universität wien

DISSERTATION

Titel der Dissertation

Direct Regulation of the *Pax5* Promoter by the Transcription Factor EBF1 in Early B Cell Development

Verfasserin

Anca-Sarmiza Tigan, M.Sc.

angestrebter akademischer Grad

Doktorin der Naturwissenschaften (Dr.rer.nat.)

Wien, im Juni 2012

Studienkennzahl lt. Studienblatt:	A 091 490
Dissertationsgebiet lt. Studienblatt:	Molekulare Biologie
Betreuer:	Univ. Prof. Dr. Meinrad Busslinger

CONTENTS

.

ABSTRACT

B cells derive from hematopoietic stem cells in a stepwise process, driven mainly by transcription factors that play an important role in activating lineage-specific genes and controlling commitment by repression of other developmental options. The transcription factor Pax5 is the key player of the transcriptional network leading to the commitment of progenitors to the B cell lineage and is involved in both suppression of lineage-inappropriate genes and activation of B cell-specific genes. In the absence of Pax5, B cell development is arrested at the early pre-pro-B cell stage. However, it is still largely unknown which transcription factors are responsible for the activation of the *Pax5* gene at the onset of B cell development. Recent work from the Busslinger laboratory has identified all B cell-specific regulatory elements of the *Pax5* locus and a few transcription factors that interact with these elements. Interestingly, the promoter region of *Pax5* contains two EBF1-binding sites located in the DNaseI hypersensitive sites 6 and 7 (HS6 and HS7). This caught my attention since EBF1 acts genetically upstream of Pax5. Furthermore, in $Ebf1^{-/-}$ progenitor cells, the DNaseI hypersensitive regions of the *Pax5* promoter are not formed and the repressive chromatin mark H3K27me3 is present. Based also on previous reports that identify EBF1 as a pioneer transcription factor able to drive chromatin remodeling to a transcription-permissive state, I was interested to investigate how the transcription of *Pax5* is affected by removal of the two EBF1-binding sites in the gene promoter.

I generated targeting vectors for deletion of either of the two sites individually or in combination. To allow comparison of the changes in the level of *Pax5* expression upon the aforementioned EBF1-binding sites deletion, I targeted an ES cell line that carries an IRES-humanCD2 reporter gene in the 3' untranslated region of the *Pax5* locus ($Pax5^{ihCD2/+}$). This allows monitoring the changes in expression of *Pax5* upon disruption of promoter binding by EBF1 in vivo at single-cell resolution. Targeting experiments resulted in the generation of $Pax5^{\Delta EHS7/+}$ and $Pax5^{\Delta EHS6+7/+}$ mice that lack the EBF1-binding site in HS7 or both EBF1-binding sites in HS6 and HS7, respectively. Analysis of the $Pax5^{\Delta EHS7/+}$ and $Pax5^{\Delta EHS7/\Delta EHS7}$ mice revealed no influence of the EBF1-binding site at HS7 on *Pax5* expression levels or B cell development. Since EBF1 binding at this site is disrupted in the analyzed mice, I conclude that EBF1 action on

the *Pax5* promoter is not carried out through binding at HS7 alone. However, *Pax5* expression (monitored by the hCD2 reporter) is two-fold reduced at the early stages of B cell development in the bone marrow of $Pax5^{\Delta EHS6+7/+}$ and $Pax5^{\Delta EHS6+7/\Delta EHS6+7}$ mice. Notably, expression levels of *Pax5* are not affected by EBF1-binding sites disruption in mature B cells. This demonstrates that the two EBF1-binding sites in the *Pax5* promoter act in cooperation in order to maintain the correct expression levels of *Pax5* from the pro-B cell to the immature B cell stage. Further analysis is required to identify what changes in promoter architecture, chromatin landscape, and transcription factor binding accompany the disruption of the two EBF1-binding sites in the *Pax5* promoter, and these experiments are discussed in the final part of the thesis.

ZUSAMMENFASSUNG

B Zellen stammen von hämatopoetischen Stammzellen ab. Ihre Entstehung ist ein stufenweiser Prozess, der hauptsächlich von Transkriptionsfaktoren gesteuert wird, deren Aufgabe zum einen darin besteht, B-Zelllinien-spezifische Gene zu aktivieren, und zum anderen, die Zugehörigkeit zur B-Zelllinie zu kontrollieren, indem weitere Entwicklungsmöglichkeiten verhindert werden. Der Transkriptionsfaktor Pax5 spielt eine Schlüsselrolle in dem transkriptionellen Netzwerk, das die Zugehörigkeit von Vorläuferzellen zur B-Zelllinie steuert. Pax5 kann sowohl B-Zell-spezifische Gene aktivieren, als auch nicht-B-Zell-spezifische Gene hemmen. In der Abwesenheit von Pax5 ist die B-Zellentwicklung im frühen pre-pro B Zellstadium blockiert. Bisher ist es jedoch weitgehend unklar, durch welche Transkriptionsfaktoren die Aktivierung des *Pax5* Gens beim Eintritt in das pro-B Zellstadium erfolgt. In vorherigen Arbeiten des Busslinger Labors konnten B-Zell-spezifische regulatorische Elemente im *Pax5* Genlokus beschrieben werden, sowie verschiedene Trankriptionsfaktoren, die mit diesen Elementen interagieren. In den DNaseI hypersensitiven Stellen 6 und 7 (HS6 and HS7) des *Pax5* Promotors befinden sich zwei Bindestellen für den Transkriptionsfaktor EBF1. Dies ist besonders interessant, da EBF1 epistatisch oberhalb von Pax5 agiert. Weiters sind in *Ebf1*$^{-/-}$ Vorläuferzellen die DNaseI hypersensitiven Regionen des *Pax5* Promotors nicht ausgebildet, sondern mit der repressiven Histonmodifizierung H3K27me3 dekoriert. Zusätzlich wurde gezeigt, dass EBF1 als Pioniertranskriptionsfaktor in der Lage ist Chromatinremodelierung hin zu einem aktiven Chromatinstatus zu steuern. Mein Interesse lag darin zu untersuchen, wie die Transkription von *Pax5* durch den Verlust beider EBF1-Bindungsstellen innerhalb des Pomotors beeinflusst wird.

Ich habe Targetingvektoren zur Deletion beider Bindungsstellen, einzeln und in Kombination miteinander, hergestellt. Um eventuelle Änderungen in der *Pax5* Expression durch die beschriebenen Deletionen der EBF1-Bindungsstellen zu bestimmen, habe ich eine ES Zellline für homologe Rekombination verwendet, die ein IRES-humanCD2 Reportergen innerhalb der 3' untranslatierten Region des *Pax5* Gens trägt (*Pax5*$^{ihCD2/+}$). Dadurch wird es ermöglicht, Änderungen der *Pax5* Expression durch Verlust der Bindung von EBF1 an der Promotorregion *in vivo* in einzelnen B-Zellen zu erfassen. Die Targetingexperimente resultierten in der Herstellung

von *Pax5*$^{\Delta EHS7/+}$ und *Pax5*$^{\Delta EHS6+7/+}$ Mäusen, denen entweder die EBF1-Bindungsstelle in HS7 oder beide EBF1-Bindungsstellen in HS6 und HS7 fehlen. Die Analyse der *Pax5*$^{\Delta EHS7/+}$ und *Pax5*$^{\Delta EHS7/\Delta EHS7}$ Mäuse zeigte, dass die EBF1-Bindungsstelle in HS7 keinen Einfluss auf die *Pax5* Expression oder die B-Zellentwicklung hat. Da die Bindung von EBF1 an die HS7 Stelle in den analysierten Mäusen zerstört wurde, komme ich zu dem Schluss, dass die Aktivität von EBF1 am *Pax5* Promotor nicht alleine von der Bindung an HS7 abhängt. Allerdings, ist die *Pax5* Expression (gemessen mit dem hCD2 Reporter) in den Anfangsstadien der B-Zellentwicklung im Knochenmark der *Pax5*$^{\Delta EHS6+7/+}$ und *Pax5*$^{\Delta EHS6+7/\Delta EHS6+7}$ Mäuse zweifach reduziert. In reifen B-Zellen hingegen ist die *Pax5* Expression durch das Fehlen der zwei EBF1-Bindungsstelle nicht beeinflusst. Dies zeigt, dass die zwei EBF1-Bindungsstellen im *Pax5* Promotor in Kooperation agieren, um die angemessenen *Pax5* Expressionslevel vom pro-B zum unreifen B-Zellstadium aufrechtzuerhalten. Weitere Analysen sind erforderlich um festzustellen, welche Änderungen in der Promotorarchitektur, Chromatinlandschaft und Transkriptionsfaktorbindung mit der Disruption der beiden EBF1-Bindungsstellen am *Pax5* Promotor einhergehen. Diese Experimente werden im letzten Kapitel der Thesis diskutiert.

INTRODUCTION

1. Overview of early hematopoiesis

Hematopoietic stem cells (HSCs) are responsible for the continuous generation of all blood cell types throughout life. Their characteristic features are the potential to self-renew and the ability to differentiate in multiple blood cell lineages (Weissman 2000). These two essential attributes are retained upon transfer into a recipient host (Weissman 2000; Orkin and Zon, 2002). Formation of blood cells (hematopoiesis) originates at the ventral mesoderm of the developing embryo from where the first precursor cells migrate to the yolk sac and give rise to primitive red blood cells at embryonic day E7.5 (Orkin and Zon, 2002; Mikkola and Orkin, 2006). This initial wave of hematopoiesis is followed by three other types of hematopoietic activities, which are progressively more complex. Some of them are independent of each other and originate from distinct anatomical places (Dzierzak and Speck, 2008). Primitive, pro-definitive (myeloid), meso-definitive (lymphoid-myeloid), and meta-definitive (neonatal repopulating) classes of hematopoietic stem cells are limited in lifespan, engraftment potency and multilineage differentiation abilities when compared to adult HSCs. However, they are extremely important to provide the necessary maturation signals to the developing secondary hematopoietic territories (fetal liver, thymus, spleen). These territories will be finally colonized by adult HSCs, which originate from the embryo proper aorta-gonad-mesonephros (AGM) region and are independently generated in a fifth, definitive wave of hematopoiesis at E10.5 (Dzierzak and Speck, 2008). Immediately after definitive hematopoiesis, adult repopulating HSCs are expanding in the yolk sac and placenta, and finally in the fetal liver (which is the main reservoir of HSCs during the second part of embryonic development). By E12, the numbers of HSCs are 14 times increased and at this time HSC activity in the AGM region ceases. Adult HSCs are fully potent for long-term multilineage reconstitution of irradiated hosts and constitute the original hematopoietic stem cell pool that will colonize the bone marrow just before birth at E17 (Dzierzak and Speck 2008; Boisset and Robin, 2012).

When compared to fetal liver HSCs, adult HSCs located in the bone marrow divide rarely, can be quiescent for long periods of time, and are metabolically inactive. The different properties of HSCs in distinct sites of hematopoiesis are also a result of their interaction with the microenvironment, the so-called “stem cell niche” (Wilson and Trumpp, 2006; Orkin and Zon, 2008). This specific environment (Figure 1) in which HSCs are maintained throughout the life of an individual regulates their self-renewal and differentiation. In the bone marrow, hematopoietic stem cells are found in trabecular bone, in close

connection to osteoblasts that are directly involved in stem cell maintenance and regulation. This is called the endosteal bone marrow niche and it hosts mostly quiescent HSCs as a long-term "reserve pool". When needed, HSCs migrate from the endosteal bone surface to the center of the bone marrow to the vascular niche (located in the proximity of sinusoids) where they participate in hematopoiesis (Wilson and Trumpp, 2006). Furthermore, specialized cells that express high amounts of CXCL12 (CAR cells) are located throughout the bone marrow and seem to have an important contribution to the niches as well as to play a role in HSCs retention (Sugiyama et al, 2006). FoxP3 regulatory T cells have been recently shown to participate in the endosteal bone marrow niche (Fujisaki et al, 2011), which suggests the existence of a state of immune privilege that protects HSCs against the adverse effects of inflammatory responses.

Besides its function in providing an immune-privileged location for HSCs maintenance, the bone marrow also supports a plethora of cells that are essential for immune responses (Mercier et al, 2012) and represents the main site of B lymphocyte development (Figure 1).

B lymphopoiesis is integrated in the broader process of generation of blood cells, which consists of defined progenitor stages. This model for blood cell development also referred to as the "hematopoietic tree" (Zandi et al, 2010) has been long regarded as a paradigm for the study of developmental trajectories involved in the specification of distinct cell types (Mercer et al, 2011; Medina and Singh, 2005). It starts from the long-term hematopoietic stem cell (LT-HSC) – discussed above as adult HSC – which gives rise to the short-term hematopoietic stem cell (ST-HSC) upon encounter of the first differentiation signals. ST-HSCs retain multipotency but display limited self-renewal capacity. Their limited ability to proliferate as multipotent stem cells supports reconstitution of the hematopoietic system for only six weeks (Kondo, 2010). The sequential stage in this developmental lineage is the multipotent progenitor (MPP), which has completely lost the ability to self-renew but can still generate all blood cell types. LT-HSCs, ST-HSCs and MPP constitute the LSK population (lineage^{-}, ckit^{hi}, Sca-1^{+}) (Kondo, 2010). The MPP can differentiate either into the erythroid and megakaryocyte lineages through a common MEP (megakaryocyte/erythroid progenitor) precursor stage or into the lymphoid/myeloid lineages through the lymphoid primed multipotent progenitor (LMPP). The transcription factors GATA-1 and PU.1 antagonize each other to segregate the MEP (GATA-1 activity) and LMPP (PU.1 activity) cell fates (Mercer et al, 2011). LMPPs are characterized by the expression of the tyrosine kinase receptor Flt3 (Adolfsson et al, 2001) and can give rise to monocytes, granulocytes, B, T, and natural killer (NK) cells (Mercer et al, 2011). Gradual increase in the levels of Flt3 specifies the lymphoid cell fate at the expense of the myeloid one (Mansson et al, 2007).

The main developmental paths for the generation of blood cells as well as the important transcription factors that drive them are summarized in Figure 2. Granulocyte-macrophage progenitors (GMP) can originate from either LMPP or MPP compartments and give rise to the myeloid lineage. Early T cell progenitors (ETP) are derived from the bone marrow and seed the thymus at an early stage in their development. Although they still retain the ability to differentiate into myeloid, NK and dendritic cells (DC), the concerted activity of Notch signaling, E proteins, GATA-3 and Bcl11b progressively commits them to the T cell fate (Mercer et al, 2011).

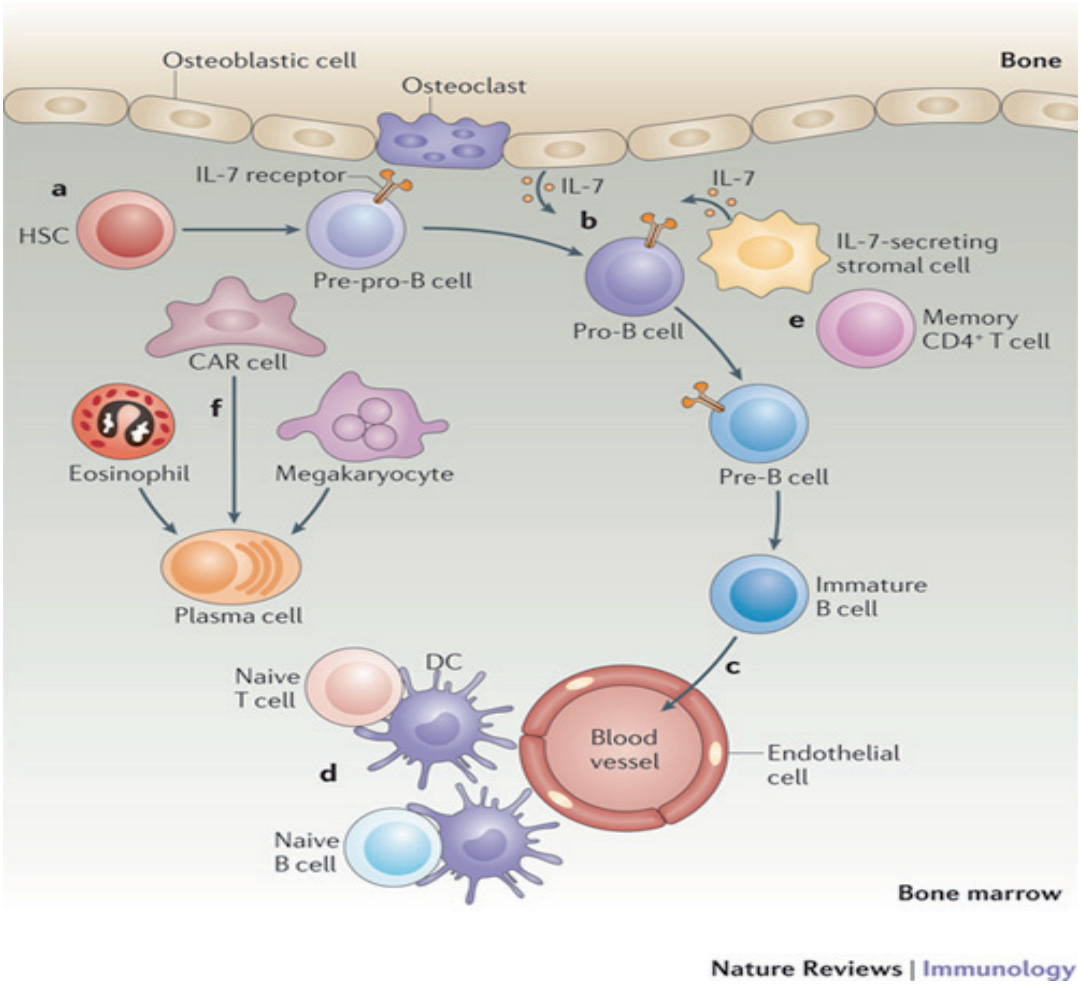


Figure 1. The bone marrow offers a unique environment for HSC maintenance, B cell development, and immune cells survival. (a) HSCs and pre-pro-B cells are in contact with CAR cells (b) pro-B cells are associated with interleukin-7 (IL-7)-secreting stromal cells (c) immature B cells are later associated with endothelial cells (d) naive B and T cells are found in a perivascular niche made by dendritic cells (DCs) – this allows a quick response against blood-borne pathogens (e) quiescent memory CD4+ T cells are found next to IL-7-secreting stromal cells (f) plasma cells are also found in the bone marrow and their engraftment and survival signals come from CAR cells, eosinophils and megakaryocytes.
(from Mercier et al, 2012)

