies of gated populations.

g-i, GFP+ FL HSCs give rise to significantly more TCR-Vγ3+ T-cells in FTOC assays compared to Tom+ FL HSCs. Percentage of FTOC cultures described in (d) that contained TCR-Vγ3$^+$ cells (g). Representative flow cytometry plots (h) and quantification (i) of donor-

derived contribution to TCR-Vγ3$^+$ cells in FTOCs described in (d). Data are shown as individual data points representing individual cultures. *, P < 0.05.

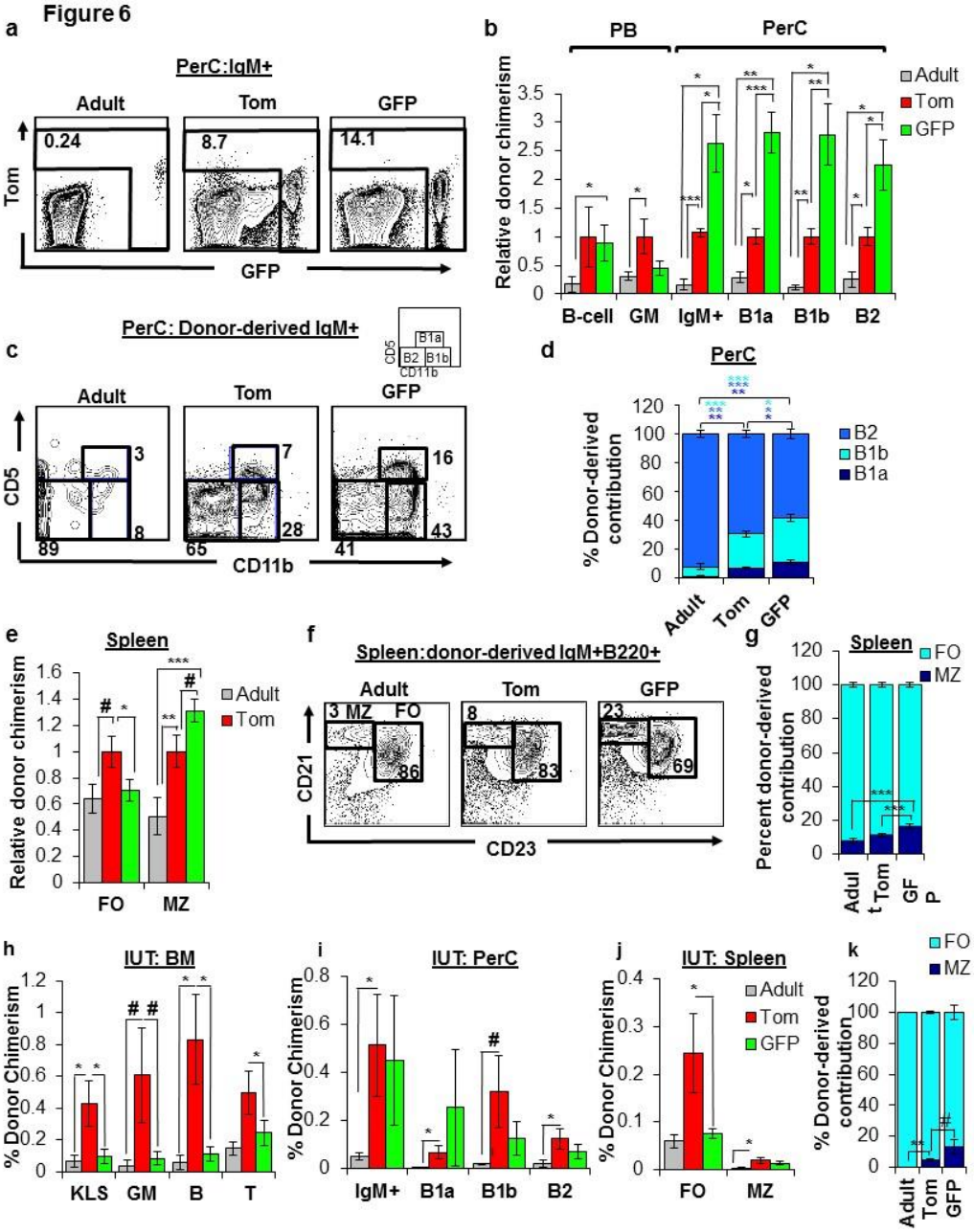


Figure 6. GFP+ fetal HSCs efficiently generate innate-like B-cells in vivo.

