UNIVERSITY OF CALIFORNIA, SAN DIEGO

SAN DIEGO STATE UNIVERSITY

The Role of E2A in Hematopoietic Development:

Establishment of a Multipotent Progenitor Cell Line to Identify Genetic and Epigenetic Regulatory Programs in Hematopoiesis

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy

in

Biology

by

Elinore McClelland Mercer

Committee in charge:

University of California, San Diego

Professor Cornelis Murre (Chair)
Professor Ananda Goldrath (Co-chair)
Professor Michael G. Rosenfeld

San Diego State University

Professor Sanford Bernstein
Professor Robert Zeller

2010

The Dissertation of Elinore McClelland Mercer is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____

Co-Chair

_____

Chair

University of California, San Diego

San Diego State University

2010

DEDICATION

To my husband, Andrew, for everything, and everything that is to be,

and

To my daughter Ruby, for adding profound depth and joy,

and

To my parents, for unflinching support.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to acknowledge and thank Kees Murre for giving me the opportunity to do research in his lab, and for always being enthusiastic and positive about my work. I am grateful to thank Kees for trusting me and allowing me tremendous flexibility in my work scheduling, which made the attainment of work-life balance possible.

I would also like to acknowledge all of the members of the Murre lab, past and present, for their friendship and intellectual support. I would especially like to thank Kristina Beck, Craig Semerad, Yin Lin, and Suchit Jhunjhunwala. I also wish to thank Christopher Benner for providing excellent bioinformatic support on demand.

The text of Chapter II, in full, is a reprint of the material as it appeared in *PNAS* (2009). Craig L. Semerad, Elinore M. Mercer, Matthew A. Inlay, Irving L. Weissman, and Cornelis Murre. 2009 Feb 10;106(6):1930-5. Craig Semerad was the primary author of the work. I contributed to experimental design, Figures 1and 3, as well as significant portions of the Supplement. Matthew Inlay performed FACS analysis for Figure 2 in the laboratory of Irving Weissman. The writing in Chapter II was a collaboration between Cornelis Murre and Craig Semerad.

Chapters III and IV are currently being prepared for submission for publication of the material. The working title is "Inducible Id2 expression allows for expansion of a long-term culture of multipotent hematopoietic progenitors and epigenetic studies of lineage commitment." I was the primary investigator and author of this material. Co-authors are Yin C. Lin, Suchit Jhunjhunwala., Christopher Benner, Christopher Glass, and Cornelis Murre.

VITA

EDUCATION

2010	Doctor of Philosophy, Biology University of California, San Diego and San Diego State University Advisor: Cornelis Murre, PhD
2002	Masters of Biotechnology Texas A&M University
2000	Bachelors of Science, Microbiology Texas A&M University

RESEARCH AND WORK EXPERIENCE

2005-2010	University of California, San Diego PhD Research (Advisor: Cornelis Murre, PhD)
2004-2005	Kalypsys, Inc. Research Associate (Advisor: Guy Oshiro, PhD)
2002-2004	Genoptix, Inc. Senior Research Associate (Advisor: Ilona Kariv, PhD)
1999-2000	Texas A&M University, College Station Undergraduate Research (Advisor: Deborah Siegele, PhD)
1997-1999	Texas A&M University, College Station Undergraduate Research (Advisor: William Park, PhD)

PUBLICATIONS

Mercer, E.M., Chun, L.Y, Jhunjhunwala, S., Benner, C., Glass, C., Murre, C. "Inducible Id2 expression allows for expansion of a long-term culture of multipotent hematopoietic progenitors and epigenetic studies of lineage commitment." Manuscript in preparation.

Semerad, C.L., Mercer, E.M., Inlay, M.A., Weissman, I.L., Murre, C.M. "E2A proteins maintain the hematopoietic stem cell pool and promote the maturation of myelolymphoid and myeloerythroid progenitors." *PNAS*, 2009 Feb 10;106(6):1930-5.

Wang, M.M., Tu, E., Raymond, D.E., Yang, J.M., Zhang, H., Hagen, N., Dees, B., Mercer, E.M., Forster, A.H., Kariv, I., Marchand, P.J., Butler, W.F. "Microfluidic Sorting of Mammalian Cells by Optical Force Switching." *Nature Biotechnology*, 2005 Jan;23(1):83-7

Forster, A.H., Wang, M.M., Butler, W.F., Chachisvilis, M., Chung, T.D.Y, Esener, S.C., Hall, J.M., Kibar, O., Lykstad, K., Marchand, P.J., Mercer, E.M., Pestana, L.M., Sur, S., Tu, E., Yang, R., Zhang, H., Kariv, I. "Use of Moving Optical Gradient Fields for Analysis of Apoptotic Cellular Responses in a Chronic Myeloid Leukemia Cell Model." *Analytical Biochemistry*. 2004 Apr 1;327(1):14-22.

Schnabel, C., Diver, J., Forster, A.H., Mercer, E.M., Hall, J., Nova, T. "Method of Using Optical Interrogation to Determine a Biological Property of a Cell or Population of Cells." U.S. Patent Pending, 2002.

ABSTRACT OF THE DISSERTATION

The Role of E2A in Hematopoietic Development:

Establishment of a Multipotent Progenitor Cell Line to Identify Genetic and Epigenetic Regulatory Programs in Hematopoiesis

by

Elinore McClelland Mercer

Doctor of Philosophy in Biology

University of California, San Diego, 2010
San Diego State University, 2010

Professor Cornelis Murre, Chair

Professor Ananda Goldrath, Co-chair

Over 55,000 bone marrow transplants are performed every year to treat patients suffering from bone and blood cancers, anemia, and auto-immune disorders. The success of these transplants is highly correlated to a high number of hematopoietic stem cells (HSCs) and/or early progenitor cells injected into the patient. Obtaining high numbers of these early progenitors from donors is extremely difficult, making *ex vivo* expansion of these cells an attractive option. Currently, however, there is still much uncertainty as to how to expand these cells while maintaining their pluripotency. In our lab, we have shown that mouse E47

knockout progenitor cells can be expanded almost indefinitely while retaining their pluripotency when cultured on S17 stromal cells in the presence of IL-7, SCF and Flt-3. Here, we overexpress Id2 under the control of a tetracycline-responsive promoter to expand multipotent progenitors that retain differentiating ability both in vitro and in vivo. This is a novel and robust method for ex vivo expansion of HSCs.

We then exploited the in vitro growth capacity and multipotent differentiation potential of these Id2-HPCs to study both the genetic and epigenetic regulation of B and myeloid development in mice. We analyzed gene expression patterns and genome-wide binding of histone modifications, namely H3K4me1 and H3K4me3, to generate a predictive model of cis-regulatory elements at active enhancers in multipotent Id2-HPCs, B cells, or myeloid cells. We also utilized the Id2-HPCs to study the temporal changes in gene expression, epigenetic modifications and transcription factor binding in the pre-pro-B to pro-B cell transition. This analysis has confirmed the role of E2A to directly bind to nearly half of the regulated genes in B-cell development and, in conjunction with H3K4me1 binding patterns, has allowed the prediction of novel enhancers that are activated by specific combinations of transcription factors in a temporal fashion.

I

General Introduction

Hematopoiesis in murine adult bone marrow

The differentiation of hematopoietic stem cells into the varied mature cells of the immune system is a tightly regulated process that controls the development of all the mature blood cell lineages at the proper numbers and ratios to maintain functional immunity and blood cell homeostasis. All of the cells of the immune system are derived from the long-term hematopoietic stem cell (LT-HSC). In the bone marrow, LT-HSCs have the ability to both self-renew and reconstitute the entire immune system for the life of the animal (Spangrude et al., 1988). LT-HSCs can differentiate into short-term hematopoietic stem cells (ST-HSCs), which also maintain multipotency, but only have self-renewal capability for a limited time. Short-term HSCs then differentiate into multipotent progenitors (MPPs), which are multipotent, but lack self-renewal capability (Akashi et al., 2000; Kondo et al., 1997).

Symmetric, self-renewing divisions of LT-HSCs occur quite rapidly in fetal development to ensure proper seeding of blood and immune compartments. However, in adults, LT-HSCs are often found in quiescent, not dividing states (Morrison and Kimble, 2006). The self-renewal capability of these cells is maintained through complex regulation of multiple pathways, including the Notch and Wnt signaling pathways, Hox transcription factors, and the cell cycle regulators INK4A, INK4C, p21, p27, and PTEN (Tsiftsoglou et al., 2009; Zon, 2008). LT-HSCs are identified through the absence of cell surface markers of differentiated immune cells (lineage negative), and the presence of cKit, Sca1 (LSK), and CD150. Downregulation of CD150 marks differentiation into the ST-HSC and MPP, and a subsequent decline of self-renewal activity (Kiel et al., 2005; Papathanasiou et al., 2009). All three of these subsets have the ability to differentiate into the erythroid or megakaryocyte lineage through a pre-megakaryocyte/erythroid progenitor, which has low to intermediate

levels of CD105 and CD34, and low levels of CD41, or into the lymphoid/myeloid lineage through the lymphoid primed MPP (LMPP), which has high levels of Flt3. LMPPs differentiate into the myeloid lineage through the FcγR postiive granulocyte/macrophage progenitor (GMP), or into the lymphoid lineage through the IL7R positive common lymphoid progenitor (CLP) (Dias et al., 2008a; Murre, 2009; Semerad et al., 2009).

From the CLP, early B-cell development proceeds through an ordered pathway starting with the Ly6D positive B-primed progenitor (BLP), then into a B220 intermediate and CD43 high pre-pro-B cell, a B220 and CD19 high pro-B cell, and finally into a CD25 positive pre-B cell (Borghesi et al., 2004). Many of these early lineages are defined based on rearrangement of the antigen receptor genes IgH, IgL and IgK (Chen and Alt, 1993). Rearrangement of the D-J joint of the IgH gene has been documented as early as the CLP stage, although resolution of IgH gene rearrangement in Ly6D-positive versus Ly6D-negative CLPs is still unclear. Upon expression of CD19 at the pro-B cell stage, V-DJ rearrangement of IgH has been completed. The cells then proliferate to a CD25 positive pre-B cell stage, where rearrangement of the BCR light chain occurs (Inlay et al., 2009; Mansson et al., 2010; Melchers et al., 1994).

Precise control of rearrangement of the BCR genes and differentiation through the B-cell pathway is controlled through expression of specific transcription factors. Ikaros is expressed in LT-HSCs throughout B-cell development, and Ikaros knockout animals show a severe block in B, T, natural killer (NK) and dendritic cell development. At the branchpoint of the megakaryocyte/erythroid and myeloid/lymphoid pathways, GATA-1 and PU.1 play opposing roles in promoting differentiation into the pre-MegE or LMPP, respectively. PU.1 can then control activation of CEBPα and AP-1 to direct myeloid differentiation. In contrast,

lymphoid differentiation is controlled by the joint activities of Ikaros, STAT5 and PU.1 and E2A. E2A directs the upregulation of EBF1, which, in concert with PAX5, commits cells to the B-cell lineage. E2A, EBF1, and PAX5 all play roles in upregulating other required B-cell genes such as Bcl11a, Ly6D, and Foxo1, although direct regulation of downstream targets by individual or combinations of transcription factors is still unclear (Busslinger, 2004; Dias et al., 2008b; Murre, 2009).

E and Id Proteins in Hematopoiesis and B-cell Development

The E2A gene encodes the splice variants E12 and E47, collectively referred to as E2A. E2A belongs to the class I family of helix-loop-helix (HLH) transcription factors. HLH proteins contain a conserved helix-loop-helix dimerization domain flanked by a basic DNA binding region, and two upstream activation domains which serve to recruit co-activators and co-repressors. The mammalian class I HLH protein family also includes E2-2 and HEB, and their various splice variants. These proteins all share the conservered bHLH structure, and act as homo- or heterodimers with other E- or HLH proteins (Massari and Murre, 2000). E2A was originally identified based on its ability to bind to the κ enhancer region in a B-cell lymphoma line (Murre et al., 1989).

There are seven classes of HLH factors which are segregated based on their tissue expression patterns, dimerization capabilities and ability to bind DNA. Class I HLH factors are expressed in numerous different tissues, whereas class II factors, such as Myc or NeuroD, have tissue restricted expression. Class III factors have the addition of a leucine-zipper motif near the basic DNA binding domain, class IV factors dimerize with Myc, class V factors lack a basic DNA binding domain, class VI factors have a central proline in the DNA binding

domain, and class VII factors have a PAS domain (Massari and Murre, 2000). The class V factors, known as the Id proteins, heterodimerize with class I and II factors to inhibit their ability to bind DNA (Benezra et al., 1990).

The function of E-proteins, and specifically E2A, has been best characterized in B-cell development. Animals that are lacking the E2A gene have a complete block in B-cell development at a BLP stage before DJ rearrangement of the IgH gene has occurred (Bain et al., 1994; Zhuang et al., 1994). Many of the B-cell specific genes, such as lambda-5, mb-1, VpreB and Pax5 are not expressed in E2A deficient BLPs or pre-pro-B cells, suggesting that E2A may control the expression of the genes, either by a direct or indirect mechanism. HEB and E2-2 also have roles in early B-cell development, as double heterozygotes show a perturbation in B-cell development at the pro-B cell stage (Zhuang et al., 1996). In addition, when HEB is knocked-in to the E2A site, HEB can rescue the severe B-cell phenotype seen in E2A knockout animals (Zhuang et al., 1998). This suggests that while E2A is essential for normal B-cell development, other class I HLH factors also have roles.

Recently, it has been demonstrated that E2A has a distinct role prior to B-cell development in the hematopoietic stem cell and LMPP stage of hematopoiesis, and even in megakaryopoiesis. There is a significant decrease in the number of LT-HSCs in E2A knockout animals, which can partially be explained by an increase in cell-cycling of these HSCs and a decrease in expression of cell cycle regulators Gfi-1, Mpl and p21(Semerad et al., 2009). There is also a significant decrease in the numbers of LMPPs in E2A heterozygous and knockout animals (Dias et al., 2008a; Semerad et al., 2009; Yang et al., 2008). Interestingly, E2A knockouts also have a reduction in preMegE and pre CFU-E progenitors, which may indicate control of the erytrhroid specific gene GATA1 by E2A (Semerad et al., 2009).

E2A is directly regulated by the class V HLH inhibitors Id1, Id2, Id3 and Id4. The Id proteins were identified through a screen to find new HLH proteins, and were subsequently shown to inhibit the DNA binding ability of the bHLH proteins (Benezra et al., 1990). Overexpression of the Id proteins inhibits B-cell development both in vitro and in vivo (Ji et al., 2008). Downregulation of Id2 and Id3 during pre-pro-B to pro-B cell development is crucial for proper differentiation, and this downregulation is at least indirectly dependent on the activity of EBF1 (Thal et al., 2009).

Epigenetics and hematopoiesis

The precisely controlled expression of the E- and Id genes, as well as all other genes involved in hematopoiesis and B-cell differentiation, mediates proper development and prevents aberrant cell growth or differentiation. Both direct and indirect binding of transcription factors, as well as epigenetic changes such as DNA methylation or histone modifications, control this process (Fisher, 2002; Hagman and Lukin, 2006). Histones are proteins that wind genomic DNA in 146 bp segments to compact the genome and control expression. Each unit of condensed DNA and histones is called a nucleosome, and is made up of two each of the four core histone subunits-H2A, H2B, H3 and H4. Histone tails can be post-translationally modified by acetylation, methylation, ubiquitination, and phosphorylations (Talbert and Henikoff, 2010). Recent work has illuminated specific roles for many of these modifications, including methylation of histone 3 lysine 4 (H3K4).

Histone 3 can be mono-, di- or trimethylated at lysine 4. It is becoming apparent that the presence of one of these marks at a promoter or enhancer can be predictive in promoter/enhancer activity and thus gene expression. H3K4 trimethylation (H3K4me3) is

typically concentrated at promoters, and seems to "mark" active or primed promoters. In contrast, H3K4 monomethylation (H3K4me1) is typically concentrated at promoter distal regions, and has been shown to correlate with p300 binding to mark active enhancers in the genome (Heintzman et al., 2009; Heintzman et al., 2007). The presence of H3K4 dimethylation (H3K4me2) is not as telling as the H3K4me1 or H3K4me3 nucleosome marks, although dimethylation is often found in the coding region of actively transcribed genes (Bernstein et al., 2002; Robertson et al., 2008).

Conclusion

This thesis will demonstrate the work completed on three projects: 1-the identification of the role of E2A in the HSC, LMPP and preMegE stages of hematopoiesis, 2-the creation and long-term establishment of a multipotent hematopoietic progenitor cell line through the overexpression of the Id2 protein, and 3-the subsequent genome-wide analyses during in vitro differentiation of this cell line into both B and myeloid lineages. This work delineates novel roles for E2A in direct gene regulation, a robust method for ex vivo expansion of hematopoietic progenitor cells, and predictions of specific enhancer elements and cis-regulatory factors involved in the activation and repression of genes during the lineage commitment choice.

References

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

Bain, G., Maandag, E. C. R., Izon, D. J., Amsen, D., Kruisbeek, A. M., Weintraub, B. C., Krop, I., Schlissel, M. S., Feeney, A. J., van Roon, M., *et al.* (1994). E2A proteins are required for proper B cell development and initiation of immunoglobulin gene rearrangements. Cell *79*, 885-892.

Benezra, R., Davis, R. L., Lockshon, D., Turner, D. L., and Weintraub, H. (1990). The protein Id: A negative regulator of helix-loop-helix DNA binding proteins. Cell *61*, 49-59.

Bernstein, B. E., Humphrey, E. L., Erlich, R. L., Schneider, R., Bouman, P., Liu, J. S., Kouzarides, T., and Schreiber, S. L. (2002). Methylation of histone H3 Lys 4 in coding regions of active genes
10.1073/pnas.082249499. Proceedings of the National Academy of Sciences of the United States of America *99*, 8695-8700.

Borghesi, L., Hsu, L.-Y., Miller, J. P., Anderson, M., Herzenberg, L., Herzenberg, L., Schlissel, M. S., Allman, D., and Gerstein, R. M. (2004). B Lineageâ€“specific Regulation of V(D)J Recombinase Activity Is Established in Common Lymphoid Progenitors 10.1084/jem.20031800. The Journal of Experimental Medicine *199*, 491-502.

Busslinger, M. (2004). Transcriptional Control of Early B Cell Development1 doi:10.1146/annurev.immunol.22.012703.104807. Annual Review of Immunology *22*, 55-79.

Chen, J., and Alt, F. W. (1993). Gene rearrangement and B-cell development. Current Opinion in Immunology *5*, 194-200.

Dias, S., Månsson, R., Gurbuxani, S., Sigvardsson, M., and Kee, B. L. (2008a). E2A Proteins Promote Development of Lymphoid-Primed Multipotent Progenitors. Immunity *29*, 217-227.

Dias, S., Xu, W., McGregor, S., and Kee, B. (2008b). Transcriptional regulation of lymphocyte development. Current Opinion in Genetics & Development
Differentiation and gene regulation *18*, 441-448.

Fisher, A. G. (2002). Cellular identity and lineage choice. *2*, 977-982.

Hagman, J., and Lukin, K. (2006). Transcription factors drive B cell development. Current Opinion in Immunology
Lymphocyte development / Tumour immunology *18*, 127-134.

Heintzman, N. D., Hon, G. C., Hawkins, R. D., Kheradpour, P., Stark, A., Harp, L. F., Ye, Z., Lee, L. K., Stuart, R. K., Ching, C. W., *et al.* (2009). Histone modifications at human enhancers reflect global cell-type-specific gene expression. *459*, 108-112.

Heintzman, N. D., Stuart, R. K., Hon, G., Fu, Y., Ching, C. W., Hawkins, R. D., Barrera, L. O., Van Calcar, S., Qu, C., Ching, K. A., *et al.* (2007). Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *39*, 311-318.

Inlay, M. A., Bhattacharya, D., Sahoo, D., Serwold, T., Seita, J., Karsunky, H., Plevritis, S. K., Dill, D. L., and Weissman, I. L. (2009). Ly6d marks the earliest stage of B-cell specification and identifies the branchpoint between B-cell and T-cell development 10.1101/gad.1836009. Genes & Development *23*, 2376-2381.

Ji, M., Li, H., Suh, H. C., Klarmann, K. D., Yokota, Y., and Keller, J. R. (2008). Id2 intrinsically regulates lymphoid and erythroid development via interaction with different target proteins
10.1182/blood-2008-01-133504. Blood *112*, 1068-1077.

Kiel, M. J., Yilmaz, Ö. H., Iwashita, T., Yilmaz, O. H., Terhorst, C., and Morrison, S. J. (2005). SLAM Family Receptors Distinguish Hematopoietic Stem and Progenitor Cells and Reveal Endothelial Niches for Stem Cells. Cell *121*, 1109-1121.

Kondo, M., Weissman, I. L., and Akashi, K. (1997). Identification of Clonogenic Common Lymphoid Progenitors in Mouse Bone Marrow. *91*, 661-672.

Mansson, R., Zandi, S., Welinder, E., Tsapogas, P., Sakaguchi, N., Bryder, D., and Sigvardsson, M. (2010). Single-cell analysis of the common lymphoid progenitor compartment reveals functional and molecular heterogeneity
10.1182/blood-2009-08-236398. Blood *115*, 2601-2609.

Massari, M. E., and Murre, C. (2000). Helix-Loop-Helix Proteins: Regulators of Transcription in Eucaryotic Organisms
10.1128/MCB.20.2.429-440.2000. Mol Cell Biol *20*, 429-440.

Melchers, F., Haasner, D., Grawunder, U., Kalberer, C., Karasuyama, H., Winkler, T., and Rolink, A. G. (1994). Roles of IGH and L Chains and of Surrogate H and L Chains in the Development of Cells of the B Lymphocyte Lineage
doi:10.1146/annurev.iy.12.040194.001233. Annual Review of Immunology *12*, 209-225.

Morrison, S. J., and Kimble, J. (2006). Asymmetric and symmetric stem-cell divisions in development and cancer. *441*, 1068-1074.

Murre, C. (2009). Developmental trajectories in early hematopoiesis 10.1101/gad.1861709. Genes & Development *23*, 2366-2370.

Murre, C., McCaw, P. S., and Baltimore, D. (1989). A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, MyoD, and myc proteins. Cell *56*, 777-783.

Papathanasiou, P., Attema, J. L., Karsunky, H., Xu, J., Smale, S. T., and Weissman, I. L. (2009). Evaluation of the Long-Term Reconstituting Subset of Hematopoietic Stem Cells with CD150. Stem Cells *27*, 2498-2508.

Robertson, A. G., Bilenky, M., Tam, A., Zhao, Y., Zeng, T., Thiessen, N., Cezard, T., Fejes, A. P., Wederell, E. D., Cullum, R., *et al.* (2008). Genome-wide relationship between histone H3 lysine 4 mono- and tri-methylation and transcription factor binding 10.1101/gr.078519.108. Genome Research *18*, 1906-1917.

Semerad, C. L., Mercer, E. M., Inlay, M. A., Weissman, I. L., and Murre, C. (2009). E2A proteins maintain the hematopoietic stem cell pool and promote the maturation of myelolymphoid and myeloerythroid progenitors 10.1073/pnas.0808866106. Proceedings of the National Academy of Sciences *106*, 1930-1935.

Spangrude, G., Heimfeld, S., and Weissman, I. (1988). Purification and characterization of mouse hematopoietic stem cells 10.1126/science.2898810. Science *241*, 58-62.

Talbert, P. B., and Henikoff, S. (2010). Histone variants [mdash] ancient wrap artists of the epigenome. *11*, 264-275.

Thal, M. A., Carvalho, T. L., He, T., Kim, H.-G., Gao, H., Hagman, J., and Klug, C. A. (2009). Ebf1-mediated down-regulation of Id2 and Id3 is essential for specification of the B cell lineage 10.1073/pnas.0802550106. Proceedings of the National Academy of Sciences *106*, 552-557.

Tsiftsoglou, A. S., Bonovolias, I. D., and Tsiftsoglou, S. A. (2009). Multilevel targeting of hematopoietic stem cell self-renewal, differentiation and apoptosis for leukemia therapy. Pharmacology & Therapeutics *122*, 264-280.

Yang, Q., Kardava, L., St. Leger, A., Martincic, K., Varnum-Finney, B., Bernstein, I. D., Milcarek, C., and Borghesi, L. (2008). E47 Controls the Developmental Integrity and Cell Cycle Quiescence of Multipotential Hematopoietic Progenitors. J Immunol *181*, 5885-5894.

Zhuang, Y., Barndt, R. J., Pan, L., Kelley, R., and Dai, M. (1998). Functional Replacement of the Mouse E2A Gene with a Human HEB cDNA. Mol Cell Biol *18*, 3340-3349.

Zhuang, Y., Cheng, P., and Weintraub, H. (1996). B-lymphocyte development is regulated by the combined dosage of three basic helix-loop-helix genes, E2A, E2-2, and HEB. Mol Cell Biol *16*, 2898-2905.

Zhuang, Y., Soriano, P., and Weintraub, H. (1994). The helix-loop-helix gene E2A is required for B cell formation. Cell *79*, 875-884.

Zon, L. I. (2008). Intrinsic and extrinsic control of haematopoietic stem-cell self-renewal. *453*, 306-313.

II

E2A Proteins Maintain the Hematopoietic Stem Cell Pool and Promote the Maturation of Myelolymphoid and Myeloerythroid Progenitors

E2A proteins maintain the hematopoietic stem cell pool and promote the maturation of myelolymphoid and myeloerythroid progenitors

Craig L. Semerad[a], Elinore M. Mercer[a], Matthew A. Inlay[b], Irving L. Weissman[b], and Cornelis Murre[a,1]

[a]Division of Biological Sciences, 0377, University of California at San Diego, La Jolla, CA 92093; and [b]Institute of Stem Cell Biology and Regenerative Medicine, Stanford University School of Medicine, Stanford, CA 94305

Edited by Inder M. Verma, The Salk Institute for Biological Studies, La Jolla, CA, and approved December 5, 2008 (received for review September 8, 2008)

Hematopoiesis is a tightly controlled process maintained by a small pool of hematopoietic stem cells (HSCs). Here, we demonstrate that the LT-HSC, MPP, premegakaryocytic/erythroid, Pre CFU-E, Pre GM, MkP, and granulocyte-macrophage compartments were all significantly reduced in E2A-deficient bone marrow. Despite a severe depletion of erythroid progenitors, the erythrocyte and megakaryocyte compartments were equivalent in E2A-deficient bone marrow as compared with wild-type mice. E2A-deficient HSCs also failed to efficiently maintain the HSC pool on serial transplantation, and we demonstrate that the E2A proteins regulate cell cycle progression of HSCs by regulating the expression of $p21^{Cip1}$, $p27^{Kip1}$, and the thrombopoietin receptor, known regulators of HSC self-renewal activity. Based on these observations, we propose that the E2A proteins promote the developmental progression of the entire spectrum of early hematopoietic progenitors and to suppress an erythroid specific program of gene expression in alternative cell lineages. Last, the data mechanistically link E2A, cell cycle regulators, and the maintenance of the HSC pool in a common pathway.

E47 | lymphoid/myeloid versus erythroid/megakaryocyte development | self-renewal

Hematopoiesis is a tightly regulated process maintained by a small pool of hematopoietic stem cells (HSC) uniquely capable of undergoing self-renewal and generating mature progeny of all of the hematopoietic cell lineages. To sustain the proper levels of blood cells, HSCs must continuously monitor and regulate the balance between self-renewal and lineage differentiation. Following the decision to differentiate, hematopoiesis proceeds in a step-wise manner from the primordial long-term (LT)-HSCs. LT-HSCs possess the ability to self-renew and the capacity for long-term reconstitution of lethally irradiated hosts (1). Upon differentiation, LT-HSCs lose their capacity for self-renewal and give rise to a population of short-term (ST)-HSCs. The ST-HSCs, limited to a transient ability to self-renew and reconstitute lethally irradiated hosts, differentiate into a multipotent progenitor (MPP) population. The MPPs lack the capacity to undergo self-renewal, but retain multipotentcy. From the MPP population develops a series of intermediate progenitors that give rise to the assorted hematopoietic lineages. In the classical pathway of hematopoiesis, these intermediates include the common lymphoid progenitors (CLPs) that differentiate into lymphoid, but not myeloid, progeny, and the common myeloid progenitors (CMPs), which retain full erythro-myeloid potential (2, 3). The CMPs further differentiate to form the granulocyte/macrophage progenitors (GMPs) that differentiate to the myelomonocytic lineage and the megakaryocytic/erythrocyte progenitors (MEPs) that eventually differentiate to form red blood cells and platelets.

Other studies have suggested alternative pathways for the differentiation of the megakaryocytic/erythroid versus lymphoid/myeloid cell lineages. Specifically, these studies have indicated that the separation of the myeloid versus megakaryocytic/erythroid cell lineages may occur at an earlier branch point before the development of CMPs. This work has suggested the presence of lymphoid-primed (L) MPPs in the bone marrow, which have the ability to develop into lymphoid and myeloid progeny, but cannot give rise to erythroid and megakaryocytic cells (4). Also, more recent studies have used additional markers, including CD105 and CD41, to separate the myeloerythroid progenitors into Pre GM, GMP, premegakaryocytic/erythroid (Pre MegE), Pre CFU-E, CFU-E, and MkP compartments. These studies suggested a hierarchy, in which the Pre GM population gives rise to the GMP compartment, whereas the Pre MegE progenitors act upstream of the MkP and the Pre CFU-E (5).

Transcriptional regulation is a key mechanism controlling HSC homeostasis, development, and lineage commitment (6). For example, the commitment of hematopoietic progenitors to the B cell lineage and their development to mature B cells depends on combined activities of the transcription factors E2A, EBF, and Pax5 (7, 8). E2A is a member of the E-protein family of basic helix–loop–helix (bHLH) proteins. The E2A gene encodes 2 E proteins, E12 and E47, which are generated by differential splicing of the exon encoding the DNA binding and dimerization domain (9). Along with E12 and E47, the E-proteins include E2–2, HEB, and the *Drosophila* gene product Daughterless. E-proteins have the ability to bind canonical E-box [G(orA)CAXXTGG(orA)] elements as either homodimers or heterodimers with other members of the bHLH family (10, 11). Within the hematopoietic compartment, the E2A proteins form heterodimers with SCL. SCL becomes expressed in mesodermal cells that develop into embryonic blood cells and continues to be expressed in fetal and adult HSCs (12). Although SCL is not required for the maintenance of HSC self-renewal, it is critical for proper erythroid and megakaryocyte development in the adult (13).