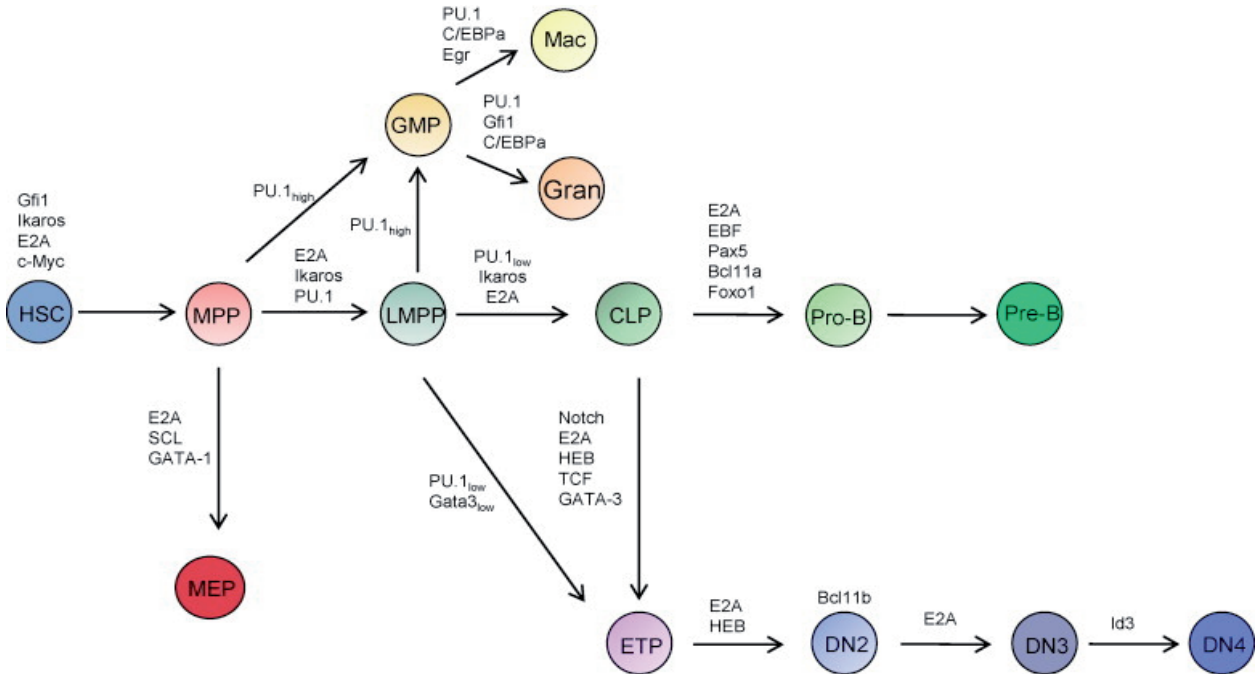


Figure 2. Developmental pathways in early hematopoiesis and the most important transcription factors involved in their specification. Abbreviations not detailed in text body: CLP – common lymphoid progenitors; Mac – macrophages; Gran – granulocytes. (from Mercer et al, 2011)

2. B cell development

The "hematopoietic tree" outlined in Figure 2 initially includes progenitors that display patterns of gene expression associated with mixed lineages (Miyamoto et al, 2002). Cell-type specification in this context requires resolution of these mixed lineage genetic programs. This effect is mediated by transcription factors and signaling pathways, regulatory modules that can be assembled into genetic networks. The pathway for B cell development stands out as a leading model for the study of such genetic networks that orchestrate distinct developmental transitions, which culminate in the appearance of committed cells (Medina and Singh, 2005).

Activation of Flt3 receptor in LMPPs favors differentiation towards B and T cell fates (Adolfsson et al, 2005) and induces the expression of the Interleukin 7 receptor (IL-7R) (Borge et al, 1999). IL-7R expression is a hallmark of the subsequent developmental stage, the common lymphoid progenitor (CLP), from which the different lymphoid cell types branch out (Figure 2). Recent work by Weissman and colleagues has further subdivided the CLP into two distinct stages. The $Ly6d^-$ ALP (all-lymphoid progenitor) can still generate all lymphocyte cell types and seed the thymus to give rise to the T cell lineage. In contrast, the $Ly6d^+$ BLP (B-cell-biased lymphoid progenitor) already displays the

expression signature of B cell progenitors, and is thus specified to the B cell lineage (Inlay et al, 2009). Furthermore, an increase in the levels of the Recombination Activating Gene-1 enzyme (Rag-1), which is essential for the gene rearrangement events that result in functional B and T cell receptors (BCR, TCR), correlates with a reduced differentiation potential towards the NK cell lineage (Mansson et al, 2010). These findings suggest that B-lymphocyte committed precursors arise by a gradual loss firstly of NK and then of T cell potential (Figure 3).

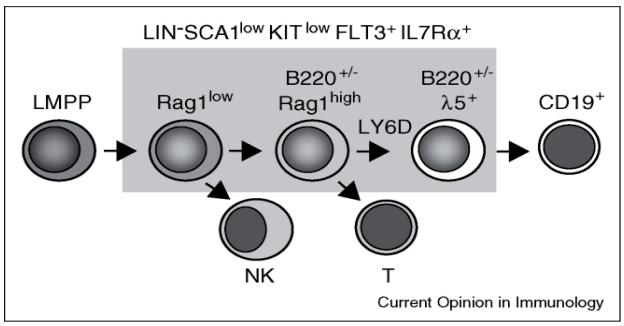


Figure 3. Gradual differentiation of the CLP towards a B-lineage restricted cell.
Within the $LIN^{-}KIT^{low}SCA1^{low}FLT3^{+}IL7^{+}$ compartment in the bone marrow, cells first loose the potential to become NK cells. Once the Rag1 expression is upregulated, cells differentiate either into T cells or upregulate Ly6d expression on the way to becoming B cells. (from Bryder and Sigvardsson, 2010)

From the BLP onwards, B lymphopoiesis follows an ordered series of events defined by the sequential expression of surface markers and the progressive rearrangement of the immunoglobulin heavy and light chain genes (IgH and IgL). The B cell developmental pathway is orchestrated by a defined network of transcription factors (discussed in the next sections), of which Pax5 (Paired box protein) plays a crucial role as the B cell-specific commitment factor. Figure 4 summarizes the main stages of B lymphopoiesis alongside with the functions of Pax5 at each step. Initially, the activity of the Rag genes targets the IgH locus at the pre-pro-B and pro-B cell stages. Successful IgH rearrangement results in the expression of the BCR heavy chain together with the surrogate light chain as the pre-B cell receptor (pre-BCR). Signaling from the pre-BCR controls the transition from pro-B to pre-B cell stage and induces allelic exclusion at the IgH locus. Pre-BCR signaling also triggers several rounds of clonal expansion followed by differentiation to small pre-B cells that start to recombine the light chain locus. Upon successful immunoglobulin rearrangements, cells expressing a functional BCR are termed immature B cells. They migrate from the bone marrow to the peripheral lymphoid organs (Figure 1) and differentiate into transitional B cells (Busslinger 2004). In the spleen transitional B cells further develop into quiescent mature B cells that recirculate through the blood and secondary lymphoid organs (Carsetti et al, 2004). Encounter with T cell dependent antigens generally triggers B cells to become

activated and differentiate to antibody secreting plasma cells or to enter germinal centers where they receive T cell help and undergo somatic hypermutation and immunoglobulin class switching. B cells that exit the germinal centers can either terminally differentiate into antibody secreting plasma cells or become memory cells (Tarlinton, 2006).

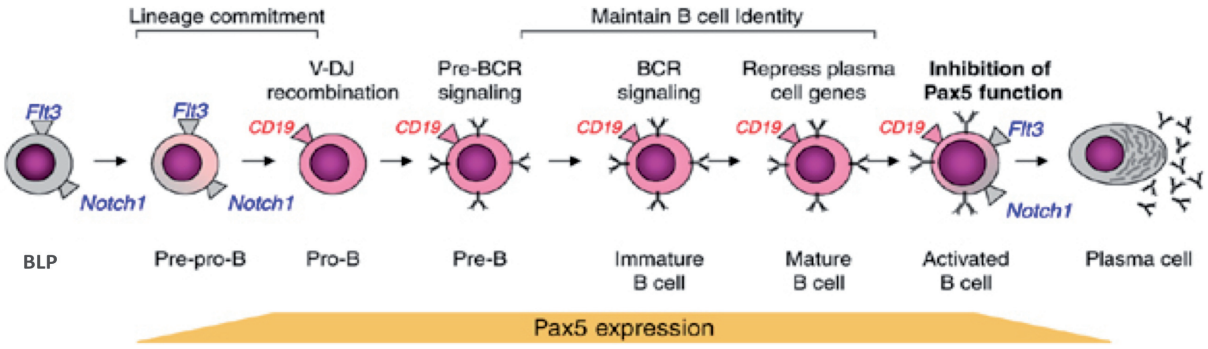


Figure 4. Stages of the B cell development.
Developmental stages are defined by the expression of Flt3 and CD19 cell surface receptors, Notch1 and Pax5 transcription factors, and immunoglobulin genes. From the point when *Pax5* expression begins at the pre-pro-B cell stage, cells are committed to the B cell lineage. Pax5 orchestrates lineage commitment – it activates the B cell-specific receptor CD19, and represses Notch1 (important for T cell development) and Flt3. Expression of *Pax5* remains constant up to the plasma cell stage when its transcription is silenced. During this time the main function of Pax5 is to maintain B cell identity; its functions at each developmental point are outlined in the upper part of the figure. Activated B cells use post-translational regulation mechanism to inhibit Pax5 function; *Pax5* expression is down regulated at the plasma cell stage, when a new genetic program (previously repressed by Pax5) is activated. (Adapted from Holmes et al, 2008)

In addition to conventional B cells described above (also termed follicular B cells or B-2 cells) the mature B cell pool contains two other subsets with particular functions (Samitas et al, 2010). Marginal zone (MZ) B cells are distinctly located in the marginal sinus region of the spleen, at the border between the red and the white pulp. They are specialized to react to blood-borne pathogens, being potent antigen presenting cells (APC). They also have the ability to quickly differentiate into plasma cells in response to T cell independent antigens (Pillai et al, 2005).

B1 cells are predominantly located in the pleural and peritoneal cavities and are responsible for the production of most of the natural antibodies (IgM in the spleen and serum, IgA in the intestinal lamina propria and other mucosal sites) (Samitas et al, 2010). The uniqueness of B-1 cells, which possess a characteristic limited BCR repertoire, stems from their ability to self-renew in contrast to conventional B cells. Their activation seems to be independent of the BCR specificity and mostly driven by innate immune signaling such as Toll-like receptor interactions (Baumgarth 2011).

In the context of hematopoiesis, B lymphopoiesis has been intensely studied as a model of cellular differentiation. More than for any other mammalian system, the transcription factors involved have been identified (Figure 2) and their functions and interactions are being elucidated (Zandi et al, 2010). My study focuses on mechanistic interactions of B cell-specific transcription factors that result in the activation of *Pax5* expression at the onset of B cell development (Figure 4). The far-reaching aims of such an undertaking concern elucidation of general mechanisms of gene regulation and how they influence cell-type specification. To this end, a general outline of epigenetic mechanisms involved in gene expression is necessary.

3. Epigenetic mechanisms

The consecutive developmental cell stages that assemble the "hematopoietic tree" are subject to tight regulation to ensure activation of lineage-specific genetic programs and repression of inappropriate genes. On the one hand this regulation is achieved through cell-stage specific networks of transcription factors (TF) and feedback loops between them resulting in "self-propagating transcriptional states" that represent an example of epigenetic regulatory mechanism acting in trans (Bonasio et al, 2010). On the other hand, epigenetic states that act in cis are associated with chromatin. Their establishment is often the result of TF binding to the DNA in response to developmental cues or other environmental signals (Bonasio et al, 2010). Chromatin structure itself helps the assembly of a hierarchy of epigenetic regulation mechanisms.

Firstly, the DNA sequence can be directly modified by cytosine methylation. Being one of the most stable modifications, DNA methylation has a strong influence on developmental decisions. DNA methylation patterns are initially erased during embryogenesis, only to be *de novo* established at implantation. DNA methyltransferases methylate the entire genome except CpG islands (predominantly associated with promoters of housekeeping genes), which are protected due to particular sequences such as Sp1 elements (Brandeis et al, 1994). As a consequence, CpG islands are found in an accessible configuration as opposed to promoters of tissue- and lineage-specific genes that adopt a closely-packaged chromatin conformation (Cedar and Bergman, 2011).

Secondly, the DNA is wrapped around protein histones yielding the basic unit of chromatin – the nucleosome. It is made of a histone core containing two copies of each H2A, H2B, H3, H4, around which 147 base pairs (bp) of DNA wrap. At this second level, epigenetic modifications include covalent modifications of histone tails at specific amino acids (methylation, acetylation, phosphorylation or

sumoylation) and chromatin remodeling by protein complexes. Remodeling complexes use ATP hydrolysis to change the location of nucleosomes along the DNA. These modifications change the mode/extent of chromatin packaging and thus they influence the access of TF or the transcription machinery to the DNA (Parra 2009; Campos and Reinberg, 2009). Histone tail modifications are the result of specific enzymatic activities that either add or remove the respective chemical group. Based on the way they influence DNA accessibility and on their ability to recruit TF or other DNA modifying enzymes, histone tail modifications can generally be grouped in either activating or repressing marks (Jenuwein and Allis, 2001; Turner 2007). Although post-translational modifications of histones have been shown to recruit effector molecules, which mediate distinct, predictable outcomes ("the histone code" hypothesis), additional layers of specificity seem to mediate context-dependent outcomes as well (Sims and Reinberg, 2008).

Nucleosome positioning correlates with energetically favorable/unfavorable sequence contexts as well as with the activity of remodelers or the binding of specific proteins. For example, in the hematopoietic system the lineage-determining factor PU.1 has been shown to initiate nucleosome remodeling followed by the deposition of histone modifications such as H3K4me1. This then serves as a platform for binding of cell-type specific transcription factors that influence transcription (Heinz et al, 2010).

Finally, histone modifications act in tandem with DNA-binding chromatin regulators to establish higher-order structures of compacted chromatin. For example, nuclear-lamina associated domains are made up of large regions of heterochromatin that contain transcriptionally silenced genes (Zhou et al, 2011). A schematic representation of the current thinking on the different levels of chromatin organization is depicted in Figure 5.

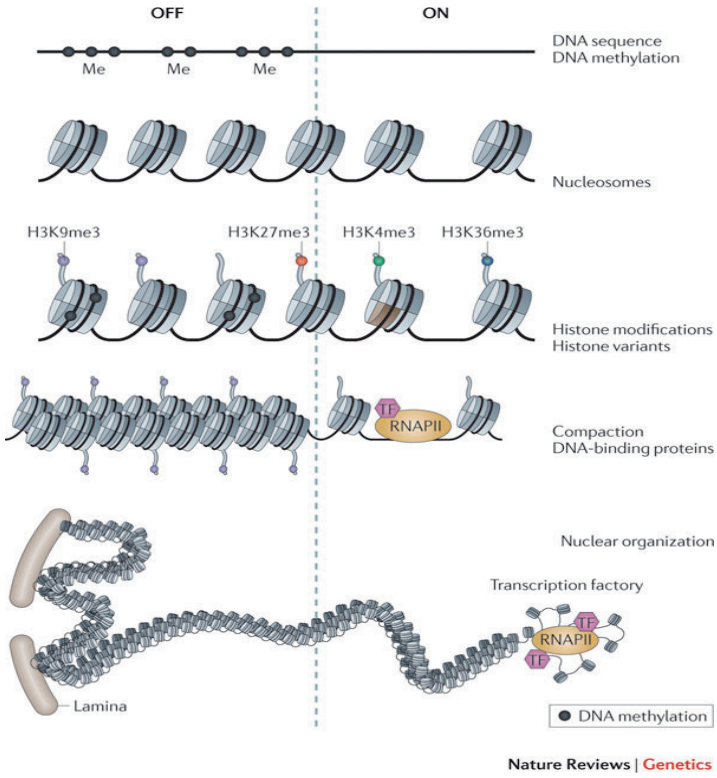


Figure 5. Levels of chromatin organization.
Levels of chromatin organization are divided according to their association with inactive (off) or active (on) transcription. Different histone modifications are represented by coloured dots and histone variants such as H2A.Z are brown. Higher-order structures such as nuclear lamina-associated domains and transcription factories are also depicted. (from Zhou et al, 2011)

Histone marks have been instrumental for the identification of functional elements in the genome. Particular combinations of modifications have been used to identify gene promoters, transcribed regions, and enhancers (Figure 6). The transcriptional regulator CTCF has been associated with boundary elements (that prevent the spread of heterochromatin), with insulators (DNA elements that block the function of enhancers), and mediation of long-range DNA interactions together with cohesin (Zhou et al, 2011).