a-d, GFP+ FL HSCs generate peritoneal cavity (PerC) B cells with greater efficiency than Tom+ FL HSCs or adult HSCs.

a, Representative flow cytometry plots indicating donor-derived IgM+ cells within the peritoneal cavity in secondary recipients of adult HSCs, Tom+ FL HSCs, or GFP+ FL HSCs. Donor-derived cells from adult and Tom+ donors can be either Tom+ or GFP+; donor-derived cells from GFP+ donors are GFP+.

b, Donor chimerism of PB B-cells and GM cells, and peritoneal cavity IgM+ cells, B-1a (IgM+ CD5+ Cd11b+), B-1b (IgM+ CD5- CD11b+) and B2 (IgM+ CD5- CD11b-) B-cells in secondary recipients of either Tom+ or GFP+ FL KLS cells, or adult HSCs (KLSF-). Results are displayed relative to recipients of Tom+ FL HSCs; Extended data Figure 9a displays the results as percent donor chimerism.

c-d, Representative flow cytometry plots (c) and accompanying quantification (d) of donor-derived contribution to peritoneal cavity B-cell subsets in secondary recipients of adult HSCs, Tom+ FL HSCs, or GFP+ FL HSCs.

e-g, GFP+ FL HSCs give rise to a higher proportion of marginal zone B cells in the spleen compared to Tom+ FL HSCs or adult HSCs.

e, Donor chimerism of splenic marginal zone (MZ; $CD21^{hi}$ CD23-) and follicular zone (FO, $CD21^{lo}$ CD23+) B-cells among secondary recipients of Tom+ or GFP+ FL KLS cells or comparable recipients of adult KLS cells. Results are displayed relative to recipients of Tom+ FL HSCs; Extended data Figure 9b displays the results as percent donor chimerism.

f-g, Representative flow cytometry plots (f) and quantification (g) of donor-derived contribution to splenic B-cell subsets from the same mice as in (e).

N = 5-17 per group in at least three independent experiments. Numbers indicate P<0.1. *, P<0.05. **, P< 0.01. *** P< 0.001.

h-k, Contribution of Tom+ and GFP+ FL or adult HSCs to BM, PerC, and spleen cells upon *in utero* transplantation.

h, BM donor chimerism in progenitor and mature cells among fetal recipients transplanted in utero with Tom+ or GFP+ FL KLS cells or adult KLS cells. Chimerism was analyzed 12-16 weeks post-transplantation in mice that exhibited STMR or LTMR (Fig. 3d). Cell populations are defined in the online methods. N = 3-10 group. Data are shown as mean ± SEM. #, P<0.10. *, P<0.05.

i, Donor chimerism of total PerC IgM+ cells, and PerC B-1a, B-1b and B2 B-cells among fetal recipients of Tom+ or GFP+ FL KLS cells or adult KLS cells 12-16 weeks post- in utero transplantation in the same mice as in (h).

j, Donor chimerism of splenic MZ and FO B-cells among fetal recipients of Tom+ or GFP+ FL KLS cells or adult KLS cells 12-16 weeks post-in utero transplantation in the same mice as in (h).

k, Quantification of donor-derived contribution to splenic B-cell subsets among fetal recipients of Tom+ or GFP+ FL KLS cells 12-16 weeks post-in utero transplantation in the same mice as in (h).

References

1 Medvinsky, A., Rybtsov, S. & Taoudi, S. Embryonic origin of the adult hematopoietic system: advances and questions. *Development* **138**, 1017-1031 (2011).

2 Dzierzak, E. & Speck, N. A. Of lineage and legacy: the development of mammalian hematopoietic stem cells. *Nature immunology* **9**, 129-136 (2008).