Here, we show the E2A proteins are ubiquitously expressed in HSCs and in subsets of hematopoietic progenitor cells, but that their expression levels are dynamic. We show that the E2A proteins act to promote the developmental progression of the entire spectrum of early hematopoietic progenitors. The observations also mechanistically connect the E2A proteins, cell cycle regulators, and the maintenance of the HSC pool in a common pathway.

Results

Reduced Numbers of HSCs in E2A-Deficient Bone Marrow. To examine E2A expression in HSCs and early hematopoietic progenitors,

Author contributions: C.L.S., E.M.M., M.A.I., I.L.W., and C.M. designed research; C.L.S., E.M.M., M.A.I., and C.M. performed research; C.L.S., E.M.M., M.A.I., I.L.W., and C.M. analyzed data; and C.L.S., M.A.I., and C.M. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence should be addressed. E-mail: murre@biomail.ucsd.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0808866106/DCSupplemental.

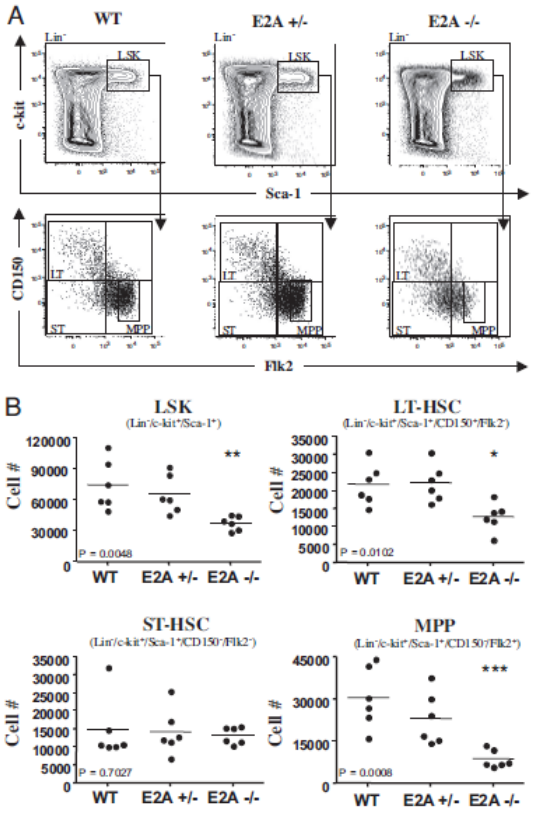


Fig. 1. Effects of E2A deletion on the numbers of HSCs in adult mouse bone marrow (BM). Total BM cells from wild-type, E2A +/−, and E2A −/− mice were harvested and prepared for analysis by flow cytometry. (*A*) Representative staining profiles for LSK cells and the LT-HSC, ST-HSC, and MPP subpopulations. The small gate in the MPP quadrant is representative of the LMPP population. (*B*) Reduced HSC numbers in the BM of E2A −/− mice. Shown are the absolute numbers of the LSK, LT-HSC, ST-HSC, and MPP populations in the BM of wild-type, E2A +/−, and E2A −/− mice ($n = 6$). Horizontal bars show the mean values. Statistical significance determined by unpaired *t* test, 2-tailed, between E2A −/− and wild type.

we used a knock-in mouse mutant, in which the coding sequence for GFP was fused to the C terminus of the E2A gene in-frame through homologous recombination (14). As previously reported, E2A levels are high in the LSK ($Lin^-/c\text{-}kit^+/Sca\text{-}1^+$) population [supporting information (SI) Fig. S1] (14). Within the LSK population, the LMPP compartment exhibited higher levels of E2A expression than the LT- and ST-HSC populations (Fig. S1). Given that E2A is expressed throughout the HSC compartment, we investigated whether loss of E2A affects the numbers of the various HSC populations in the bone marrow. By using multiparameter flow cytometry, we compared the number of HSCs in the bone marrow of wild-type and E2A mutant mice. We analyzed mice <2 months of age, because, at later ages, E2A −/− mice readily develop lymphoma (15).

We observed a significant decrease in the LSK compartment of E2A −/− mice, compared with wild type (2.5-fold) (Fig. 1 *A* and *B*). The LSK compartment can be further separated on the basis of CD150 and Flk2 expression (16–19). The $CD150^+/Flk2^-$ population comprises the LT-HSC compartment, the $CD150^-/Flk2^-$ population the ST-HSC and the $CD150^-Flk2^+$ compartment consists of MPPs. Loss of E2A activity resulted in a 2-fold decrease in the number of LT-HSCs (Fig. 1*B*). The cellularity of MPPs in the bone marrow of E2A −/− mice was reduced 4-fold, compared with wild-type mice (Fig. 1*B*). To determine whether the fraction of LMPPs was affected by the absence of E2A, LSK cells were examined for the expression of Flk2 as described previously (4, 20). Strikingly, the proportion and cellularity of Flk2 expressing cells was severely reduced in E2A −/− bone marrow (Fig. 1*A*). We note that Flk2 expression is not down-regulated at the CLP or Pre GM stage in E2A −/− bone marrow, indicating that Flk2 is not a direct target of E2A. Collectively, these data indicate that E2A acts during early hematopoiesis to maintain the HSC compartment and to promote the development of MPPs as well as LMPPs.

Reduced Numbers of Lymphoid and Myeloid Progenitors in E2A-Deficient Bone Marrow. To determine whether the loss of E2A also affects the numbers of hematopoietic progenitors, we examined bone marrow cells for the presence of lymphoid and myeloerythroid progenitors. Consistent with previous observations, loss of E2A resulted in a significant decrease in the number of CLPs ($Lin^-/IL\text{-}7R\alpha^+/ckit^{int}/Sca\text{-}1^{int}$) (Fig. S2 *A* and *B*). Also, we examined bone marrow cells for the presence of erythroid/megakaryocyte and GMPs by using markers recently described (5). The Pre MegE compartment ($Lin^-/IL\text{-}7R\alpha^-/c\text{-}kit^+/Sca\text{-}1^-/CD105^{lo}/CD150^{hi}/CD41^-$) was substantially decreased (Fig. 2 *A* and *B*). Similarly, the megakaryocyte progenitor (MkP) immediately downstream of the Pre MegE was also reduced (Fig. 2 *A* and *B*). However, the erythroid progenitor immediately downstream of the Pre MegE, the Pre CFU-E, was almost absent. Despite the near absence of the Pre CFU-E population, the CFU-E ($Lin^-/c\text{-}kit^+/Sca\text{-}1^-/IL\text{-}7R\alpha^-/CD105^{hi}/CD150^{lo}/Ter119^-$) compartment immediately downstream was not reduced (Fig. 2 *A* and *B*, and Fig. 3*D*).

The GMPs were examined as described previously, with the addition of Flk2 (5). Although E2A −/− bone marrow showed no overall change in the Pre GM ($Lin^-/c\text{-}kit^+/Sca\text{-}1^-/CD105^-/CD150^-/Fc\gamma R^-$) compartment, the population was skewed toward the $Flk2^-$ side, such that the number of $Flk2^-$ Pre GM was increased and $Flk2^+$ Pre GM decreased, compared with wild type (Fig. 2 *A* and *B*). The GMP compartment was reduced as well (Fig. 2*B*). Also, we performed an in vitro colony-forming unit-granulocyte (CFU-G) assay. In methylcellulose cultures, bone marrow from E2A −/− mice demonstrated a statistically significant decrease in the number of CFU-G (Fig. S3). Collectively, these data show that the E2A proteins perform a wide spectrum of activities during early hematopoiesis to modulate the developmental progression of megakaryocyte/erythroid and myeloid progenitors.

Defective Long-Term Repopulating Ability of E2A-Deficient HSCs. To determine whether loss of E2A affects the functional capabilities of HSCs, we evaluated the repopulating capacity of purified LSK cells by using a competitive reconstitution assay. For this purpose, FACS-purified LSK cells from wild-type ($CD45.1^+$) and E2A −/− ($CD45.2^+$) mice were mixed at a 1:1 ratio (1 WT:1 E2A −/− mixed chimera) and transplanted into lethally irradiated $CD45.1^+$ recipient mice. Also, separate cohorts of $CD45.1^+$ recipient mice were transplanted with LSK cells purified from wild-type, E2A +/−, or E2A −/− ($CD45.2^+$) mice, respectively. FACS-purified LSK cells were used instead of whole bone marrow for this assay to control for differences in the number of bone marrow mononuclear cells and the distribution of lineage positive subsets in the bone marrow of wild-type and E2A −/− mice (Fig. S4 *A* and *B*). Although the overall number of LT-HSCs was reduced in E2A −/− bone marrow, the percentage of LT-HSCs within the LSK gate were the same between mutant and wild type; thus, the number of LT-HSCs transplanted between mutant and wild-type were equivalent. After transplantation, peripheral blood from the transplanted recipients

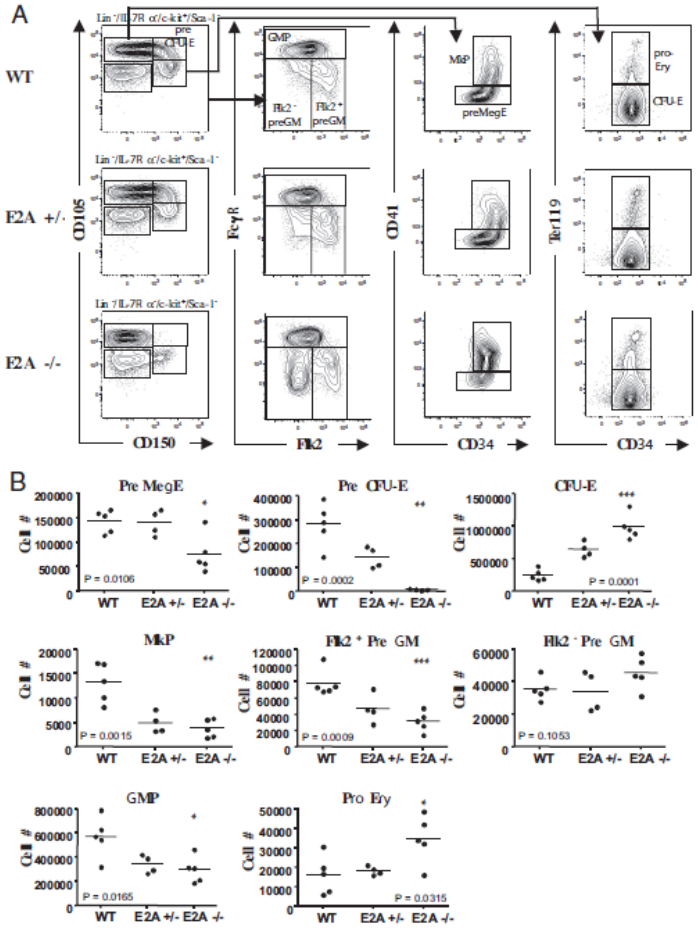


Fig. 2. Effects of E2A deletion on the numbers o myeloerythroid progenitors in adult mouse BM. Tota BM cells from wild-type, E2A +/−, and E2A −/− mic were harvested and prepared for analysis by flow cytometry. (*A*) Representative staining profiles for my eloerythroid progenitors. (*B*) Reduced numbers of my eloerythroid progenitor subsets in the BM of E2A −/− mice. Shown are the absolute numbers of the myelo erythroid progenitor subsets in the BM of wild-type E2A +/−, and E2A −/− mice ($n = 5$). Horizontal bar show the mean values. Statistical significance deter mined by unpaired *t* test, 2-tailed, between E2A −/− and wild type.

was analyzed monthly by flow cytometry to assess the contribution of $CD45.2^+$ versus $CD45.1^+$ to the myeloid and lymphoid cell lineages. As expected, wild-type or E2A +/− LSK cells alone successfully reconstituted the myeloid and lymphoid cell lineages in irradiated recipients, whereas E2A −/− LSK cells, whether alone or in competition with wild-type LSK cells, demonstrated a considerable defect in their ability to contribute to the lymphoid cell lineage (Fig. S5*A* and Fig. S6*A*). Conversely, E2A −/− LSK cells made a significant contribution to reconstitution of the myeloid compartment. At 4 and 8 weeks posttransplantation, the myeloid lineage repopulating ability of E2A −/− LSK cells was near predicted levels in the 1 WT: 1 E2A −/− mixed chimeras (Fig. S5*A*). However, analysis of the myeloid lineage in the 1 WT:1 E2A −/− mixed chimeras at later time points showed a significant decrease in the long-term repopulating ability of E2A −/− LSK cells (Fig. S5*A*).

Because of the decrease in GMPs in the E2A −/− bone marrow, the granulocyte output of E2A −/− LSKs may be reduced; thus, skewing an accurate assessment of LSK chimerism in a competitive setting. Therefore, we examined the LSK compartment in the bone marrow of the transplanted recipient mice to directly measure LSK chimerism and self-renewal. At 24 weeks posttransplantation, mice transplanted with E2A −/− LSK cells alone demonstrated a significant reduction in the number of LSK cells in the bone marrow, compared with mice transplanted with only wild-type o E2A +/− LSK cells (Fig. S5 *B* and *C*). Also, in the 1 WT:1 E2A −/− mixed chimeras, the contribution of E2A −/− derived ($CD45.2^+$ cells was 25 ± 4% of a theoretically possible 50% of cells in the LSK compartment (Fig. S5*C*). These data indicate that the E2A protein have a critical role in the maintenance of the LSK compartment.

To examine the serial reconstituting ability of E2A −/− LSK cell we performed secondary bone marrow transplants. Within 2 weeks, 6 of 8 secondary recipients transplanted with bone marrow from primary E2A −/− LSK cell transplants died (Fig. S5*D*). Importantly, none of the transplant recipients receiving E2A −/− bone marrow cells demonstrated any evidence of developing thymic lymphoma during this study. In contrast, all secondary recipients transplanted with bone marrow from primary wild-type LSK cell transplants survived the 24 weeks (Fig. S5*D*). Similar to the primary LSK cell transplants, analysis of peripheral blood indicated wild-type and E2A +/− bone marrow cells efficiently re populated the lymphoid and myeloid lineages of the secondary transplants, whereas E2A −/− bone marrow cells made minimal contribution to the myeloid and lymphoid compartment

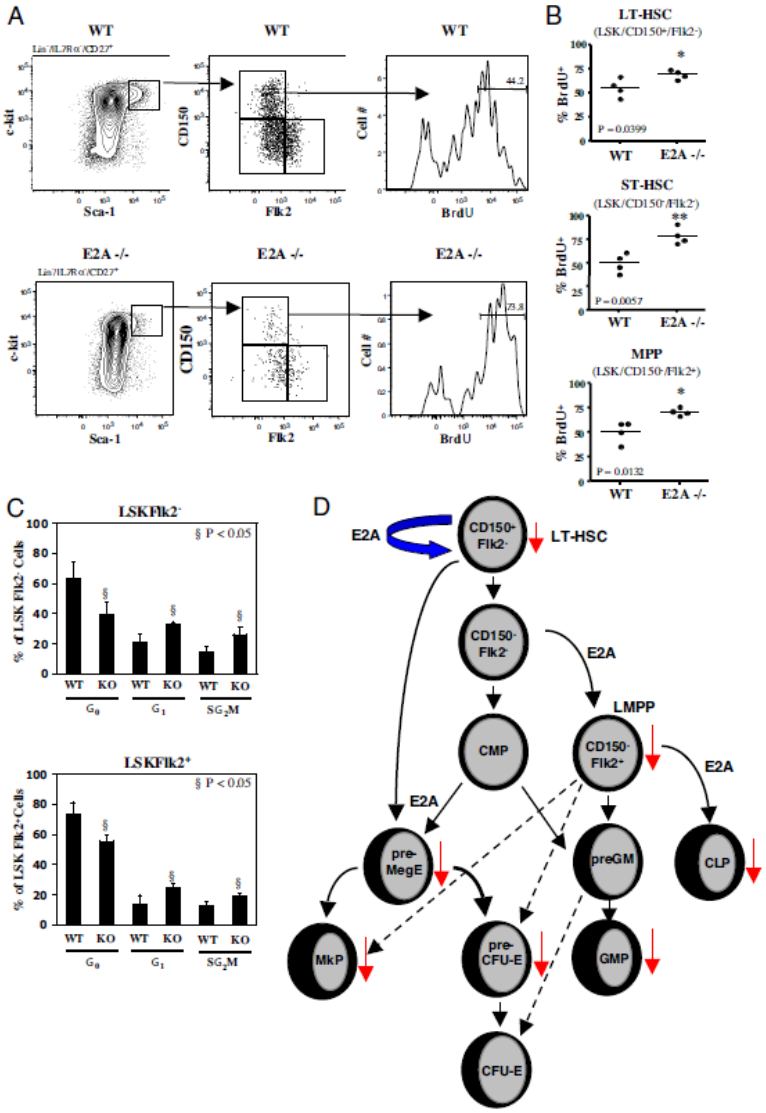


Fig. 3. Increased cycling by E2A-deficient HSCs. To investigate cycling of E2A −/− HSCs, BrdU was administered to 6 week-old wild-type and E2A −/− mice and BrdU incorporation by the LT-HSC, ST-HSC, and MPP cell fractions was analyzed by flow cytometry. (*A*) Representative staining profiles for the LT-HSC (LSK/CD150$^+$/Flk2$^-$) fraction. (*B*) Increased incorporation of BrdU by E2A −/− HSCs. Shown are the percentages of BrdU incorporation by the LT-HSC (LSK/CD150$^+$/Flk2$^-$), ST-HSC (LSK/CD150$^-$/Flk2$^-$), and MPP (LSK/CD150$^-$/Flk2$^+$) fractions in the bone marrow of wild-type and E2A −/− mice ($n = 4$). Statistical significance determined by unpaired *t* test, 2-tailed, between E2A −/− and wild type. (*C*) Cell cycle distribution in E2A −/− HSCs. Shown are the percentages of LSKFlk2$^+$ and LSKFlk2$^-$ fractions in G_0, G_1, and SG_2M in the BM of wild-type and E2A −/− mice ($n = 3$). Statistical significance determined by 2-sided Student's *t* test, E2A −/−, compared with wild-type. (*D*) Schematic representation depicting the roles of E2A in early hematopoiesis. The blue arrow indicates the importance of E2A proteins in HSC self-renewal. Decreases in the indicated hematopoietic populations detected in E2A −/− mice are shown by solid red down arrows. In the absence of E2A, significant decreases in LMPPs, CLPs, and GMPs are detected. Also, the Pre MegE, Pre CFU-E, and MkP compartments are significantly decreased. However, E2A −/− mice have near wild-type levels of Pre GMs and CFU-Es. We propose that the E2A proteins act to promote the development of the LMPPs and the Pre MegE progenitors, and to suppress the development of the Pre GM stage into CFU-Es. Data represent the mean ± SD. Statistical significance determined by unpaired *t* test, 2-tailed, between E2A −/− and wild type.

(Fig. S5*E* and Fig. S6*B*). Also, analysis of the LSK compartment in the bone marrow of the 2 remaining mice transplanted with only E2A −/− bone marrow (70 ± 29% primary to 17 ± 4% secondary), and the mice that received bone marrow from the 1 WT:1 E2A −/− mixed chimeras (25 ± 4% primary to 13 ± 4% secondary) demonstrated a significant reduction in E2A −/− contribution to the LSK population (Fig. 5*F*). To determine whether the absence of E2A affects the number of hematopoietic progenitors in the bone marrow on serial transplantation, the bone marrow of the primary and secondary transplant recipients was evaluated for contribution of E2A −/− cells to the lineage-committed hematopoietic progenitor subsets. Similar to the HSC compartment, the contribution from E2A −/− bone marrow to the hematopoietic progenitor compartment was significantly reduced (Fig. S7 *A–D*). Collectively, these data support a major role for the E2A proteins in the self-renewal activity of HSCs.

The E2A Proteins Modulate the Expression of Cell Cycle Regulators to Control Cell Cycle Progression of the HSC Compartment. To address the mechanism by which E2A maintains the HSC pool, we examined E2A −/− hematopoietic progenitors for their ability to migrate to the bone marrow after transplantation. However, the reconstitution defect observed in E2A −/− cells is not due to the inability of E2A −/− HSCs to home to the hematopoietic niches of the bone marrow, because homing of E2A −/− HSC subsets to the bone marrow of lethally irradiated recipients is comparable with wild-type HSCs (Fig. S8). A

number of studies have indicated that the maintenance of HSC self-renewal activity is mediated in part by cell cycle regulators (21–24). E2A has been shown to modulate cell cycle progression by regulating the expression of cell cycle regulators, raising the possibility that E2A maintains HSC self-renewal activity by modulating cell cycle entry (15, 25). As a first approach to this question, we examined the cell cycle state of HSCs in wild-type and E2A −/− bone marrow. Wild-type and E2A −/− mice were labeled with BrdU, bone marrow was isolated and immunostained for BrdU, and the LT-HSC, ST-HSC, and MPP subsets analyzed by flow cytometry. Interestingly, the proportion of cycling HSCs was significantly increased in E2A −/− bone marrow (Fig. 3*A* and *B*). In agreement with the increased uptake of BrdU by E2A −/− HSCs, we show, by using Ki-67 and DAPI staining, that E2A −/− mice have fewer LSKFlk2$^-$ HSCs in G_0 (Fig. 3*C* and Fig. S9). Also, the LSKFlk2$^+$ MPP subset displayed a similar pattern of enhanced cell cycling (Fig. 3*C*). Based on these observations, we propose that the E2A proteins act in murine bone marrow by regulating the rate and capability of cell cycle progression in the HSC compartment.

To determine how the E2A proteins act to modulate cell cycle progression and self-renewal in LSK cells, we examined the expression of a number of known E2A targets and/or regulators of cell cycle progression. To accomplish this objective, we purified LSK, LSKFlk2$^-$, and LSKFlk2$^+$ cells from the bone marrow of wild-type and E2A −/− mice and analyzed the expression of $p18^{INK4C}$, $p19^{INK4D}$, $p21^{Cip1}$, $p27^{Kip1}$, Cdk6, Bmi-1, Gfi-1, HoxB4, Mpl, and Notch-1 by quantitative PCR (Fig. S10). Interestingly, $p21^{Cip1}$ and Gfi-1 levels were decreased significantly in LSK cells derived from E2A −/− bone marrow and have been demonstrated to modulate the self-renewal activity of HSCs (Fig. S10 *A* and *B*) (21–24). Also, previous studies have demonstrated that the E2A proteins directly regulate $p21^{Cip1}$ and Gfi-1 expression (15, 26). In addition to their significantly decreased expression in LSK cells, $p21^{Cip1}$, Gfi-1, and Notch1 demonstrated decreased expression in purified E2A −/− LSKFlk2$^-$ and LSKFlk2$^+$ subsets, compared with wild type (Fig. S10*C*). Together, these data suggest that the E2A proteins mechanistically regulate HSC self-renewal by modulating the expression of genes involved in HSC cell proliferation.

Discussion

Previous data indicate that the initial stages of B- and T-lineage development require the activities of E-proteins (27–29). Also, recent studies have revealed a role for the E-proteins in the CLP compartment and it has been suggested that E2A becomes transcriptionally active in the CLP cell stage (30). Once E2A is activated, it acts in concert with PU.1 and IL7Rα-mediated signaling to induce the expression of EBF transcription, which, in turn, activates Pax5 gene expression (31). In addition to the activation of a B-lineage specific program of gene expression, E2A also acts to repress the expression of genes involved in the commitment of alternative cell fates, including GATA-1 and GATA-3 (32). Thus, the E2A proteins initiate and maintain a B-lineage specific program of gene expression and repress transcription of non-B lineage specific genes. However, although these studies revealed a role for the E2A proteins in B cell specification, it has remained unclear whether and how the E2A proteins act in HSCs and in hematopoietic progenitors.

Roles of Id1 and E2A in Early Hematopoiesis. Recent studies demonstrated that Id1 acts to constrain myeloid commitment. Specifically, it was demonstrated that Id1-ablated LSK cells showed a premature induction of the myeloid differentiation program (33). These data are consistent with our observations indicating a defect in myeloid maturation in E2A −/− hematopoietic progenitors. Thus, high levels of E2A would favor commitment toward the myeloid cell lineage, whereas inactivation of E2A DNA binding would suppress the development toward the myeloid/lymphoid lineages. Surprisingly, however, are the observations that both E2A and Id1 deficiencies result in higher levels of GATA-1 expression within the LSK compartment (data not shown) (33). Similarly, the expression of Gfi-1, Bmi-1, and Hoxb4 were slightly decreased in both E2A- and Id1-deficient bone marrow (33). How can we reconcile these observations? We propose that the dosage of E- and Id-proteins is carefully calibrated in developing hematopoietic progenitors. The loss of one component of the E/Id ensemble of proteins may dramatically affect the equilibrium of interactions. For example, it is conceivable that in the absence of Id1, HEB and/or E2–2 might be released from their inhibitor to complex with E2A, converting its ability from a transcriptional repressor into an activator. Such a scenario is not that unlikely, because the E2A proteins can act to both activate and repress transcription, depending on the recruitment of a coactivator or corepressor (34). Such a model may also explain the strong dosage effects that we observed during early hematopoiesis in E2A +/− mice.

Role of E2A in Specification of the Granulocyte/Macrophage and Lymphoid Lineages. Striking abnormalities in E2A −/− bone marrow were observed within the Pre GM progenitors, which were further separated on the basis of Flk2 expression. Pre GM progenitors in the E2A−/− bone marrow were heavily skewed toward Flk2 negativity, such that the absolute number of Flk2$^-$ cells was increased in these mice, compared with wild-type, and the Flk2$^+$ were decreased. These data bring into question how the E2A proteins act to promote developmental progression of both lymphoid and myeloid progenitors from a common MPP or LMPP precursor. The decrease in the number of intermediates like the CLP and the Flk2$^+$ Pre GM populations may simply reflect a decrease in the number of CD150$^-$Flk2$^+$ progenitors. Thus, the E2A proteins may act at the CD150$^-$Flk2$^+$ cell stage to induce, at low levels, the expression of a myeloid and common lymphoid, but not B and T cell specific, program of gene expression. In MPPs the E2A proteins act to induce the expression of common lymphoid genes and to suppress the expression of genes, such as GATA-1, that are involved in erythroid differentiation (32). The priming of a program of myeloid specific gene expression is likely to be induced by PU.1, which has been demonstrated to act at the branch point separating the MegE and lymphoid/myeloid cell lineages (35). Thus, based on these previous observations and the data described here, we suggest that E2A and PU.1 both act in the CD150$^-$Flk2$^+$ compartment to induce the priming of a common myeloid and lymphoid program of gene expression.

Role of the E2A in Erythroid and Megakaryocytic Development. It has been previously proposed that the LSK CD150$^-$Flk2^{lo} compartment may directly feed into the Pre MegE population (5). The Pre MegE, in turn, feeds into the CFU-E via the Pre CFU-E. However, the Pre CFU-E is almost absent in E2A −/− bone marrow, but the CFU-E population is present at elevated levels. How are erythrocytes generated in E2A −/− bone marrow in the absence of erythroid and megakaryocytic progenitors? It is conceivable that the Pre GM cells are contributing to the CFU-E. In such a scenario, the E2A −/− Flk2$^-$ Pre GM would function as a CMP. In the absence of E2A, how does the Pre GM differentiate into the erythroid lineage? We have previously shown that the E2A proteins in MPPs act to suppress the transcription of GATA-1 (32). GATA-1 is a critical transcriptional regulator for erythroid/megakaryocyte development, because GATA-1-ablated mice lack the erythroid/megakaryocytic lineages (36, 37). Also, GATA-1 exerts instructive signals for erythroid/megakaryocytic lineage commitment and has the ability to antagonize the transactivation capacity of PU.1 to promote erythroid differentiation at the expense of myeloid maturation (38–41). Thus, although the induction of GATA-1 expression in the absence of E2A is merely 2-fold, it is conceivable that an increase of such abundance in conjunction with the induction of additional regulators yet to be identified would promote the development of CFU-E cells through a pathway independent of the Pre MegE and Pre CFU-E intermediates. Little if any GATA-1 expression is present in wild-type Pre GM cells, compared with Pre MegE and Pre CFU-E (5). Therefore, we propose that E2A suppresses GATA-1 expression in the Pre GM, preventing this

population from differentiating down the erythroid lineage. We also note that, within the $CD150^{+}CD105^{-}$ gate, which contains both Pre MegE and MkP progenitors, a higher percentage within this gate were MkP progenitors in E2A −/− mice. This suggests that the Pre MegE population in the E2A −/− mice is skewed toward megakaryocyte development, possibly again mediated by increased levels of GATA-1 expression.

The role of the E2A Proteins in Maintaining the HSC Pool. E2A −/− bone marrow also shows a diminished steady-state abundance of LSKs and a defect in the ability to self-renew in serial transplantation assays. How do the E2A proteins act to maintain the HSC compartment? Our data show that the absence of E2A activity results in a substantial increase in the rate of cell cycle progression of the HSC compartment. Previous observations have indicated that the maintenance of the stem cell pool is controlled, among others, by the thrombopoietin receptor (Mpl), Gfi-1, $p21^{Cip1}$, and $p27^{Kip1}$ (21–24). We demonstrate here that the expression of Mpl, Gfi-1, $p21^{Cip1}$, and $p27^{Kip1}$ is significantly decreased in E2A −/− LSK cells (Fig. S10). It is unlikely that the reduction in p27 levels is responsible for the increase in cycling HSCs, because p27 has been demonstrated not to modulate HSC cell cycling or self-renewal, but rather controls the expansion and pool size of progenitors (21). In contrast, $p21^{Cip1}$ and Mpl have been shown to have a critical role in maintaining the HSC pool. Thus, we propose that the E2A proteins mechanistically enforce the maintenance of HSCs by directly modulating the expression of Mpl and $p21^{Cip1}$.