Prediction of enhancers has proven particularly thrilling since the function of these DNA elements is to positively regulate transcription at distally located promoters (Visel et al, 2009a). To this end, they have been shown to recruit the transcriptional machinery, regulators of chromatin, transcription factors, and have been shown to associate with the co-activator p300 (Visel et al, 2009b). Interestingly, it was recently shown that RNA polymerase II (RNAPII) transcribes short non-coding RNAs at enhancers (termed eRNAs) the synthesis of which require interaction with the gene promoter; the levels of eRNAs mirror the levels of gene activity (Kim et al, 2010).

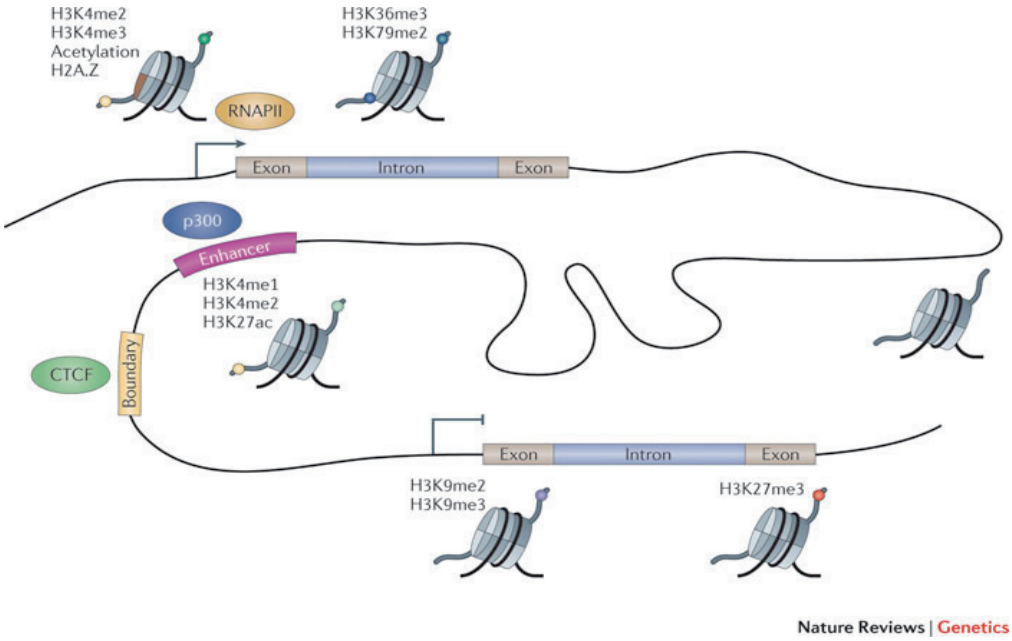


Figure 6. Histone modifications mark functional elements in the genome. Promoters, enhancers, gene bodies, and a boundary element are schematically shown Active promoters are enriched for histone H3 lysine 4 dimethylation (H3K4me2), H3K4me3, acetylation (ac), and H2A.Z. Transcribed regions are marked by H3K36me3 and H3K79me2. Repressed genes are found in regions with H3K9me2 and/or H3K9me3 or H3K27me3. Enhancers show enrichment for H3K4me1, H3K4me2, H3K27ac and p300. CTCF marks boundary elements and insulators. (from Zhou et al, 2011)

The epigenetic architecture of the genome changes during development to accommodate the complex processes of transcription activation and repression that accompany the establishment of tissue- and cell type-specific gene expression. Several patterns of regulation have been identified so far and they can be distinguished by the epigenetic signals present at gene promoters during development.

In embryonic stem cells, genes encoding transcription factors involved in development and lineage specification are marked by “bivalent chromatin domains” that contain both H3K27me3 (repressive) and H3K4me3 (activating) histone modifications. These domains resolve during development as the gene becomes either repressed or activated (by losing one of the two histone marks). Thus their function seems to be the maintenance of a “poised” chromatin state that would allow rapid activation or repression in response to subsequent developmental cues (Bernstein et al, 2006). HSCs have also been shown to have a large number of genes marked by a bivalent chromatin structure; lineage differentiation is associated with a reduction in their number (Cui et al, 2009; Weishaupt et al, 2010). Furthermore, the promoters and enhancers of genes that will become activated during differentiation

by losing H3K27me3 are pre-marked with additional modifications in HSCs (see Figure 7), whereas bivalent genes that carry the H3K9me3 mark will be silenced during differentiation (Cedar and Bergman, 2011).

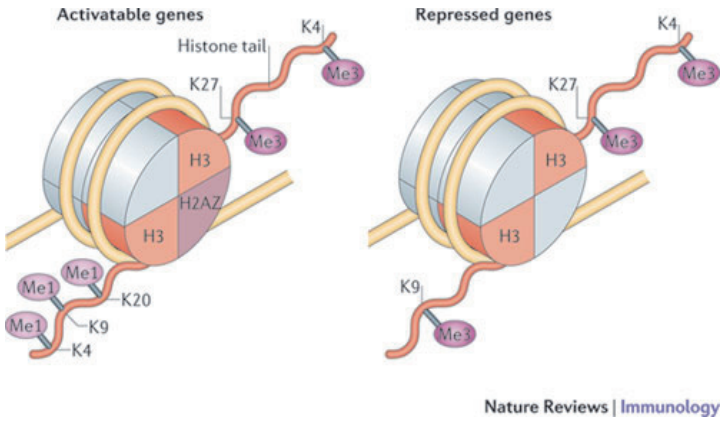


Figure 7. Gene priming with the help of epigenetic marks. H3 lysine 4 (H3K4me3) and H3K27me3 mark the promoters of bivalent genes. In HSC, two types of such bivalent promoters are present – those that will be activated later and those that will be subsequently silenced. These two types of promoters can be distinguished by the nature of histone modifications that accompany the classical bivalent pattern: histone H2AZ, H3K2me1, H3K4me1 and H3K9me1 mark activatable genes whereas H3K9me3 shows repressed genes. (from Cedar and Bergman, 2011)

The H3K4me2 histone mark constitutes an alternative method of poising developmentally important genes, such as *Rag2* (expressed in the lymphoid system) or *Gata1* (encoding a transcription factor specific for the erythroid lineage). Genes that contain H3K4me2 in early precursors (at the promoter or in the gene body) can be either demethylated and thus repressed during differentiation (e.g. *Rag2* in erythroid cells) or they can be further marked to H3K4me3 and become activated (e.g. *Rag2* in lymphoid cells) (Orford et al, 2008; Cedar and Bergman, 2011).

Genes that become devoid of chromatin marks (for example by demethylation of H3K4) are generally repressed by DNA methylation. In the hematopoietic system, *de novo* methylation is used to silence pluripotency genes as the cells differentiate to more specified fates (Cedar and Bergman, 2011). The reverse process (demethylation of genes required at later stages of differentiation) was also demonstrated to be important as it specifically targets genes in a certain cell type. For example, T-cell specific demethylation of *Lck* kinase that initiates signaling downstream the TCR has been reported (Ji et al, 2010).

Interestingly, a recent study of genome-wide methylation maps in hematopoietic progenitors identified marked differences in specification of myeloid versus lymphoid cell fates. Whereas myeloid

differentiation relies more on the loss of methylation and thus on an end-point pattern of reduced global DNA methylation, lymphopoiesis is characterized by the acquisition of DNA methyl marks, consistent with repression of specific genes (Ji et al, 2010). This phenomenon is also illustrated by the fact that induced pluripotent stem cells (iPSCs) derived from lymphocytes retain a degree of residual methylation. This "lymphocyte memory" correlates with increased blood-forming potential, feature that distinguishes lymphocyte-derived iPSCs from cells derived from the myeloid compartment (Kim et al, 2010b).

4. Pax5 regulates B cell commitment, development, and leukemogenesis

Pax5 is one of the most studied transcription factors of the B cell lineage (Cobaleda et al, 2007a). It controls the identity and function of B cells from the pro-B to the mature B cell stages (Figure 4). The question of how *Pax5* expression is regulated is the point of this study, and the importance for a tightly controlled regulation of this gene will be addressed below (Sections 4.2, 4.3 and 4.4). However, it is also interesting to look at Pax5 function in gene regulation, as a model of how transcription factor mediated effects on gene expression are achieved in the B cell lineage (Section 4.1).

4.1 Pax5-mediated gene regulation in the B cell lineage

Pax5 belongs to the paired box transcription factors (Pax) family, the members of which are important during embryogenesis and beyond for tissue specification and cellular differentiation (Robson et al, 2006). There are nine members of the family described in mammals, with orthologues identified in many species such flies, worms, frogs, fish, and birds (Dahl et al, 1997).

Pax5 is expressed in adult testis (Barberis et al, 1990) and at the midbrain-hindbrain boundary (Ye et al, 2001). It is also the only member of the family expressed in the hematopoietic system, where its expression is exclusive to B cells (Fuxa and Busslinger, 2007). Here, Pax5 tailors a B cell-specific genetic program through its ability to function in both gene repression and activation (Cobaleda et al, 2007a). Most of its functions at promoters of target genes are carried out through versatile protein-protein interactions with a variety of partners (Medvedovic et al, 2011).

The N-terminal DNA binding paired domain of Pax5 consists of 128 amino acids structured into two subdomains (each containing a homeodomain-like helix-turn-helix motif), which bind to adjacent major grooves in the DNA. This results in a degenerate recognition sequence, as each of the two sites makes

an independent contribution to the total affinity (Czerny et al, 1993). The two subdomains are connected by a flexible linker that makes DNA contacts in the minor groove (Xu et al, 1999). This N-terminal domain has been shown to recruit Ets family proteins to an element located in the *CD79a (mb-1)* gene (Fitzsimmons et al, 1996). Moreover, the same domain has been implicated in both synergistic and antagonistic interactions with other transcription factors (Figure 8).

However, the protein interaction motifs located centrally and C-terminally determine most of the Pax5 partners. The partial homeodomain has been shown to interact with the TATA-box binding protein of TFIID and the cell-cycle suppressor retinoblastoma protein (Eberhard and Busslinger, 1999). The C-terminal part of the transcriptional corepressor Daxx also interacts with the homeodomain; interestingly, this interaction can also result in transcriptional activation via recruitment of CREB-binding protein (Emelyanov et al, 2002). The octapeptide domain of Pax5 recruits members of the Groucho family of transcriptional repressors (Eberhard et al, 2000). The C-terminal domain of Pax5 is rich in prolines and hydroxylated serines, threonines and tyrosines (Adams et al, 1992); 55 aminoacids of this region form a transactivation domain, whereas the C-terminal amino acids confer inhibitory function (Dorfler and Busslinger, 1996). The transactivation domain has been shown to interact with histone acetyltransferases like CBP or the SAGA complex (Barlev et al, 2003). Pax5 also interacts with the Brg1 component of the BAF chromatin-remodeling complex (Barlev et al, 2003) and with the histone acetyltransferase p300 (He et al, 2011).

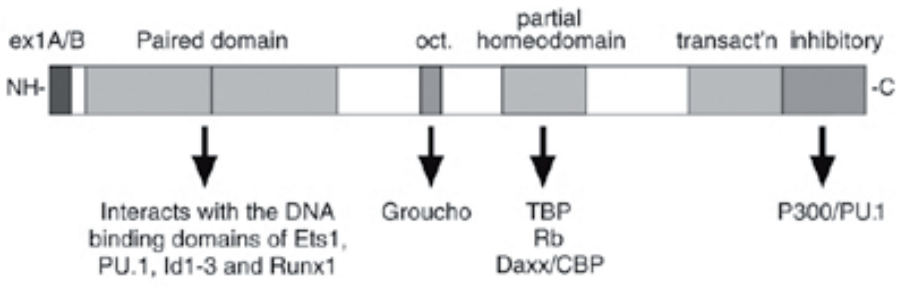


Figure 8. Pax5 structural domains and interaction partners. Pax5 is transcribed from two separate promoters (1A/1B), which result in two different mRNAs arisen from splicing of alternative 5' exons on the common coding sequence. The locations of the paired domain, octapeptide, partial homeodomain, transactivation and inhibitory domains are shown. Alongside, interaction partners of each of the domains are indicated below. TBP-TATA-box binding protein; Rb – retinoblastoma protein; CBP – CREB-binding protein (from Holmes et al, 2008)

Upon its expression, Pax5 restricts progenitors to the B cell fate, which led to its being named the "B cell commitment factor" (Cobaleda et al, 2007a). Commitment is achieved in part by repression of B lineage-inappropriate genes, which suppresses alternative lineage options (Delogu et al, 2006). At the same time Pax5 activates B lineage-specific genes that control processes such as adhesion, migration, signal transduction downstream of the BCR, germinal center formation, antigen presentation. Thus Pax5 enforces a B cell-specific developmental program (Schebesta et al, 2007). It was recently shown by streptavidin pulldown ChIP-chip experiments in pro-B cells expressing in vivo biotinylated Pax5 that 24% of previously-identified Pax5-repressed genes and 44% of activated genes are direct targets of Pax5 since it binds to their promoters and/or enhancers (McManus et al, 2011). This study identified a role for Pax5 as an epigenetic regulator that controls the chromatin state of its target genes. Pax5 binding is associated with the induction of active chromatin marks at regulatory elements of its target genes and the elimination of these marks at repressed genes (McManus et al, 2011). This is additionally supported by reports that identified various Pax5-interaction partners including chromatin remodelers (the BAF complex), histone methyltransferases and activators of transcription (two subunits of the MLL complex), transcriptional activators (p300, CBP, TFIID) and repressors (NcoRI) (Medvedovic et al, 2011). Figure 9 summarizes the interactions that Pax5 utilizes to regulate its target genes.

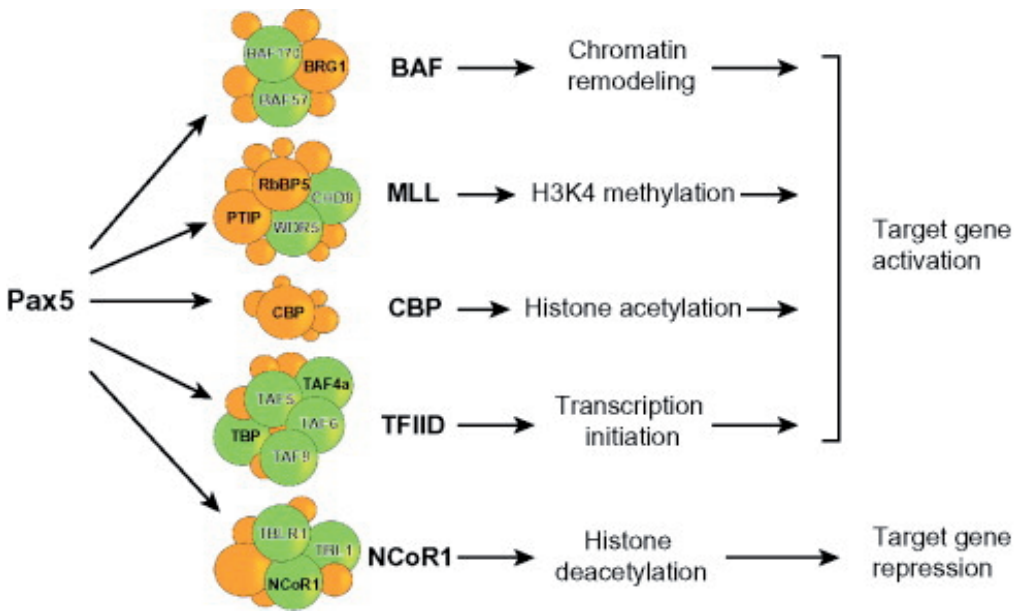


Figure 9. Characteristics and functions of the complexes Pax5 interacts with to regulate its target genes. See text for details. (from Medvedovic et al, 2011)

4.2 Pax5 in B lineage commitment and development

Most of the knowledge about Pax5 functions comes from the analysis of *Pax5*-deficient mice, either throughout development (constitutive *Pax5* deletion) – where B cell development is arrested at an early precursor stage (Urbanek et al, 1994), or after conditional inactivation (Horcher et al, 2001). In the fetal liver, absence of Pax5 results in a complete lack of B cell progenitors (defined by the surface marker B220). However, in the bone marrow of adult *Pax5*$^{-/-}$ mice B cell development proceeds to the pro-B cell stage (cKit^{+}B220^{+}) where early B cell-specific genes are already expressed (Nutt et al, 1997; Nutt et al, 1998).