3 McGrath, K. E. & Palis, J. Hematopoiesis in the yolk sac: more than meets the eye. *Experimental hematology* **33**, 1021-1028, doi:10.1016/j.exphem.2005.06.012 (2005).

4 Gomez Perdiguero, E. *et al.* Tissue-resident macrophages originate from yolk-sac-derived erythro-myeloid progenitors. *Nature* **518**, 547-551, doi:10.1038/nature13989 (2015).

5 Hoeffel, G. *et al.* C-myb(+) erythro-myeloid progenitor-derived fetal monocytes give rise to adult tissue-resident macrophages. *Immunity* **42**, 665-678, doi:10.1016/j.immuni.2015.03.011 (2015).

6 Samokhvalov, I. M., Samokhvalova, N. I. & Nishikawa, S. Cell tracing shows the contribution of the yolk sac to adult haematopoiesis. *Nature* **446**, 1056-1061 (2007).

7 Gothert, J. R. *et al.* In vivo fate-tracing studies using the Scl stem cell enhancer: embryonic hematopoietic stem cells significantly contribute to adult hematopoiesis. *Blood* **105**, 2724-2732 (2005).

8 Kumaravelu, P. *et al.* Quantitative developmental anatomy of definitive haematopoietic stem cells/long-term repopulating units (HSC/RUs): role of the aorta-gonad-mesonephros (AGM) region and the yolk sac in colonisation of the mouse embryonic liver. *Development* **129**, 4891-4899 (2002).

9 North, T. E. *et al.* Runx1 expression marks long-term repopulating hematopoietic stem cells in the midgestation mouse embryo. *Immunity* **16**, 661-672 (2002).

10 Medvinsky, A. & Dzierzak, E. Definitive hematopoiesis is autonomously initiated by the AGM region. *Cell* **86**, 897-906 (1996).

11 Muller, A. M., Medvinsky, A., Strouboulis, J., Grosveld, F. & Dzierzak, E. Development of hematopoietic stem cell activity in the mouse embryo. *Immunity* **1**, 291-301 (1994).

12 Matsuoka, S. *et al.* CD34 expression on long-term repopulating hematopoietic stem cells changes during developmental stages. *Blood* **97**, 419-425 (2001).

13 Christensen, J. L. & Weissman, I. L. Flk-2 is a marker in hematopoietic stem cell differentiation: a simple method to isolate long-term stem cells. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 14541-14546, doi:10.1073/pnas.261562798 (2001).

14 Bowie, M. B. *et al.* Hematopoietic stem cells proliferate until after birth and show a reversible phase-specific engraftment defect. *The Journal of clinical investigation* **116**, 2808-2816, doi:10.1172/JCI28310 (2006).

15 Rebel, V. I., Miller, C. L., Eaves, C. J. & Lansdorp, P. M. The repopulation potential of fetal liver hematopoietic stem cells in mice exceeds that of their liver adult bone marrow counterparts. *Blood* **87**, 3500-3507 (1996).

16 Morrison, S. J., Hemmati, H. D., Wandycz, A. M. & Weissman, I. L. The purification and characterization of fetal liver hematopoietic stem cells. *Proceedings of the National Academy of Sciences of the United States of America* **92**, 10302-10306 (1995).

17 Ikuta, K. *et al.* A developmental switch in thymic lymphocyte maturation potential occurs at the level of hematopoietic stem cells. *Cell* **62**, 863-874 (1990).

18 Ghosn, E. E. *et al.* Distinct B-cell lineage commitment distinguishes adult bone marrow hematopoietic stem cells. *Proceedings of the National Academy of Sciences*

of the United States of America **109**, 5394-5398, doi:10.1073/pnas.1121632109 (2012).

19 Hardy, R. R. & Hayakawa, K. A developmental switch in B lymphopoiesis. *Proceedings of the National Academy of Sciences of the United States of America* **88**, 11550-11554 (1991).

20 Havran, W. L. & Allison, J. P. Developmentally ordered appearance of thymocytes expressing different T-cell antigen receptors. *Nature* **335**, 443-445, doi:10.1038/335443a0 (1988).