Materials and Methods

Competitive Repopulation Assay. Lethally γ-irradiated (single dose of 1,000 cGy) CD45.1 congenic C57/B6 mice at 8–10 weeks of age were used as recipients. We injected 1,000 sorted LSK cells from wild-type ($CD45.2^{+}$), E2A +/− ($CD45.2^{+}$), E2A −/− ($CD45.2^{+}$), or a 1 to 1 ratio of wild-type ($CD45.1^{+}$) to E2A −/− ($CD45.2^{+}$) i.v. into each lethally irradiated recipient mouse. To establish secondary transplants, single-cell suspensions from bone marrow of primary LSK cell transplant recipients were prepared. Bone marrow cells were resuspended in α-MEM containing 10% FBS. Lethally irradiated $CD45.1^{+}$ secondary recipients were injected intravenously with a total of 2×10^6 cells.

Quantitative Real-Time PCR. Total RNA was extracted from sorted LSK cells by using an Rneasy kit (Qiagen) according to manufacturer's protocol. For real-time PCR, cDNA synthesis was performed by using SuperScript III (Invitrogen). Real-time PCR was performed by using SYBR Green Master Mix (Stratagene) and analyzed by Mx3005P instrumentation (Stratagene). The reactions were performed in triplicate at 95 °C for 10 min, followed by 40 cycles of 95 °C for 30 s, 55 °C for 1 min, and 72 °C for 30 s. The primer sequences used are available on request.

Cell Cycle Assays. To examine BrdU incorporation, mice were given a single i.p. injection of BrdU (Sigma) (1 mg/6 g) in PBS and maintained on 1 mg/mL of BrdU in the drinking water for 72 h. After 72 h, the bone marrow was harvested and stained with antibodies against lineage markers, c-kit, Sca-1, CD150, and Flk2. BrdU incorporation was measured by using a FITC BrdU Flow Kit (BD PharMingen).

To analyze the cell cycle status of the LSK, $LSKFlk2^{-}$, and $LSKFlk^{2}$ HSC subsets, bone marrow cells from wild-type and E2A −/− mice were initially stained with antibodies against Lin^{+} cells, c-kit, Sca-1, and Flk2. After incubation with the cell surface antibodies, the cells underwent fixation with a Cytofix/Cytoperm kit (BD Biosciences). After fixation, the cells were incubated with FITC-anti-Ki-67, washed, and stained with DAPI. Analysis was performed on a FACS LSRII (BD Biosciences).

For more details, see *SI Materials and Methods*.

ACKNOWLEDGMENTS. We thank Yuan Zhuang (Duke University Medical Center, Durham, NC) for providing E2A/GFP mice, and members of the C.M. and I.L.W. laboratories for reagents and stimulating discussions. C.M. and I.L.W. were supported in part by grants from the National Institutes of Health and the California Institute for Regenerative Medicine (CIRM); C.L.S. by a Hemoglobin and Blood Training grant and by CIRM; E.M.M. by an Endocrine Training grant; and M.A.I. by CIRM Fellowship T1-00001.

1. Spangrude GJ, Heimfeld S, Weissman IL (1988) Purification and characterization of mouse hematopoietic stem cells. *Science* 241:58–62.
2. Akashi K, Traver D, Miyamoto T, Weissman IL (2000) A clonogenic common myeloid progenitor that gives rise to all myeloid lineages. *Nature* 404:193–197.
3. Kondo M, Weissman IL, Akashi K (1997) Identification of clonogenic common lymphoid progenitors in mouse bone marrow. *Cell* 91:661–672.
4. Adolfsson J, *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.
5. Pronk CJ, *et al.* (2007) Elucidation of the phenotypic, functional, and molecular topography of a myeloerythroid progenitor cell hierarchy. *Cell Stem Cell* 1:428–442.
6. Pimanda JE, *et al.* (2007) Gata2, Fli1, and Scl form a recursively wired gene-regulatory circuit during early hematopoietic development. *Proc Natl Acad Sci USA* 104:17692–17697.
7. Busslinger M (2004) Transcriptional control of early B cell development. *Annu Rev Immunol* 22:55–79.
8. Singh H, Medina KL, Pongubala JM (2005) Contingent gene regulatory networks and B cell fate specification. *Proc Natl Acad Sci USA* 102:4949–4953.
9. Murre C, McCaw PS, Baltimore D (1989) A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, MyoD, and myc proteins. *Cell* 56:777–783.
10. Lazorchak A, Jones ME, Zhuang Y (2005) New insights into E-protein function in lymphocyte development. *Trends Immunol* 26:334–338.
11. Sun XH (2004) Multitasking of helix-loop-helix proteins in lymphopoiesis. *Adv Immunol* 84:43–77.
12. Shivdasani RA, Mayer EL, Orkin SH (1995) Absence of blood formation in mice lacking the T-cell leukaemia oncoprotein tal-1/SCL. *Nature* 373:432–434.
13. Mikkola HK, *et al.* (2003) Haematopoietic stem cells retain long-term repopulating activity and multipotency in the absence of stem-cell leukaemia SCL/tal-1 gene. *Nature* 421:547–551.
14. Zhuang Y, Jackson A, Pan L, Shen K, Dai M (2004) Regulation of E2A gene expression in B-lymphocyte development. *Mol Immunol* 40:1165–1177.
15. Schwartz R, Engel I, Fallahi-Sichani M, Petrie HT, Murre C (2006) Gene expression patterns define novel roles for E47 in cell cycle progression, cytokine-mediated signaling, and T lineage development. *Proc Natl Acad Sci USA* 103:9976–9981.
16. Adolfsson J, *et al.* (2001) Upregulation of Flt3 expression within the bone marrow Lin(−)Sca1(+)c-kit(+) stem cell compartment is accompanied by loss of self-renewal capacity. *Immunity* 15:659–669.
17. Chen CZ, *et al.* (2002) Identification of endoglin as a functional marker that defines long-term repopulating hematopoietic stem cells. *Proc Natl Acad Sci USA* 99:15468–15473.
18. Christensen JL, Weissman IL (2001) Flk-2 is a marker in hematopoietic stem cell differentiation: A simple method to isolate long-term stem cells. *Proc Natl Acad Sci USA* 98:14541–14546.
19. Yang L, *et al.* (2005) Identification of Lin(-)Sca1(+)kit(+)CD34(+)Flt3- short-term hematopoietic stem cells capable of rapidly reconstituting and rescuing myeloablated transplant recipients. *Blood* 105:2717–2723.
20. Yoshida T, Ng SY, Zuniga-Pflucker JC, Georgopoulos K (2006) Early hematopoietic lineage restrictions directed by Ikaros. *Nat Immunol* 7:382–391.
21. Cheng T, Rodrigues N, Dombkowski D, Stier S, Scadden DT (2000) Stem cell repopulation efficiency but not pool size is governed by p27(kip1). *Nat Med* 6:1235–1240.
22. Cheng T, *et al.* (2000) Hematopoietic stem cell quiescence maintained by p21cip1/waf1. *Science* 287:1804–1808.
23. Hock H, *et al.* (2004) Gfi-1 restricts proliferation and preserves functional integrity of haematopoietic stem cells. *Nature* 431:1002–1007.
24. Walkley CR, Fero ML, Chien WM, Purton LE, McArthur GA (2005) Negative cell-cycle regulators cooperatively control self-renewal and differentiation of haematopoietic stem cells. *Nat Cell Biol* 7:172–178.
25. Engel I, Murre C (2004) E2A proteins enforce a proliferation checkpoint in developing thymocytes. *EMBO J* 23:202–211.
26. Prabhu S, Ignatova A, Park ST, Sun XH (1997) Regulation of the expression of cyclin-dependent kinase inhibitor p21 by E2A and Id proteins. *Mol Cell Biol* 17:5888–5896.
27. Bain G, *et al.* (1997) Both E12 and E47 allow commitment to the B cell lineage. *Immunity* 6:145–154.
28. Barndt RJ, Dai M, Zhuang Y (2000) Functions of E2A-HEB heterodimers in T-cell development revealed by a dominant negative mutation of HEB. *Mol Cell Biol* 20:6677–6685.
29. Heemskerk MH, *et al.* (1997) Inhibition of T cell and promotion of natural killer cell development by the dominant negative helix loop helix factor Id3. *J Exp Med* 186:1597–1602.
30. Borghesi L, *et al.* (2005) E47 is required for V(D)J recombinase activity in common lymphoid progenitors. *J Exp Med* 202:1669–1677.
31. Pongubala JM, *et al.* (2008) Transcription factor EBF restricts alternative lineage options and promotes B cell fate commitment independently of Pax5. *Nat Immunol* 9:203–215.
32. Ikawa T, Kawamoto H, Wright LY, Murre C (2004) Long-term cultured E2A-deficient hematopoietic progenitor cells are pluripotent. *Immunity* 20:349–360.
33. Jankovic V, *et al.* (2007) Id1 restrains myeloid commitment, maintaining the self-renewal capacity of hematopoietic stem cells. *Proc Natl Acad Sci USA* 104:1260–1265.
34. Murre C (2005) Helix-loop-helix proteins and lymphocyte development. *Nat Immunol* 6:1079–1086.
35. Scott EW, Simon MC, Anastasi J, Singh H (1994) Requirement of transcription factor PU. 1 in the development of multiple hematopoietic lineages. *Science* 265:1573–1577.
36. Fujiwara Y, Browne CP, Cunniff K, Goff SC, Orkin SH (1996) Arrested development of embryonic red cell precursors in mouse embryos lacking transcription factor GATA-1. *Proc Natl Acad Sci USA* 93:12355–12358.
37. Shivdasani RA, Fujiwara Y, McDevitt MA, Orkin SH (1997) A lineage-selective knockout establishes the critical role of transcription factor GATA-1 in megakaryocyte growth and platelet development. *EMBO J* 16:3965–3973.
38. Iwasaki H, *et al.* (2003) GATA-1 converts lymphoid and myelomonocytic progenitors into the megakaryocyte/erythrocyte lineages. *Immunity* 19:451–462.
39. Kulessa H, Frampton J, Graf T (1995) GATA-1 reprograms avian myelomonocytic cell lines into eosinophils, thromboblasts, and erythroblasts. *Genes Dev* 9:1250–1262.
40. Nerlov C, Querfurth E, Kulessa H, Graf T (2000) GATA-1 interacts with the myeloid PU. 1 transcription factor and represses PU. 1-dependent transcription. *Blood* 95:2543–2551.
41. Rekhtman N, Radparvar F, Evans T, Skoultchi AI (1999) Direct interaction of hematopoietic transcription factors PU. 1 and GATA-1: Functional antagonism in erythroid cells. *Genes Dev* 13:1398–1411.

This chapter, in full, is a reprint of the material as it appears in PNAS (2009) Craig L. Semerad, Elinore M. Mercer, Matthew A. Inlay, Irving L. Weissman, and Cornelis Murre. 2009 Feb 10;106(6):1930-5. Dr. Craig L. Semerad is the primary author of this work.

III

Creation of Multipotent Cell Line with Self-Renewal Capability

Summary

Although success rates of bone marrow transplants have improved radically over the last 30 years, the survival rates after transplantation still hover between thirty and fifty percent. The main causes for transplant failure are a lack of engraftment and/or a lack of long-term reconstitution of the recipient. Both of these problems have been linked to the limited number of early progenitor cells available for injection. As such, there is a great deal of research in the field of ex vivo expansion of hematopoietic stem cells. Here, we overexpress the transcriptional inhibitor, Id2, in a tetracycline-responsive expression cassette to expand multipotent progenitor cells ex vivo. These cells, Id2-HPCs, have the ability both to reconstitute irradiated recipients in a competitive transplantation setting and to be differentiated into multiple lineages in vitro, allowing for in depth genetic studies of lineage choice.

Introduction

Over 50,000 bone marrow transplants (BMT) are performed annually to treat diseases such as lymphomas, leukemias, auto-immunities, and lympho-suppressive syndromes. Although the success of BMTs has increased dramatically in the past decade, long-term survival rates after BMT are only between thirty and fifty percent. This is due to a myriad of factors including failure of the initial engraftment of the transplant, rejection of the graft or the host, and a failure of long-term reconstitution of the host (Gratwohl et al., 2010). The success of a bone marrow transplant is directly related to the number of hematopoietic stem cells and early progenitor cells that are injected, and as such, there is a great deal of research into methods to expand hematopoietic stem cells ex vivo (Kelly et al., 2009). Although there has

been significant progress in the field of ex vivo expansion, there are still limitations to nearly every method being tested (Sauvageau et al., 2004).

Hematopoiesis proceeds in an ordered, developmental progression in adults beginning with the long-term hematopoietic stem cell (LT-HSC). LT-HSCs have the ability to self-renew for the life of the organism as well as to maintain pluripotent differentiation potential (Spangrude et al., 1988). LT-HSCs differentiate into short-term hematopoietic stem cells (ST-HSCs) and multipotent progenitors (MPPs), which still have pluripotent potential but have a marked reduction or loss of self-renewal activity, respectively. MPPs then differentiate into more lineage restricted progenitor cells, namely the megakaryocyte/erythroid restricted pre-megakaryocyte/erythrocyte (preMegE) or the lymphoid primed multipotent progenitor (LMPP) (Adolfsson et al., 2005; Murre, 2009). These cells then proceed along different pathways to form the all the cells of the mature immune system.

E2A, one of four E-proteins in mammals, plays a critical role at multiple stages of hematopoietic differentiation. E2A is a basic helix-loop-helix transcription factor which homo- or herterodimerizes with other E-proteins to bind the canonincal CANNTG E-box sequence and regulate transcription. The E-proteins are inhibited by the helix-loop-helix Id proteins, which heterodimerize with E-proteins to abolish their DNA binding ability ((Benezra et al., 1990; Massari and Murre, 2000; Murre et al., 1989). The role of E2A is best-characterized during B-cell development, as E2A knockout animals display a severe block at the BLP stage of development (Bain et al., 1994; Inlay et al., 2009; Zhuang et al., 1994). B cell development proceeds from the LMPP to the common lymphoid progenitor (CLP), and then into a pre-pro-B cell. Pre-pro-B cells express intermediate levels of B220 and are Ly6D negative, and upon upregulation of Ly6D the cells commit to the B-cell lineage as a BLP cell (Inlay et al., 2009; Mansson et al., 2010). These cells then differentiate into CD19 positive

pro-B cells, which can give rise to CD25 positive pre-B cells. Pre-B cells differentiate to the more mature IgM positive, and eventually IgD positive circulating B cells (Morita et al., 2010; Murre, 2009).

One of the hallmarks of B-cell development is the rearrangement of the IgH, IgL and IgK chains of the B-cell receptor. Rearrangement of the D-J joint of the IgH gene has been reported as early as the CLP and BLP, and V-DJ rearrangement is completed by the pro-B cell stage of development. IgL and IgK rearrangement occurs in preB cells (Borghesi et al., 2004; Chowdhury and Sen, 2004; Schatz et al., 1992).

When E2A -/- lineage negative cells are cultured in B-cell supportive conditions in vitro, the cells differentiate to the pre-pro-B cell stage of development, at which point further development is blocked, and then self-renew indefinitely. These self-renewing pre-pro-B cells have completed DJ rearrangement at the IgH locus, but have not recombined the V-DJ joint. Upon proper cytokine stimulation in vitro, or upon injection into irradiated recipients, these E2A -/- cells display multipotency in that they can differentiate into myeloid, erythroid, NK, T, and dendritic cell lineages (Ikawa et al., 2004). EBF -/- lineage negative cells have similar potential, yet the block in E2A -/- cells can be rescued by forced EBF expression, suggesting that E2A is upstream of EBF in B-cell development (Hagman and Lukin, 2006; Lazorchak et al., 2005; Seet et al., 2004). Evidence shows that E2A and EBF are both upstream of the B-cell commitment factor Pax5, and the Foxo1 and Bcl11a gene products are thought to be downstream of the E2A-EBF-Pax5 regulation of the B cell lineage (Decker et al., 2009; Dengler et al., 2008; Liu et al., 2003; Roessler et al., 2007).

Since E2A -/- cells have the unique ability to self-renew and maintain multipotency, we asked if inhibition of E2A in wild-type cells would be sufficient for ex vivo expansion of multipotent progenitor cells. Here, we show that overexpression of the E-protein inhibitor Id2

from a TetOff promoter in bone marrow progenitor cells allows for the expansion of a progenitor cell line that maintains multipotency both in vivo and in vitro. We then harnessed the unlimited growth capacity and broad differentiation potential of these cells to study the genetic changes involved in the lineage choice between B- and myeloid cell commitment.

Results

TetOff_hId2 infected cells proliferate in vitro

Previous experiments show that E2A -/- progenitor cells can self-renew in vitro and retain pluripotency in vivo (Ikawa et al., 2004). To determine if overexpression of Id proteins would effectively inhibit E2A and thus allow in vitro expansion of multipotent progenitor cells, we created lentiviral vectors utilizing the tetracycline transactivator tTA-2S which is constitutively expressed by the hPGK promoter. The Tet07 promoter drives expression of the MCS-IRES-GFP cassette, into which we cloned human Id2, allowing for overexpression of hId2 in the absence of doxycycline (Figure 3.1A).

Lineage depleted bone marrow from CD45.1 congenic C57/Bl6 mice was infected with either TetOff_Null or TetOff_hId2 lentivirus, and the cells were expanded in the presence of IL-3, IL-6, WEHI and SCF. Initially, both the controls and hId2 infected cells rapidly expanded, but by 4-5 weeks, the Control-HPCs (C-HPCs) had ceased growing while the Id2-HPCs continued to divide. The Id2-HPCs were then cultured in IL-7, Flt3L, and SCF on sub-confluent S17 feeder cells. In these conditions, the cells continued to expand over time. The Id2-HPCs overexpress hId2 mRNA, and this hId2 expression can be robustly turned off in the presence of doxycycline (Figure 3.1B).

Id2-HPCs do not express markers for T cell and NK cell lineage, and express intermediate levels of B220, high levels of CD43, and are CD25, Ly6D, IgM and IgD negative (Figure 3.2A). Id2-HPCs have a successfully rearranged the IgH DJ joint, but have failed to undergo IgH V-DJ, Igλ or Igκ rearrangement (Figure 3.2B), suggesting that these cells represent early pro-B cells (Busslinger, 2004). Interestingly, a small percentage of the cells (1-5%) differentiate to express CD11b, and another 1%-5% downregulate Id2 and differentiate to CD19 and/or CD25 positive cells during normal culture conditions. When cultures are depleted of CD11b, CD19, and CD25, the phenotypes re-appear in 1-2 days, and are maintained at a low percentage.

To further characterize the Id2-HPCs, we performed microarray analysis using purified RNA from freshly isolated LT-HSCs, ST-HSCs, CLPs, LMPPs, pre-B cells, pro-B cells, and cultured E2A -/-, EBF -/-, and Id2-HPCs. Hierarchical clustering analysis of genes with greater than 2-fold changes in expression was performed using d-Chip software. Eight expression patterns were identified from the 3,330 genes with changes in expression: (I) genes upregulated in pre-B cells, (II) genes upregulated in pro-B cells, (III) genes upregulated in EBFKO cells, (IV) genes upregulated in E2AKO cells, (V) genes upregulated in Id2-HPCs, (VI) genes upregulated in all cell lines, (VII) genes upregulated in pluripotent progenitors, (VIII) genes upregulated in CLPs (Figure 3.3A and Table 3.1).

Gene ontology was performed on the 8 clusters, revealing that many of the genes upregulated in Id2-HPCs, E2A -/-, and EBF -/- cells (cluster 6) are involved in biopolymer metabolic processes, including lipid biosynthesis and mitochondrial maintenance, suggesting that the highest similarities between the three cell lines are genes necessary for regular division and growth in vitro. Of the 59 genes are upregulated in the Id2-HPCs (cluster V)

approximately 11% of these represent genes involved in extracellular contact and processes, such as Cdh2, Gjal and Spp1. Genes required for B-cell development such as Foxo1, Pou2af1, IgL, Ly6D, and Vpreb1 have equivalent expression in the Id2-HPCs relative to stem cells and LMPPs, but lower expression relative to pro- and pre-B cells. In contrast, genes such as CD34, Kit, Tie1, Slamf1, and Tal1 are upregulated in stem cells and CLPs relative to the Id2-HPCs. These gene expression patterns suggest that the Id2-HPCs represent cells which have not yet upregulated B-cell or other lineage commitment genes, but have downregulated "stemness" genes. Vertical clustering of cell types confirms that the Id2-HPCs cluster on the same branch as E2A -/- and EBF -/- cells, and are closer in gene expression pattern to pre-B and pro-B cells than to earlier progenitors (Figure 3.3B).

Id2-HPCs can Reconstitute Multiple Immune Lineages in vivo

To determine if Id2-HPCs maintain pluripotency and repopulating ability in vivo, cells were injected into irradiated CD45.2 recipients in a competitive repopulating assay. Since in the absence of E2A, B cells do not develop past the pro-B cell stage and alternate lineage development is perturbed, (Bain et al., 1994), we administered doxycycline via food pellets to recipient animals to turn off Id2 expression and thus allow restoration of E-protein activity in the Id2-HPCs. The mice received doxycycline food pellets 24 hours before injection and throughout the remainder of the experiment. Recipient animals were lethally irradiated 24 hours before tail vein injection of Id2-HPCs (CD45.1) mixed with freshly harvested CD45.2 bone marrow cells. At 6 weeks post-transplant, Id2-HPCs had successfully reconstituted the bone marrow, thymus, and spleen of irradiated recipients to varying degrees, albeit at a greatly reduced ability compared to wild type bone marrow (Figure 3.4). This may

be due to an impaired homing capability, as the Id2-HPCs home to the bone marrow with an approximate four-fold impairment relative to E2A -/- cells. Upwards of 95% of the CD45.1 cells were GFP negative, showing the robust sensitivity of the tetracycline transactivator to doxycycline in vivo (Figure 3.5).

In the bone marrow, Id2-HPCs reconstitute both B cell and myeloid lineages successfully (Figure 3.5A). There are increased relative numbers of CD45.1 B220+ cells and decreased numbers of CD45.1 CD11b+ cells relative to their CD45.2 counterparts, suggesting a propensity for Id2-HPCs to commit to the B cell lineage, which might be expected due to their pre-pro B cell phenotype in vitro (Figure 3.5B). The Id2-HPCs also successfully migrate to the peripheral immune organs. In the thymus, Id2-HPCs reconstitute the double negative, double positive and single positive populations at comparable ratios to wild-type cells (Figure 3.5C). There is also successful migration of Id2-HPCs from the bone marrow and thymus to reconstitute the T, B and myeloid lineages in the spleen (Figure 3.5A). In the spleen, IgM is expressed on approximately 20% of CD45.1 cells, demonstrating that Id2-HPCs can successfully progress past the pre-pro-B cell stage that is seen in the in vitro culture. Thus, Id2-HPCs can both self-renew in vitro and reconstitute multiple immune lineages in vivo when hId2 expression is down-regulated.

In vitro differentiation of Id2-HPCs into myeloid and B cell lineages

Since the Id2-HPCs were able to repopulate different immune lineages in vivo, we hypothesized that we could induce differentiation in vitro to study hematopoietic development in a reproducible manner using a primary cell line. To stimulate myeloid differentiation, the Id2-HPCs were cultured in IL3, Flt3-L, GM-CSF and M-CSF either with or without

doxycycline. Cells were analyzed by FACS and Q-PCR. In the absence of doxycycline, the cells rapidly died, and surviving cells showed only modest CD11b expression. Conversely, in the presence of doxycycline, the cells expressed significant levels of CD11b, although similarly failed to proliferate once they had differentiated into the myeloid lineage (Figure 3.6). Interestingly, the mean fluorescence intensity of B220 was increased in the myeloid cultures in all three clones, similar to that seen in other B/myeloid differentiation cultures (Xie et al., 2004).

To stimulate Id2-HPCs to differentiate into pro-B cells, the cells were cultured on S17s in the presence of IL-7 and SCF either with or without doxycycline. The cells that were cultured in the absence of doxycycline continued to divide but maintained the pre-pro-B cell phenotype. Cells that were grown in the presence of doxycycline rapidly downregulated GFP and showed upregulation of the pro-B cell marker CD19. Interestingly, Ly6D was upregulated by 48 hours after doxycycline addition, and then downregulated by 120 hours (Figure 3.7A). Ly6D is a recently reported marker present in early B-lineage primed progenitors prior to the pro-B cell stage, but characterization of its expression past the BLP in B cell development is not yet been documented (Inlay et al., 2009; Mansson et al., 2008). Importantly, the Id2-HPCs also rearrange V-DJ joints after an 8-day culture with doxycycline, showing that Id2-HPCs can fully differentiate to the pro-B cell stage (Figure 3.7B) (Matthias and Rolink, 2005).

Q-PCR analysis showed robust downregulation of hId2 and upregulation of the B cell genes EBF and Foxo1 (Figure 3.8). EBF is a well-known regulator of B-cell commitment and there is strong evidence that EBF also directly regulates Foxo1. Collectively these studies

demonstrate that Id2-HPCs recapitulate key steps in early B cell development (Medina et al., 2004; Zandi et al., 2008).

Monoclonal Analysis of Id2-HPCs

Single cell cultures were initiated to determine the clonal properties of Id2-HPCs. Individual cells from the highest 4%, lowest 4%, or intermediate 10% of GFP mean fluorescence were sorted onto pre-seeded S17s in 96-well plates and cultured as before. Clones from each of the conditions (Low GFP= L, Intermediate GFP= I, High GFP= H) were expanded, and one clone from each group was chosen for further analysis. Interestingly, the three clones had varying levels of cell surface marker expression, including GFP (Figure 3.9A). The cells were then differentiated in vitro either into B-cells or into myeloid cells, as described previously. The three lines had surprisingly different abilities to develop into either CD19 or CD11b positive cells under B or myeloid instructive conditions. H1, the clone with the highest GFP mean fluorescence, did not generate B220+ CD19+ B cells as efficiently as L1 and I2, yet all three cell lines developed into CD11b+ cells at similar rates (Figure 3.9A-C).

Correspondingly, the levels of lineage specific gene expression also varied between the single cell clones (Figure 3.9D). Again, H1 has the lowest upregulation of the B-cell specific genes EBF and Foxo1 under B-cell differentiating conditions. In contrast, all three cell lines show similar upregulation of the myeloid specific genes Itgam and Csf1r.

Gene expression analysis in B and myeloid cell development

The Id2-HPC cell line provides an excellent system for studying discrete events in hematopoietic cell differentiation in a primary cell line. To this end, we differentiated the Id2-HPCs either into CD19+ B cells or CD11b+ myeloid cells and performed microarray gene expression analysis. We hierarchically clustered the expression of genes with a two-fold or more change in absolute expression levels between any two groups. Four patterns emerged: I) genes upregulated in CD19+ B cells, II) genes repressed in CD11b+ myeloid cells, III) genes activated at the 0-hour timepoint, prior to differentiation, and IV) genes that are activated in myeloid development (Figure 3.11).

Not surprisingly, many of the required B-cell differentiation factors, including Ebf1, Foxo1, Bcl11a, Vpreb1, and Rag1 are upregulated in B cells. While these genes are strongly upregulated in the B cell culture, they are also slightly downregulated in the myeloid conditions, indicating that there is B-lineage priming at the 0-hour timepoint. Gene ontology of cluster II, genes that are downregulated in the myeloid conditions, shows overrepresentation of genes involved in cell cycle and metabolic processes (Table 3.2). Since the Id2-HPCs fail to blast in the myeloid culture conditions, it is not surprising to see the downregulation of genes involved in cell cycle and chromosome rearrangement. Interestingly, Ly6D expression falls into this cluster, supporting previous data implicating Ly6D as a specific marker for early B cell development (Inlay et al., 2009; Mansson et al., 2010).

Cluster III represents genes that were repressed in both the B and myeloid differentiated cell stages. Genes present in this cluster include genes required for erythroid differentiation, such as Hbb-b1, and in early hematopoietic development, such as CD34, Gfi1b, and several of the Hoxa family of genes. Finally, genes that were activated in the myeloid cultures include the myeloid specific genes CD68 and Csf1r, as well as the Irf family

of genes. Interestingly, several genes involved in T-cell development were also upregulated in the CD11b cells relative to the CD19 B cells or 0-hour progenitor cells. Collectively, these data indicate that Id2-HPCs express a lineage-primed signature of gene expression prior to differentiation, and repress the expression of genes associated with alternate cell lineages (Tables 3.3 and 3.4).

Discussion

Here we have shown the development of a robust self-renewing and multipotent cell line, Id2-HPC. Id2-HPCs show in vivo reconstitution potential and in vitro differentiation potential. By overexpressing a transcriptional inhibitor, namely Id2, instead of a transcription factor, we avoid any aberrant activation of off-target genes and, upon elimination of the inhibitor, differentiation can be induced. This is a novel method for ex vivo expansion of hematopoietic progenitors, as most prior methods involve the overexpression or exposure to transcription factors in concert with proper stimulatory cytokines (Kelly et al., 2009; Sauvageau et al., 2004). The development of this cell line was based on the pluripotency of both E2A and EBF knockout pre-pro-B cells (Ikawa et al., 2004). We report here the first genome-wide expression analysis of these cells. RNA expression analysis findings indicate that, as predicted by the expression cell-surface markers such as B220, Ly6D and CD19, the E2A -/-, EBF -/-, and Id2-HPC cell lines associate more closely with a B-cell lineage than a multipotent lineage (Figure 3.3). However, even though the multipotent cell lines cluster with other B-cell lineages, the expression of B-cell specific genes such as Foxo1, Ly6D and Vpreb1 is lower in Id2-HPCs relative to pro-B cells. In contrast, the Id2-HPCs seem to have downregulated many genes involved in “stemness”, such as Tal1 and Kit. This suggests that

the Id2-HPCs, and the E2A and EBF -/- cell lines represent a B-cell "primed", yet still multipotent state of development. Indeed, these cell lines all maintain some degree of multipotency and self-renewal both in vivo and in vitro.

Upon injection into irradiated recipients, the Id2-HPCs successfully reconstitute multiple lineages not only in the bone marrow, but also the peripheral immune organs, such as the spleen and the thymus. As the Id2-HPCs have undergone DJ rearrangement at the IgH locus, it is somewhat surprising that they reconstitute not only the developmental compartments in the thymus, but can also proceed through positive selection and travel to the spleen. It will be interesting to determine if these thymic repopulating cells are descended from a small, non-DJ rearranged subset of Id2-HPCs, or if they maintain IgH DJ rearrangements as circulating, mature T cells.