Interestingly, *Pax5*$^{-/-}$ cells from adult bone marrow can be propagated in culture in the presence of stromal cells and Interleukin 7 (IL-7) thus displaying a self-renewal capacity (Nutt et al, 1997) similar to the self-renewal ability of wild-type HSCs (Schaniel et al, 2002a). Furthermore, if IL-7 is substituted with cytokines appropriate for other lineages, *Pax5*$^{-/-}$ cells can differentiate in vitro into granulocytes, macrophages, dendritic cells, osteoclasts, NK or T cells (Nutt et al, 1999a; Rolink et al, 1999; Hoflinger et al, 2004). When transplanted into recipient mice, these cells home to the bone marrow and develop into all the hematopoietic cells except B cells (Schaniel et al, 2002b). In order to differentiate into B cells, *Pax5* expression needs to be restored in *Pax5*$^{-/-}$ progenitors (Nutt et al, 1999a).
Conditional inactivation of *Pax5* in pro-B cells reverts lineage commitment and converts them into multipotent progenitors (Mikkola et al, 2002). Therefore, *Pax5* expression is required continuously in order to maintain B cell identity and prevent de-differentiation to an uncommitted cell stage. This is further supported by the fact that the physiological downregulation of *Pax5* expression in plasma cells is accompanied by the expression of many Pax5-repressed target genes (Delogu et al, 2006). Finally, conditional inactivation of *Pax5* in mature B cells results in cells de-differentiating to an uncommitted stage from which they can reconstitute T cell development (Cobaleda et al, 2007b).

Taken together, these studies prove that Pax5 is the B cell commitment factor and its expression needs to be maintained from the pro-B to the mature B cell stages. Presence of this transcription factor is crucial for the generation of the B cell lineage, and thus for the integrity of the antibody-mediated immune response. Cells in which *Pax5* expression is lost, or cells of the *Pax5*$^{-/-}$ genotype are uncommitted lymphoid progenitors with a latent myeloid potential (Cobaleda et al, 2007a). Consistent with its function, the expression pattern of Pax5 (monitored with an IRES-human CD2 reporter inserted in the 3' untranslated region of the *Pax5* gene) is initiated in a small proportion of uncommitted pre-pro-B but it is only fully expressed at the committed pro-B cell stage (Fuxa and Busslinger, 2007).

Expression of *Pax5* is maintained at a stable level throughout B cell development and is finally downregulated during plasma cell differentiation. During this time, not only does Pax5 maintain B cell identity, but it also fulfills crucial roles in the regulation of the immunoglobulin heavy-chain assembly and pre-BCR signaling (Figure 4).

Despite its crucial role in B cell function and essential contribution to lineage commitment, Pax5 cannot be considered a master regulator of B cell development (Cobaleda et al, 2007a) since panhematopoietic expression of *Pax5* under the control of the *Ikaros* locus (Souabni et al, 2002) or retroviral-mediated *Pax5* expression in HSCs (Cotta el al, 2003) does not interfere with the development of myeloid cell types in vivo. However, *Pax5* expression skews the developmental potential of lymphoid progenitors towards the B cell lineage at the expense of T lymphopoiesis. Therefore, *Pax5* expression is not sufficient to divert the fate of HSCs and erythro-myeloid progenitors to the B cell pathway. It relies on the collaboration of other transcription factors of the lymphoid lineage in order to induce B cell lymphopoiesis in progenitor cells (Cobaleda et al, 2007a).

4.3 Pax5 in leukemogenesis

In order to ensure correct activation of the B cell-specific gene expression program at the pro-B cell stage, as well as downregulation of *Pax5* expression at the onset of plasma cell development (Kallies et al, 2007), a tight regulation of *Pax5* expression needs to be in place at all developmental stages of the B cell lineage. Deviations from the dosage or physiological expression pattern of *Pax5* result in detrimental situations such as the development of various types of cancers. Pax5 can function both as an oncoprotein and as a tumor suppressor (Medvedovic et al, 2011; Cobaleda et al, 2007a).

In the case of aggressive B cell non-Hodgkin's lymphoma (arising from germinal center B cells), recurrent transformations between the *Pax5* and the IgH loci have been reported (Poppe et al, 2005). They can either bring the Pax5-coding region under the control of the strong Eμ enhancer (Busslinger et al, 1996) or under the control of the Sμ promoter in the *IgH* locus (Morrison et al, 1998), which results in general increased levels of Pax5 and continuous expression in plasma cells.

In the case of the B-progenitor acute lymphoblastic leukaemia, oncogenic Pax5 fusion proteins (Coyoud et al, 2010), as well as monoallelic deletions or point mutations that result in hypomorphic alleles or *Pax5* haploinsufficiency have been identified, indicating a tumor suppressor function of Pax5 (Mullighan et al, 2007). Pax5 fusion partners vary from transcription factors to kinases, chromatin regulators or proteins of unknown function. All fusion products retain the DNA binding paired domain of

Pax5 and most of them seem to act as constitutive repressors that antagonize the function of the wild-type Pax5 protein (Coyoud et al, 2010; Medvedovic et al, 2011). The role of monoallelic loss of *Pax5* in leukemogenesis is less well understood. B cells develop normally in heterozygous *Pax5*$^{-/+}$ mice (Nutt et al, 1999b); however, *Cd19-Cre Pax5*$^{-/fl}$ mice develop aggressive progenitor cell lymphomas upon dedifferentiation of B cells to an uncommitted progenitor, caused by the biallelic loss of *Pax5* (Cobaleda et al, 2007b).

4.4 Regulation of *Pax5* expression

Since *Pax5* expression controls B cell fate (Section 4.2) and its deregulation results in leukemias and lymphomas (Section 4.3), the question arises which upstream regulators control the activation of the *Pax5* gene in early B lymphopoiesis. Equally interesting is the question of how *Pax5* is repressed, both to prevent premature entry in the B cell differentiation program, and to ensure correct activation of the plasma cell program. Plasma cell development requires reactivation of several Pax5-repressed genes as well as a change in gene expression program that is coordinated by the transcription factor Blimp1 (Shaffer et al, 2002).

4.4.1 The transcriptional network upstream of Pax5

The question of how *Pax5* expression is activated in B lymphopoiesis provides a good starting platform for understanding the regulatory interactions between transcription factors in genetic networks. Studies so far have focused on identification of the main players (Figure 2) and on placing them at distinct developmental stages. Ikaros, PU.1, and E2A are mainly responsible for transcriptional priming prior to B cell commitment, whereas EBF1 fulfills multiple roles in specification and commitment immediately upstream of Pax5.

At the LMPP cell stage, the Kruppel-like zinc finger protein Ikaros is important, as *Ikaros*$^{-/-}$ mice have defects in self-renewal of HSCs, fail to upregulate Flt3 and thus have no CLP or other cells downstream of LMPPs (Nichogiannopoulou et al, 1999).

The Ets family transcription factor PU.1 has been shown to be important at multiple stages of B cell development. At the LMPP cell stage levels of PU.1 specify GMP versus CLP specification, whereas at the CLP stage it controls expression of the *Il7R*α chain gene and of *Flt3* (Carotta et al, 2010). While it continues to be expressed throughout the B cell lineage, PU.1 seems to be dispensable until later stages

where it functions in concert with IRF transcription factors and contributes to the control of plasma cell differentiation (Carotta et al, 2010).

E2A proteins have also been implicated in the induction of *Il7R*α and other B cell-specific target genes, among which those coding for the transcription factors EBF1, Foxo1, Pax5, IRF4 and IRF8 are particularly noteworthy (Kee and Murre, 1998; Murre 2007; Lin et al, 2010). The E2A protein is a basic helix-loop-helix transcription factor and it occurs in two splice variants, E12 and E47 (Sun and Baltimore, 1991). Although both variants are present in LMPPs and CLPs, only E47 is necessary for B lineage specification (Beck et al, 2009). E2A proteins are involved in HSC maintainance and lymphoid lineage priming in LMPPs. They are not required for the expression of Ikaros and PU.1 but they activate genes at the LMPP stage in synergy with these transcription factors (Ramirez et al, 2010). Mice deficient in E2A (*E2a*$^{-/-}$) have a complete block at the ALP stage before any *IgH* locus rearrangements occur (Bain et al, 1994; Zhuang et al, 1994). E2A directs the B cell-specific program due to its activation of the transcription factors EBF1 and Pax5, but it is also important for B cell maturation in the bone marrow and germinal center formation (Kwon et al, 2008).

Recently, Foxo1 and Bcl11a have been identified as members of the transcription factors network that leads to lineage commitment. Bcl11a deficiency results in absence of transcripts for *Il7r*, *Ebf1*, *Pax5* and *CD19* and a block of B cell development similar to that of *E2a*$^{-/-}$ mice; thus it was proposed that Bcl11a and E2A might act in synergy (Liu et al, 2003; Medina and Singh, 2005). *Foxo1* deficiency results in a developmental arrest at the pro-B cell stage (Dengler et al, 2008) and this transcription factor has been recently shown to participate, in concert with E2A and EBF1, in the early activation of a B-lineage program of gene expression (Lin et al, 2010).

Finally, EBF1 is a transcription factor that contains novel DNA binding ("the Zinc knuckle") and helix-loop-helix domains (Hagman et al, 2012). Its expression is initiated at the ALP, increases through the BLP and pre-pro-B stages, and reaches highest levels at the pro-B cell stage. *Ebf1* is downregulated during plasma cell differentiation (Vilagos et al, 2012). In the absence of EBF1, B cells are developmentally arrested at the pre-pro-B cell stage (Lin and Grosschedl, 1995). *Ebf1*-deficient cells fail to upregulate important B cell-specific transcripts like *CD79a*, *CD79b*, *Igll1*, and do not initiate *IgH* rearrangements. Thus EBF1 is essential for the specification of the B cell lineage (Hagman et al, 2012). EBF1 contributes to the commitment of B cells by activating *Pax5* (discussed below); it has also been demonstrated to have a Pax5-independent contribution to commitment, by repression of the myeloid fate determinants C/EBPα and PU.1, and the E2A antagonists Id2 and Id3 (Pongubala et al, 2008; Thal et al, 2009).

Importantly, dose-dependent effects of EBF1 have been highlighted in fetal liver B cell development of *Ebf1* heterozygous mice (O'Riordan and Grosschedl, 1999; Lukin et al, 2010). These effects were exacerbated by compound haploinsufficiencies of *E2a* (O'Riordan and Grosschedl, 1999) and *Runx1* (Lukin et al, 2010), suggesting coordinate activities of these factors. This was also confirmed by a high frequency of co-occupancy of E2A binding sites by EBF1 (Lin et al, 2010). $Ebf^{+/-}Runx1^{+/-}$ mice also display increased expression of NK-lineage markers, which are repressed upon enforced expression of EBF1 (Lukin et al, 2011), thus further demonstrating the necessity for normal levels of EBF1 in order to maintain B cell identity. Furthermore, EBF1 target genes throughout B cell development include components of pre-BCR and PI3K signaling, as well as molecules important for cell adhesion and migration (Lin et al, 2010; Treiber et al, 2010). This is well illustrated by the strict requirement of EBF1 for the maintenance of pro-B cells and the transition to pre-B cells (Vilagos et al, 2012). Very importantly, EBF1 and Pax5 have non-redundant functions in B cell commitment, since ectopic *Pax5* expression from the *Ikaros* locus does not rescue *Ebf1* deficiency, and conversely *Ebf1* expression from the *Rosa26* locus does not overcome *Pax5* deficiency. In the late B cell compartment, MZ and B1 cells are lost upon conditional EBF1 inactivation, whereas follicular and germinal center B cells of the spleen can partially tolerate the loss of EBF1 (Vilagos et al, 2012).

A clear hierarchical picture emerges from the above-described experiments, however direct interactions between transcription factors are not elucidated by such approaches. Studies of the direct action of DNA-binding factors on expression of their target genes at each developmental stage are needed to complement the knowledge gained so far. For example, the activation of the *Ebf1* gene has received some attention, and the findings are noteworthy. *Ebf1* is regulated by a distal (α) and a proximal (β) promoter. The *Ebf1*α promoter is directly bound by E2A (Smith et al, 2002) and also regulated by STAT5 mediated IL-7 signaling. Once EBF1 is expressed, it also binds *Ebf1*α promoter in an autoregulatory mechanism (Roessler et al, 2007). *Pax5* expression is dependent on EBF1 (Decker et al, 2009) and, once activated, Pax5 binds the *Ebf1*β promoter together with PU.1 and Ets1 (Roessler et al, 2007). Thus, Pax5 exerts a feedback regulation on *Ebf1* in order to consolidate commitment to the B cell lineage. The combinatorial mechanism of *Ebf1* gene activation also ensures amplification of B cell-specific expression of EBF1 (Nutt and Kee, 2007).

4.4.2 Direct upstream regulators of *Pax5* expression

In the case of *Pax5*, genetic experiments have placed the transcription factors E2A and EBF1 upstream of this gene because *Pax5* transcripts are not present in progenitors deficient in either E2A or EBF1 (Bain et al, 1994; Lin and Grosschedl, 1995). Since these findings did not clarify the mechanism of action of EBF1 and E2A on the *Pax5* gene, a bottom-up approach was necessary in order to identify the *cis*-regulatory elements of the *Pax5* locus.

The *Pax5* gene spans 190 kb and is transcribed from two separate promoters (1A and 1B) which result in two different mRNAs arisen from splicing of alternative 5' exons onto the common coding sequence (exons 2-10). A transgenic reporter approach identified a 435 bp enhancer in the 5' flanking region of the gene, which was sufficient to direct the correct expression of *Pax5* at the mid-hindbrain boundary (Pfeffer et al, 2000). However, this enhancer did not give rise to B lymphocyte-specific reporter expression and thus showed that the regulatory elements necessary to drive expression of *Pax5* in the lymphoid system are located elsewhere. A transgenic analysis combined with deletion mapping identified an enhancer in intron 5 of the *Pax5* gene (Decker et al, 2009). Together with the promoter region (consisting of 8 DNaseI hypersensitive sites), the B cell-specific enhancer (which contains 3 DNaseI hypersensitive sites) was enough to recapitulate the *Pax5* expression pattern in vivo in a transgenic assay. Therefore, all regulatory elements that drive the B cell-specific expression of the *Pax5* gene are located in the enhancer located in intron 5 and the promoter region (Figure 10). Chromatin immunoprecpitation (ChIP) experiments identified EBF1-binding sites in hypersensitive sites 6 and 7 (HS6 and HS7) of the *Pax5* promoter and an E2A-binding site in HS6 (Figure 10). The presence of these binding sites was confirmed by electrophoretic mobility shift assays (EMSA), ChIP-Seq and in vivo DMS footprinting (Decker et al, 2009; Hiromi Tagoh, personal communication; Busslinger laboratory, unpublished observations). E2A is already expressed in *Ebf1*$^{-/-}$ progenitor cells, however in these cells the *Pax5* promoter is still repressed by H3K27me3 and is not hypersensitive to DNaseI digestion (Decker et al, 2009). Consistent with these observations, E2A binding at HS6 is first detected in *Pax5*$^{-/-}$ cells, in which EBF1 is already expressed (Busslinger laboratory, unpublished data). Thus, E2A binding is secondary to EBF1 action in this context.

Although it was previously published that EBF1 regulates *Pax5* by binding to a site 1110 bp upstream of the transcription start site of the distal promoter (O'Riordan and Grosschedl, 1999), this binding site could not be confirmed in the studies mentioned above. The same holds true for the overlapping binding site for STAT5, which was shown to act in cooperation with EBF1 in activating *Pax5* transcription

(Goetz et al, 2005; Hirokawa et al, 2003). The absence of STAT5-binding sites in the *Pax5* promoter and enhancer regions is consistent with the observation that *Pax5* is normally expressed in Bcl2-rescued STAT5-deficient pro-B cells (Malin et al, 2010).

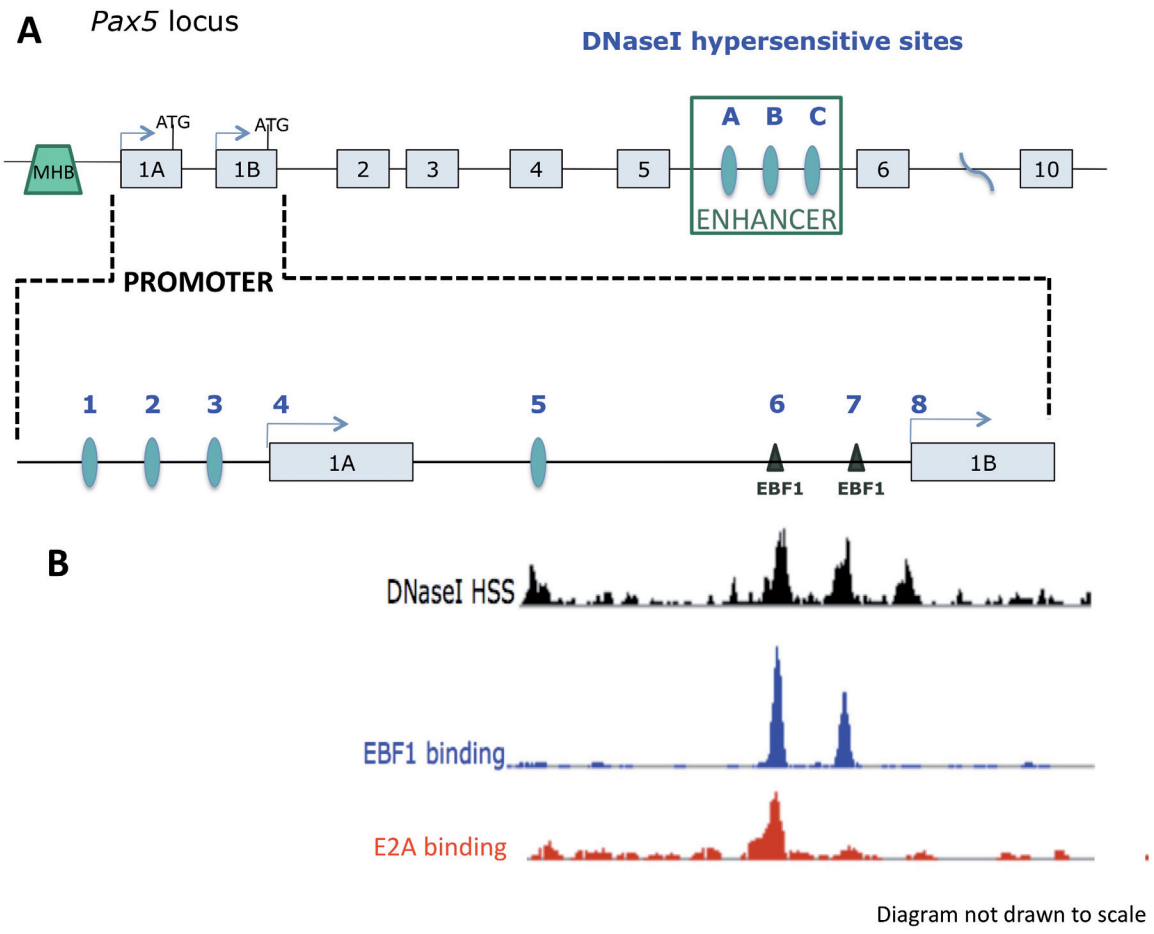


Figure 10. Schematic representation of the *Pax5* promoter and enhancer and direct binding of EBF1 and E2A to the *Pax5* promoter. A. Schematic representation of DNaseI hypersensitive sites (blue ovals) in the *Pax5* promoter (blow-up) and enhancer (marked in intron 5) regions. The numbering for the DNase I hypersensitive sites is made on top, in blue. EBF1-binding sites, as identified by Decker et al (2009) and Busslinger laboratory (unpublished observations) are also represented as dark triangles. The previously identified mid-hindbrain enhancer is shown upstream of exon 1A (MHB, green trapeze). **B.** EBF1- and E2A-binding peaks (ChIP-Seq) and DNaseI hypersensitive sites identified by genome-wide approaches (in $Rag2^{-/-}$ cells) at DNaseI hypersensitive sites 6 and 7.