21 Barber, C. L., Montecino-Rodriguez, E. & Dorshkind, K. Reduced production of B-1-specified common lymphoid progenitors results in diminished potential of adult marrow to generate B-1 cells. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 13700-13704, doi:10.1073/pnas.1107172108 (2011).

22 Herzenberg, L. A. & Herzenberg, L. A. Toward a layered immune system. *Cell* **59**, 953-954 (1989).

23 Mold, J. E. *et al.* Fetal and adult hematopoietic stem cells give rise to distinct T cell lineages in humans. *Science* **330**, 1695-1699, doi:10.1126/science.1196509 (2010).

24 Boyer, S. W., Beaudin, A. E. & Forsberg, E. C. Mapping stem cell differentiation pathways from hematopoietic stem cells using Flk2/Flt3L lineage tracing. *Cell Cycle* **11**, 3180-3188 (2012).

25 Boyer, S. W., Schroeder, A. V., Smith-Berdan, S. & Forsberg, E. C. All hematopoietic cells develop from hematopoietic stem cells through Flk2/Flt3-positive progenitor cells. *Cell stem cell* **9**, 64-73, doi:10.1016/j.stem.2011.04.021 (2011).

26 Buza-Vidas, N. *et al.* FLT3 expression initiates in fully multipotent mouse hematopoietic progenitor cells. *Blood* **118**, 1544-1548, doi:10.1182/blood-2010-10-316232 (2011).

27 Kiel, M. J. *et al.* SLAM family receptors distinguish hematopoietic stem and progenitor cells and reveal endothelial niches for stem cells. *Cell* **121**, 1109-1121, doi:10.1016/j.cell.2005.05.026 (2005).

28 Kim, I., He, S., Yilmaz, O. H., Kiel, M. J. & Morrison, S. J. Enhanced purification of fetal liver hematopoietic stem cells using SLAM family receptors. *Blood* **108**, 737-744, doi:10.1182/blood-2005-10-4135 (2006).

29 Buza-Vidas, N. *et al.* FLT3 receptor and ligand are dispensable for maintenance and posttransplantation expansion of mouse hematopoietic stem cells. *Blood* **113**, 3453-3460, doi:10.1182/blood-2008-08-174060 (2009).

30 Zou, Y. R., Kottmann, A. H., Kuroda, M., Taniuchi, I. & Littman, D. R. Function of the chemokine receptor CXCR4 in haematopoiesis and in cerebellar development. *Nature* **393**, 595-599, doi:10.1038/31269 (1998).

31 Nagasawa, T. *et al.* Defects of B-cell lymphopoiesis and bone-marrow myelopoiesis in mice lacking the CXC chemokine PBSF/SDF-1. *Nature* **382**, 635-638, doi:10.1038/382635a0 (1996).

32 Sugiyama, T., Kohara, H., Noda, M. & Nagasawa, T. Maintenance of the hematopoietic stem cell pool by CXCL12-CXCR4 chemokine signaling in bone marrow stromal cell niches. *Immunity* **25**, 977-988, doi:10.1016/j.immuni.2006.10.016 (2006).

33 Nie, Y., Han, Y. C. & Zou, Y. R. CXCR4 is required for the quiescence of primitive hematopoietic cells. *The Journal of experimental medicine* **205**, 777-783, doi:10.1084/jem.20072513 (2008).

34 Benz, C. *et al.* Hematopoietic stem cell subtypes expand differentially during development and display distinct lymphopoietic programs. *Cell stem cell* **10**, 273-283, doi:10.1016/j.stem.2012.02.007 (2012).

35 Bowie, M. B. *et al.* Identification of a new intrinsically timed developmental checkpoint that reprograms key hematopoietic stem cell properties. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 5878-5882, doi:10.1073/pnas.0700460104 (2007).

36 Hayashi, S. *et al.* Mixed chimerism following in utero hematopoietic stem cell transplantation in murine models of hemoglobinopathy. *Experimental hematology* **31**, 176-184 (2003).