As Id2-HPCs readily differentiate in vitro to either the B or myeloid lineage, we studied the changes in gene expression that allow either the maintenance of the self-renewing and multipotent state or direct lineage commitment choices. As mentioned before, although Id2-HPCs have lower expression levels of "stemness" like genes relative to CLPs and LMPPs, the cells further down-regulate expression of genes such as CD34 and the Hoxa family upon differentiation into either the myeloid or B cell lineage. This indicates that although the Id2-HPCs do not genetically represent true multipotent cells, they still maintain moderate transcript levels of many stem-like genes, and this is sufficient for their ability to self-renew.

As many investigations are in progress to expand hematopoietic stem and pluripotent stem cells ex vivo for therapeutic pursuits, the Id2-HPCs represent a paradigm shift in the field of stem cell expansion. Here, for the first time, overexpression of a transcriptional inhibitor

instead of a transcription factor has permitted the unlimited growth capacity and maintenance of multipotency of hematopoietic progenitors. As such, studies are underway to overexpress Id2 in human cord blood cells to determine the therapeutic potential of this expansion procedure for human stem cell transplantation.

Materials and Methods

Viral Vectors and Virus Production:

hId2 was amplified from 293T cDNA with the addition of BamH1 sites at both 3' and 5' ends, and then cloned into pRRL.sin-18.PPT.PGK.MCS.IRES.GFP.pre, which was kindly provided by Dr. Irving Weissman (Stanford University, Stanford, CA). The hId2_IRES_GFP expression cassette was cut out with AgeI and BsrG1, and replaced the eGFP gene in pBob_TA1_R2_corrected120803, which was kindly provided by Dr. Inder Verma (Salk Institute, La Jolla, CA). hId2 was then cut out of the resulting vector to create an empty vector control. These vectors were then called TetOff_hId2 and TetOff_null, respectively.

Viral vectors were produced by the transient calcium phosphate mediated transfection of TetOff_hId2 or null, pMD2.VSVG and pCMVdR8.74 into 293T cells. The packaging plasmids were kindly provided by Dr. Irving Weissman (Stanford University, Stanford, CA). Virus was harvested at 24 and 48 hours, and then concentrated by ultracentrifugation in a Beckman Sw28 rotor at 19,400 r.p.m. for 2 hours at 20°C. The virus was resuspended in PBS + 1% BSA and aliquoted for freezing at -80°C.

Transduction and Cell Culture:

CD45.1 congenic animals (8-12 weeks of age) were injected with 250mg/kg of 5-fluorouracil 4 days before bone marrow was harvested and lineage depleted by auto-MACS. Purified lineage negative cells were cultured overnight in expansion media (DMEM + 15% FBS in the presence of 10ng/mL IL-3, 10ng/ml IL-6, 1:200 SCF and 1:20 WEHI). After 12-18 hours in culture, cells were pelleted and resuspended in fresh expansion media plus TetOff_hId2 or TetOff_Null lentivirus in the presence of 4ug/mL polybrene and the above cytokines. Cells were spin infected twice for 1.5 hours at 2500rpm, 30° C, with a 12-18 rest in between spin infections. After the second spin, cells were resuspended in fresh expansion media and cultured for 4-5 weeks. Cells were then moved into IMDM + 10% FCS, 2% PSG and 2μL β-me on subconfluent S17 feeder cells in the presence of 1:100 IL-7, 1:100 Flt3L and 1:200 SCF. S17 feeder cells were maintained in α-MEM + 10% FBS and 2% PSG. 22D6 cells were maintained in RPMI + 10% FBS and 2% PSG.

In Vitro Differentiation Experiments

Id2-HPCs were depleted of small (<1-5%) numbers of CD19, CD25 and CD11b positive cells by auto-MACS. For myeloid differentiation, cells were cultured for up to 6 days in IMDM + 10% FCS and β-me in the presence of IL3, Flt3L, GMCSF and MCSF and in the presence or absence of 1ug/mL doxycycline. For B-cell differentiation cells were cultured for up to 10 days in IMDM + 10% FCS and β-me in the presence of 1:100 IL-7 and 1:200 SCF in the presence of absence of 1ug/mL doxycycline. Fresh doxycycline and cytokines were added every two days.

Flow Cytometry

Flow cytometry was performed on either a FACSCalibur or LSR II Flow Cytometer (Becton Dickinson, Franklin Lakes, NJ) and analyzed with FlowJo (Tree Star, Ashland, OR). Single cells were stained with antibodies purchased from eBioscience (San Diego, CA).

Bone Marrow Chimeras

8-10 week old 45.2 congenic recipients were lethally irradiated (1000 rads λ irradiation from a Cs source) and started receiving doxycycline food pellets (Bio-Serv, Frenchtown, NJ) 24 hours before injection. Id2-HPCs were depleted of small (<1-5%) numbers of CD19, CD25 and CD11b positive cells by auto-MACS, and 1 million cells were injected via tail vein injection along with 1 million freshly isolated wild type 45.2 bone marrow cells. Mice continued to receive doxycycline food throughout the study.

Microarray Profiling

All RNA was prepared using RNeasy Columns (Qiagen, Seattle, WA). Gene expression profiling depicted in Figure 3, comparing 9 different cell types, was performed by hybridization to the Affymetrix MOE430 2.0 gene expression array according to the manufacturer's instructions. Gene expression profiling depicted in Figures 7 was performed by amplification and hybridization to the Illumina Mouse WG-6 v1.1 gene expression array according to the manufacturer's instructions.

Quantitative PCR

RNA was prepared using RNeasy Columns (Qiagen, Seattle, WA) and cDNA was prepared using Superscript III (Invitrogen, Carlsbad, CA). Real-time PCR was performed on a Stratagene Mx3500p using Brilliant Sybr Green Master Mix (Stratagene, La Jolla, CA). Gene specific primers are available upon request.

Chapter III is currently being prepared for submission for publication of the material. The working title is "Inducible Id2 expression allows for expansion of a long-termculture of multipotent hematopoietic progenitors and epigenetic studies of lineage commitment." I was the primary investigator and author of this material. Co-authors are Yin C. Lin, Suchit Jhunjhunwala., Christopher Benner, Christopher Glass, and Cornelis Murre.

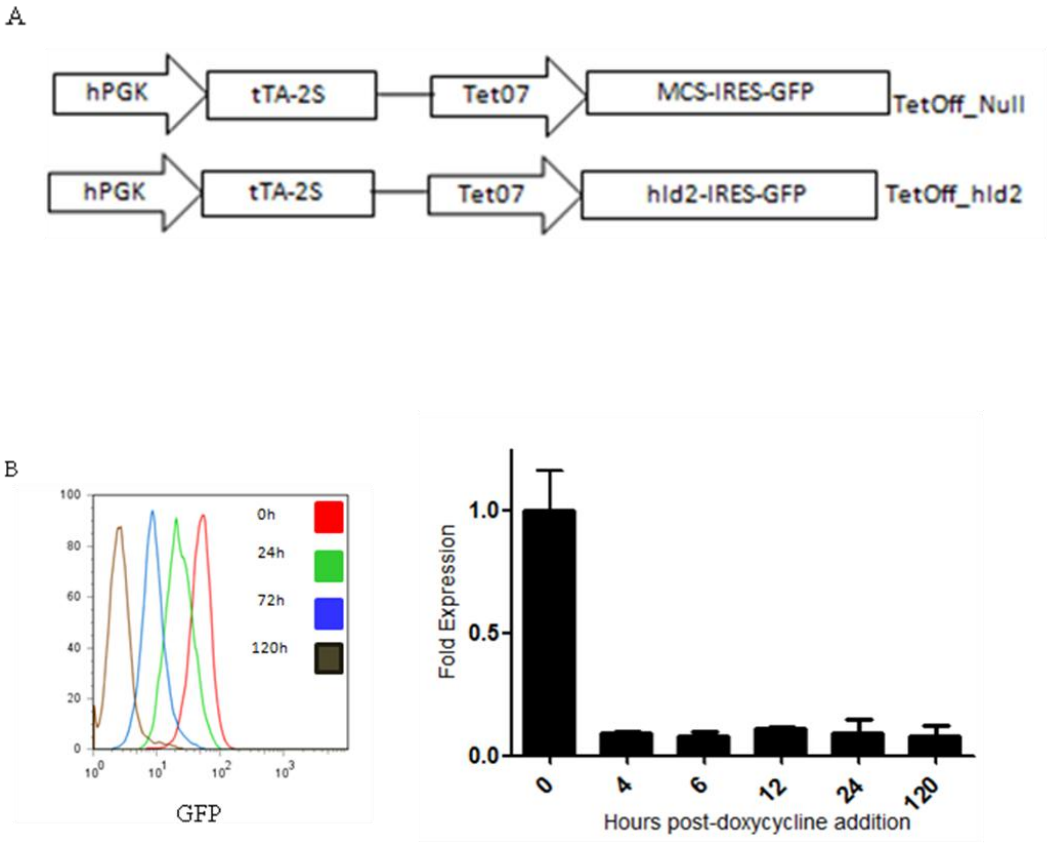


Figure 3.1

Experimental Design

A. Lentiviral constructs. Expression of hId2-IRES-GFP or MCS-IRES-GFP cassette is downstream of the Tet07 promoter, which is inactivated in the presence of doxycycline.

(B) Rapid and robust downregulation of GFP and hId2 after doxycycline addition in vitro. FACS plots of GFP (left panel) at 0, 24, 72 and 120 hours post-doxycycline addition. Real-time PCR analysis of hId2 expression at 0, 4, 6, 12, 24 and 120-hours post-doxycycline addition.

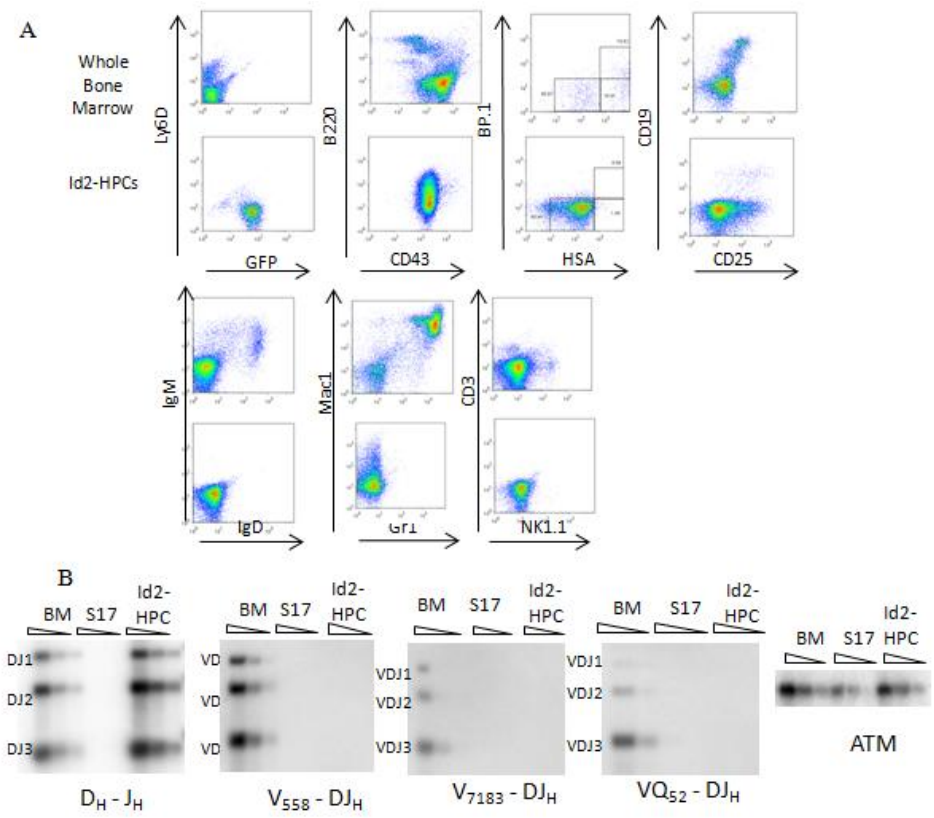


Figure 3.2

Cell Surface Phenotyping and IgH Rearrangement Analysis of Id2-HPCs

(A) Phenotypic analysis of Id2-HPC cells. Id2-HPCs were cultured on S17 feeder cells with IL-7, SCF and Flt3-L. Cells were analyzed by FACS for the expression of B220, Ly6D, CD43, CD25, CD19, Mac1, Gr1, CD3 and NK1.1 after 3 months in culture. The upper panel shows staining from purified wild type bone marrow and the lower panel shows staining from expanded Id2-HPCs in the absence of doxycycline.

(B) IgH gene rearrangement analysis of Id2-HPCs. DNA was isolated from wild type bone marrow, S17 feeder cells, and Id2-HPCs and analyzed by Southern Blot for the presence of IgH DJ and V-DJ rearrangements. ATM was used as a loading control.

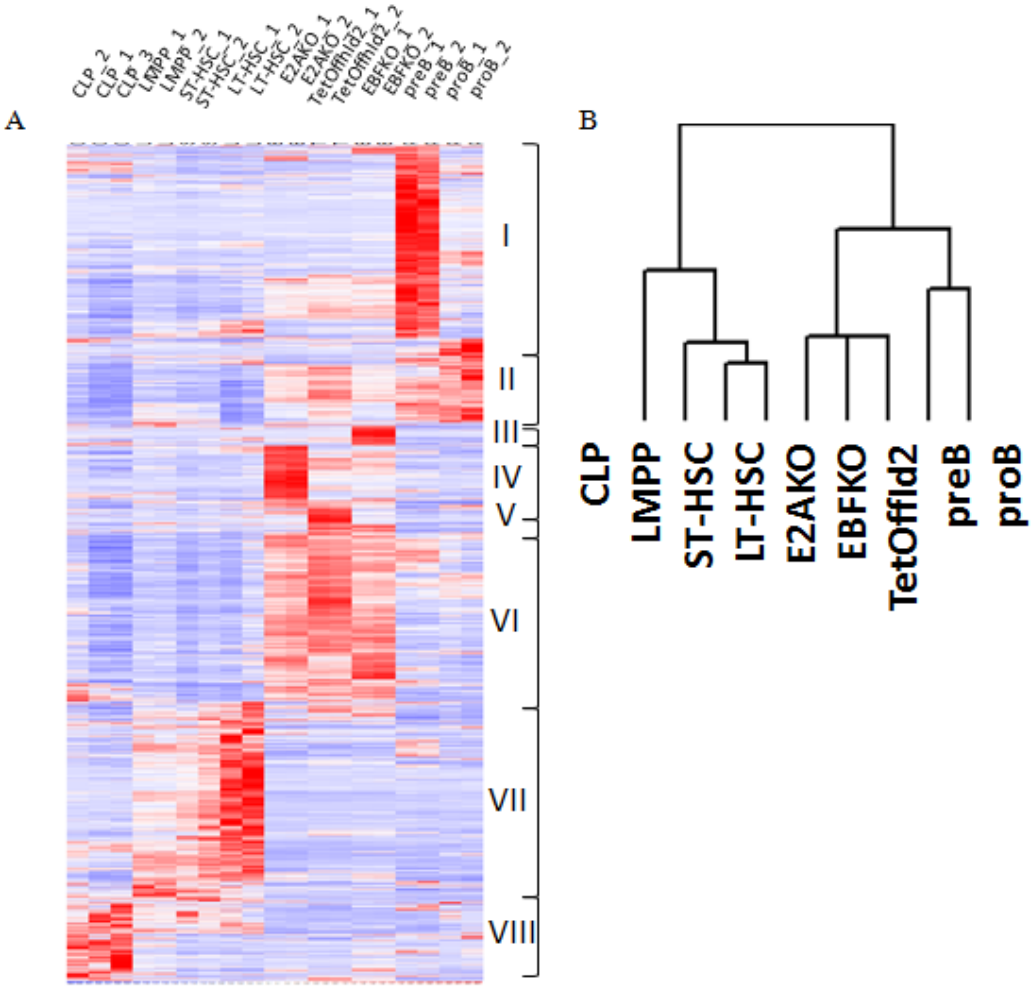


Figure 3.3

Microarray Analysis of Id2-HPC vs. Other Hematopoietic Lineages

(A) Microarray analysis of LT-HSCs, ST-HSCs, LMPPs, CLPs, pre-B cells, pro-B cells, E2A-/- cell line, EBF-/- cell line, and Id2-HPC cell line. Clustering was performed with d-Chip software on genes with 2-fold or greater change in expression between 7% or more of the samples.

(B) Vertical clustering of cell types. Clustering of cell types into branches was performed using d-Chip software.

Table 3.1

Gene Ontology Terms for Hierarchical Clusters from Id2-HPCs vs. Other Hematopoietic Lineages Microarray

Cluster	Top BP Gene Ontology Clusters
I	Lymphocyte activation
	Apoptosis
	Metabolic processes
II	Mitosis
	Cell cycle processes
	DNA replication
III	Lymphocyte activation
	Immune system processes
	Cell migration
IV	Immune response
	Chemotaxis/response to stimulus
	Biopolymer metabolic processes
V	Blood vessel development
	Negative regulation of kinase activity
	Apoptosis
VI	Metabolic processes
	Lipid synthesis
	Metabolic processes/protein modiciation
VII	Developmental processes
	Apoptosis
	Ion homeostasis
VIII	Metabolic processes
	Protein metabolic processes
	System development

Cluster	Top CC Gene Ontology Clusters
I	Intracellular organelles
	Lytic vacuole
	Cytoskeleton
II	Chromosome
	Intracellular organelle
	Membrane bound organelle
III	Cytoplasmic vesicle
	Plasma membrane
	Extracellular space
IV	Lytic vacuole
	Extracellular space
	Membrane
V	Extracellular matrix
	Extracellular space
	Intracelluar
VI	Intracellular part
	Mitochondria
	Organelle part
VII	Intracellular
	Cytoplasm
	Extracellular space
VIII	Nucleus
	Intracellular organelles
	Cytoskeleton

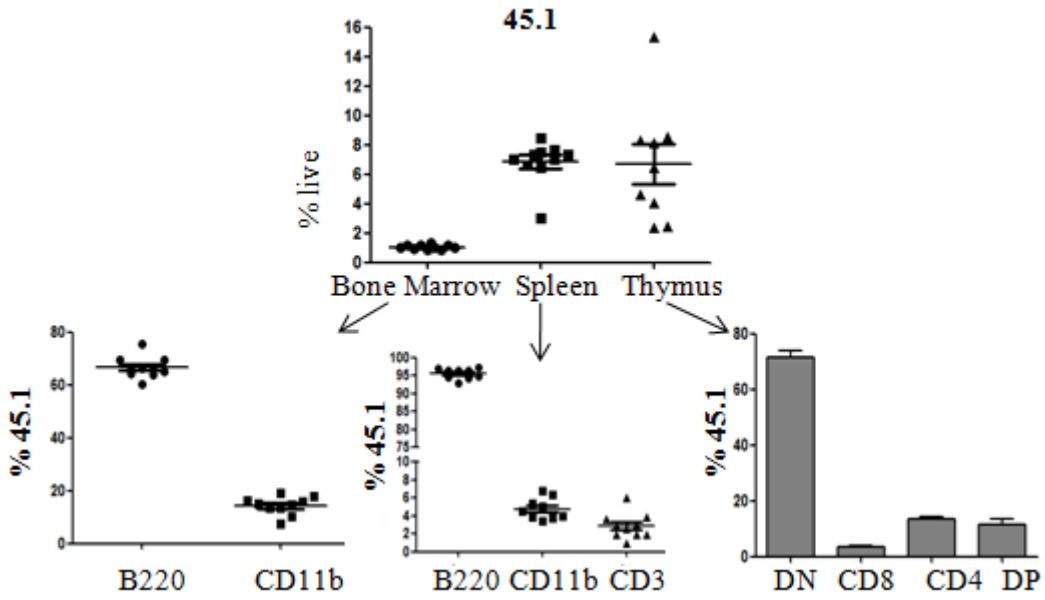


Figure 3.4

Id2-HPCs can Reconstitute Irradiated Recipients

Id2-HPCs can reconstitute multiple lineages in irradiated recipients. A competitive reconstitution assay was used to assess the ability of Id2-HPCs to reconstitute irradiated recipients. One million Id2-HPCs and one million freshly isolated wild type bone marrow cells were injected into lethally irradiated recipients receiving doxycycline feed. Reconstitution was analyzed 6-weeks later. The data is representative of two experiments.

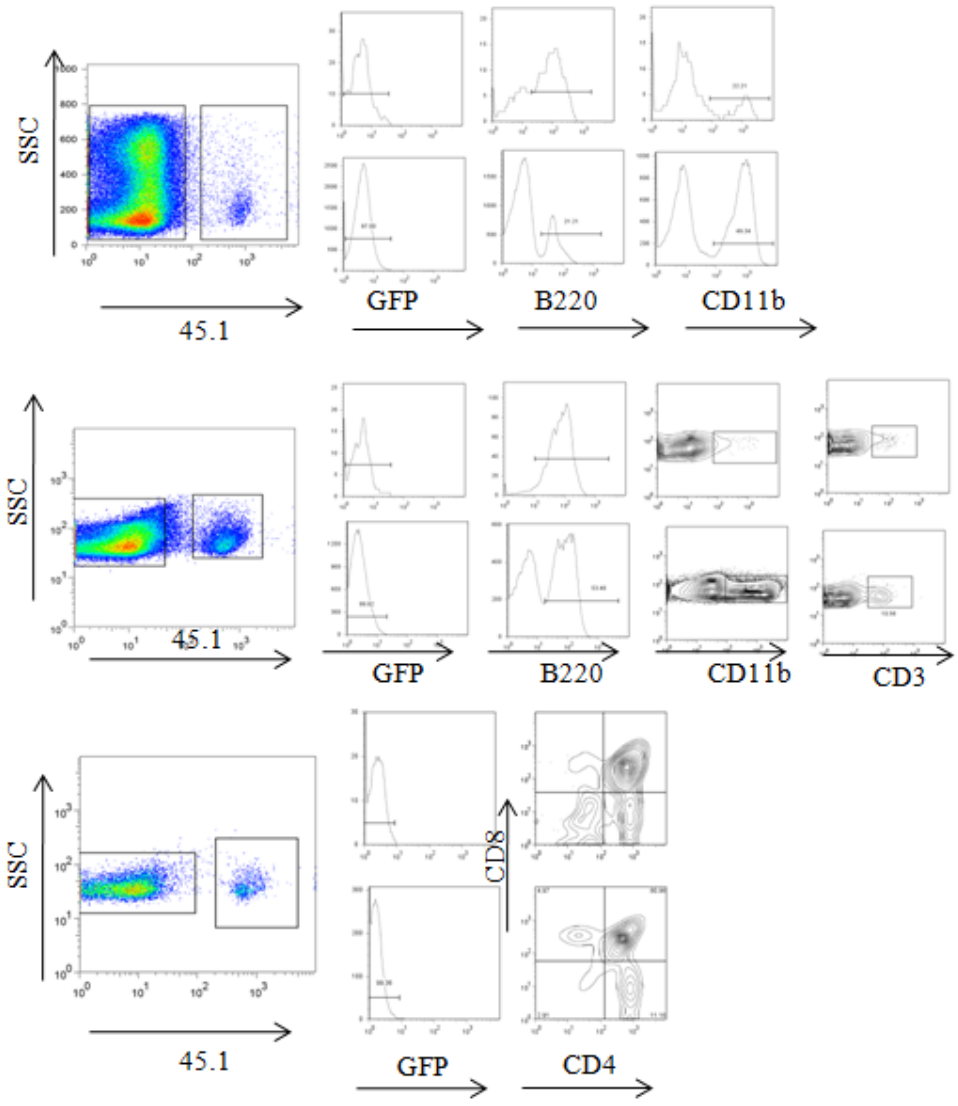


Figure 3.5

Id2-HPC Reconstitution of Bone Marrow, Spleen and Thymus

(A) Id2-HPCs successfully reconstitute the bone marrow of recipients. Bone marrow cells were analyzed by FACS analysis for expression of GFP, B220 and CD11b.

(B) Id2-HPCs successfully reconstitute the spleen of recipients. Bone marrow cells were analyzed by FACS analysis for expression of GFP, B220, CD11b and CD3.

(C) Id2-HPCs successfully reconstitute the thymus of recipients. Bone marrow cells were analyzed by FACS analysis for expression of GFP, CD4 and CD8.

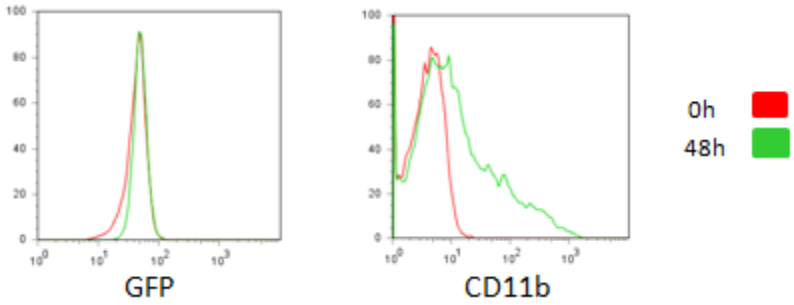


Figure 3.6

In Vitro Differentiation of Id2-HPCs into Myeloid Lineage

In vitro differentiation of Id2-HPCs into myeloid cells. Cells were cultured in IL-3, SCF, GMCSF and MCSF in the presence or absence of doxycycline for 2 days in vitro and analyzed by FACS at day 0, day 4 and day 6 for CD11b.

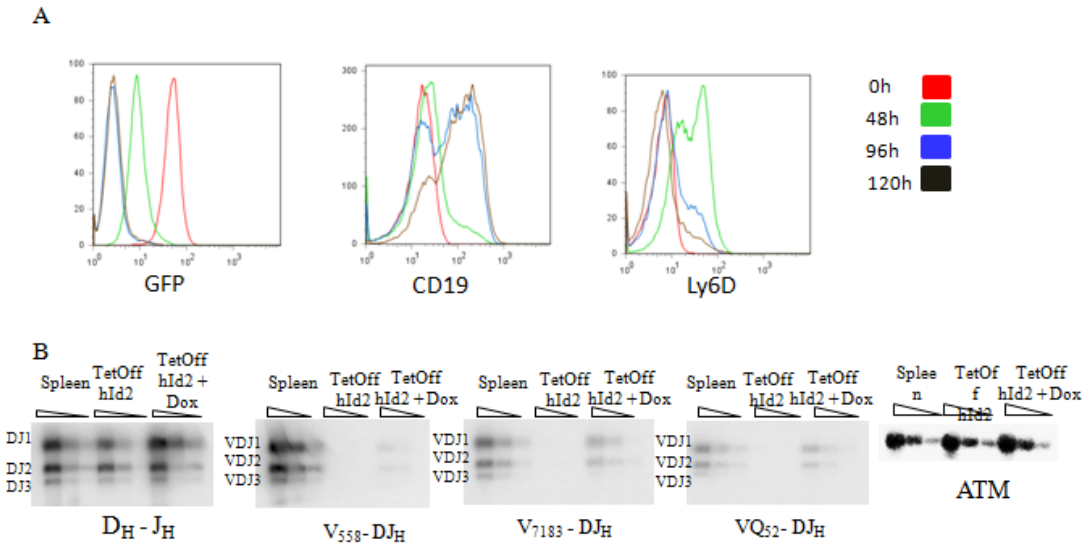


Figure 3.7

In Vitro Differentiation of Id2-HPCs into B Lineage

(A) In vitro differentiation of Id2-HPCs to pro-B cells. Cells were cultured in IL-7, SCF in the presence or absence of doxycycline for 6 days in vitro and analyzed at day 0, day 4 and day 6 for GFP, CD19, and Ly6D expression.

(B) IgH gene rearrangement in differentiated Id2-HPCs. DNA was isolated from wild type bone marrow, Day 0 Id2-HPCs and Day 8 + doxycycline Id2-HPCs and analyzed by Southern blotting for IgH DJ and V-DJ rearrangements. ATM was used as a loading control.

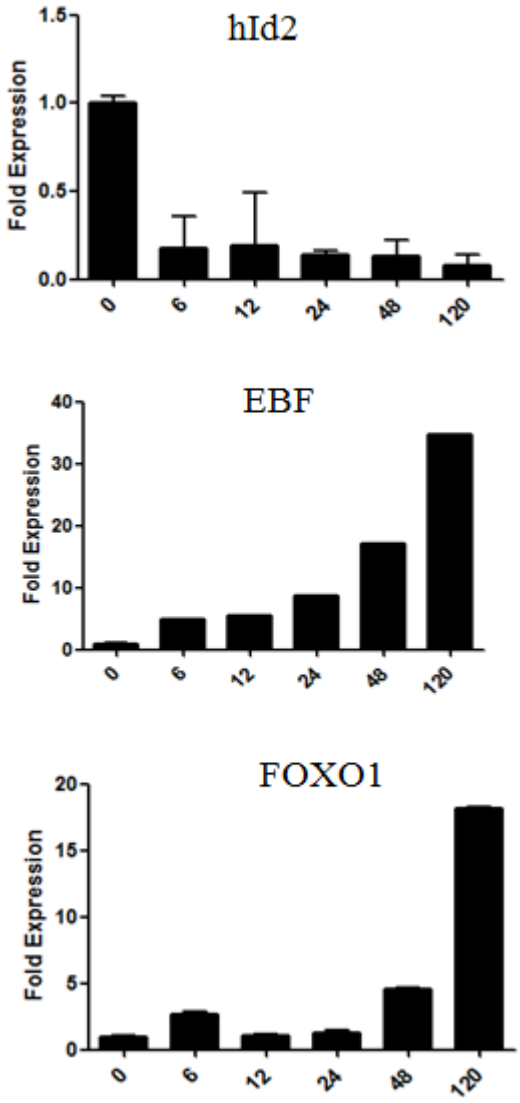


Figure 3.8

Gene Expression Analysis of Id2-HPCs During In Vitro B-cell Differentiation

RNA was isolated from Id2-HPCs at 6 time-points during in vitro B-cell differentiation, and real time PCR was performed on hId2, Ebf1, and Foxo1 transcripts.