The *Pax5* promoter (as well as the *Ebf1* promoter) is "poised" in HSCs by the presence of both repressive (H3K27me3) and active (H3K4me3) chromatin marks (Adli et al, 2010). These marks are deposited by the Polycomb repressive complex 2 (PRC2) and Trithorax group proteins respectively, in order to ensure the possibility for rapid activation at subsequent developmental stages (Ringrose and Paro, 2004; Spivakov and Fisher, 2007). However, the Polycomb repressive complex 1 (PRC1) exerts an additional layer of regulation through its component Bmi1, which maintains the self-renewal properties of adult stem cells, HSCs and MPPs. Bmi1 has been shown to directly bind the *Pax5* and *Ebf1* promoters in order to maintain it in a bivalent state prior to expression. Absence of Bmi1 results in premature resolution of bivalent domains followed by expression of *Pax5* and *Ebf1* in MPPs and an accelerated specification of the lymphoid lineage (Oguro et al, 2010).

The hypersensitive site B (HS-B) in the newly discovered *Pax5* enhancer has been shown to contain binding sites for the transcription factors PU.1, IRF4, IRF8, and NF-kB, which are sequentially expressed during HSC differentiation to mature B cells. This is consistent with the finding that the enhancer (which is silenced by DNA methylation in embryonic stem cells) is already activated in MPPs and thereafter during the entire B cell development (Decker et al, 2009). In contrast, the promoter region is silenced by the presence of PRC1 and PRC2 in ES cells, MPPs, and non-B cells. The promoter is only open in the presence of the transcription factor EBF1 when DNaseI hypersensitive sites in the promoter are formed and active histone modifications (H3K9Ac) are induced (Decker et al, 2009). These findings led to a model proposing stepwise activation of the *Pax5* gene at the onset of B cell development, by enhancer activation in MPPs and subsequent opening of the promoter region in pre-pro-B cells in an EBF1 dependent manner (Decker et al, 2009).

5. EBF1 functions as an activator of gene expression

Although both E2A and EBF1 bind the *Pax5* promoter, it is the presence of EBF1 that results in the induction of active chromatin states and DNaseI hypersensitive sites at the *Pax5* promoter (Decker et al, 2009; Treiber et al, 2010). This is in agreement with previous reports that identified EBF1 as a "pioneer transcription factor", a term used to describe factors that are among the first to access a locus in a tissue-specific manner, induces changes in chromatin architecture and histone modifications, recruit additional transcription factors and modifiers, and thus initiate a cascade that eventually results in transcriptional activation (Hagman and Lukin, 2005; Smale 2010).

This view of early EBF1 action at gene promoters has been expanded by genome-wide profiling of histone marks and transcription factor binding. However, the function of EBF1 at its target genes is just being unraveled and a clear view of the mechanisms employed (as available for Pax5, see section 4.1) is not yet available. While some mechanisms might be common to the two transcription factors, examples so far hint at a versatile usage of epigenetic mechanisms by EBF1 (see below).

Treiber and colleagues (2010) identified three classes of EBF1 target genes whose histone mark signature changes upon EBF1 expression at the transition from pre-pro-B cells to pro-B cells: activated (first occupied by H3K4me2, followed by acquisition of H3K4me3 and H3K9Ac), poised (H3K4me2 presence at genes for which EBF1-dependent transcription will occur at later mature B cell stages), and repressed (maintenance of H3K27me3 and loss of active marks) (Treiber et al, 2010).

However, this work challenges the hypothesis of EBF1 being a "pioneer transcription factor". Treiber and colleagues demonstrate that EBF1 binding at its target genes requires a previously-established permissive chromatin context, since it does not bind its adipocyte- or neuronal-specific target genes in pro-B cells, nor does EBF1 bind its pro-B cell-specific targets in a fibroblastic cell line (Treiber et al, 2010). An interesting model for the elucidation of the mechanisms of gene activation in early B lymphopoiesis is the B cell-specific *Cd79a (mb-1)* promoter. Prior to its activation, the *mb-1* promoter is found in an inaccessible state, characterized by DNA methylation and inactive histone modifications (H3K27me3) (Gao et al, 2009; Treiber et al, 2010). This repressive chromatin signature is a substrate for the Mi-2/NuRD complex, known to remove active histone modifications and assemble nucleosomes in a compact structure (Ramirez and Hagman, 2009). The promoter is pre-activated in CLPs and pre-pro-B cells, when binding of E2A induces H3K4me1 (Lin et al, 2010). Following its expression, EBF1 together with RUNX1 assembles complexes (that include E2A) on the *mb-1* promoter. Importantly, the presence of EBF1 at the *mb-1* promoter is sufficient to initiate DNA demethylation (Maier et al, 2004) and

chromatin remodeling by the direct recruitment of the SWI/SNF complex, which antagonizes the Mi-2/NuRD complex (Gao et al, 2009). Finally, binding of Pax5 leads to the propagation of DNA demethylation, exclusion of Mi-2/NuRD, induction of active histone modifications (H3K4me3, H3K9Ac), and transcription. Pax5 binds the *mb-1* promoter in cooperation with Ets-1 but their actions are dependent on prior demethylation of the promoter is response to EBF1 action (Maier et al, 2003). The involvement of EBF1 at later stages of activation could involve the recruitment of the histone acetyltransferase domains of p300 (He et al, 2011). p300 can also directly acetylate Pax5, thus increasing its activity (He et al, 2011).

In a study that investigated the mechanism of CD19 activation in B lymphocytes, Walter and colleagues demonstrated that the transcription factors E2A and then EBF1 are involved in priming of the gene enhancer, prior to Pax5 expression and Pax5-dependent transcription (Walter et al, 2008).
Thus EBF1 is clearly involved in the epigenetic priming of regulatory elements of B cell-specific genes.

6. Aim of this work

The developmental pathway that leads to B cell commitment has long been regarded as a paradigm for the study of cell-fate specification. Elucidation of the molecular mechanisms that govern this process might provide insight into the general principles of gene regulation and cellular differentiation. To achieve this, one can start from understanding how Pax5, the commitment factor of the B cell lineage, is regulated. This knowledge will also provide a handle on the development of lymphomas and lead to an improvement of current therapeutic strategies.

Whereas information is accumulating on the transcriptional network upstream of Pax5, as well as on the molecular events downstream of it, the molecular mechanisms that lead to *Pax5* gene activation remain elusive. A step forward was made by identification of the *cis*-regulatory elements in the *Pax5* locus and some of the transcription factors that bind them (Decker et al, 2009). However, the identification of the exact contribution and importance of each of the identified regulatory sequences in the *Pax5* locus, as well as their cooperation at various points in the development of B cells, remains a challenge for the future. The focus of this study is to mechanistically explain EBF1 activity at the *Pax5* promoter, mediated by binding to the hypersensitive sites 6 and 7 (HS6 and HS7). In my thesis I am addressing, with a series of in vivo mutagenesis experiments, the hypothesis that these two hypersensitive sites are the main elements through which the activation of *Pax5* transcription is initiated.

EBF1 and Pax5 have non-redundant functions in B cell commitment. However, their interplay is highly complex as they co-regulate many genes important for B cell function. Also, both EBF1 and Pax5 have been shown to act as epigenetic regulators that change the chromatin status of their target genes. Recent genome-wide studies have identified their direct targets, but have not provided full insight into the relationship between them. We know that EBF1 acts genetically upstream of Pax5 and that changes of the *Pax5* promoter to a transcriptionally-permissive state only happen when EBF1 is expressed (Decker et al, 2009). Once activated, Pax5 binds the *Ebf1*β promoter, resulting in a feedback amplification of *Ebf1* expression (Roessler et al, 2007; Nutt and Kee, 2007). The question that immediately arises is whether EBF1 directly activates *Pax5* expression, thus indirectly regulating its own levels and ensuring B cell commitment. There are only two identified EBF1-binding sites in the *Pax5* promoter (Figure 10). What is their function? Is EBF1 binding at these two sites, accompanied by EBF1-dependent chromatin remodeling, enough to drive *Pax5* transcription?

To answer the question of the contribution of direct EBF1 binding at HS6 and HS7 to the activation of *Pax5* expression, I have undertaken a knock-out approach by which I mutated the EBF1-binding sites in the endogenous *Pax5* locus. This allowed investigation of their contribution to the expression of *Pax5* independent of disturbances in B lymphopoiesis that results from *Ebf1* inactivation (Vilagos et al, 2012). To monitor changes in the level of *Pax5* expression upon deletion of EBF1-binding sites in vivo at single-cell resolution, targeting experiments were performed in an ES cell line that carries an IRES-human CD2 reporter gene in the 3' untranslated region of the *Pax5* locus (Fuxa and Busslinger, 2007). I report here the generation of $Pax5^{\Delta EHS7/+}$ and $Pax5^{\Delta EHS6+7/+}$ mice that lack the EBF1-binding site in HS7 or both EBF1-binding sites in HS6 and HS7, respectively.

RESULTS

1. Strategy for the generation of *Pax5*$^{\Delta EHS7/+}$ and of *Pax5*$^{\Delta EHS6+7/+}$ mice

The AP5 embryonic stem (ES) cell line was isolated from blastocytes of a cross of *Pax5*$^{ihCD2/ihCD2}$ (C57BL/6) and a wild-type mouse (129/Sv) and is therefore of mixed 129/Sv x C57BL/6 genetic background. Targeting experiments were thus conducted in a heterozygous *Pax5*$^{ihCD2/+}$ ES cell line, with the intention to preferentially target the C57BL/6 allele that carries the IRES-human CD2 reporter (IRES-hCD2) in the 3' untranslated region of the *Pax5* locus (Fuxa and Busslinger, 2007). After identification of clones correctly targeted with knock-out constructs that delete the EBF1-binding sites in the *Pax5* promoter, this approach requires an additional screening of the correctly targeted ES cell clones for the occurrence of homologous recombination on the desired *Pax5-IRES-hCD2* allele. The immediate advantage of the heterozygous *Pax5*$^{ihCD2/+}$ ES cell line stems from the fact that one functional allele of *Pax5* is enough to drive normal B cell development (Nutt et al, 1999b). If the mutations of EBF1-binding sites in the *Pax5* promoter of the IRES-hCD2 carrying allele were to result in abrogation of gene expression, the other functional allele would be enough to drive B lymphopoiesis and thus allow analysis of *Pax5* expression (monitored by hCD2 expression on cell surface) already in the peripheral blood of chimeric mice.

1.1 Generation and colony establishment for *Pax5*$^{\Delta EHS7/+}$ mice

For historical reasons, the first to be targeted was the EBF1-binding site in HS7 of the *Pax5* promoter, in order to generate *Pax5*$^{\Delta EHS7/+}$ mice. To this end, a strategy was designed that, following deletion of the neomycin resistance gene, would result in the replacement of 17 bp containing the EBF1-binding site in HS7 with an Frt and a loxP site covering a total of 106 bp (Figure 11A). Before targeting experiments were conducted, the Pax5$^{\Delta EHS7+neo}$ targeting vector was sequenced throughout (not shown). Using a Southern blotting screening that allows direct digestion of genomic DNA in 96-well format, I identified seven ES cell clones, in which one of the two *Pax5* alleles was correctly targeted (Figure 11A and 11B). Five of these clones were injected into blastocysts, four contributed to the birth of chimeric mice, and three resulted in germline transmission.

These three chimeric lines were subsequently tested for targeting on the *Pax5* allele that carries the IRES-hCD2 reporter. The immediately available way to address this was analysis of the F1 generation

from a cross between chimeric mice and C57BL/6 wild-type mice. If both the knock-in modification of interest and the IRES-hCD2 reporter were present on the same *Pax5* allele (as illustrated in Figure 11A), they would co-segregate in offspring mice of the F1 generation. However, if the two modifications are present on different *Pax5* alleles, offspring mice will only have one or the other gene modification. Genotyping for the presence of the IRES-hCD2 reporter gene was carried out using previously reported polymerase chain reaction (PCR) primers (Fuxa and Busslinger, 2007). To detect the targeted *Pax5* allele, a Southern blotting strategy was employed that uses a probe internal to the targeting vector (Figure 12A). As shown in Figure 12B, one of the three chimeric lines (resulted from ES cell clone 5E) showed co-segregation of the two modifications. Analysis of peripheral blood of chimeric mice of clone 5E showed a very mild reduction in *Pax5* expression (monitored by hCD2) in comparison to the expression levels of clone 5F - $Pax5^{\Delta EHS7+neo}$ targeted on the "other" *Pax5* allele and thus containing a wild-type $Pax5^{ihCD2}$ allele (Figure 12C).

Next, clone 5E was crossed with a mouse carrying Flpase (FlpE) transgene (Rodriquez et al, 2000) in order to remove the neomycin resistance cassette. The offspring of this cross was analyzed by Southern blot, with a strategy based on the one presented in Figure 12A. The new Southern strategy uses a double digestion (BglII and BamHI) instead of the original digestion scheme (BglII; Figure 12A), and the same probe. Since the BamHI site of interest is present in the loxP site, it is possible to monitor successful removal of the neomycin resistance cassette (Figure 13A). The results for four offspring mice of clone 5E are presented in Figure 13B, upper panel. At the same time, I designed a PCR strategy for further genotyping of the neomycin-deleted $Pax5^{\Delta EHS7/+}$ mice. It amplifies across the EBF1-binding site in HS7 and thus results in bands of different sizes, for the wild-type and the knock-in *Pax5* alleles (Figure 13B, lower panel). Since the PCR results confirmed the Southern blot genotyping, further mice were only genotyped with a combination of PCR reactions that detect the knock-in at HS7 and the IRES-hCD2 reporter. Mice labeled from here on as $Pax5^{\Delta EHS7/+}$ therefore represent the neomycin deleted knock-in mice that have the EBF1 site in HS7 of the *Pax5* promoter replaced with Frt-plus-loxP on the allele carrying an IRES-hCD2 reporter in the 3'untranslated region of *Pax5*.

After backcrossing 5 times into the C57BL/6 background, $Pax5^{\Delta EHS7/+}$ mice were crossed to obtain the homozygous genotype $Pax5^{\Delta EHS7/\Delta EHS7}$. At this point, the genotyping PCR band was sequenced, and the results confirmed the expected configuration (Figure 13C).