37 Bowie, M. B., Kent, D. G., Copley, M. R. & Eaves, C. J. Steel factor responsiveness regulates the high self-renewal phenotype of fetal hematopoietic stem cells. *Blood* **109**, 5043-5048, doi:10.1182/blood-2006-08-037770 (2007).

38 Busch, K. *et al.* Fundamental properties of unperturbed haematopoiesis from stem cells in vivo. *Nature*, doi:10.1038/nature14242 (2015).

39 Sankaran, V. G., Xu, J. & Orkin, S. H. Advances in the understanding of haemoglobin switching. *Br J Haematol* **149**, 181-194 (2010).

40 Yoshimoto, M. *et al.* Embryonic day 9 yolk sac and intra-embryonic hemogenic endothelium independently generate a B-1 and marginal zone progenitor lacking B-2 potential. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 1468-1473, doi:10.1073/pnas.1015841108 (2011).

41 Bendelac, A., Bonneville, M. & Kearney, J. F. Autoreactivity by design: innate B and T lymphocytes. *Nature reviews. Immunology* **1**, 177-186, doi:10.1038/35105052 (2001).

42 Carey, J. B., Moffatt-Blue, C. S., Watson, L. C., Gavin, A. L. & Feeney, A. J. Repertoire-based selection into the marginal zone compartment during B cell development. *The Journal of experimental medicine* **205**, 2043-2052, doi:10.1084/jem.20080559 (2008).

43 Christensen, J. L., Wright, D. E., Wagers, A. J. & Weissman, I. L. Circulation and chemotaxis of fetal hematopoietic stem cells. *PLoS biology* **2**, E75, doi:10.1371/journal.pbio.0020075 (2004).

44 de Bruijn, M. F., Speck, N. A., Peeters, M. C. & Dzierzak, E. Definitive hematopoietic stem cells first develop within the major arterial regions of the mouse embryo. *Embo J* **19**, 2465-2474 (2000).

45 Burt, T. D. Fetal regulatory T cells and peripheral immune tolerance in utero: implications for development and disease. *Am J Reprod Immunol* **69**, 346-358, doi:10.1111/aji.12083 (2013).

46 Montecino-Rodriguez, E. & Dorshkind, K. B-1 B cell development in the fetus and adult. *Immunity* **36**, 13-21, doi:10.1016/j.immuni.2011.11.017 (2012).

47 Epelman, S. *et al.* Embryonic and adult-derived resident cardiac macrophages are maintained through distinct mechanisms at steady state and during inflammation. *Immunity* **40**, 91-104, doi:10.1016/j.immuni.2013.11.019 (2014).

48 Hashimoto, D. *et al.* Tissue-resident macrophages self-maintain locally throughout adult life with minimal contribution from circulating monocytes. *Immunity* **38**, 792-804, doi:10.1016/j.immuni.2013.04.004 (2013).

49 Smith-Berdan, S. *et al.* Robo4 cooperates with CXCR4 to specify hematopoietic stem cell localization to bone marrow niches. *Cell stem cell* **8**, 72-83, doi:10.1016/j.stem.2010.11.030 (2011).

50 Forsberg, E. C., Serwold, T., Kogan, S., Weissman, I. L. & Passegue, E. New evidence supporting megakaryocyte-erythrocyte potential of flk2/flt3+ multipotent hematopoietic progenitors. *Cell* **126**, 415-426, doi:10.1016/j.cell.2006.06.037 (2006).

51 Nijagal, A., Le, T., Wegorzewska, M. & Mackenzie, T. C. A mouse model of in utero transplantation. *Journal of visualized experiments : JoVE*, doi:10.3791/2303 (2011).

52 Forsberg, E. C. *et al.* Differential expression of novel potential regulators in hematopoietic stem cells. *PLoS Genet* **1**, e28, doi:10.1371/journal.pgen.0010028 (2005).

53 Beaudin, A. E., Boyer, S. W. & Forsberg, E. C. Flk2/Flt3 promotes both myeloid and lymphoid development by expanding non-self-renewing multipotent hematopoietic progenitor cells. *Experimental hematology* **42**, 218-229 e214, doi:10.1016/j.exphem.2013.11.013 (2014).