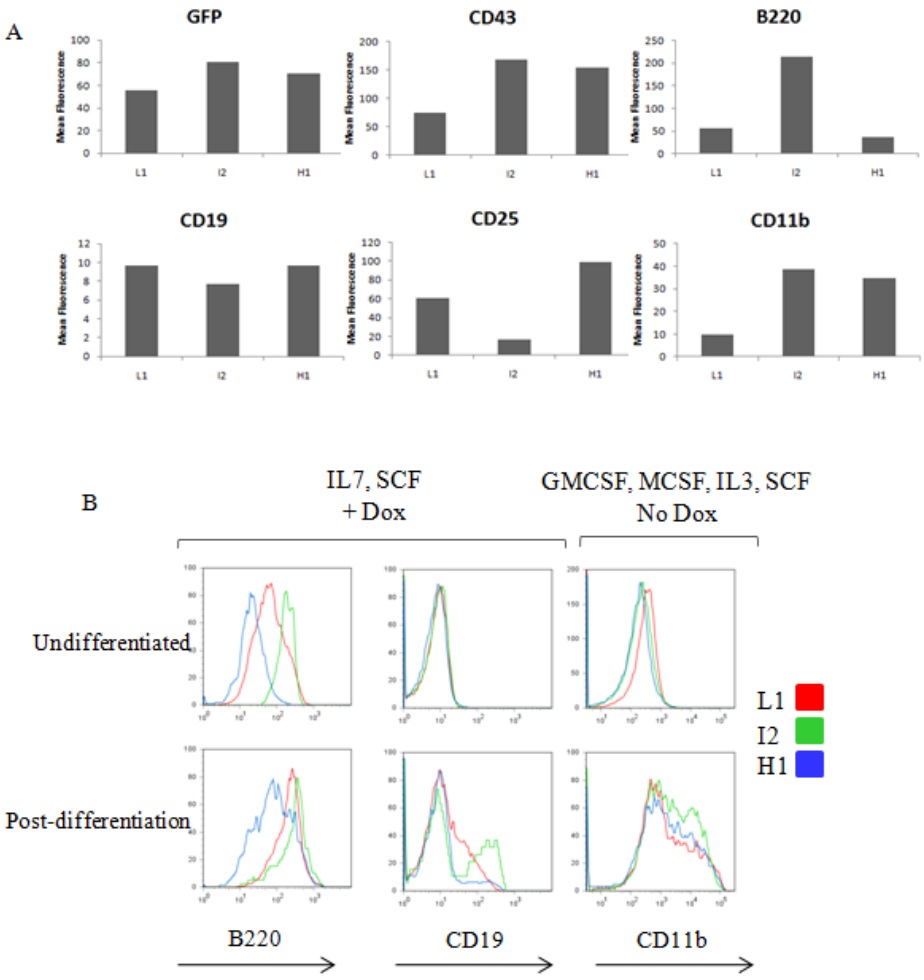


Figure 3.9

Phenotypic Analysis and Differentiation Abilities of Monoclonal Id2-HPCs

(A) Comparison of GFP levels and lineage specific cell surface markers expression in single cell clones. Cells are gated on live cells, and mean fluorescence intensity of each signal or marker is represented.

(B) Single-cell clones of Id2-HPCs differentiate into B and myeloid lineages at varying degrees. Cells were cultured either in B-cell differentiation conditions (IL7, SCF, Doxycycline on S17 feeders) or myeloid differentiation conditions (GMCSF, MCSF, IL3, SCF) for 2-5 days and then analyzed for expression of B220, CD19 and CD11b.

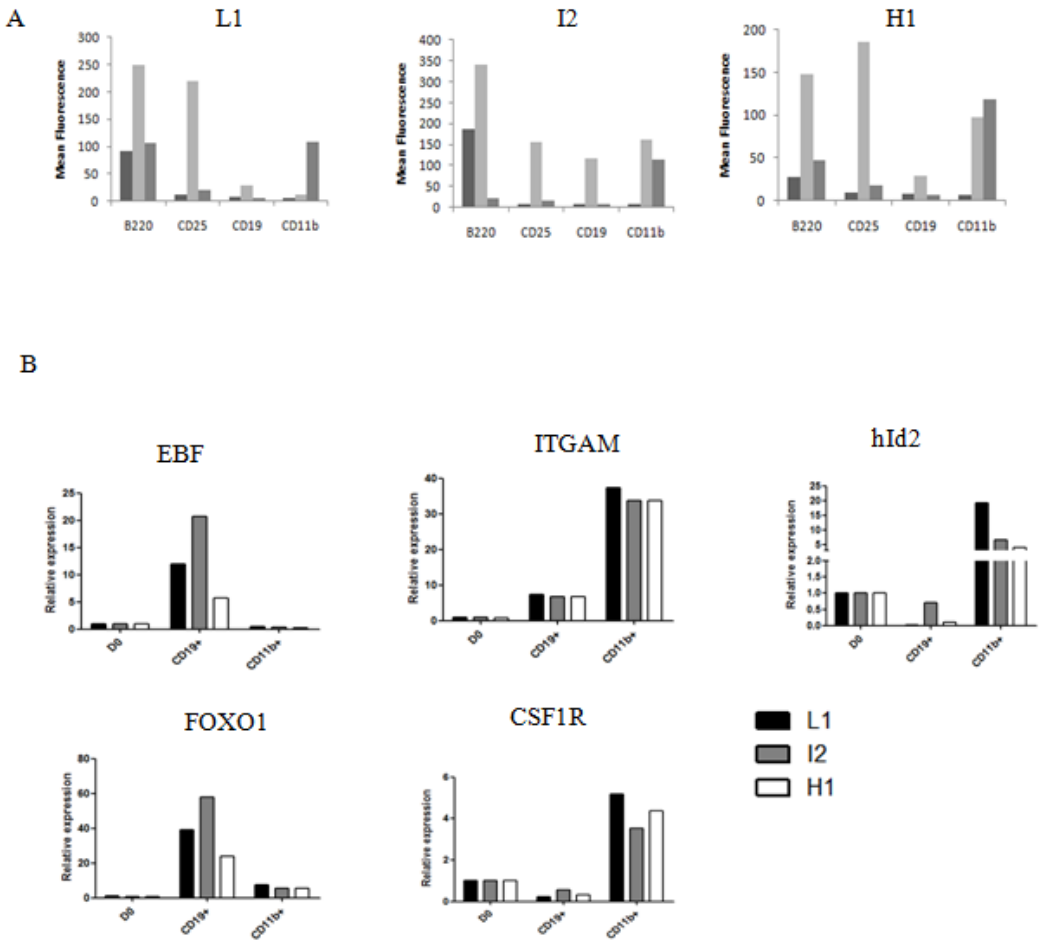


Figure 3.10

Phenotypic and Gene Expression Analysis of B and Myeloid Differentiation of Monoclonal Id2-HPCs

(A) Mean fluorescent intensity of cell surface markers on L1, I2, and H1 cells at the 0-hour progenitor timepoint, the B-cell terminal timepoint, or the myeloid cell terminal timepoint.

(B) Lineage specific gene expression in alternate differentiation pathways. Gene expression in the L1, I2, and H1 single cell clones at the 0-hour timepoint, B-cell CD19+ timepoint, or CD11b+ timepoint was analyzed by real-time PCR.

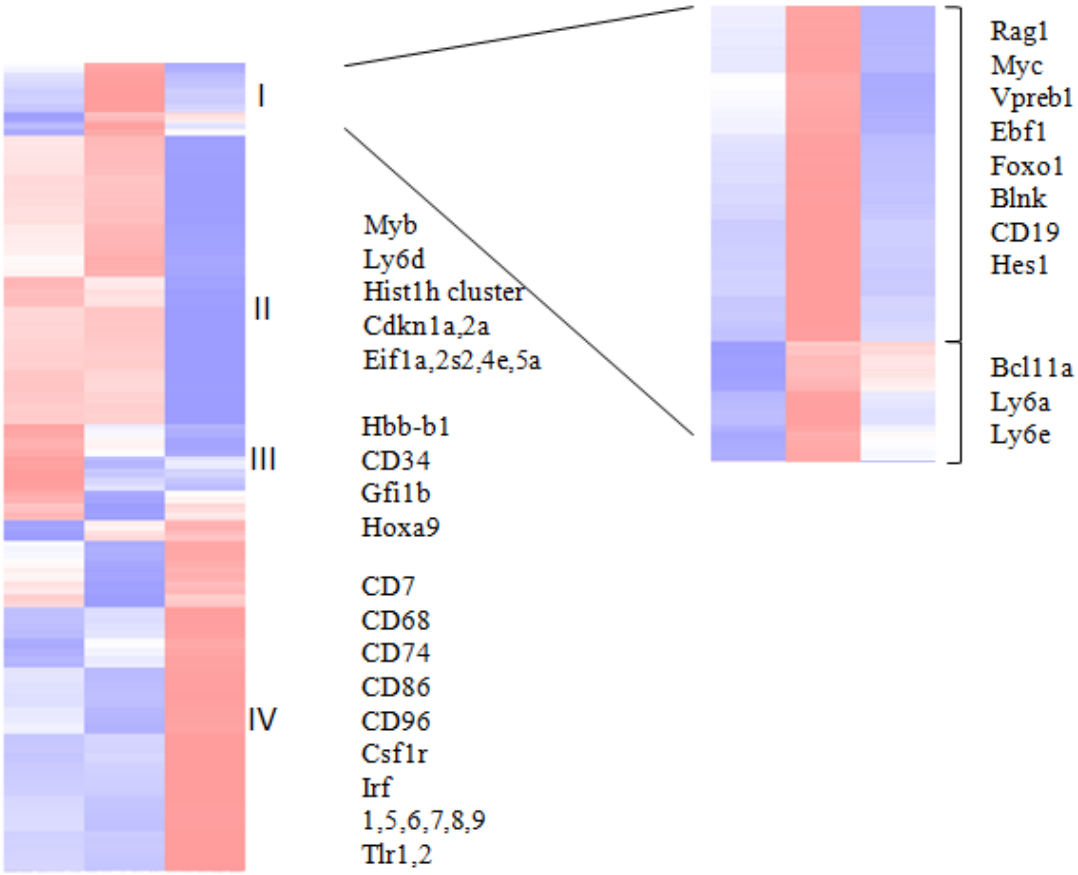


Figure 3.11

Microarray Analysis of Id2-HPCs, CD19+ B-cells, or CD11b Myeloid Cells

Id2-HPCs were differentiated into either CD19+ B cells or CD11b+ myeloid cells, and gene expression at 0 hours, B, or M-cell timepoints was analyzed by microarray. Clustering was performed with d-Chip software on genes with 2-fold or greater change in expression between any two groups.

Table 3.2

Gene Ontology Terms from Hierarchical Clusters from Differentiation Microarray

Cluster	Gene Ontology Clusters
I	Immune system processes
	Positive regulation of cell proliferation
	Tube development
	Developmental processes
II	Cell cycle
	Cellular metabolic processes/DNA replication
	Cellular response to stress/stimulus
	Chromosome organization
III	Oxidation reduaction
	Immune system processes
	Cellular hometostatic processes
	Response to external stimulus
IV	Immune system processes
	Leukocyte activation
	Immune system processes
	Response to wounding

Table 3.3

Top 50 Repressed or Activated Genes in CD19+ cells vs. Id2-HPCs

Top repressed genes B:0 (log2)		Top activated genes B:0 (log2)	
Mpo	-5.7295	Vpreb1	5.3565
Rsad2	-5.276	LOC100047815	5.185
Tox	-3.7425	Pou2af1	4.8305
Clec4d	-3.4555	Vpreb3	4.753
Slc35d3	-3.42	Cd79b	4.676
Tyki	-3.387	Lta	4.5545
Sp7	-3.332	5730488B01Rik	4.5065
Itgb7	-3.192	Crhbp	4.231
LOC100048346	-3.179	Pde2a	4.2235
Tnni1	-3.152	Vat1l	4.052
Ccl5	-3.0005	Atp1b1	3.9435
LOC545732	-2.997	Blnk	3.789
LOC333331	-2.84	Ctla4	3.6
Wipi1	-2.776	Ankrd22	3.5875
Gria3	-2.7335	LOC100044439	3.5065
Lpxn	-2.7135	Mmp11	3.397
Megf10	-2.7095	Vpreb2	3.3705
Hspb3	-2.709	Rag1	3.2805
Speer3	-2.669	6430550H21Rik	3.261
Ahnak	-2.6375	Cd74	3.205
Lmna	-2.5725	Txnrd3	3.15
Hmgn3	-2.5375	LOC100045567	3.109
Tsc22d1	-2.534	Hmha1	3.037
Dgkg	-2.525	LOC638935	3.0365
Samsn1	-2.517	Foxo1	3.0055
Hoxa5	-2.5145	Chst3	2.9945
OTTMUSG00000000971	-2.4145	Chchd10	2.9865
Speer6-ps1	-2.404	6430571L13Rik	2.978
Whrn	-2.3895	Slco4a1	2.9765
Expi	-2.37	Apbb1	2.957
Lat	-2.31	Rnase12	2.9255
Rgs10	-2.292	H2-Aa	2.8635
Xcl1	-2.2175	Amotl1	2.845
Ctsw	-2.19	2010001M09Rik	2.734
Klhl30	-2.18	Ebf1	2.7235
P2rx1	-2.1605	1700027J07Rik	2.72
Il18r1	-2.117	Slc5a9	2.6615
3110040M04Rik	-2.1015	Gng4	2.63
Aqp9	-2.0925	Bcar3	2.63
Hoxa9	-2.079	Bfsp2	2.629
C030048H21Rik	-2.0485	Stambpl1	2.6145
Casz1	-2.006	Sdc4	2.6025
Ccdc113	-1.999	Ly6d	2.599
Tmem51	-1.986	Cyp1b1	2.5795
Acadm	-1.9705	Blk	2.565
Rapsn	-1.952	1700011H14Rik	2.5645
Hoxa7	-1.9075	4930513E20Rik	2.5175
Kcnh2	-1.906	P2rx3	2.5095
Igsf10	-1.8895	St5	2.4345
Ctla2b	-1.8845	Hes1	2.43

Table 3.4

Top 50 Repressed or Activated Genes in CD11b+ cells vs. Id2-HPCs

Top repressed genes M:0 (log2)	
Lmna	-4.855
Lgals3	-4.021
Sgk1	-3.949
Rsad2	-3.743
LOC630470	-3.68
Clec4d	-3.597
Ly6d	-3.568
2010001M09Rik	-3.547
Spp1	-3.428
Gja1	-3.401
LOC100047856	-3.344
Eif4ebp1	-3.315
1190002H23Rik	-3.307
Hist1h3a	-3.14
Asns	-3.099
Rpl3l	-3.063
Megf10	-2.985
Hist1h3c	-2.913
Hist2h2be	-2.897
Creb5	-2.881
LOC640696	-2.876
Hist1h3d	-2.812
Hist1h2ab	-2.807
Phgdh	-2.76
LOC666036	-2.725
LOC245892	-2.709
Hist1h3e	-2.666
Dusp2	-2.648
Mic2l1	-2.641
LOC234081	-2.635
3110013H01Rik	-2.613
Hbb-b1	-2.594
Myl4	-2.594
Hist1h4b	-2.59
F2r	-2.552
Rnu6	-2.538
Prdx4	-2.53
Hist1h4c	-2.524
Fbxo10	-2.513
Hist1h4k	-2.483
S100a6	-2.452
Pitpnm2	-2.451
St3gal6	-2.448
3110040M04Rik	-2.443
LOC100044948	-2.431
Crisp1	-2.429
Mylc2pl	-2.424
Prl2c2	-2.423
Anxa3	-2.384
Oxct1	-2.363

Top activated genes M:0 (log2)	
Hmha1	7.142
Asb2	6.225
Ccr6	5.853
Cd7	5.213
Rorc	4.643
Rnase6	4.635
Ffar2	4.538
1810011H11Rik	4.489
Dpep2	4.374
Gpr18	4.272
Gpr114	4.166
Klrd1	4.158
Kmo	4.052
Ifi27	4.014
Pld4	3.937
Ly86	3.893
Mycl1	3.877
Card11	3.831
Cfp	3.822
Tbxas1	3.741
Cd74	3.733
Gpr68	3.733
Ltb	3.706
H2-Aa	3.668
Ccl5	3.629
Cd52	3.627
Il4i1	3.609
Il17re	3.604
LOC100039742	3.537
Ptgs1	3.528
Blk	3.436
Hgfac	3.405
Cmya5	3.387
Tbc1d8	3.374
Il10ra	3.351
Srr	3.347
Dusp7	3.334
Tmem176b	3.328
Aatk	3.32
Cd82	3.32
Itgb2	3.313
Rogdi	3.302
Unc93b1	3.296
2410004L22Rik	3.256
Ier3	3.25
Itpr1	3.229
Id2	3.224
Ctla4	3.134
Xcl1	3.123
Rtn4rl1	3.094

References

Adolfsson, J., Månsson, R., Buza-Vidas, N., Hultquist, A., Liuba, K., Jensen, C. T., Bryder, D., Yang, L., Borge, O.-J., Thoren, L. A. M., *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.

Bain, G., Maandag, E. C. R., Izon, D. J., Amsen, D., Kruisbeek, A. M., Weintraub, B. C., Krop, I., Schlissel, M. S., Feeney, A. J., van Roon, M., *et al.* (1994). E2A proteins are required for proper B cell development and initiation of immunoglobulin gene rearrangements. Cell *79*, 885-892.

Benezra, R., Davis, R. L., Lockshon, D., Turner, D. L., and Weintraub, H. (1990). The protein Id: A negative regulator of helix-loop-helix DNA binding proteins. Cell *61*, 49-59.

Borghesi, L., Hsu, L.-Y., Miller, J. P., Anderson, M., Herzenberg, L., Herzenberg, L., Schlissel, M. S., Allman, D., and Gerstein, R. M. (2004). B Lineage€"specific Regulation of V(D)J Recombinase Activity Is Established in Common Lymphoid Progenitors 10.1084/jem.20031800. The Journal of Experimental Medicine *199*, 491-502.

Busslinger, M. (2004). Transcriptional Control of Early B Cell Development1 doi:10.1146/annurev.immunol.22.012703.104807. Annual Review of Immunology *22*, 55-79.

Chowdhury, D., and Sen, R. (2004). Regulation of immunoglobulin heavy-chain gene rearrangements. Immunological Reviews *200*, 182-196.

Decker, T., Pasca di Magliano, M., McManus, S., Sun, Q., Bonifer, C., Tagoh, H., and Busslinger, M. (2009). Stepwise Activation of Enhancer and Promoter Regions of the B Cell Commitment Gene Pax5 in Early Lymphopoiesis. Immunity *30*, 508-520.

Dengler, H. S., Baracho, G. V., Omori, S. A., Bruckner, S., Arden, K. C., Castrillon, D. H., DePinho, R. A., and Rickert, R. C. (2008). Distinct functions for the transcription factor Foxo1 at various stages of B cell differentiation. *9*, 1388-1398.

Gratwohl, A., Baldomero, H., Aljurf, M., Pasquini, M. C., Bouzas, L. F., Yoshimi, A., Szer, J., Lipton, J., Schwendener, A., Gratwohl, M., *et al.* (2010). Hematopoietic Stem Cell Transplantation: A Global Perspective 10.1001/jama.2010.491. JAMA *303*, 1617-1624.

Hagman, J., and Lukin, K. (2006). Transcription factors drive B cell development. Current Opinion in Immunology
Lymphocyte development / Tumour immunology *18*, 127-134.

Ikawa, T., Kawamoto, H., Wright, L. Y. T., and Murre, C. (2004). Long-Term Cultured E2A-Deficient Hematopoietic Progenitor Cells Are Pluripotent. Immunity *20*, 349-360.

Inlay, M. A., Bhattacharya, D., Sahoo, D., Serwold, T., Seita, J., Karsunky, H., Plevritis, S. K., Dill, D. L., and Weissman, I. L. (2009). Ly6d marks the earliest stage of B-cell specification and identifies the branchpoint between B-cell and T-cell development 10.1101/gad.1836009. Genes & Development *23*, 2376-2381.

Kelly, S. S., Sola, C. B. S., de Lima, M., and Shpall, E. (2009). Ex vivo expansion of cord blood. *44*, 673-681.

Lazorchak, A., Jones, M. E., and Zhuang, Y. (2005). New insights into E-protein function in lymphocyte development. Trends in Immunology *26*, 334-338.

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

Mansson, R., Zandi, S., Anderson, K., Martensson, I.-L., Jacobsen, S. E. W., Bryder, D., and Sigvardsson, M. (2008). B-lineage commitment prior to surface expression of B220 and CD19 on hematopoietic progenitor cells
10.1182/blood-2007-11-125385. Blood *112*, 1048-1055.

Mansson, R., Zandi, S., Welinder, E., Tsapogas, P., Sakaguchi, N., Bryder, D., and Sigvardsson, M. (2010). Single-cell analysis of the common lymphoid progenitor compartment reveals functional and molecular heterogeneity
10.1182/blood-2009-08-236398. Blood *115*, 2601-2609.

Massari, M. E., and Murre, C. (2000). Helix-Loop-Helix Proteins: Regulators of Transcription in Eucaryotic Organisms
10.1128/MCB.20.2.429-440.2000. Mol Cell Biol *20*, 429-440.

Matthias, P., and Rolink, A. G. (2005). Transcriptional networks in developing and mature B cells. *5*, 497-508.

Medina, K. L., Pongubala, J. M. R., Reddy, K. L., Lancki, D. W., DeKoter, R., Kieslinger, M., Grosschedl, R., and Singh, H. (2004). Assembling a Gene Regulatory Network for Specification of the B Cell Fate. *7*, 607-617.

Morita, Y., Ema, H., and Nakauchi, H. (2010). Heterogeneity and hierarchy within the most primitive hematopoietic stem cell compartment 10.1084/jem.20091318. The Journal of Experimental Medicine, -.

Murre, C. (2009). Developmental trajectories in early hematopoiesis 10.1101/gad.1861709. Genes & Development *23*, 2366-2370.

Murre, C., McCaw, P. S., and Baltimore, D. (1989). A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, MyoD, and myc proteins. Cell *56*, 777-783.

Roessler, S., Gyory, I., Imhof, S., Spivakov, M., Williams, R. R., Busslinger, M., Fisher, A. G., and Grosschedl, R. (2007). Distinct Promoters Mediate the Regulation of Ebf1 Gene Expression by Interleukin-7 and Pax5 10.1128/MCB.01192-06. Mol Cell Biol *27*, 579-594.

Sauvageau, G., Iscove, N. N., and Humphries, R. K. (2004). In vitro and in vivo expansion of hematopoietic stem cells. Oncogene *23*, 7223-7232.

Schatz, D. G., Oettinger, M. A., and Schlissel, M. S. (1992). V(D)J Recombination: Molecular Biology and Regulation doi:10.1146/annurev.iy.10.040192.002043. Annual Review of Immunology *10*, 359-383.

Seet, C. S., Brumbaugh, R. L., and Kee, B. L. (2004). Early B Cell Factor Promotes B Lymphopoiesis with Reduced Interleukin 7 Responsiveness in the Absence of E2A 10.1084/jem.20032202. J Exp Med *199*, 1689-1700.

Spangrude, G., Heimfeld, S., and Weissman, I. (1988). Purification and characterization of mouse hematopoietic stem cells 10.1126/science.2898810. Science *241*, 58-62.

Xie, H., Ye, M., Feng, R., and Graf, T. (2004). Stepwise Reprogramming of B Cells into Macrophages. *117*, 663-676.

Zandi, S., Mansson, R., Tsapogas, P., Zetterblad, J., Bryder, D., and Sigvardsson, M. (2008). EBF1 Is Essential for B-Lineage Priming and Establishment of a Transcription Factor Network in Common Lymphoid Progenitors. J Immunol *181*, 3364-3372.

Zhuang, Y., Soriano, P., and Weintraub, H. (1994). The helix-loop-helix gene E2A is required for B cell formation. Cell *79*, 875-884.

IV

Establishment of lineage-specific enhancer repertoires in developing hematopoietic progenitors

Summary

The ability of hematopoietic stem cells to maintain proper homeostasis in the immune system, in both health and infection, is a remarkable example of the precise control of gene expression required for cells to either preserve a pluripotent, quiescent cell state, or to differentiate and divide into appropriate downstream lineages. This precision is maintained by the transcription factors that regulate gene expression and the epigenetic state of the chromatin to allow transcription factor accessibility. Here, we harness the power of an inducible multipotent cell line, Id2-HPC, to study the changes in histone 3 lysine 4 (H3K4) methylation patterns in the lineage commitment choice from pluripotent progenitor to either the B or myeloid lineage, and to probe temporal control of early B cell development. This analysis has allowed the prediction of novel and functional enhancer elements in immune development

Introduction

In the adult bone marrow, long-term hematopoietic stem cells (LT-HSCs) have the ability to both self-renew and reconstitute the entire immune system for the life of the organism (Spangrude et al., 1988). LT-HSCs have the ability to differentiate into short-term HSCs (ST-HSCs), which give rise to multipotent progenitors (MPPs) and then differentiate into the lymphoid-primed multipotent progenitors (LMPPs) (Adolfsson et al., 2005). LMPPs differentiate into common lymphoid progenitors (CLPs) or granulocyte/macrophage progenitors (GMPs) which give rise to a lymphoid or myeloid restricted pathway, respectively (Inlay et al., 2009; Kondo et al., 1997).

During the past decade, transcription factors have been identified that play critical roles in early B cell development. These include the E2A proteins, EBF1, FOXO1 and PAX5. The E2A proteins are members of the E-protein transcription factor family that also include

HEB, E2-2 and E2A. E-proteins belong to a family of transcription factors that is characterized by the presence of a helix-loop-helix (HLH) dimerization domain. The E2A locus encodes for two proteins, E12 and E47. E12 and E47 arise from differential splicing of exons that encode for their DNA binding and dimerization domains. In mice deficient for E2A, B-cell development is blocked at the CLP cell stage (Bain et al., 1997; Bain et al., 1994; Zhuang et al., 1994). A similar block in B cell development is observed in EBF-, FOXO1- and PAX5-deficient mice (Dengler et al., 2008; Lin and Grosschedl, 1995; Nutt et al., 1997).

Previous observations have established that, when cultured in B-cell supportive conditions *in vitro*, E2A-deficient bone marrow cells self-renew indefinitely without losing multipotent differentiation potential (Ikawa et al., 2004). B-cell differentiation in E2A -/- cells can be rescued by forced EBF expression, suggesting that the E2A proteins act upstream of EBF in B-cell development (Hagman and Lukin, 2006; Lazorchak et al., 2005; Seet et al., 2004). Once EBF expression is induced, E2A and EBF, in turn, activate the expression of Pax5, Foxo1, and Bcl11a to promote the B-cell fate (Decker et al., 2009; Dengler et al., 2008; Liu et al., 2003; Roessler et al., 2007). The precisely controlled expression of these genes mediates proper development and prevents aberrant cell growth or differentiation. This precision is controlled by direct and indirect binding of transcription factors, as well as epigenetic changes such as DNA methylation or histone modifications (Fisher, 2002; Hagman and Lukin, 2006). Histone tails can be modified by acetylation, methylation, ubiquitination, and phosphorylations. Recent work has illuminated specific roles for many of these modifications, especially methylation of H3K4. Trimethylation of H3K4 typically marks active promoters, while H3K4 monomethylation is an accurate mark of active enhancer regions (Heintzman et al., 2009; Heintzman et al., 2007) .

We showed previously that overexpression of the E-protein inhibitor, Id2, from a TetOff promoter in bone marrow progenitor cells allows for the expansion of a progenitor cell line that maintains multipotency both *in vivo* and *in vitro* (Chapter III). Here, we harnessed the unlimited growth capacity and broad differentiation potential of these cells to monitor how enhancer repertoires are established in order to specify the B or myeloid cell fate. Briefly, we have examined genome-wide E2A occupancy as well as patterns of H3K4me1 and H3Kme3 during developmental progression. The data indicate how distinct combinations of a few transcriptional regulators can transiently activate or suppress a spectrum of lineage- and stage-specific enhancer repertoires as cells progress to the committed cell stage.

Results

Genome wide analysis of histone methyl marks in differentiating cells

The gene expression patterns uncovered by microarray analysis in the Id2-HPCs indicate genome-wide changes in transcriptional activity (Figure 3.7). We are now faced with the question as to how these changes are established. Recent studies have demonstrated a tight correlation between transcriptionally active promoters and H3K4 trimethylation, whereas H3K4 monomethylation has been associated with enhancer activity (Pokholok et al., 2005; Heintzman et al., 2007). To determine whether the changes in the gene expression pattern correlate with changes in H3K4me3 as well as H3K4me1, cell lysates derived from Id2-HPCs as well as differentiated myeloid and B-lineage cells were immunoprecipitated with antibodies directed against H3K4me1and H3K4me3 and analyzed by ChIP-Sequencing (Barski et al., 2007). Over 15,000 regions were identified as containing H3K4me3 in Id2-HPCs, as well as differentiated B and myeloid cells. The number of H3K4me1 islands varied from 20,813 (Id2-

HPCs) to 19,905 (myeloid cells) to 16,091 (B-lineage cells) (Table 4.1). As expected, the majority of the H3K4me3 peaks are promoter proximal, while most of the H3K4me1 peaks are promoter distal (Figure 4.1A).

The presence of H3K4 mono- or trimethylation, either proximally or distally, correlated with an increase in absolute gene expression at all timepoints (Figure 4.1B). To compare the lineage specific changes in H3K4 methylation to gene expression on a global scale, we plotted tag counts from high confidence H3K4me1 distal peaks and H3K4me3 proximal peaks in scatter plots to determine peaks that had changes in tag count between varying timepoints. We then correlated peaks with changes in tag counts to mRNA expression associated with the nearest gene (Figure 4.2 and 4.3). We observe that in all cases, when H3K4 mono- or trimethylation increases between two timepoints, so does the average relative level of expression.

Enhancer regions are H3K4me1 marked in hematopoietic progenitors prior to differentiation

To explore the possibility that enhancer regions are primed in Id2-HPCs prior to differentiation, we mapped H3K4me1 marks for a subset of cell-type specific genes including the B-lineage specific genes Foxo1, EBF and VpreB3. We also analyzed myeloid specific gene such as CSFR1 and CEBPa, and we examined the patterns of H3K4me1 in genes that were repressed upon specification into either the B or myeloid cell lineages, such as Thy1 (Figure 4.4). Interestingly, H3K4me1 marks are already present in Id2-HPCs prior to differentiation and are markedly elevated in genes whose expression is activated upon progressing towards a differentiated state. In contrast, the levels of H3K4 monomethylation

decline significantly in loci whose expression levels decrease upon specification to a distinct cell lineage, consistent with the global analysis described above (Figure 4.2). Taken together, these findings indicate that in Id2-HPCs, H3K4 monomethylation patterns mark a wide spectrum of enhancers primed to become either activated or suppressed during developmental progression.