A

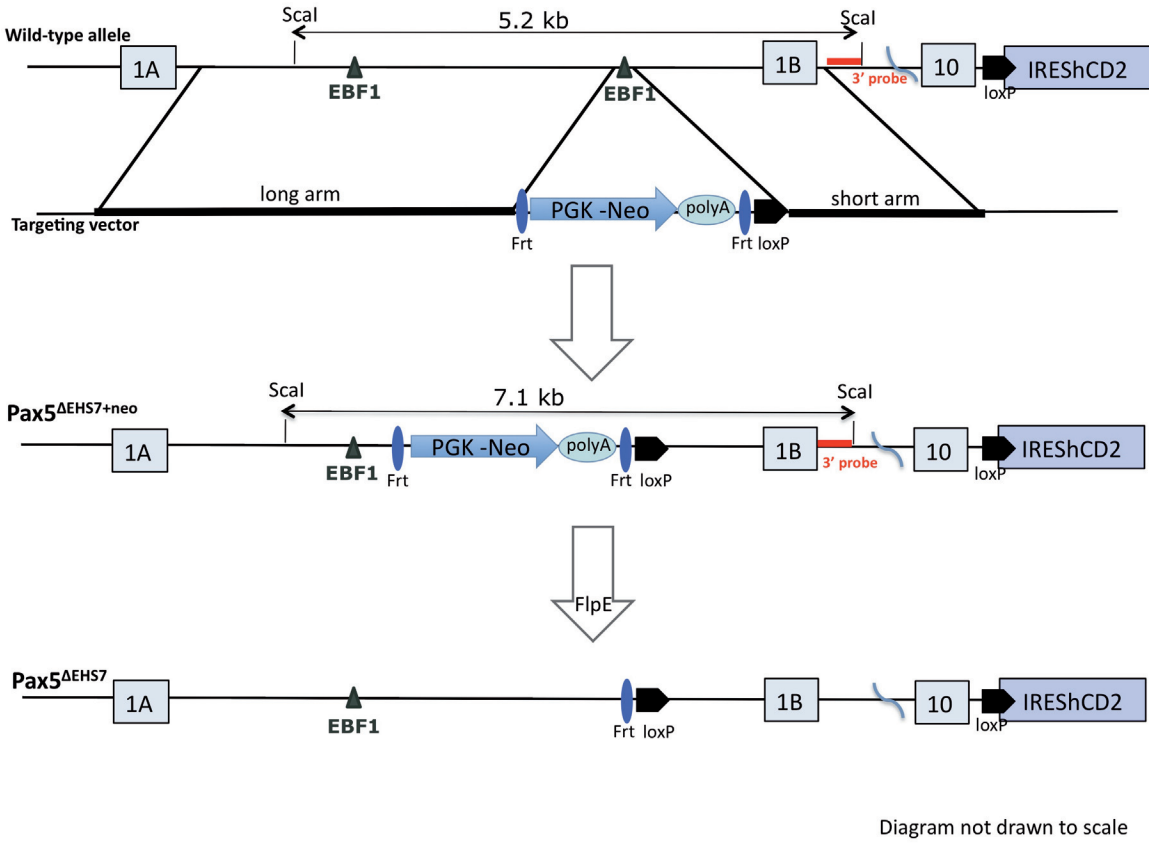


Diagram not drawn to scale

B

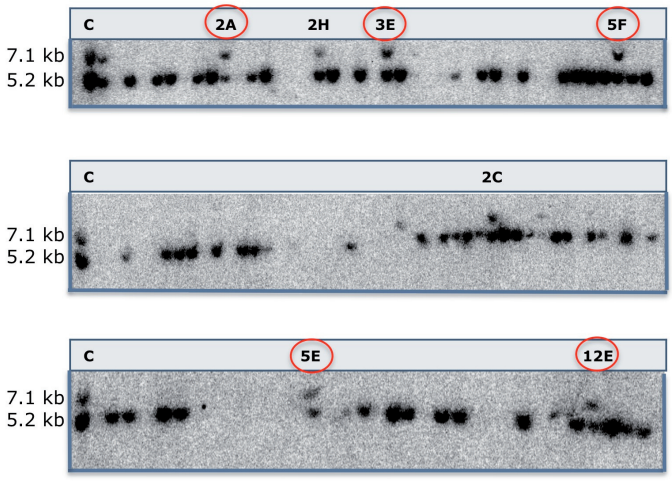


Figure 11. ES cell targeting strategy and screening to obtain the *Pax5*$^{\Delta EHS7/+}$ mice.
(see explanation on next page)

Figure 11. ES cell targeting strategy and screening to obtain the *Pax5$^{\Delta EHS7/+}$* mice. (see previous page) **A.** Schematic representation of Pax5$^{\Delta EHS7+neo}$ targeting vector and targeted Pax5 alleles (before and after neomycin resistance cassette removal). All inserted modules are represented, together with the span of homology arms. Homologous recombination following electroporation in ES cells results in a targeted allele that can be distinguished from the wild-type allele by a Southern blotting strategy. Restriction sites for the enzyme used, the length of the respective fragments, as well as the position of the probe used for Southern detection are indicated. FlpE mediated deletion of the neomycin (Neo) cassette can be achieved by crossing the chimeric mice with mice carrying the FlpE transgene (Rodriguez et al, 2000).

B. Southern blot screening of picked ES cell clones, using the strategy depicted in A. The five clones circled in red were selected for injection into blastocysts.

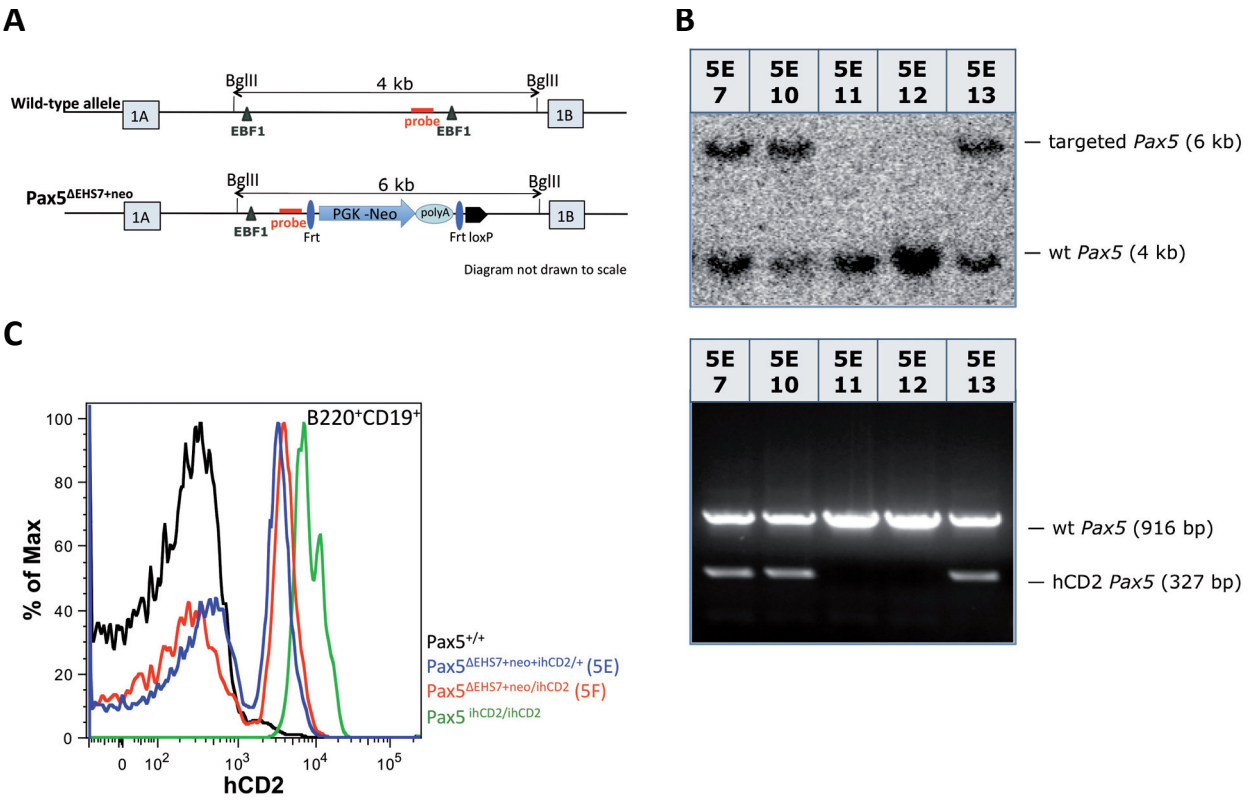


Figure 12. Clone 5E (*Pax5*$^{\Delta EHS7+neo/+}$) is targeted on the allele with the IRES-hCD2 reporter and expresses *Pax5* at almost normal levels in peripheral blood.

A. B. Five mice from the F1 generation of a cross of a 100% chimeric 5E male with a wild-type (wt) female were genotyped for the presence of both IRES-hCD2 and the knock-in modification introduced with the Pax5$^{\Delta EHS7+neo}$ targeting. **A.** Schematic representation of the restriction digest strategy used for Southern blotting detection in panel B, top. Restriction sites used are depicted as well as the position of the detection probe. **B.** Analysis of littermates from the cross between *Pax5*$^{\Delta EHS7+neo/+}$ chimeric 5E male and a wt female. Southern blotting detection for the Pax5$^{\Delta EHS7+neo}$ knock-in (upper panel, according to strategy in A.) and PCR detection for the IRES-hCD2 (bottom panel, according to Fuxa and Busslinger, 2007; Table 4). In all the mice analyzed, the two modifications co-segregate and thus illustrate that clone 5E is targeted on the desired *Pax5* allele.

C. Chimeric mice of clone 5E (Pax5$^{\Delta EHS7+neo}$ targeted on the Pax5-IRES-hCD2 allele, see A and B of this figure) and clone 5F (Pax5$^{\Delta EHS7+neo}$ targeted on the wild-type Pax5 allele, data not shown) were bled from the tail vein and the blood was analyzed by Fluorescence Activated Cell Sorting (FACS) for B220^{+}CD19^{+} cells. Expression levels of hCD2 were monitored in this population. The hCD2 negative cells correspond to the non-targeted B cells of the chimeric mice. As controls, *Pax5*$^{+/+}$ and *Pax5*$^{ihCD2/ihCD2}$ mice were used. Notice the two-fold reduction in hCD2 levels between mice that have one reporter-carrying allele (experimental chimeric mice) and two alleles with the reporter (control *Pax5*$^{ihCD2/ihCD2}$ mice).

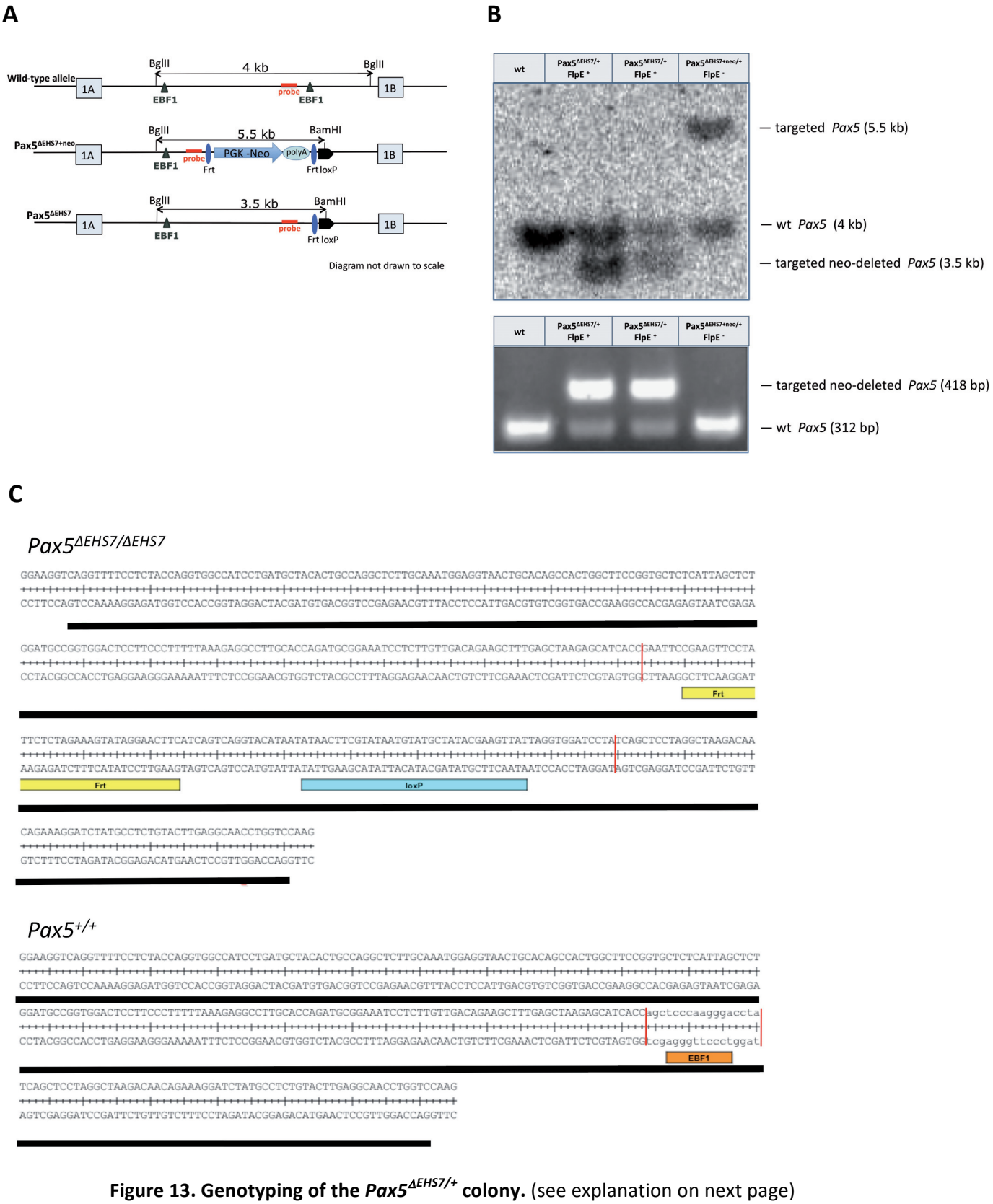


Figure 13. Genotyping of the *Pax5*$^{\Delta EHS7/+}$ colony. (see explanation on next page)

Figure 13. Genotyping of the *Pax5*$^{\Delta EHS7/+}$ colony. (see previous page)

A. Southern blot strategy to confirm deletion of neomycin resistance following mating of chimeric mice with FlpE mice**;** restriction sites used are depicted and the position of the detection probe.

B. Offspring mice of a cross of a 100% chimeric 5E male with a FlpE female were genotyped by Southern blot (top panel) and PCR (bottom panel). From left to right the genotype of analyzed mice is: *Pax5*$^{+/+}$, *Pax5*$^{\Delta EHS7\text{-}neo/+}$, *Pax5*$^{\Delta EHS7\text{-}neo/+}$, *Pax5*$^{\Delta EHS7+neo/+}$. Genotyping primers are summarized in Table 4.

C. Sequencing of genotyping PCR band for one *Pax5*$^{\Delta EHS7/\Delta EHS7}$ and one *Pax5*$^{+/+}$ mice resulted from the cross of two *Pax5*$^{\Delta EHS7\text{-}neo/+}$ mice (from here on simply termed *Pax5*$^{\Delta EHS7/+}$). Primers are in Table 4. The black line shows the coverage of the sequencing result for which no mismatch from the expected sequence could be identified (100% overlap to expected sequence). The coverage of the Frt, loxP, and EBF1-binding site are shown as filled bars below the sequence. The red vertical lines mark the positions in between which the sequences differ between the wild-type and the knock-in situations.

1.2 Generation and colony establishment for *Pax5*$^{\Delta EHS6+7/+}$ mice

The second targeting strategy developed aimed to generate *Pax5*$^{\Delta EHS6+7/+}$ mice (in which both EBF1-binding sites at HS6 and HS7 are removed), and *Pax5*$^{\Delta EHS6/+}$ mice (in which only the EBF1-binding site at HS6 is removed) in a single targeting experiment. A strategy was developed, in which the neomycin resistance gene was placed upstream of the region of interest, and the desired mutations were introduced in the short arm as mismatches as follows: the EBF1-binding site at HS6 was inactivated by the insertion of an NdeI restriction site, whereas the EBF1-binding site at HS7 was replaced by a ClaI restriction site (Figure 14A). Targeting experiments with this vector can result in different outcomes in individual ES cell clones, depending on the position where homologous recombination occurs (Figure 14B). Before targeting, the Pax5$^{\Delta EHS6+7+neo}$ vector was sequenced throughout (not shown). Screening by Southern blotting following direct digestion of genomic DNA in 96-well format was used to identify correctly targeted clones (Figure 14B and 14C).

Surprisingly, the majority of the ES cell clones picked appeared to be targeted, probably due to an unspecific band contaminating the Southern detection. This called for further screening of candidate clones before injection into blastocysts. This screen aimed to investigate in which of the clones the targeting of EBF1-binding sites in the *Pax5* promoter took place on the Pax5-IRES-hCD2 allele. For this purpose, a loxP site introduced in the targeting vector was used; the previous targeting for introduction of the IRES-hCD2 reporter gene also generated a loxP site in *Pax5* exon 10. Thus Cre-mediated deletion of the sequences between the two loxP sites can be used to investigate which ES cell clones carry the desired configuration at the *Pax5* locus (Figures 14B and 15A). Nine of the candidate clones were expanded (chosen to represent the diversity of positively-targeted ES cell clones based on different band intensity ratios in the screening Southern blot) and in vitro infected with an adenovirus that expresses the Cre recombinase (AdenoCre). Since AdenoCre infection results in the deletion of the entire sequence between them, it is possible to detect a PCR product with primers placed immediately upstream of the loxP site in intron 1 and downstream of the loxP site in exon 10 (Figure 15A). This PCR band is only detectable when the two loxP sites are on the same *Pax5* allele, and therefore its presence indicates targeting of Pax5$^{\Delta EHS6+7+neo}$ on the allele carrying the IRES-hCD2 reporter gene. Two out of the nine clones (2B and 9G) were PCR-positive after AdenoCre infection, and thus these clones were chosen for further analysis and subsequent injection into blastocysts (Figure 15B). Firstly, a Southern blot strategy with a different digestion scheme (SphI) and a probe that hybridizes upstream (5') of the targeting vector was used (Figure 15C) to confirm that the two clones of interest were correctly targeted (Figure 15D).

Secondly, it was crucial at this point to investigate whether homologous recombination introduced the desired mutations at both HS6 and HS7 or at HS6 alone (see Figure 14B). For this, two PCR reactions were carried out, across each of the sites of interest (EBF1-binding site in HS6 and EBF1-binding site in HS7 – see Figure 15E). The PCR products were subsequently digested with the restriction enzyme that would have been introduced at the respective site in case of favorable homologous recombination (NdeI and ClaI respectively). This experiment demonstrated that clone 9G contained the desired mutations at both HS6 and HS7, whereas clone 2B contained two wild-type EBF1-binding sites, most likely due to homologous recombination between the neomycin cassette and the mutated EBF1-binding site at HS6. Therefore, clone 9G was injected into blastocysts for generation of chimeric mice of the *Pax5*$^{\Delta EHS6+7+neo/+}$ genotype. Blastocysts injection of clone 9G initially did not result in chimeric mice.

After reinjection, three chimeric males were born (of 50-75% fur color chimerism), two of which were smaller in size. The normal size mouse was first crossed to a wild-type C57BL/6 female, and the offspring of one litter was analyzed by Southern blotting (Figure 16A) to investigate germline transmission. Three out of nine mice from this crossing displayed the targeted Southern blot band, as did the control parent 9G chimeric mouse (Figure 16B). However, in all experiments using mouse tail genomic DNA, the targeted band showed up as a double band. This can be due to incomplete digestion or to alternative usage of the three SphI sites, which are present in relative close proximity to each other within the targeting vector (only one is shown in the schematic representation of Southern blot strategy in Figure 16A). To confirm the presence of the correct targeting event in the offspring, males numbers 4 and 8 were re-analyzed using a different Southern blot digestion scheme (NsiI) with the same detection probe (Figure 16C). Re-analysis did not confirm the presence of a double-band, and correctly showed the wild-type and the desired targeted band (Figure 16D). Therefore, these mice (number 4 and number 8) were further used to establish the colony. First, *Pax5* expression was monitored in their peripheral blood (Figure 16E); only a very mild reduction in hCD2 levels was monitored in this experiment, most probably due to the presence of the neomycin resistance gene in the locus. The two mice were further crossed to FlpE mice to remove the neomycin resistance cassette. Following the removal of the neomycin resistance gene, the presence of the desired restriction site mutations at HS6 and HS7 was re-confirmed by performing PCR amplification followed by restriction digestion, as previously described (Figure 16F and 16G). In addition, a PCR strategy that amplifies across the Frt site that remains in the locus after FlpE mediated deletion was designed; the resulting PCR bands of a *Pax5*$^{\Delta EHS6+7\text{-}neo/+}$*FlpE*$^{+}$ mouse (from here on *Pax5*$^{\Delta EHS6+7/+}$) were sequenced (one wild-type band and one band containing the Frt site). The results verified the removal of the neomycin resistance cassette (Figure 16H).

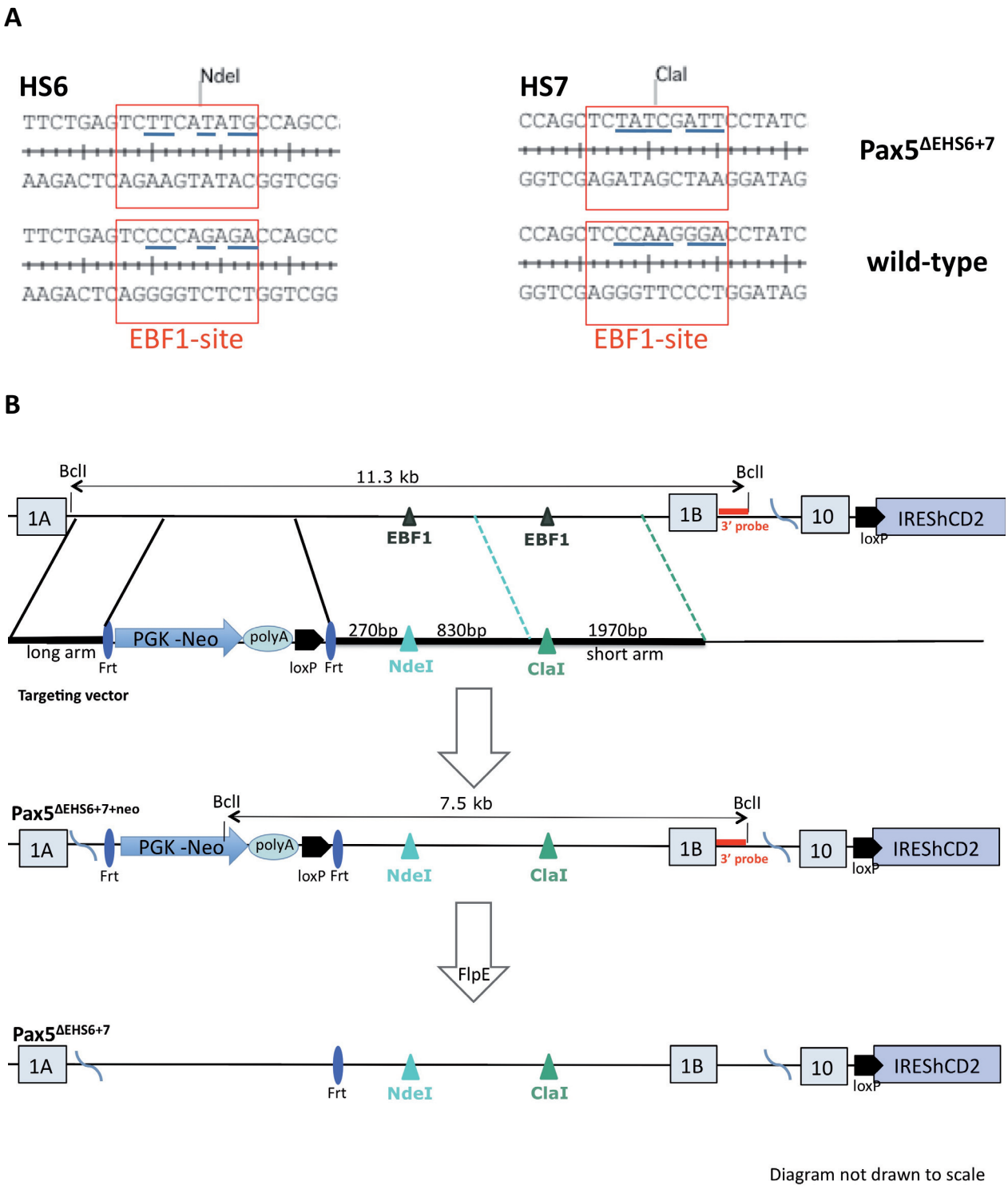


Figure 14. ES cell targeting strategy and screening to obtain the *Pax5*$^{\Delta EHS6+7/+}$ and *Pax5*$^{\Delta EHS6/+}$ mice.
(see explanation on next page)

C

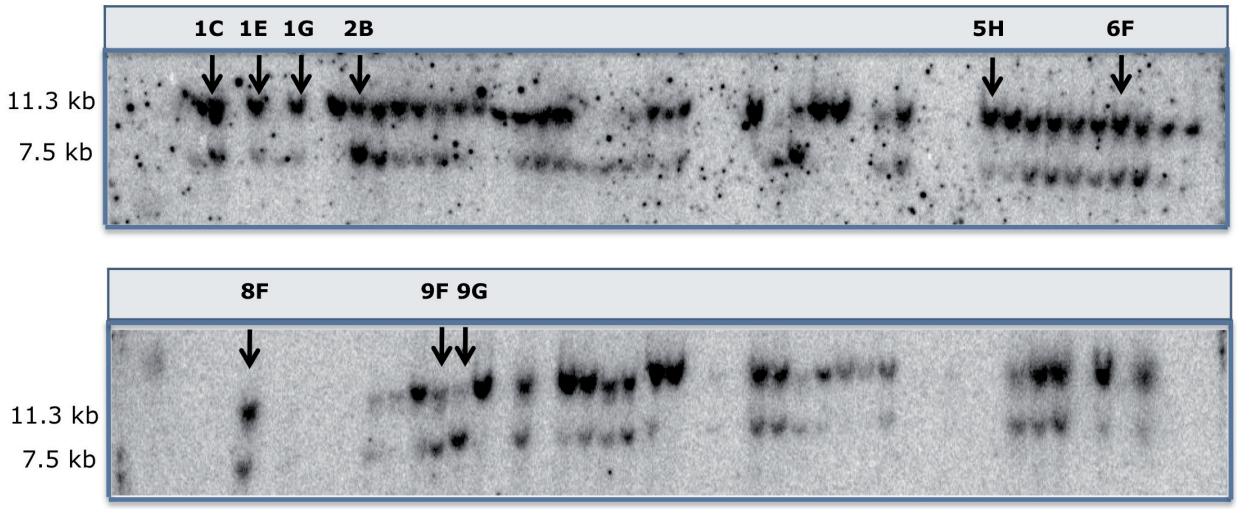


Figure 14. ES cell targeting strategy and screening to obtain the *Pax5*$^{\Delta EHS6+7/+}$ and *Pax5*$^{\Delta EHS6/+}$ mice. (see also previous page)

A. Sequences of the EBF1-binding sites at HS6 and HS7 and the mutations that inactivate them. The sequences of interest are marked by the red rectangle. The bases underlined in blue are mutated. The EBF1 binding consensus is TCCCNNGGGA (Vilagos et al, 2012).

B. Schematic representation of Pax5$^{\Delta EHS6+7+neo}$ targeting vector and targeted *Pax5* alleles (before and after neomycin removal). All inserted modules are represented, together with the span of homology arms. FlpE mediated deletion of the neomycin (Neo) cassette can be achieved by crossing the resulted chimeric mice with mice carrying the FlpE transgene. Homologous recombination following electroporation in ES cells can result in two outcomes, since the two mutations of interest were introduced in the short arm as mismatches: EBF1-binding site in HS6 was replaced by an NdeI restriction site; EBF1-binding site in HS7 was replaced by a ClaI restriction site. If homologous recombination stops before the second mismatch, then only the EBF1 site in HS6 will be mutated (shown by the light green dotted line); this will generate *Pax5*$^{\Delta EHS6/+}$ mice. If homologous recombination spans the whole of the homology short arm regardless of the mismatches, then both EBF1 sites in HS6 and HS7 will be mutated (shown by the dark green dotted line); this will generate *Pax5*$^{\Delta EHS6+7/+}$ mice. It is theoretically unlikely that the homologous recombination stops at the first mismatch (and thus no modifications are introduced), since only 260 bp of homology exist before this mismatch. A minimum of 500 bp of homology in the short arm is necessary for homologous recombination to occur (Muller, 1999). The targeted allele can be distinguished from the wild-type allele by a Southern blotting strategy. Restriction sites for the enzyme used, the length of the respective fragments, as well as the position of the probe used for detection are indicated (same probe as for the *Pax5*$^{\Delta EHS7/+}$ screening).

C. Southern blot screening of picked ES cell clones, using the strategy depicted in A. The nine clones marked were picked for further analysis.

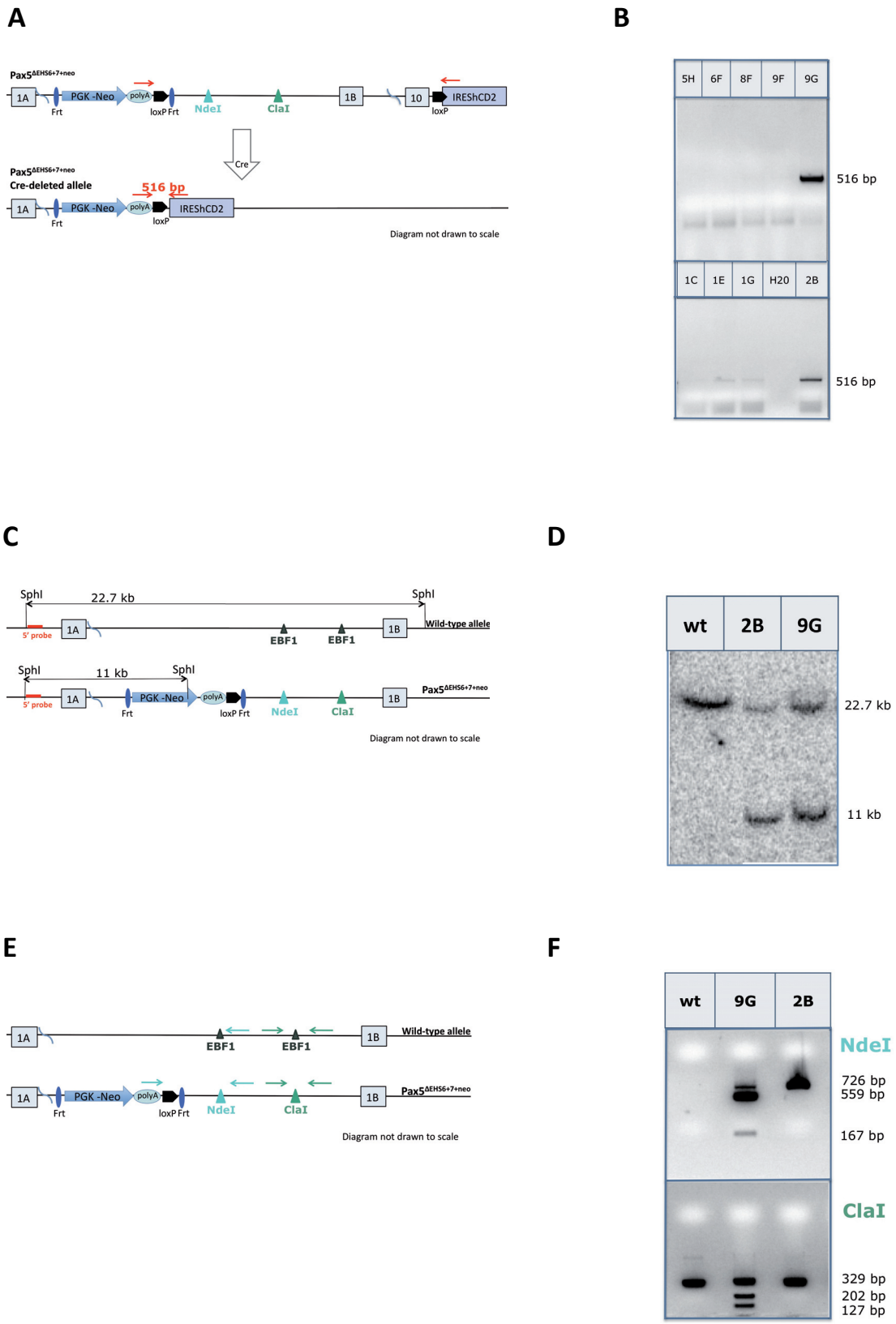


Figure 15. Identification of the correct *Pax5*$^{\Delta EHS6+7+neo/+}$ clone. (see explanation on next page)

Figure 15. Identification of the correct *Pax5*$^{\Delta EHS6+7+neo/+}$ clone. (see previous page)

A. Nine ES cell clones were expanded (marked in Figure 14B) and infected in vitro with an AdenoCre virus. Infection results in the deletion of sequences between loxP sites, if the Pax5$^{\Delta EHS6+7+neo}$ targeting and the IRES-hCD2 reporter are located on the same *Pax5* allele. In this case, the new configuration of the *Pax5* locus allows detection of a PCR product with the primers marked by red arrows (see Table 5).

B. PCR result on DNA isolated after AdenoCre infection of the nine clones marked in Figure 14C. One lane with no DNA (water control, marked H2O) was randomly introduced among the PCR reactions. Two clones (2B and 9G) showed bands of the expected size.

C. Schematic representation of the Southern blotting strategy used to confirm that the two clones identified in B are correctly targeted. SphI restriction sites used for digestion are represented, as well as the binding site for the probe used for detection.

D. Southern blot analysis of the two ES cell clones of interest with the strategy depicted in C. Wild-type genomic DNA was used as negative control.

E. Positions of PCR primers (light/dark green for the respective site) used to amplify across the two EBF1/restriction sites of interest; subsequent digestion of PCR products with the respective restriction enzyme (NdeI, ClaI) identifies whether the desired substitutions of EBF1-binding sites with restriction sites were introduced in the positive ES cell clones analyzed. Primers MB10991, MB10999, MB11993, MB11994 in Table 6.

F. Results of PCR amplification followed by restriction digestion as exemplified in E. PCR products from clone 2B are not digested by either restriction enzyme, thus clone 2B contains a recombination error (EBF1-binding sites at HS6 and HS7 are in the wild-type configuration). Both PCR bands of clone 9G are digested, thus clone 9G is *Pax5*$^{\Delta EHS6+7+neo/+}$. Digestion with the NdeI enzyme was incomplete and thus resulted in an undigested band as well. ClaI digestion was complete, but in this case the PCR primers detect the wild-type band as well, and therefore show the wild-type undigested band.