Cis-regulatory codes and developmental progression

The data described above raise the question as to whether changes in the levels of H3K4 monomethylation in the B versus myeloid cell lineage correlate with the presence of distinct cis-regulatory codes. To address this question, we examined H3K4me1-associated sites for enriched DNA sequences. To correlate the differences in H3K4me1 between B and Id2-HPCs, we looked for the presence of distinct cis-regulatory codes using a *de novo* motif finder algorithm, HOMER, to identify potential sequence determinants localized within ±1000 bp from the H3K4me1 peaks. The ranking of such associated sequences is based on their enrichment as compared to background genomic DNA sequences. We searched for motifs based on the scatter plots in Figure 4.2 (shown again Figure 4.5). As expected, EBF, ETS and E2A were ranked among the top-scoring motifs associated with patterns of H3K4me1 in cells with at least a two-fold increase in tag counts in B cells compared to Id2-HPCs (Figure 4.5A). In Id2-HPCs compared to B cells, H3K4me1 regions were enriched for both RUNX1 and PU.1 consensus binding sites (Figure 4.5A). When comparing myeloid cells to Id2-HPCs, peaks associated the myeloid lineage were significantly enriched for the PU.1 and RUNX consensus binding sites but lacked E2A and EBF motifs as compared to both Id2-HPCs and B cells (Figure 4.5B and 4.5C). Interestingly, peaks associated with Id2-HPCs compared to B

cells also showed significant enrichment for PU.1 (Figure 4.5A). On the other hand, E2A consensus binding sites were primarily enriched in Id2-HPCs compared to myeloid cells. These data are consistent with a model in which in progenitor cells, H3K4me1 marked enhancers are primed to induce either a B- or myeloid- specific program of gene expression. In progenitor cells, activation of E2A in conjunction with EBF1 would promote the activation of a B-lineage specific transcription signature, while conditions that favor myeloid development would lead to decreased H3K4me1 association with E2A associated genes. This suggests that the establishment of the pluripotent gene program occurs at a stage of development prior to Id2 overexpression, and this program is maintained in the Id2-HPCs until the cells are directed into alternate pathways. Upon addition of doxycycline or myeloid specific cytokines, the cis-regulatory code involved with active enhancers changes to favor B or myeloid specific signatures.

Genome wide enhancer analysis during the pre-pro-B to pro-B differentiation

One of the main advantages of the Id2-HPCs is that it allows for analysis of intermediate events in differentiation, not just the endpoints. We thus used the Id2-HPCs to analyze intermediate changes during the pre-pro-B to pro-B transition. To this end, we differentiated Id2-HPCs into CD19+ pro-B cells and isolated RNA at 7 different timepoints during developmental progression. We then analyzed gene expression by microarray analysis and performed hierarchical clustering on genes that had a 2-fold or greater change in expression at any timepoint during development. From this clustering, we observed three main expression patterns: I) transcript levels that were elevated, II) transcript levels that were initially increased but eventually declined, and III) transcript levels that declined during

developmental progression (Figure 4.6). We note that enrichment of genes in clusters I and III (showing elevated and declining transcript levels, respectively) is similar as described above (Figure 4.6). Activated genes include Ebf1, Foxo1 and Rag1 and repressed genes include genes involved in alternate lineages, such as Hbb-b1, Tox, and Tcra, as well as loci potentially involved in the maintenance of the HSC phenotype, including members of the Hoxa family (Figure 4.6, Table 4.2). Interestingly, a group of genes, including Ly6D, Bcl2l1 and Xbp1, showed a transient pattern of expression (Figure 4.6).

To identify the cis-regulatory codes that underpin these patterns of gene expression, we performed ChIP-Seq to define the H3K4me1 islands at 0-hours, 48-hours, and 120-hours post differentiation, and analyzed the enriched transcription factor motifs present at active enhancers at each of the developmental stages using the HOMER algorithm as described in above. We compared transcription factor motif association at H3K4me1 peaks that have 2-fold or more changes in tag counts between any two timepoints (Figure 4.7). We see enrichment of RUNX and ETS binding sites in 0-hour specific peaks relative to both the 48-hour and 120-hour timepoint. Additionally, there is specific enrichment of the myeloid associated motifs PU.1 and AP1 in 0-hour peaks relative to the 120-hour peaks, suggesting a loss of lineage plasticity by 120-hours. Interestingly, peaks with increased tag-counts at 48-hours relative to 0-hours have EBF and E2A enrichment, whereas peaks increased at 48-hours relative to 120-hours still maintain PU1 and RUNX DNA binding motifs. At 120 hours, there is strong enrichment of EBF and E2A relative to both 0-hour and 48-hour peaks. This suggests the coordinated activity of E2A, EBF and ETS transcription factors to activate the B cell program, and RUNX and ETS to maintain a lineage primed phenotype.

To identify the regulatory code at active enhancers for specific patterns of gene expression, we performed a combinatorial analysis on the three patterns of gene expression identified from the microarray with the H3K4me1 ChIP-Seq from 0-hours, 48-hours, and 120-hours (Figure 4.8). Distinct patterns emerged from the analysis. Cis-regulatory coding elements associated with H3K4me1 islands in Id2-HPCs were enriched for E2A binding sites but did not show linkage with EBF consensus binding sites (Figure 4.8, left panels). However, as predicted, upon differentiating into the B cell lineage, the cis-regulatory codes associated with B-cell specific peaks were significantly enriched for consensus E2A as well as EBF binding sites (Figure 4.8, top right panel). Interestingly, loci whose expression was suppressed upon developmental progression showed a similar cis-regulatory pattern (Figure 4.8, lower right panel). These data indicate that EBF-associated enhancer repertoires, for both upregulated and repressed genes, become H3K4 monomethylated upon commitment to the B cell lineage.

In contrast to the presence of the EBF motif in both upregulated and repressed genes, the FOXO consensus binding motif was only associated with H3K4me1 peaks whose related transcript levels were elevated during developmental progression (Figure 4.8, top right panel) but not with suppressed genes (Figure 4.8, lower right panel). The appearance of EBF as well FOXO associated H3K4me1 regions was only observed at a stage in which EBF and FOXO1 transcript levels are elevated (Figure 4.6).

Also of interest are the cis-regulatory codes linked with cluster II. Cluster II includes a subset of transcripts that are temporally elevated during the early stages of B cell development. The cis-regulatory codes associated with H3K4me1 regions in cluster II are distinct from those linked with clusters I and III. Most notable is their association with ETS-

RUNX composite sites and lack of E2A as well as EBF consensus motifs (Figure 4.8, middle panels). In sum, in a subset of enhancer elements linked to genes that are transiently activated, other factors, for example RUNX in conjunction with PU.1 and/or related ETS proteins such as Erg1, Ets1, Ets2 or Fli1, may act to transiently induce the expression of a distinct set of genes that include Ly6D, Bcl2l1,Jun and Fos. Additionally, the data suggest that the spectrum of enhancers for both activated and repressed genes in B-cell development are responsive to E2A. We propose that once E2A activity is restored, E2A actively promotes nucleosome remodeling at a wide spectrum of cis-regulatory elements to direct activation of both the EBF and FOXO1 genes, as well as to permit EBF and FOXO1 to bind and generate a repertoire of B-lineage specific enhancer and, in concert with EBF, to suppress enhancers associated with alternative cell lineages.

Thus, we have identified a gradient of changes in cis-regulatory binding motifs at active enhancers during B-cell development. At the multipotent pre-pro-B stage, genes maintain their expression during culture through the presence of ETS, E2A and RUNX binding motifs at active enhancers. In genes that become activated or repressed, association of both E2A and EBF motifs at active enhancers increases, and the FOXO motif appears at enhancers linked to activated genes. In contrast, the enhancers of genes that are intermediately activated retain their association with ETS/PU1 and RUNX motifs, but not E2A or EBF motifs.

Genome-wide analysis of E2A binding in CD19+ Id2-HPCs

The data described above suggest that E2A is a critical determinant in activating an enhancer repertoire that promotes the developmental progression of B-lineage cells. The observations also indicate that E2A associates with EBF and FOXO1 at a subset of regulatory elements in differentiated pro-B cells. To confirm these data, Id2-HPCs were differentiated for a five days period. Committed B-lineage cells were purified based on the expression of CD19, lysed and immunoprecipitated with an antibody directed against E2A. Immunoprecipitates were then analyzed for global E2A occupancy using ChIP-Seq. 10,889 high-confidence E2A peaks were identified at 120 hours post-doxycycline addition (Table 4.1). The majority of E2A peaks lie in the intergenic and intronic regions, but there is also significant binding in promoter and 5'UTR regions (Figure 4.9A). This is not surprising as E2A has well-defined promoter and enhancer activity for several genes (Greenbaum and Zhuang, 2002; Ikawa et al., 2006; Xu and Kee, 2007).

When we compare E2A peaks to the absolute expression levels of the nearest genes, we see that genes associated with E2A binding have a significant increase in relative expression ($p=.0095$) when compared to genes lacking E2A binding (Figure 4.9B). We observed this increase in expression with both proximal and distal E2A peaks (Figure 4.9C). Genes with one or more E2A peaks have a significant increase ($p<0.001$ for all cases) in absolute expression at 120-hour post-doxycycline addition, and there is also a significant increase ($p=0.060$) in expression of genes that have 4 or more E2A peaks relative to genes that only have one E2A peak (Figure 4.9D).

Further confirming the important role of E2A in controlling gene expression during pre-pro-B to pro-B cell differentiation, we see that nearly half of genes with changes in gene

expression have direct E2A binding. 1,110 genes showed greater than 2-fold changes in expression during the differentiation assay, and of those 517 were bound by E2A (Fig 4.10A). 593 genes show differential regulation but have no E2A binding, possibly indicating long-term long-range genomic interactions. The top 50 repressed and activated genes after 120 hours of culture in doxycycline that are bound by E2A are listed in Supplementary Table 5. Genes with known roles in B cell differentiation such as VpreB1, Ebf1, Foxo1, CD19 and Rag1 are all bound by E2A and upregulated, while genes involved in alternate lineage development such as Tox, Dlk1 and Zap70 and CEBPα are bound by E2A and repressed (Figure 4.10B and Table 4.3). This data is in agreement with previous studies identifying E2A as a direct regulator of many genes involved in lymphocyte development (Cobaleda and Busslinger, 2008; Friedman, 2002; Rothenberg and Taghon, 2005).

Top clustered biological process terms for bound and changed genes shows that genes that are bound by E2A and upregulated are involved in positive regulation of cellular processes, cellular differentiation, hematopoiesis, localization and apoptosis. Gene ontology terms for E2A bound and repressed genes include negative regulation of cell differentiation and regulation of myeloid development (Figure 4.10B and Table 4.4). Thus, our data shows new evidence for direct regulation by E2A of many genes known to be involved in lineage commitment, such as Irf8, CEBPα, CEBPβ, SpiB, Bcl11a, EBF1, and Foxo1 (Busslinger, 2004; Murre, 2005).

To determine whether E2A occupancy is associated with EBF binding sites, we again used the *de novo* motif finder algorithm HOMER to identify enriched motifs within 200bp of E2A peaks. As expected the canonical CAGCTG E2A motif was the top-scoring motif (Figure 4.10C). As predicted by the analysis described above, genomic sites with E2A

occupancy were significantly enriched for EBF consensus binding sites (Figure 4.8). Taken together, these data indicate that an increase in E2A activity in developing B-lineage cells acts in concert with EBF to induce the expression of a wide spectrum of B-lineage specific genes.

Correlation of E2A and H3K4me1 in distal elements

Since we saw prevalent E2A binding in intronic and intergenic regions and since E2A associates with many differentially regulated genes, we performed a combinatorial analysis of E2A and H3K4me1 ChIP-Seq data on a global scale. To determine the association of E2A binding and H3K4 monomethylation, we plotted the distance of distal H3K4me1 peaks relative to E2A peaks over a 6kb region (Figure 4.11A). We observed a strong bimodal distribution of H3K4me1 with a dip around E2A binding peaks (Figure 4.11B). To investigate the possibility of a specific cis-regulatory code associated with E2A-H3K4me1 peaks, we performed *de novo* motif analysis (Figure 4.11C). At E2A peaks that are associated with H3K4me1 peaks, we noticed strong enrichment of the canonical E-box, CAGCTG, as well as the CACCTG E-box. In contrast, E2A peaks that were not associated with H3K4me1 enrich a non-specific-SP1/KLF GC-box motif, as well as the CAGGTG and CACCTG E-boxes. In addition, genes associated with distal E2A peaks in proximity to H3K4me1 peaks showed a significant ($p=0.0009$) increase in expression when compared to distal E2A peaks without H3K4me1 (Figure 4.12A). Finally, distal E2A peaks in proximity to H3K4me1 showed much higher enrichment for other transcription factor binding sites such as EBF, ETS, and RUNX than E2A peaks without H3K4me1 (Figure 4.12B).

These data suggest that the recruitment of E2A to the distal, canonical CAGCTG E-boxes is strongly correlated with active enhancer marks and increases in gene expression,

most likely through cooperative binding of other transcription factors. In contrast, distal E2A peaks not associated with active enhancers are bound preferentially at non-canonical E-boxes, and the genes associated with these peaks do not show upregulation of expression levels.

Delineation of novel enhancer elements by the association of E2A and H3K4me1

Since we found correlation between the presence of E2A binding and H3K4 monomethylation and gene expression during B cell differentiation, we mapped E2A binding sites and H3K4 monomethylation to the cell specific genes Foxo1 and CEBPα loci using the UCSC Genome Browser (Figure 4.12A). The Foxo1 gene has several strong E2A peaks that correlate with an increase in H3K4 monomethylation. We then chose 4 regions in the Foxo1 locus or putative enhancer region with strong E2A binding and H3K4monomethylation for further functional studies.

Potential enhancer elements were selected as 300-400 bp regions centered on E2A peaks that correlated with increases in H3K4me1 peaks at 120-hours (Figure 4.12A). All four regions contained conserved E-boxes, suggesting direct recruitment of E2A to the sites by E2A DNA binding motifs (Figure 4.12B). These regions were cloned by PCR and inserted upstream of a basal promoter in a modified enhancer pGl3 vector and transfected, with control renilla luciferase, into the pro-B cell line 22D6. Luciferase activity was quantitated after 24 hours, and all 4 regions showed significant enhancer activity over the empty vector alone (Figure 4.12C). When E-box sites present in the regions were mutated from CANNTG to CANNCC, enhancer activity was significantly reduced, showing that E2A specific recruitment to these regions is required for proper enhancer activity. (Figure 4.12B and 4.12C). Furthermore, sites with the canonical CAGCTG E-boxes had stronger association with

H3K4me1 at 120-hours, and showed stronger luciferase activity. Thus, the co-association of E2A and H3K monomethylation in concert with the specific sequence of E-boxes allows us to predict novel enhancer elements in the genome during B-cell differentiation.

Discussion

The transcriptional and epigenetic activities that control hematopoietic differentiation must be extremely precise to ensure proper immune homeostasis. Here, we use a novel self-renewing and multipotent cell line, Id2-HPC, to study the cis-regulatory elements involved in modulation of active enhancer activity. Id2-HPCs maintain a multipotent state through the inducible overexpression of Id2; the cells can be readily differentiated *in vitro* in the presence of proper cytokine stimulation and doxycycline. As such, we differentiated Id2-HPCs into either the B or myeloid cell lineage and studied the changes in the active enhancer mark, H3K4me1 at different lineage commitment states.

H3K4me1 binding patterns on genes known to be involved in either B, T, or myeloid lineage fates was analyzed at a multipotent developmental stage, B cell committed stage, or myeloid committed stage of development. This gene specific analysis suggested that binding of H3K4me1 is already present at enhancers at a multipotent stage of development, and upon differentiation, H3K4me1 increases or decreases in concert with gene expression. This suggests that, even at a multipotent state, H3K4 monomethylation is present at lineage specific genes to prime them for future activation.

We then expanded this analysis to study the genome-wide association of H3K4 monomethylation with transcription factor motifs to predict cis-regulatory patterns involved in lineage specification. In the Id2-HPC specific peaks, we see strong enrichment for the RUNX and ETS motifs. When the cells are directed to differentiate to the myeloid cell lineage,

H3K4 monomethylation loses affinity for ETS motifs and gains affinity for the myeloid – specific PU.1 motif. In contrast, ETS, E2A and EBF motifs are strongly enriched at B-cell specific H3K4me1 peaks. This suggests a model of regulation in which, at active enhancers, RUNX and ETS factors maintain a stem-like phenotype, PU.1 replaces ETS in the myeloid lineage, and E2A and EBF replace ETS in ato promote differentiation towards the B cell lineage.

While this model of specific transcription-factor recruitment to active enhancers is notable, it is based only on the endpoints of differentiation and, as such, cannot answer questions about intermediate changes in development. To address this, we used the Id2-HPC cell line to study the changes in gene expression and H3K4 monomethylation during the progression of pre-pro-B to pro-B cell development. Microarray analysis confirms the expected upregulation of B-cell specific genes such as Ebf, Foxo1 and Bcl11a, and the downregulation of alternate lineage genes such as Cebpα, CD34, and Zap70. Surprisingly, there is a subset of genes that is upregulated at an intermediate timepoint and then re-repressed at the end of the B-cell differentiation timecourse. Several anti-apoptosis factors, such as Bcl2l1, are present in this group of genes, suggesting that maintenance of cell survival during development is a critical and temporally mediated process. Also, the plasma cell differentiation gene, Xbp1, is also upregulated during developmental progression, suggesting a novel function for Xbp1 in early B cell development as well.

To determine the cis-regulatory elements involved in this temporal regulation of genes, we studied H3K4me1 peak changes in combination with changes in gene expression. In multipotent Id2-HPCs, H3K4me1 peaks associate with RUNX, ETS and E2A motifs. Interestingly, by the 120-hour B cell timepoint, enhancers of genes with an increase in expression associate with E2A, EBF and FOXO motifs, while enhancers of genes with a

decrease in expression are only associated with E2A and EBF motifs. This suggests a regulatory mechanism in which E2A and EBF are recruited to many genes that are changed in B-cell development, but the combination of E2A, EBF and Foxo1 is specific to activated genes. Interestingly, enhancers of genes that are intermediately changed do not show association with EBF, but maintain the association with ETS and RUNX factors.

As these analyses confirm the role of E2A's role in B-cell development, we analyzed the binding of E2A to the genome in CD19+ pro-B cells. This analysis revealed that E2A directly binds to nearly half of the genes that are changed in B-cell development. This binding significantly correlates to an average increase in gene expression, although there are many specific cases in which E2A is bound to repressed genes. When we combined E2A binding with H3K4me1 binding, we found that gene expression increases when E2A and H3K4me1 are concurrently bound at an enhancer, and expression decreases if only E2A is present. Surprisingly, we see E-box binding bias for E2A depending on the presence of H3K4me1. When E2A and H3K4me1 are both bound, E2A is recruited to canonical CAGCTG E-boxes. In contrast, when E2A is not correlated with H3K4 monomethylation, it is preferentially recruited to CACCTG or CAGGTC E-boxes. This is the first reporting of E-box preference as related to gene expression. It is tempting to hypothesize that perhaps homodimers of E47 are binding to canonical E-boxes and recruiting H3K4me1, whereas perhaps E47/HEB heterodimers are binding to non-canonical E-boxes to recruit co-repressors.

Materials and Methods

Cell Culture

Id2-HPCs were maintained in IMDM + 10% FCS, 2% PSG and 2μL β-me on subconfluent S17 feeder cells in the presence of 1:100 IL-7, 1:100 Flt3L and 1:200 SCF. S17 feeder cells were maintained in α-MEM + 10% FBS and 2% PSG. 22D6 cells were maintained in RPMI + 10% FBS and 2% PSG.

In Vitro Differentiation Experiments

TetOff_hId2 expanded cells were depleted of small (<1-5%) numbers of CD19, CD25 and CD11b positive cells by auto-MACS. For myeloid differentiation, cells were cultured for up to 6 days in IMDM + 10% FCS and β-me in the presence of IL3, Flt3L, GMCSF and MCSF and in the presence or absence of 1ug/mL doxycycline. For B-cell differentiation cells were cultured for up to 10 days in IMDM + 10% FCS and β-me in the presence of 1:100 IL-7 and 1:200 SCF in the presence of absence of 1ug/mL doxycycline. Fresh doxycycline and cytokines were added every two days.

Flow Cytometry

Flow cytometry was performed on either a FACSCalibur or LSR II Flow Cytometer (Becton Dickinson, Franklin Lakes, NJ) and analyzed with FlowJo (Tree Star, Ashland, OR). Single cells were stained with antibodies were purchased from eBioscience (San Diego, CA).

ChIP-Sequencing and Data Analysis

The chromatin immunoprecipitation (ChIP) protocol was essentially the same as previously described (Agata et al., 2007). Briefly, the cells were cross-linked with 1% formaldehyde at room temperature. Chromatin was sheared and solubilized from whole-cell extract through sonication. The resulting lysate was spun down to be clear of cell debris. The lysate was further incubated with non-specific immunoglobulin (Ig) and Protein G Sepharose 4 Fast Flow (GE Healthcare) to remove non-specific bindings that might be resulted from Ig and the resin. The chromatin fragments were immunoprecipitated with specific antibodies overnight at 4 C. For E2A ChIP, pre-bound Dynal beads with E2A antibodies were used for immunoprecipitation. As for the histone mark ChIPs, the chromatin fragments were first incubated with specific histone mark antibodies overnight prior to precipitation by Protein G Sepharose. The immunoprecipitated complexes were washed, then reverse cross-linked overnight at 65 C. The DNA fragments were purified and analyzed by PCR. To end modify the fragments, fragment overhangs were converted into phosphorylated blunt ends by T4 DNA polymerase, Klenow polymerase, and T4 polynucleotide kinase. An ‘A’ base was added to the 3’ end of the blunt phosphorylated DNA fragments by Klenow fragment (3’ to 5’ exo minus). The oligonucleotide adapters were ligated to the ends of the DNA fragments by DNA ligase. The adaptor-modified DNA fragments were size selected between 150 to 250 bp by running on an 8% polyacrylamide gel. PCR amplification was performed on the gel-extracted DNA. Modified DNA fragments were hybridized onto the flow cell of the fully automated Illumina Cluster Generation Station to create clusters of approximately 1000 clonal copies. Six to eight million clusters were generated. The resulting high density array of clusters on the flow cell surface was sequenced by the fully automated Illumina Genome Analyzer. A 36 cycle sequencing kit was used. The data collection and analysis software mapped the sample sequences to known mouse genomic sequence to identify the ChIP DNA fragments. Peak

finding and subsequent ChIP-Seq genome wide analyses were performed using the HOMER software package, available at http://biowhat.ucsd.edu/homer.

Microarray Profiling

All RNA was prepared using RNeasy Columns (Qiagen, Seattle, WA). Gene expression profiling was performed by amplification and hybridization to the Illumina Mouse WG-6 v1.1 gene expression array according to the manufacturer's instructions.

Luciferase Reporter Assays

The pGL3-enhancer plasmid (Promega, Madison, WI, USA) was used for luciferase assays. Predicted enhancer elements were cloned into the construct through PCR cloning, and the plasmids were transiently transfected into the 22D6 pro-B cell line using the Trans*IT-LT1* transfection reagent (Mirus, Madison, WI, USA). The Dual-Luciferase Reporter Assay System (Promega) was used to quantitate luciferase activity as per the manufacturer's instructions.

Chapter IV is currently being prepared for submission for publication of the material. The working title is "Inducible Id2 expression allows for expansion of a long-termculture of multipotent hematopoietic progenitors and epigenetic studies of lineage commitment." I was the primary investigator and author of this material. Co-authors are Yin C. Lin, Suchit Jhunjhunwala., Christopher Benner, Christopher Glass, and Cornelis Murre.

Table 4.1

Tag and High Confidence Peak Counts from ChIP-Seq Experiments

Sample	Tag Count	Peak Count
E2A_120h	3850999	10889
H3K4me1_Id2-HPC	11548803	5953
H3K4me1_myeloid	10162929	4756
H3K4me1_B	10176709	16,691
H3K4me3_Id2-HPC	10658542	17,322
H3K4me3_myeloid	10861798	15,222
H3K4me3_B	10997422	16,078
H3K4me1_0h	10708901	14423
H3K4me1_48h	13053192	13916
H3K4me1_120h	11307451	12868

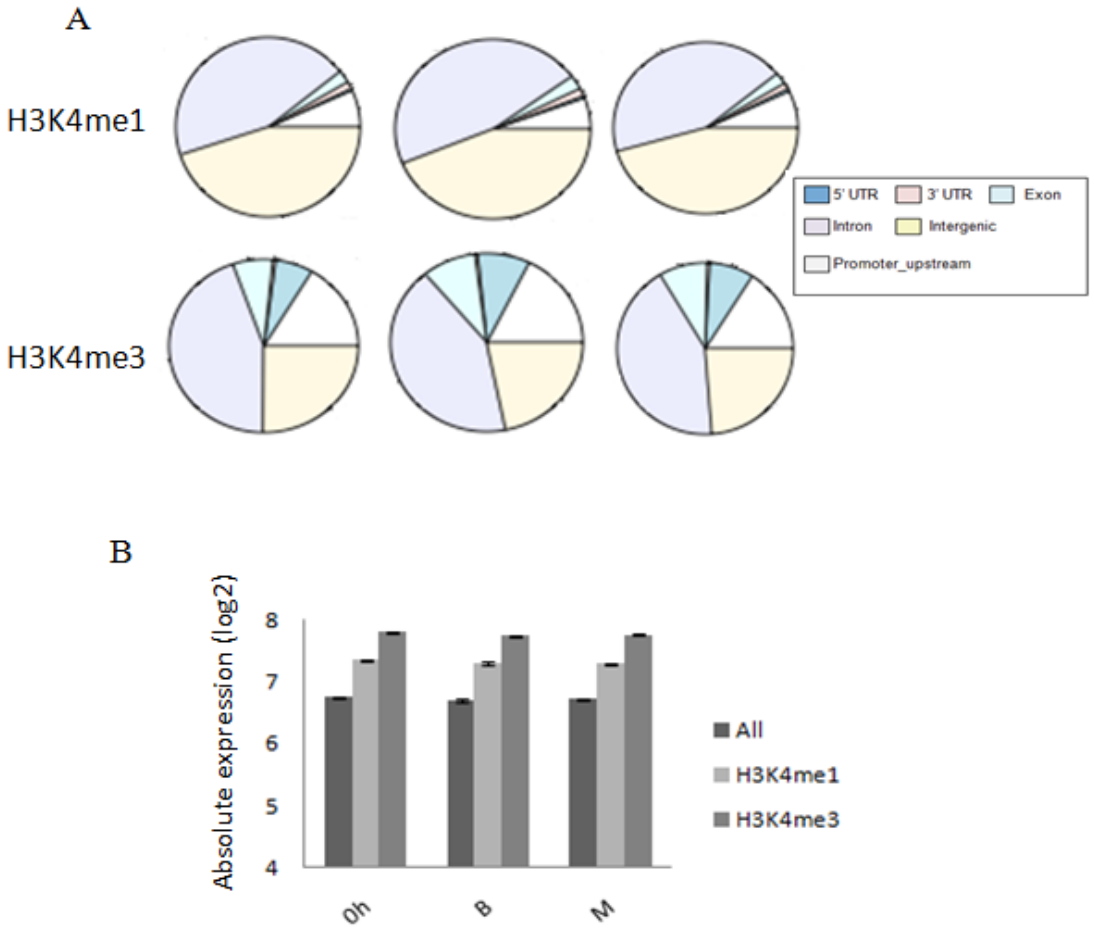


Figure 4.1

H3K4me1 and H3K4me3 binding in Id2-HPCs, B Cells, or Myeloid Cells

(A) Genomic positions of H3K4me1 and H3K4me3 at the 0-hour progenitor timepoint, the CD19+ B cell timepoint, and the CD11b+ myeloid timepoint.

(B) The presence of either H3K4me1 or H3K4me3 leads to significant increases in absolute expression level in all three cell types. Absolute expression (log2) of all genes, genes associated with H3K4me1 and genes associated with H3K4me3.

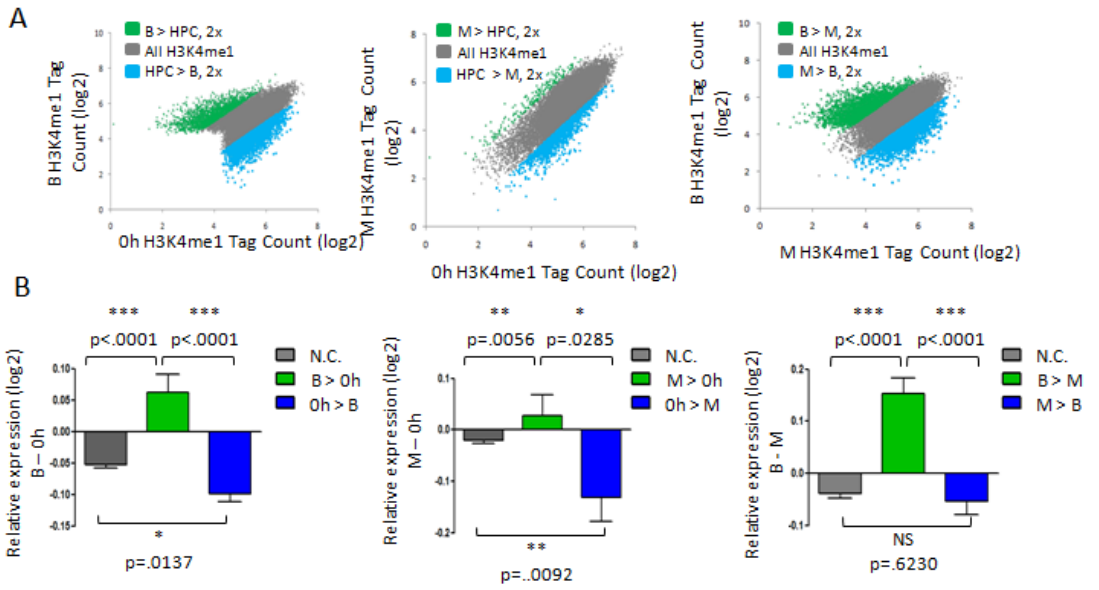


Figure 4.2

Genomic Analysis of Changes in H3K4me1 Peaks Relative to Gene Expression

(A) Comparison of tag counts (log2) of H3K4me1 0h HPC cells, CD19+ B cells, and CD11b+ M cells. Peaks with 1.5x more tags in one lineage versus another are colored green or blue, and peaks with less than a 1.5x change between the two cell types are shown in grey.

(B) Increase in proximal H3K4me1 tag counts is correlated with an increase in relative expression (log2). Expression from genes with 2x more tags in one lineage versus another are colored green or blue, and peaks with less than a 2x change in tag count are colored grey.

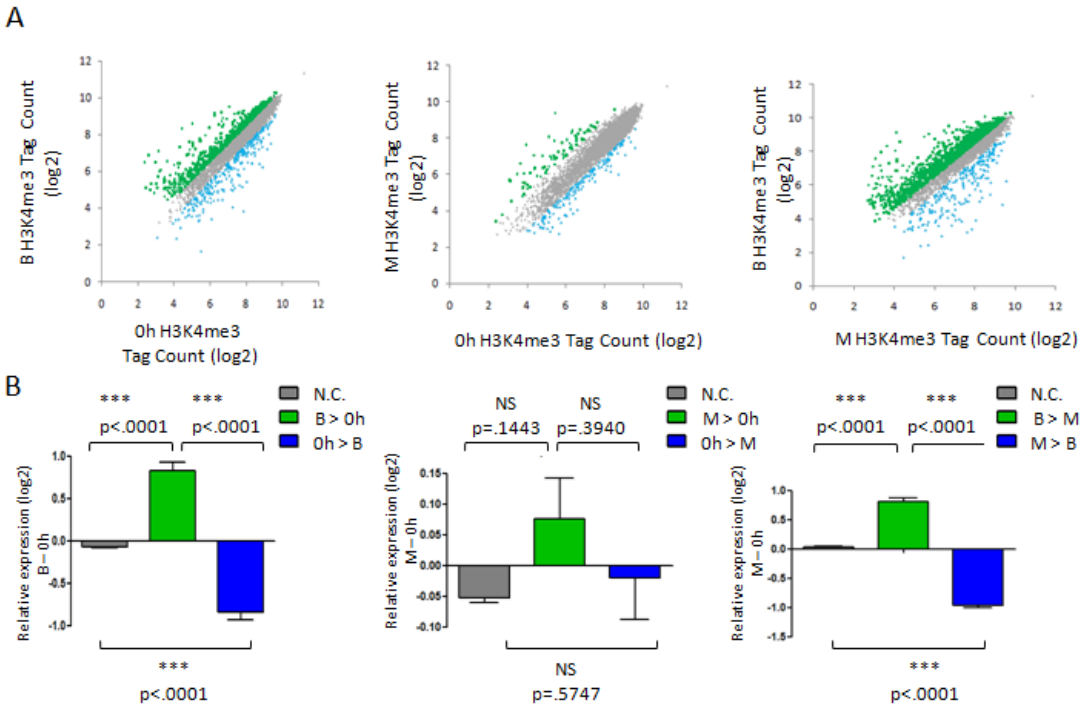


Figure 4.3

Genome-wide Analysis of Changes in H3K4me3 Peaks Relative to Gene Expression

(A) Comparison of tag counts (log2) of H3K4me3 0h Id2-HPC cells, CD19+ B cells, and CD11b+ M cells. Peaks with 1.5x more tags in one lineage versus another are colored green or blue, and peaks with less than a 1.5x change between the two cell types are shown in grey.

(B) Increase in proximal H3K4me3 tag counts is correlated with an increase in relative expression (log2). Expression from genes with 2x more tags in one lineage versus another are colored green or blue, and peaks with less than a 2x change in tag count are colored grey.

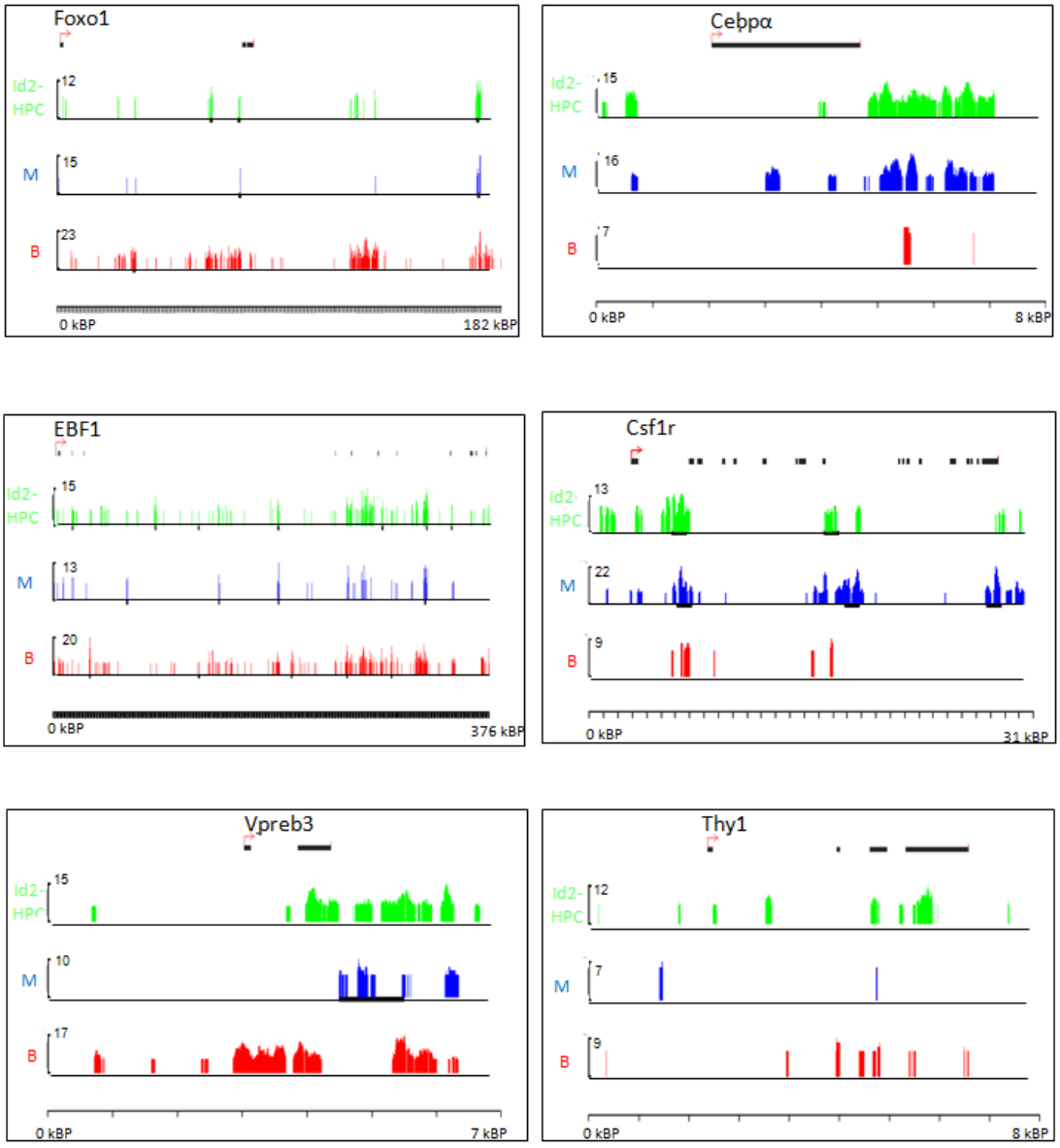


Figure 4.4

UCSC Genome Browser Images

UCSC Genome Browser images of H3K4me1 in Id2-HPCs at the undifferentiated, B-cell or myeloid cell stage at or near the Foxo1, Ebf1, Vpreb3, Cebpα, Csf1r, and Thy1 genomic positions.

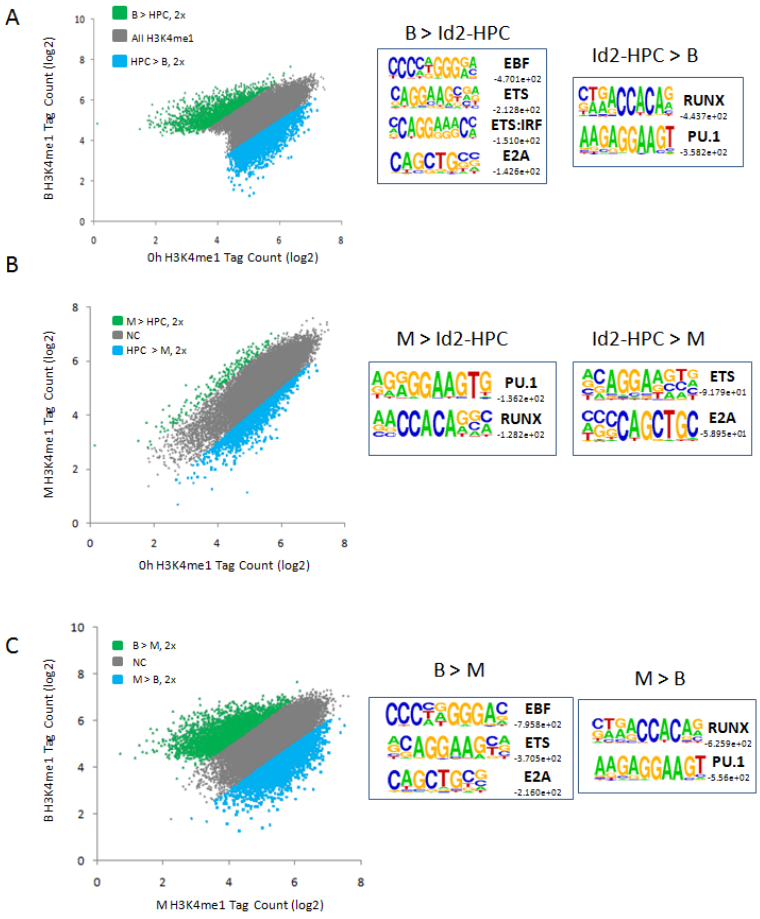


Figure 4.5

Genome-wide Analysis of Changes in Motif Enrichment at Changing H3K4me1 Peaks

(A) Analysis of changes in tag count between 0-hour progenitor peaks and CD19+ B cell peaks. Scatter plot displays tag counts from 0-hour progenitor peaks and CD19+ B cell peaks. De novo motif finding was then performed on the specific peaks.

(B) Analysis of changes in tag count between 0-hour progenitor peaks and CD11b+ M cell peaks. Scatter plot displays tag counts from 0-hour progenitor peaks and CD11b+ M. De novo motif finding was then performed on the specific peaks.

(C) Analysis of changes in tag count between CD11b+ M cell peaks and CD19+ B cell peaks. Scatter plot displays tag counts from 0-hour progenitor peaks and CD19+ B cell peaks. De novo motif finding was then performed on the specific peaks.

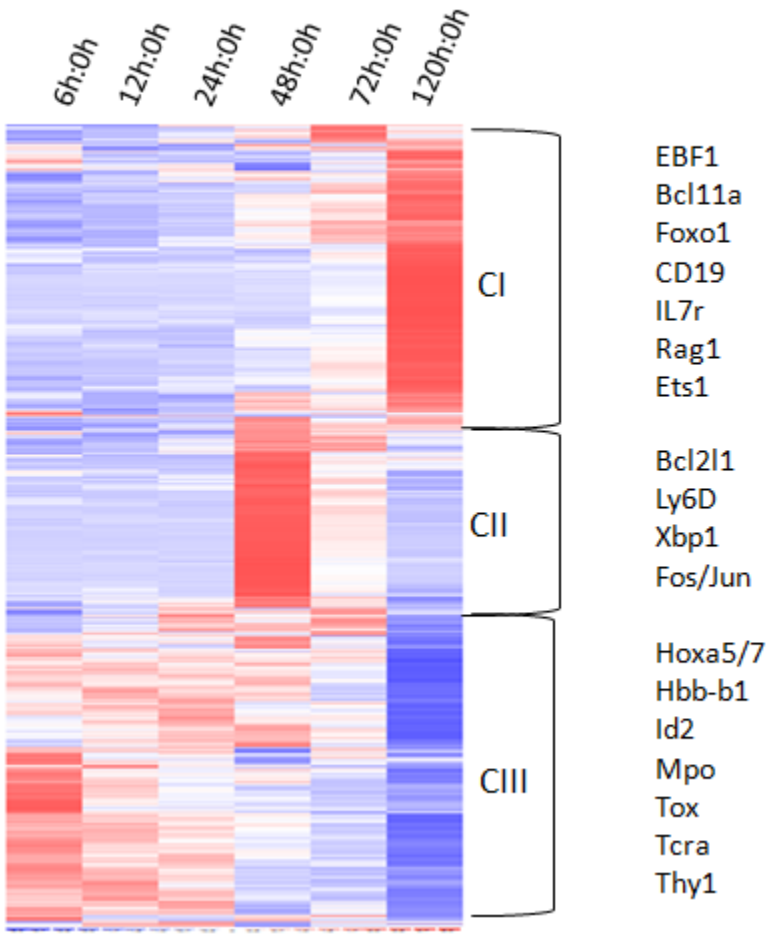


Figure 4.6

Microarray Analysis of the pre-pro-B to pro-B Transition

TetOff_hId2 cells were differentiated into pro-B cells over a 5-day timecourse, and RNA was taken at 7 timepoints for microarray analysis. Genes with a greater than 2-fold change in expression between any two timepoints were hierarchically clustered using d-Chip software

Table 4.2

Top 50 Repressed and Activated Genes at the Endpoint of the pre-pro-B to pro-B Transition

Repressed

Gene Name	F.C.
Mpo	0.015
Cox6a2	0.038
Tox	0.043
Itgb7	0.045
Xcl1	0.067
Crebbp	0.071
Il18r1	0.071
Fcer1g	0.072
Bex6	0.089
Klhl30	0.089
Samsn1	0.090
Gria3	0.091
Rnf43	0.097
Ctsw	0.098
Hbb-b1	0.098
Tsc22d1	0.105
Tmem51	0.107
Anxa2	0.108
Slc35d3	0.112
Ahnak	0.115
Fmnl3	0.117
Lmna	0.119
Tns1	0.121
Trim13	0.124
Tyrobp	0.124
LOC547380	0.125
Spns3	0.127
Slc8a1	0.127
Myo18b	0.128
Wipi1	0.129
Lpxn	0.132
Gcnt1	0.133
Nrgn	0.135
Cerkl	0.136
Glipr1	0.138
Casz1	0.141
Gcet2	0.147
Clec4d	0.150
Whrn	0.153
F2r	0.153
Tnfrsf1a	0.154
Dnajc6	0.154
8430408G22Rik	0.156
Hoxa5	0.159
A230058F20Rik	0.162
Sp7	0.166
9830130M13Rik	0.169
LOC385634	0.169
Thy1	0.169
Meis1	0.170

Activated

Gene Name	F.C.
Igl-5	109.409
Vpreb1	72.769
Nme7	60.479
Blnk	47.396
AI427515	45.722
Atp1b1	31.076
LOC100044439	29.949
Crhbp	28.478
LOC100047815	27.471
5730488B01Rik	25.265
Pou2af1	25.246
Myh10	21.677
Vpreb3	21.049
H19	21.005
Cyp11a1	20.620
Amotl1	19.045
Aqp1	19.000
Cd79b	18.820
Ly6a	18.581
Nrxn2	17.263
Apbb1	15.877
Tgm2	15.654
Mmp11	14.731
2010001M09Rik	14.637
Marcks	13.271
Uaca	13.009
Txnrd3	12.991
Lta	12.925
Clgn	12.876
Chst3	12.784
Pde2a	12.760
Ctla4	12.664
Dbndd2	12.350
Tnfsf11	12.114
Sdc4	12.031
1700027N10Rik	11.080
Cd74	10.739
Ebf1	10.428
Ng23	10.315
A730040I05Rik	9.965
Gng4	9.773
A230069K02Rik	9.497
Slc5a9	9.014
5730469M10Rik	8.620
LOC100045567	8.493
Emid2	8.199
Foxo1	8.161
Cecr2	7.970
4930513E20Rik	7.902
8430426J06Rik	7.557

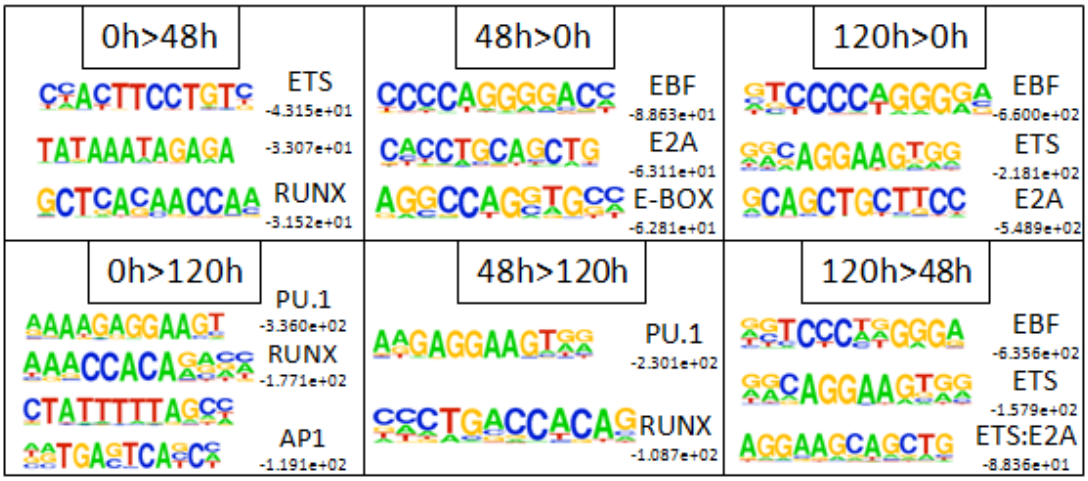


Figure 4.7

Temporal Motif Enrichment During the pre-pro-B to pro-B Cell Transition

Enriched motifs from H3K4me1 peaks that have 2-fold or greater change in tag count between 0h, 48h, and 120h during CD19+ differentiation of CD19+ cells. De novo motif finding analysis was performed on distal H3K4me1 peaks that have changes in tag counts between three timepoints in pro-B differentiation.

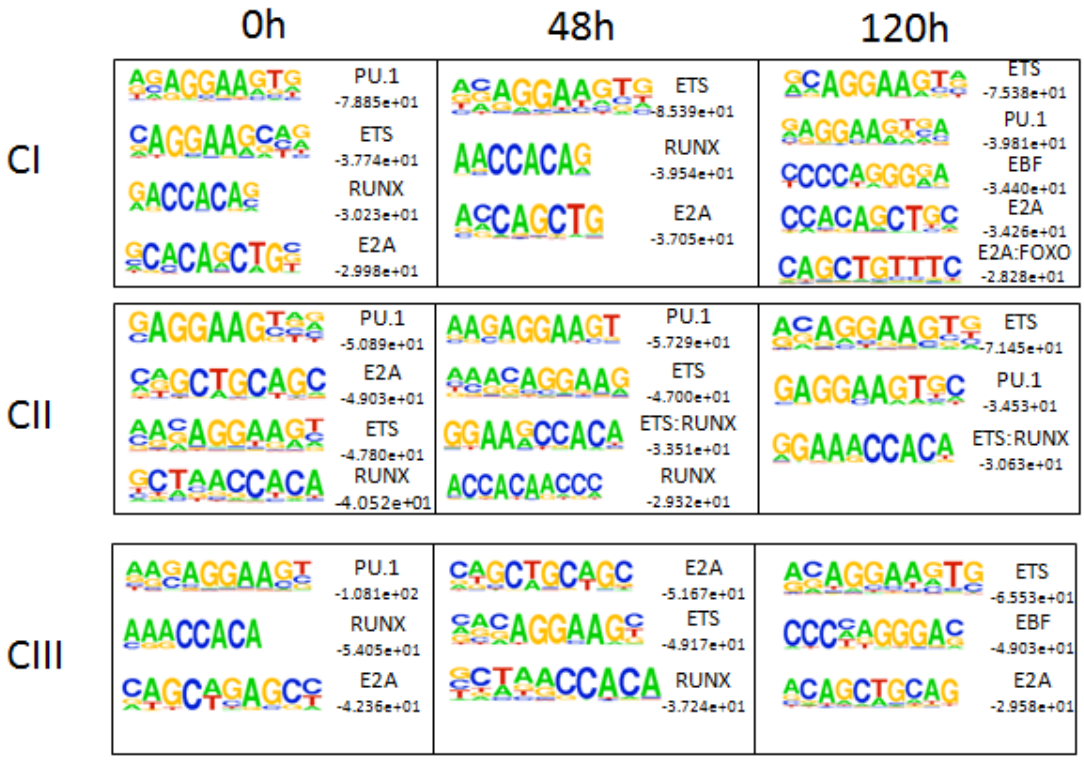


Figure 4.8

Motif Enrichment Profiles of Differentially Regulated Genes During the pre-pro-B to pro-B Transition

Enriched regulatory motifs at distal H3K4me1 sites during pro-B cell differentiation. ChIP-Sequencing was performed on TetOff_hId2 cells at 0-hours, 48-hours, and 120-hours after induction of B cell differentiation. De novo motif finding was performed to determine transcription factor binding motifs associated with active enhancers, and combinatorial analysis was done to identify cis-regulatory elements associated with active enhancers of genes that are upregulated, downregulated, or intermediately upregulated.

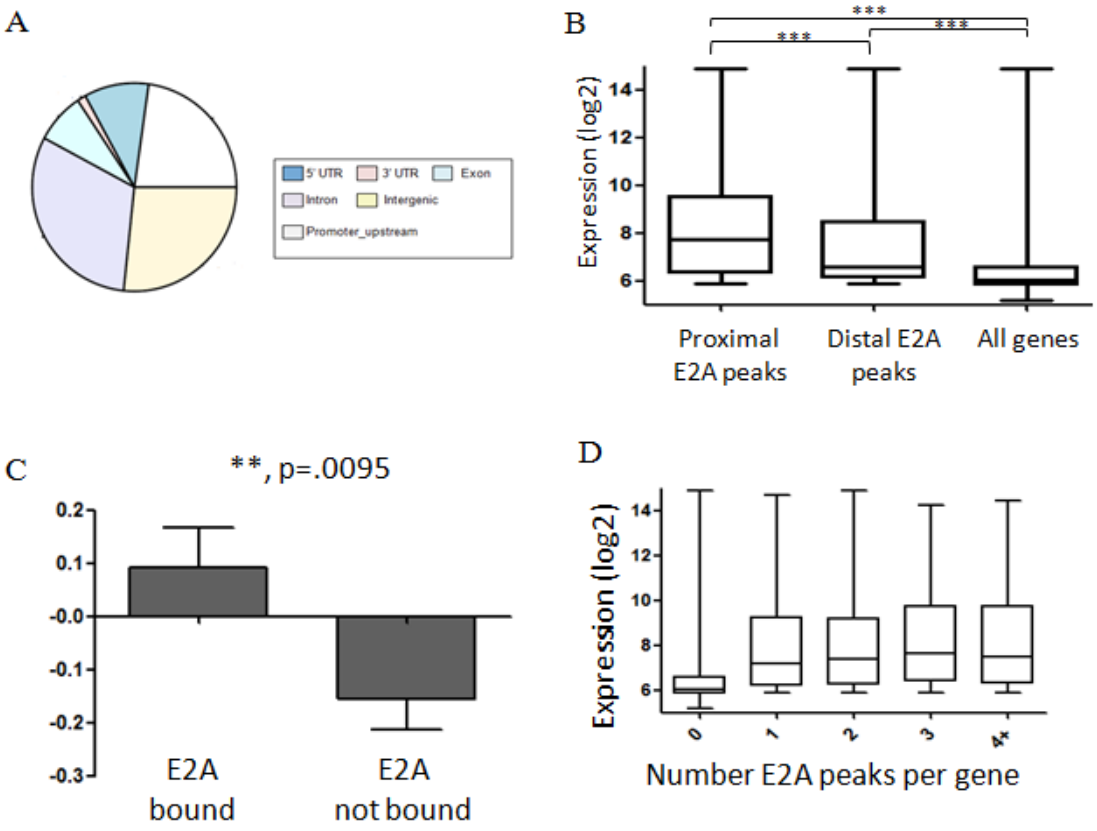


Figure 4.9

ChIP-Sequencing of E2A at CD19+ pro-B Cell Stage

(A) Genomic binding positions of E2A in Id2-HPCs at 120h post B-cell differentiation.

(B) Absolute expression of genes with proximal E2A binding, distal E2A binding, or no E2A binding.

(C) Relative expression of genes associated with E2A peaks at 120h post-B cell differentiation. Of changed genes, genes with E2A binding have a significantly higher average change in expression throughout the timecourse.

(D) Absolute expression (log2) of genes associated with no, 1, 2, 3, or 4+ E2A peaks. $p<0.0001$ for all conditions relative to 0 peaks. Significance of 4+ peaks versus 1 peak, $p= 0.095$.

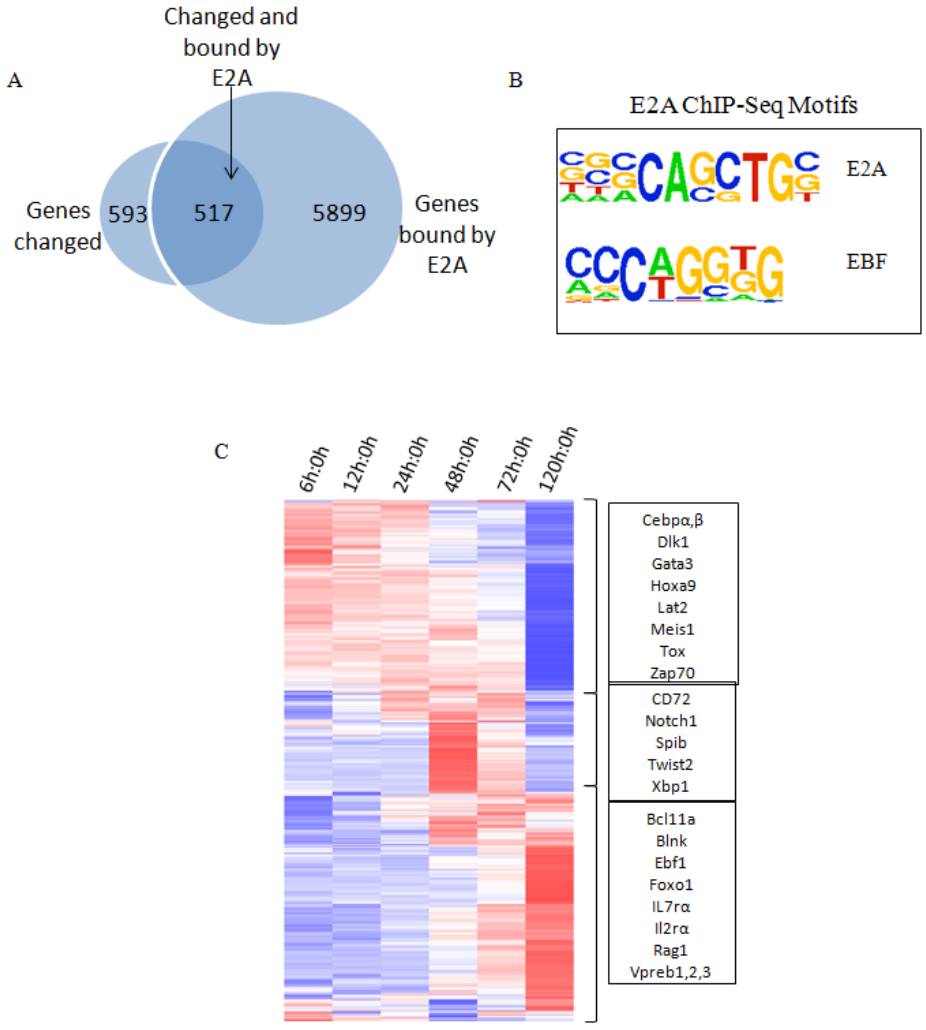


Figure 4.10

Correlation of Genome-wide E2A Binding and Changes in Gene Expression in the pre-pro-B to pro-B Transition

(A) Venn diagram displaying the overlap of genes bound by E2A that have changes in expression throughout the timecourse.

(B) Enriched sequence elements determined by de novo motif finding based on 200bp regions centered around high-confidence E2A peaks.

(C) Hierarchical clustering (by d-Chip) of microarray expression data from genes that have more than a 2-fold change in expression and are bound by E2A.

Table 4.3

Top 50 Repressed and Activated Genes that are Bound by E2A at the Endpoint of the pre-pro-B to pro-B Transition

Repressed

Gene Name	F.C.
Cox6a2	0.038
Tox	0.043
Itgb7	0.045
Xcl1	0.067
Crebbp	0.071
Il18r1	0.071
Fcer1g	0.072
Bex6	0.089
Klhl30	0.089
Ctsw	0.098
Tsc22d1	0.105
Tmem51	0.107
Ahnak	0.115
Lmna	0.119
Trim13	0.124
Tyrobp	0.124
Slc8a1	0.127
Nrgn	0.135
Whrn	0.153
F2r	0.153
Tnfrsf1a	0.154
Dnajc6	0.154
Sp7	0.166
Meis1	0.170
Dlk1	0.180
Samhd1	0.182
Sccpdh	0.191
Gng12	0.193
Ccl3	0.196
Aqp9	0.196
Ttyh2	0.197
Lat2	0.198
Cmpk2	0.205
Klrb1f	0.207
Zap70	0.208
Gpr158	0.211
Srgap3	0.222
Capn5	0.228
Aifm2	0.229
Rgs10	0.230
Zbtb7a	0.231
Ptpn11	0.233
Plec1	0.235
Arid3a	0.237
Rnase6	0.242
Ramp1	0.244
Ctr9	0.246
Hs3st1	0.247
Chn2	0.251
Cebpa	0.256

Activated

Gene Name	F.C.
Vpreb1	72.769
Blnk	47.396
Atp1b1	31.076
Vpreb3	21.049
Amotl1	19.045
Apbb1	15.877
2010001M09Rik	14.637
Marcks	13.271
Uaca	13.009
Txnrd3	12.991
Lta	12.925
Chst3	12.784
Pde2a	12.760
Ctla4	12.664
Tnfsf11	12.114
Sdc4	12.031
Rsph9	11.080
Cd74	10.739
Ebf1	10.428
Ng23	10.315
5730469M10Rik	8.620
Emid2	8.199
Foxo1	8.161
Cecr2	7.970
Cplx2	7.553
Rag1	7.357
Grb10	7.235
1700011H14Rik	7.146
Adcy6	6.561
Akap12	6.501
Enpp6	6.241
Il7r	6.091
Gna15	6.069
Cd19	6.034
Bfsp2	5.884
Chchd6	5.843
Rras2	5.834
Slc5a9	5.580
Sbk1	5.562
Serpinf1	5.526
Map4k2	5.494
Hes1	5.410
Socs3	5.345
Bach2	5.271
Pcmtd1	5.245
Prkcb	5.229
Nedd9	4.995
Trp53i11	4.892
Slco4a1	4.803
Egr1	4.700

Table 4.4

Gene Ontology of Clustered Microarray Data from Genes with Changes in Expression and E2A Binding

Array Cluster	Gene Ontology Clusters
I	Regulation of cellular processes
	Negative regulation of cell differentiation
	Immune system development
	Apoptosis
	Regulation of myeloid development
II	Multicellular development
	Biological regulation
	Tissue development
	Cell adhesion
	Negative regulation of cell proliferation
III	Positive regulation of cellular processes
	REgulation of cellular differentiation
	Cell activation and hemopoiesis
	Localization
	Apoptosis

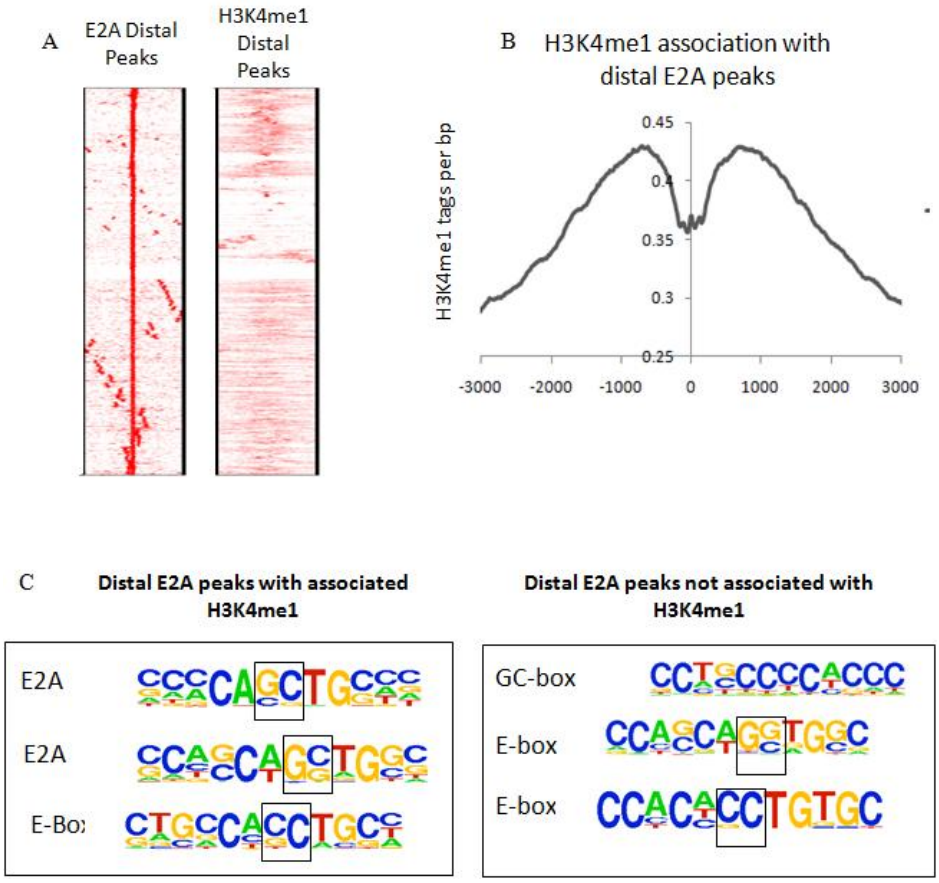


Figure 4.11

Correlation of E2A and H3K4me1 Binding in pro-B Cells

(A) Heatmap of E2A and H3K4me1 binding. Each column represents a 6-kb window based on promoter distal E2A binding at 120-hours post-doxycycline addition. The left panel shows the centered E2A peaks and the right panel shows H3K4me1 binding relative to the E2A positioning. shows all TSS distal peaks.

(B) E2A peaks correlate with H3K4me1. Genomic distances of H3K4me1 binding relative to distal E2A peaks are plotted over a 6kb window.

(C) Association of distal E2A peaks and H3K4me1 reveals E-box preference. De novo motif finding was performed for E2A peaks that have associated H3K4me1 binding within 2kb (left panel), and for E2A peaks that do not have associated H3K4me1 binding within 2kb (right panel). The central E-box sequences are highlighted.

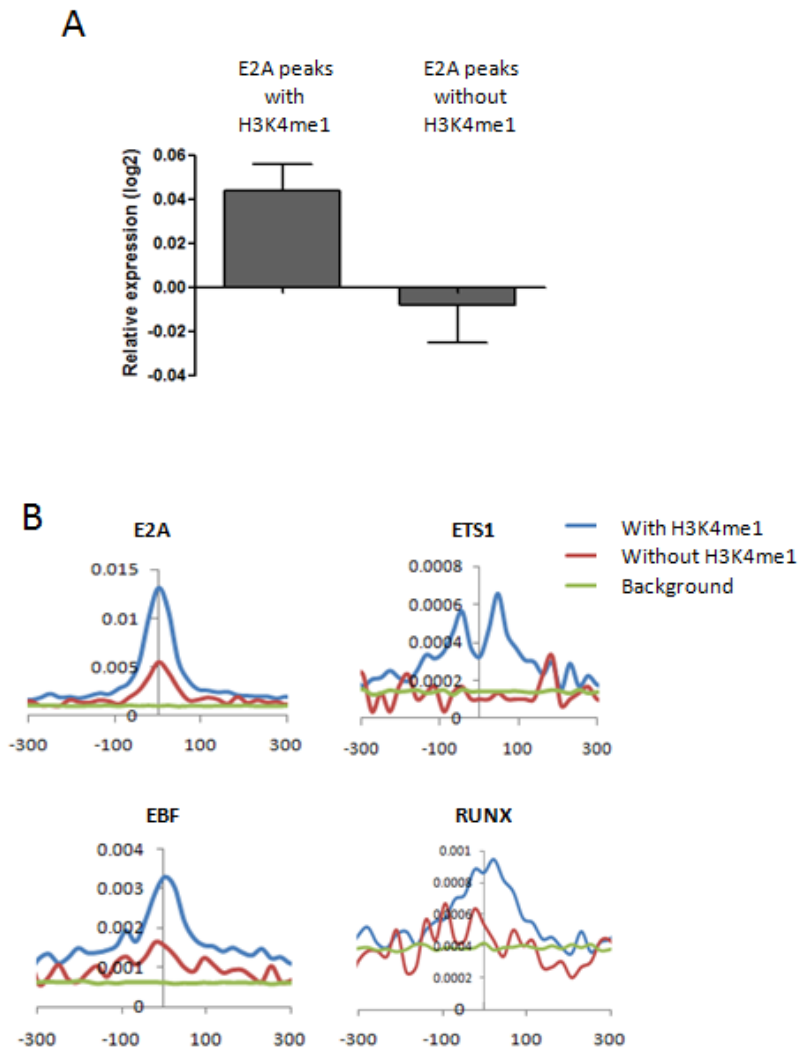
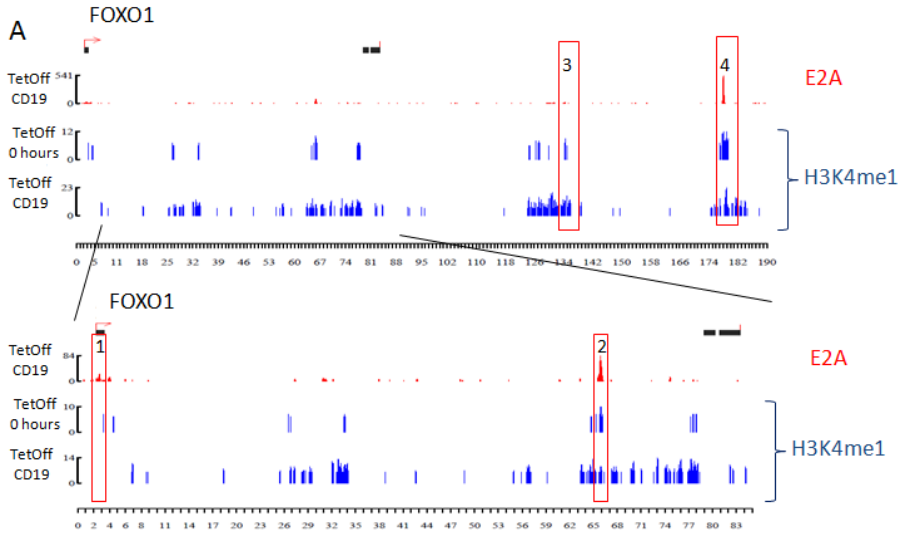


Figure 4.12

Association of E2A and H3K4me1 Binding is Correlated with Enrichment for Cis-Regulatory Motifs

(A) Association of distal E2A peaks and H3K4me1 is correlated to gene expression. The average relative expression (log2) is plotted for genes associated with distal E2A peaks either with or without H3K4me1 binding.

(B) Association of distal E2A peaks and H3K4me1 correlated with enriched cis-regulatory elements. Genomic distance of E2A, ETS, EBF and RUNX transcription factor motifs from distal E2A peaks either with or without H3K4me1 is plotted over a 300bp window. Background represents the presence of these motifs from 100,000 randomly selected genomic locations.



B

	E-box	Δ E-box
1	CACATG	CACA**CC**
	CAGATG	CAGA**CC**
2	CACCTG	
3	CAGCTG	CAGC**CC**
4	CAGCTG	CAGC**CC**
	CAGCTG	CAGC**CC**

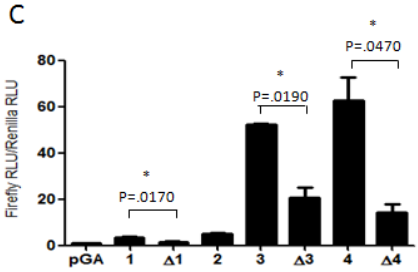


Figure 4.13

Functional Analysis of E2A and H3K4me1 Association at the FOXO1 Locus

(A) E2A and H3K4me1 binding sites in the Foxo1 locus. Numbers directly to the left of each plot indicate the maximum peak height in each plot. Red boxes indicate peaks selected for luciferase assay. Top panel includes the Foxo1 locus plus an additional 100kB to the 3' end, and the bottom panel includes only the Foxo1 locus.

(B) Regions indicated in (A) were cloned into the pGL3-ADME luciferase vector and transfected into 22D6 cells. Δ indicates that a mutation of one or more E-box sequences in these regions was performed.

(C) E2A is necessary for proper enhancer activity at the Foxo1 locus. Relative luciferase activity for the luciferase vector negative control (pGl3) and for the four selected regions and subsequent mutations is shown. Significance of increase between the pGL3 control vector and the proposed enhancer regions was calculated by Student's t-test: pGl3:1, p=0.0010, pGL3:2, p<0.0001, pGL3:3, p<.0.0001, pGL3:4, p=0.0041.

References

Adolfsson, J., Månsson, R., Buza-Vidas, N., Hultquist, A., Liuba, K., Jensen, C. T., Bryder, D., Yang, L., Borge, O.-J., Thoren, L. A. M., *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.

Agata, Y., Tamaki, N., Sakamoto, S., Ikawa, T., Masuda, K., Kawamoto, H., and Murre, C. (2007). Regulation of T Cell Receptor [beta] Gene Rearrangements and Allelic Exclusion by the Helix-Loop-Helix Protein, E47. Immunity *27*, 871-884.

Bain, G., Engel, I., Robanus Maandag, E., te Riele, H., Voland, J., Sharp, L., Chun, J., Huey, B., Pinkel, D., and Murre, C. (1997). E2A deficiency leads to abnormalities in alphabeta T-cell development and to rapid development of T-cell lymphomas. Mol Cell Biol *17*, 4782-4791.

Bain, G., Maandag, E. C. R., Izon, D. J., Amsen, D., Kruisbeek, A. M., Weintraub, B. C., Krop, I., Schlissel, M. S., Feeney, A. J., van Roon, M., *et al.* (1994). E2A proteins are required for proper B cell development and initiation of immunoglobulin gene rearrangements. Cell *79*, 885-892.

Barski, A., Cuddapah, S., Cui, K., Roh, T.-Y., Schones, D. E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. (2007). High-Resolution Profiling of Histone Methylations in the Human Genome. *129*, 823-837.

Busslinger, M. (2004). Transcriptional Control of Early B Cell Development1 doi:10.1146/annurev.immunol.22.012703.104807. Annual Review of Immunology *22*, 55-79.

Cobaleda, C., and Busslinger, M. (2008). Developmental plasticity of lymphocytes. Current Opinion in Immunology
Lymphocyte Development - Edited by Klaus Rajewsky and Harald von Boehmer / Tumour Immunology - Edited by Haruo Sugiyama *20*, 139-148.

Decker, T., Pasca di Magliano, M., McManus, S., Sun, Q., Bonifer, C., Tagoh, H., and Busslinger, M. (2009). Stepwise Activation of Enhancer and Promoter Regions of the B Cell Commitment Gene Pax5 in Early Lymphopoiesis. Immunity *30*, 508-520.

Dengler, H. S., Baracho, G. V., Omori, S. A., Bruckner, S., Arden, K. C., Castrillon, D. H., DePinho, R. A., and Rickert, R. C. (2008). Distinct functions for the transcription factor Foxo1 at various stages of B cell differentiation. *9*, 1388-1398.

Fisher, A. G. (2002). Cellular identity and lineage choice. *2*, 977-982.

Friedman, A. D. (2002). Transcriptional regulation of granulocyte and monocyte development. Oncogene *21*, 3377-3390.

Greenbaum, S., and Zhuang, Y. (2002). Identification of E2A target genes in B lymphocyte development by using a gene tagging-based chromatin immunoprecipitation system 10.1073/pnas.232299999. Proceedings of the National Academy of Sciences of the United States of America *99*, 15030-15035.

Hagman, J., and Lukin, K. (2006). Transcription factors drive B cell development. Current Opinion in Immunology
Lymphocyte development / Tumour immunology *18*, 127-134.

Heintzman, N. D., Hon, G. C., Hawkins, R. D., Kheradpour, P., Stark, A., Harp, L. F., Ye, Z., Lee, L. K., Stuart, R. K., Ching, C. W., *et al.* (2009). Histone modifications at human enhancers reflect global cell-type-specific gene expression. *459*, 108-112.

Heintzman, N. D., Stuart, R. K., Hon, G., Fu, Y., Ching, C. W., Hawkins, R. D., Barrera, L. O., Van Calcar, S., Qu, C., Ching, K. A., *et al.* (2007). Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *39*, 311-318.

Ikawa, T., Kawamoto, H., Goldrath, A. W., and Murre, C. (2006). E proteins and Notch signaling cooperate to promote T cell lineage specification and commitment 10.1084/jem.20060268. J Exp Med *203*, 1329-1342.

Ikawa, T., Kawamoto, H., Wright, L. Y. T., and Murre, C. (2004). Long-Term Cultured E2A-Deficient Hematopoietic Progenitor Cells Are Pluripotent. Immunity *20*, 349-360.

Inlay, M. A., Bhattacharya, D., Sahoo, D., Serwold, T., Seita, J., Karsunky, H., Plevritis, S. K., Dill, D. L., and Weissman, I. L. (2009). Ly6d marks the earliest stage of B-cell specification and identifies the branchpoint between B-cell and T-cell development 10.1101/gad.1836009. Genes & Development *23*, 2376-2381.

Kondo, M., Weissman, I. L., and Akashi, K. (1997). Identification of Clonogenic Common Lymphoid Progenitors in Mouse Bone Marrow. *91*, 661-672.

Lazorchak, A., Jones, M. E., and Zhuang, Y. (2005). New insights into E-protein function in lymphocyte development. Trends in Immunology *26*, 334-338.

Lin, H. H., and Grosschedl, R. (1995). Failure of B-Cell Differentiation in Mice Lacking the Transcription Factor Ebf. Nature *376*, 263-267.

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

Murre, C. (2005). Helix-loop-helix proteins and lymphocyte development. *6*, 1079-1086.

Nutt, S. L., UrbÃ¡nek, P., Rolink, A., and Busslinger, M. (1997). Essential functions of Pax5 (BSAP) in pro-B cell development: difference between fetal and adult B lymphopoiesis and reduced V-to-DJ recombination at the IgH locus.
10.1101/gad.11.4.476. Genes & Development *11*, 476-491.

Roessler, S., Gyory, I., Imhof, S., Spivakov, M., Williams, R. R., Busslinger, M., Fisher, A. G., and Grosschedl, R. (2007). Distinct Promoters Mediate the Regulation of Ebf1 Gene Expression by Interleukin-7 and Pax5
10.1128/MCB.01192-06. Mol Cell Biol *27*, 579-594.

Rothenberg, E. V., and Taghon, T. (2005). MOLECULAR GENETICS OF T CELL DEVELOPMENT
doi:10.1146/annurev.immunol.23.021704.115737. Annual Review of Immunology *23*, 601-649.

Seet, C. S., Brumbaugh, R. L., and Kee, B. L. (2004). Early B Cell Factor Promotes B Lymphopoiesis with Reduced Interleukin 7 Responsiveness in the Absence of E2A
10.1084/jem.20032202. J Exp Med *199*, 1689-1700.

Spangrude, G., Heimfeld, S., and Weissman, I. (1988). Purification and characterization of mouse hematopoietic stem cells
10.1126/science.2898810. Science *241*, 58-62.

Xu, W., and Kee, B. L. (2007). Growth factor independent 1B (Gfi1b) is an E2A target gene that modulates Gata3 in T-cell lymphomas
10.1182/blood-2006-08-043331. Blood *109*, 4406-4414.

Zhuang, Y., Soriano, P., and Weintraub, H. (1994). The helix-loop-helix gene E2A is required for B cell formation. Cell *79*, 875-884.

V

General Discussion

Role of E2A in Early Hematopoiesis

The E-proteins, E12, E47, HEB and E2-2 are critical for proper hematopoietic differentiation and immune development. In fact, mature B cell development is completely blocked before DJ rearrangement of the IgH gene has occurred in E2A knockout mice, and T-cell development is severely perturbed (Bain et al., 1997; Bain et al., 1994). In addition, three recent studies have identified crucial roles for E2A at earlier stages of hematopoietic development (Dias et al., 2008; Semerad et al., 2009; Yang et al., 2008).

In E2A -/- animals, there is a severe depletion not only of CLPs, as previously published, but also of the earlier LSK, MPP, and LT-HSC cell populations (Borghesi et al., 2004). There is also a significant block in development of the LMPP population. Interestingly, the cells present in these early progenitor populations in E2A -/- mice progress faster through the cell cycle, as indicated by BrdU staining. In fact, cell cycle analysis shows that there are fewer E2A -/- HSCs in G_0, suggesting that there is aberrant maintenance of the stem cell pool due to an impaired ability to regulate cell cycle status in E2A -/- mice. In support of this, RT-PCR analysis confirms that expression of cell-cycle regulators that have been shown to regulate the HSC pool, such as Mpl and $p21^{Cip1}$, is severely decreased in E2A -/- HSCs (Semerad et al., 2009).

Interestingly, analysis of the megakaryocyte and erythroid hematopoietic compartments revealed a surprising role for E2A in the proper differentiation of pre-MegE, pre-CFU-E, MkP, and GMP populations. E2A had not previously been implicated in Meg/E development. These data indicate novel roles for E2A in early hematopoiesis. Specifically,

the data indicate that the E2A proteins are essential to promote megakaryocyte/erythroid differentiation at intermediate stages.

Inhibition of E2A Allows for *Ex Vivo* Expansion of Multipotent Progenitors

It is now well established that in the absence of E2A, multipotent progenitors can be maintained and expanded *ex vivo* (Ikawa et al., 2004). We generated a novel and robust system in which Id2 is overexpressed in a tetracycline dependent manner to expand wild-type, multipotent hematopoietic progenitors *ex vivo*. These cells, Id2-HPCs, phenotypically resemble an undefined stage of development before commitment to the B-cell lineage but after development of the LMPP. Id2-HPCs retain both *in vitro* and *in vivo* differentiating potential. This method of progenitor cell expansion is novel as it marks the first time that multipotent progenitors can be expanded by overexpression of a transcriptional inhibitor. The use of an inhibitor, instead of a transcriptional activator, potentially prevents many of the off-target and potentially dangerous effects that can occur when overexpressing a transcriptional factor. (Sauvageau et al., 2004).

As the Id2-HPCs readily differentiate *in vitro* to either the B or myeloid lineage, we studied the changes in gene expression that allow either the maintenance of the self-renewing and multipotent state or direct lineage commitment choices. Although Id2-HPCs have lower expression levels of "stemness" like genes relative to CLPs and LMPPs, Id2-HPCs even further down-regulate the expression of genes such as CD34 and the Hoxa family upon differentiation into either the myeloid or B cell lineage. This indicates that although Id2-HPCs do not genetically represent true multipotent cells, they still maintain moderate transcript levels of many stem-like genes, and this is sufficient for their ability to self-renew.

As many investigations are in progress to expand hematopoietic stem and pluripotent stem cells *ex vivo* for therapeutic pursuits, the Id2-HPCs represent a paradigm shift in the field of stem cell expansion ((Kelly et al., 2009; Sauvageau et al., 2004). Here, for the first time, overexpression of a transcriptional inhibitor instead of a transcription factor has permitted the unlimited growth capacity and maintenance of multipotency of hematopoietic progenitors. As such, studies are underway to overexpress Id2 in human cord blood cells to determine the therapeutic potential of this expansion procedure for stem cell transplants.

Use of Id2-HPCs to Determine Cis-Regulatory Factors in Hematopoietic Development

Id2-HPCs are a unique system in which to probe genetic and epigenetic changes in hematopoietic development. Here we have described the change of cis-regulatory motif enrichment at active enhancers in myeloid and B cell development, and the temporal changes in cis-regulatory motif enrichment at active enhancers in B cell development. We then studied the direct role of E2A during B cell development, and correlated E2A binding with H34me1 binding to predict novel enhancer elements.

H3K4me1 binding patterns on genes known to be involved in either B, T, or myeloid lineage fates were analyzed at a multipotent developmental stage, B cell committed stage, or myeloid committed stage of development. This gene specific analysis suggested that binding of H3K4me1 is already present at enhancers at a multipotent stage of development, and upon differentiation, H3K4me1 levels increase or decrease in concert with gene expression. These data suggest that, even at a multipotent state, H3K4 monomethylation is present at lineage specific genes to prime them for future activation.

We then expanded this analysis to study the genome-wide association of H3K4 monomethylation with transcription factor motifs to predict cis-regulatory patterns involved in lineage specification. In the Id2-HPCs, we observed strong enrichment for the RUNX and ETS motifs within islands associated with H3K4me1. When the cells were directed to differentiate to a myeloid lineage, H3K4 monomethylation regions showed a loss in ETS motifs and gains for motifs associated with PU.1 . In contrast, ETS, E2A and EBF motifs were strongly enriched at B-cell specific H3K4me1 islands. This suggests a model of regulation in which, at active enhancers, RUNX and ETS factors maintain a stem-like phenotype, PU.1 replaces ETS in the myeloid lineage, and E2A and EBF replace ETS in the B lineage.

While this model of specific transcription-factor recruitment to active enhancers is notable, it is based only on the endpoints of differentiation and, as such, cannot answer questions about intermediate changes in development. To address this, we used the Id2-HPC cell line to study changes in gene expression and H3K4 monomethylation during the progression of pre-pro-B to pro-B cell development. Microarray analysis confirmed the expected upregulation of B-cell specific genes such as Ebf, Foxo1 and Bcl11a, and the downregulation of alternate lineage genes such as Cebpα, CD34, and Zap70 (Busslinger, 2004; Murre, 2005; Murre, 2009). Surprisingly, there was a subset of genes that is upregulated at an intermediate timepoint and then re-repressed at the end of the B-cell differentiation timecourse. Several anti-apoptosis factors, such as Bcl2l1, were present in this group of genes, suggesting that maintenance of cell survival during development is a critical and temporally mediated process. Also, the plasma cell differentiation gene, Xbp1, was also upregulated during developmental progression, suggesting a novel function for Xbp1 in early B cell development.

To determine the cis-regulatory elements involved in this temporal regulation of genes, we studied H3K4me1 levels changes in combination with alterations in gene expression levels. In multipotent Id2-HPCs, H3K4me1 peaks associated with RUNX, ETS and E2A motifs. Interestingly, by the 120-hour B cell timepoint, enhancers of genes with an increase in expression associated with E2A, EBF and FOXO motifs, while enhancers of genes with a decrease in expression were only associated with E2A and EBF motifs. This suggested a regulatory mechanism in which E2A and EBF are recruited to many genes that are changed in B-cell development, but the combination of E2A, EBF and Foxo1 is specific to activated genes. Interestingly, enhancers of genes that are intermediately changed do not show association with EBF, but maintain the association with ETS and RUNX factors.

As these analyses confirm the role of E2A in B-cell development, we analyzed the binding of E2A in a genome-wide manner in CD19+ pro-B cells. This analysis revealed that E2A directly binds to nearly half of the genes that were changed in B-cell development. This binding significantly correlates to an average increase in gene expression, although there are many specific cases in which E2A is bound to repressed genes. When we combined E2A binding with H3K4me1 binding, we observed that gene expression levels increase when E2A and H3K4me1 are concurrently present at an enhancer, and expression levels decreases if only E2A is present. Surprisingly, we observed an E-box binding bias for E2A depending on the presence of H3K4me1. When E2A and H3K4me1 are both present, E2A is recruited to canonical CAGCTG E-boxes. In contrast, when E2A is not correlated with H3K4 monomethylation, it is preferentially recruited to CACCTG or CAGGTC E-boxes. This is the first reporting of E-box preference as it relates to gene expression. It is tempting to hypothesize that perhaps homodimers of E47 are binding to canonical E-boxes and recruiting

H3K4me1, whereas perhaps E47/HEB heterodimers are binding to non-canonical E-boxes to recruit co-repressors.

In conclusion, the generation of the Id2-HPC cell line, which can either maintain a self-renewing and multipotent phenotype, or can be induced to differentiate both *in vivo* and *in vitro*, represents a significant development in the quest to expand hematopoietic progenitor cells and in the studies of early hematopoietic development. We have used Id2-HPCs to identify cis-regulatory codes that underpin the development towards the B and myeloid cell fates.

References

Bain, G., Engel, I., Robanus Maandag, E., te Riele, H., Voland, J., Sharp, L., Chun, J., Huey, B., Pinkel, D., and Murre, C. (1997). E2A deficiency leads to abnormalities in alphabeta T-cell development and to rapid development of T-cell lymphomas. Mol Cell Biol *17*, 4782-4791.

Bain, G., Maandag, E. C. R., Izon, D. J., Amsen, D., Kruisbeek, A. M., Weintraub, B. C., Krop, I., Schlissel, M. S., Feeney, A. J., van Roon, M., *et al.* (1994). E2A proteins are required for proper B cell development and initiation of immunoglobulin gene rearrangements. Cell *79*, 885-892.

Borghesi, L., Hsu, L.-Y., Miller, J. P., Anderson, M., Herzenberg, L., Herzenberg, L., Schlissel, M. S., Allman, D., and Gerstein, R. M. (2004). B Lineageâ€“specific Regulation of V(D)J Recombinase Activity Is Established in Common Lymphoid Progenitors 10.1084/jem.20031800. The Journal of Experimental Medicine *199*, 491-502.

Busslinger, M. (2004). Transcriptional Control of Early B Cell Development1 doi:10.1146/annurev.immunol.22.012703.104807. Annual Review of Immunology *22*, 55-79.

Dias, S., Månsson, R., Gurbuxani, S., Sigvardsson, M., and Kee, B. L. (2008). E2A Proteins Promote Development of Lymphoid-Primed Multipotent Progenitors. Immunity *29*, 217-227.

Ikawa, T., Kawamoto, H., Wright, L. Y. T., and Murre, C. (2004). Long-Term Cultured E2A-Deficient Hematopoietic Progenitor Cells Are Pluripotent. Immunity *20*, 349-360.

Kelly, S. S., Sola, C. B. S., de Lima, M., and Shpall, E. (2009). Ex vivo expansion of cord blood. *44*, 673-681.

Murre, C. (2005). Helix-loop-helix proteins and lymphocyte development. *6*, 1079-1086.

Murre, C. (2009). Developmental trajectories in early hematopoiesis 10.1101/gad.1861709. Genes & Development *23*, 2366-2370.

Sauvageau, G., Iscove, N. N., and Humphries, R. K. (2004). In vitro and in vivo expansion of hematopoietic stem cells. Oncogene *23*, 7223-7232.

Semerad, C. L., Mercer, E. M., Inlay, M. A., Weissman, I. L., and Murre, C. (2009). E2A proteins maintain the hematopoietic stem cell pool and promote the maturation of myelolymphoid and myeloerythroid progenitors 10.1073/pnas.0808866106. Proceedings of the National Academy of Sciences *106*, 1930-1935.

Yang, Q., Kardava, L., St. Leger, A., Martincic, K., Varnum-Finney, B., Bernstein, I. D., Milcarek, C., and Borghesi, L. (2008). E47 Controls the Developmental Integrity and Cell Cycle Quiescence of Multipotential Hematopoietic Progenitors. J Immunol *181*, 5885-5894.