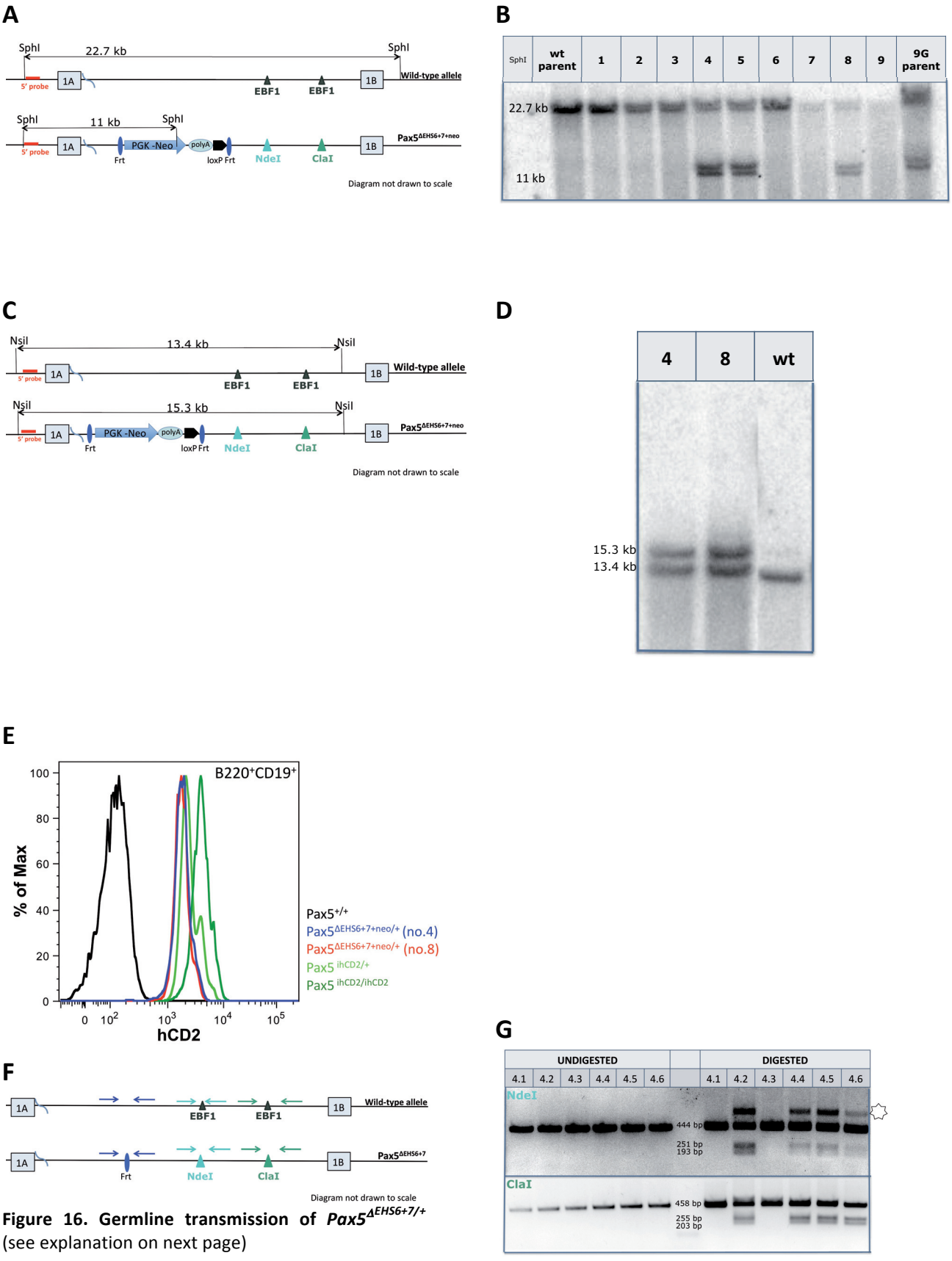


Figure 16. Germline transmission of *Pax5*$^{\Delta EHS6+7/+}$
(see explanation on next page)

H

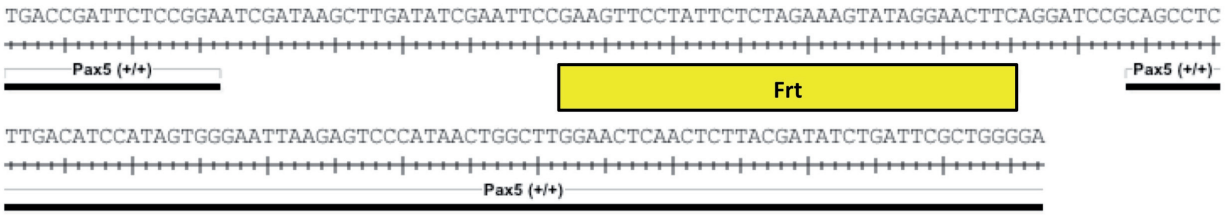


Figure 16. Germline transmission (of *Pax5*$^{\Delta EHS6+7+neo/+}$) and correct genotype of the *Pax5*$^{\Delta EHS6+7/+}$ mouse line from clone 9G. (see also previous page)

A. Pups of a cross between 9G chimeric mouse *(Pax5*$^{\Delta EHS6+7+neo/+}$*)* and a C57BL/6 wild-type mouse were analyzed by Southern blotting. The digestion (enzyme restriction sites) and detection (probe position) scheme is depicted.

B. Result of the Southern blotting experiment for nine pups, using DNA from both the chimeric and the wild-type parents as controls. This experiment shows that the targeting event of interest is transmitted through the germline. Appearance of a double band called for re-testing with the same probe, but a different enzyme (see C and D and text for details).

C. Alternative Southern blot digestion scheme for re-analysis of the mice of interest with the same detection probe.

D. Southern blotting results for pups number 4 and number 8 using the digestion scheme depicted in C.

E. Pups number 4 and number 8 (Pax5$^{\Delta EHS6+7+neo}$ targeted on the Pax5-IRES-hCD2 allele) were bled from the tail vein and the blood was analyzed by FACS for B220^{+}CD19^{+} cells. Expression levels of hCD2 were monitored in this population. As controls, *Pax5*$^{+/+}$, *Pax5*$^{ihCD2/ihCD2}$, and *Pax5*$^{ihCD2/+}$ mice were used.

F. Positions of PCR primers (light/dark green for the respective site) to amplify across the two restriction sites of interest in HS6 and HS7; this is to confirm the presence of both mutations of interest in offspring of the chimeric mice (see G). Primers MB11995-MB11998 in Table 6.

G. Six mice from the cross of mouse number 4 to a FlpE mouse, were analyzed for the correct introduction of the restriction sites of interest at HS6 and HS7 (NdeI and ClaI respectively). PCR products were digested with the respective enzyme (NdeI or ClaI). Mice 4.1 and 4.3 are *Pax5*$^{+/+}$, whereas mice 4.2, 4.4, 4.5, and 4.6 are *Pax5*$^{\Delta EHS6+7/+}$. The star marks an unexplained digestion artefact band.
From here on *Pax5*$^{\Delta EHS6+7/+}$ denominates the neomycin-deleted correctly targeted mice.

H. PCR products of primers that amplify across the Frt site (MB11730 and MB11994, Table 6 and represented as dark blue arrows in F) resulted from FlpE mediated deletion of the neomycin resistance were sequenced for *Pax5*$^{\Delta EHS6+7/+}$ mice. A wt and an Frt-containing PCR bands of different sizes were amplified and subsequently sequenced. The sequence of the region is shown (representing the 100% coverage of the PCR product for the targeted band), as well as the position of the Frt site. The span of the sequencing coverage for which no mismatch from the expected sequence could be identified (100% overlap to expected sequence) is shown for the wild-type band as a black line underlying the respective sequence.

2. Analysis of *Pax5*$^{\Delta EHS7/+}$ and of *Pax5*$^{\Delta EHS7/\Delta EHS7}$ mice

2.1 FACS analysis of *Pax5*$^{\Delta EHS7/+}$ and of *Pax5*$^{\Delta EHS7/\Delta EHS7}$ mice

To investigate whether the absence of the EBF1-binding site in HS7 in the *Pax5* locus influences the expression of the *Pax5* gene, Fluorescence Activated Cell-Sorting (FACS) experiments were conducted. Initially, heterozygous *Pax5*$^{\Delta EHS7/+}$ mice were analyzed alongside with *Pax5*$^{ihCD2/+}$ controls (mice that only contain the IRES-hCD2 reporter in the 3' untranslated region of *Pax5* – Fuxa and Busslinger, 2007) that provide the maximum (wild-type) expression levels of *Pax5,* monitored by cell surface hCD2 expression, under the same experimental conditions. The bone marrow and spleen of these mice were analyzed for all B cell compartments of interest, whereas the thymus was used as a control to show that *Pax5* expression is not aberrantly initiated in other cells of the lymphoid system. As expected and previously reported*, Pax5* expression is activated at the pro-B cell stage, remains constant through the next stages of B cell development, and is downregulated in plasma cells (Figure 17). At all B cell stages investigated, the expression levels of hCD2 and thus *Pax5* are identical between *Pax5*$^{\Delta EHS7/+}$ and *Pax5*$^{ihCD2/+}$ mice. Thus, there is no change in *Pax5* expression upon inactivation of EBF1-binding site at HS7 in a heterozygous context (presence of a second, wild-type allele of *Pax5*).

Next, the same experimental conditions were used to analyze *Pax5*$^{\Delta EHS7/\Delta EHS7}$ mice and compare them to *Pax5*$^{ihCD2/ihCD2}$ mice. Again, the bone marrow and the spleen were analyzed for B cell compartments, whereas the thymus was used as a negative control. At all B cell stages where *Pax5* is expressed, its transcript levels monitored by surface expression of hCD2 are identical between homozygous experimental and control mice (Figure 18). Therefore, no change in *Pax5* expression upon inactivation of EBF1-binding site at HS7 can be detected in either heterozygous or homozygous context.

However, the question remains whether B cell development is normal in terms of absolute cell numbers at all stages of B cell development. To this end, four age-matched mice of each genotype (*Pax5*$^{\Delta EHS7/\Delta EHS7}$, *Pax5*$^{\Delta EHS7/+}$) were analyzed alongside with control wild-type littermates (*Pax5*$^{+/+}$) and absolute cell numbers for different cellular subsets were quantified in the bone marrow, spleen, and thymus (Figure 19). No statistically significant difference was observed between the three genotypes investigated in any of the compartments analyzed.

Therefore, removal of EBF1-binding site at HS7 in the *Pax5* promoter does not affect development of the hematopoietic system (Figures 19), nor does it intervene with normal patterns and levels of *Pax5* expression (Figures 17 and 18).

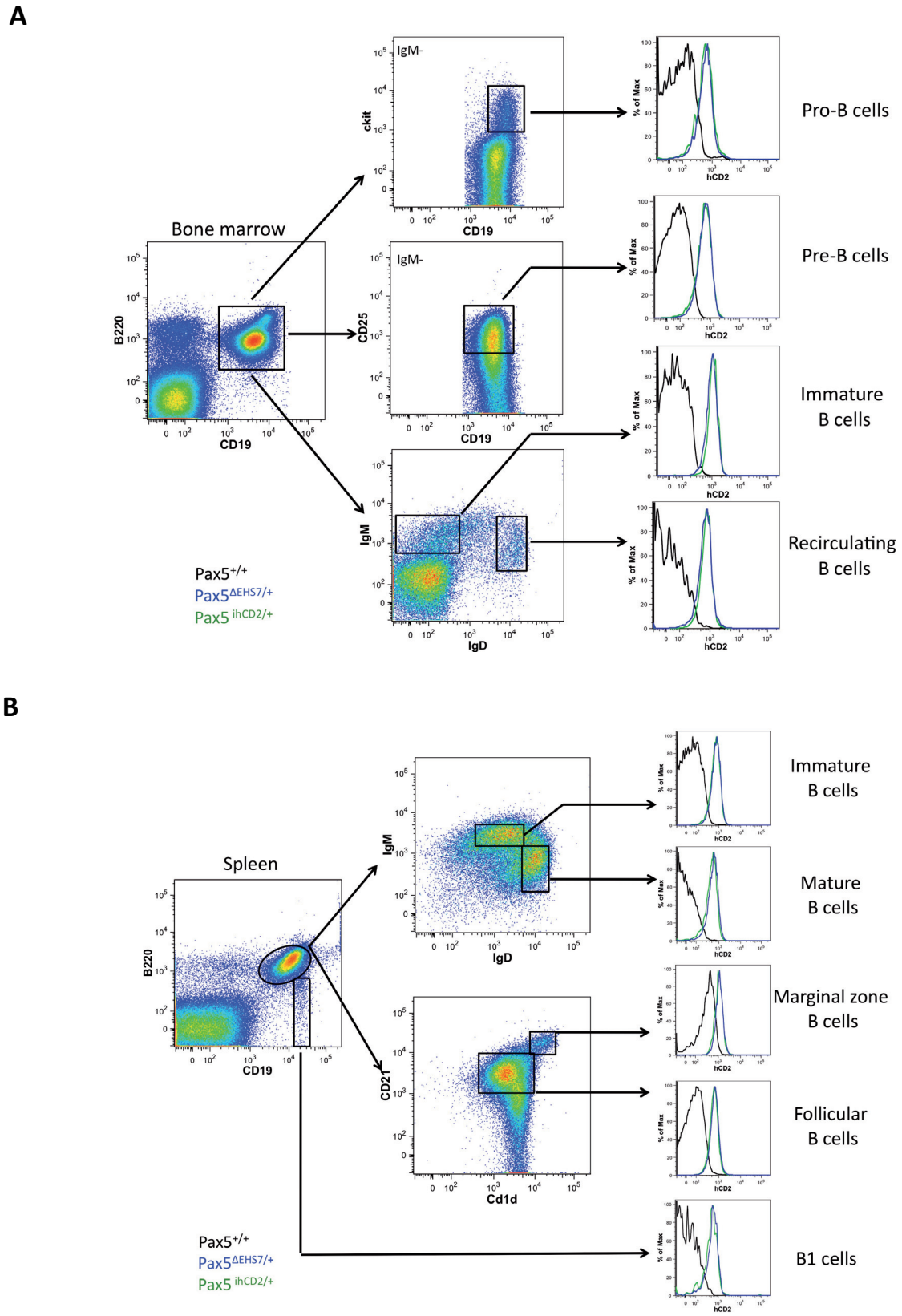
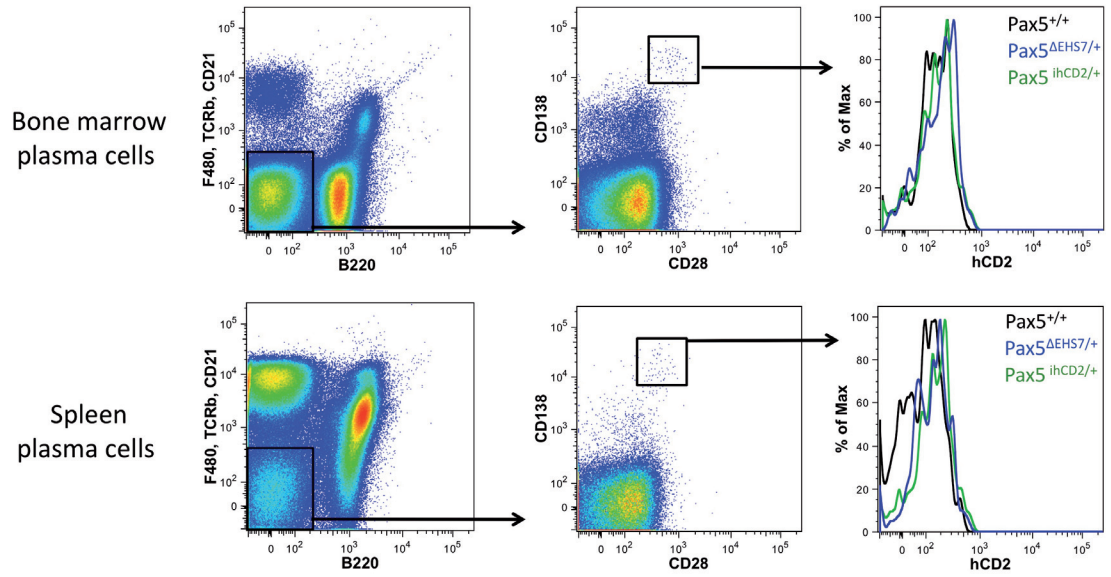


Figure 17. FACS Analysis of *Pax5*$^{\Delta EHS7/+}$ mice. (see explanation on next page)

C



D

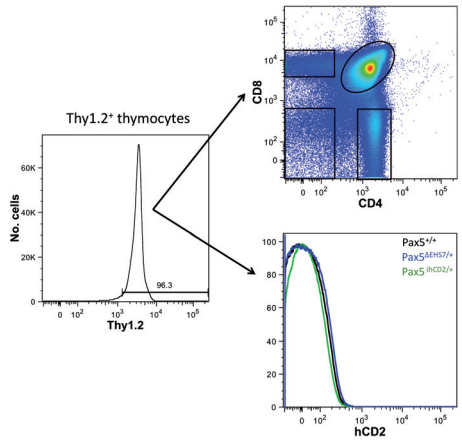


Figure 17. FACS Analysis of *Pax5$^{\Delta EHS7/+}$* mice.

Bone marrow **(A)**, spleen **(B)**, and thymus **(D)** of 8-week-old *Pax5$^{\Delta EHS7/+}$*, *Pax5$^{ihCD2/+}$*, *Pax5$^{+/+}$* were isolated and analyzed by FACS. Plasma cells in the bone marrow and spleen of non-immunized mice were also analyzed **(C)**. One illustrative gating example is shown, and expression levels of hCD2 are depicted for each gated population as overlap between the three genotypes of interest, as shown by arrows. The FACS phenotype of the different cell types was defined as following:

pro-B ($B220^+CD19^+IgM^-ckit^+$), pre-B ($B220^+CD19^+IgM^-CD25^+$),

immature B ($B220^+CD19^+IgM^+IgD^-$),

mature/recirculating B ($B220^+CD19^+IgM^-IgD^+$),

marginal zone B cells ($B220^+CD19^+CD21^{high}CD1d^{high}$),

follicular B cells ($B220^+CD19^+CD21^+CD1d^+$),

plasma cells ($B220^-CD19^-F4/80^-TCRb^-CD21^-CD138^+CD28^+$), T cells ($Thy1.2^+CD4^{+/-}CD8^{+/-}$).

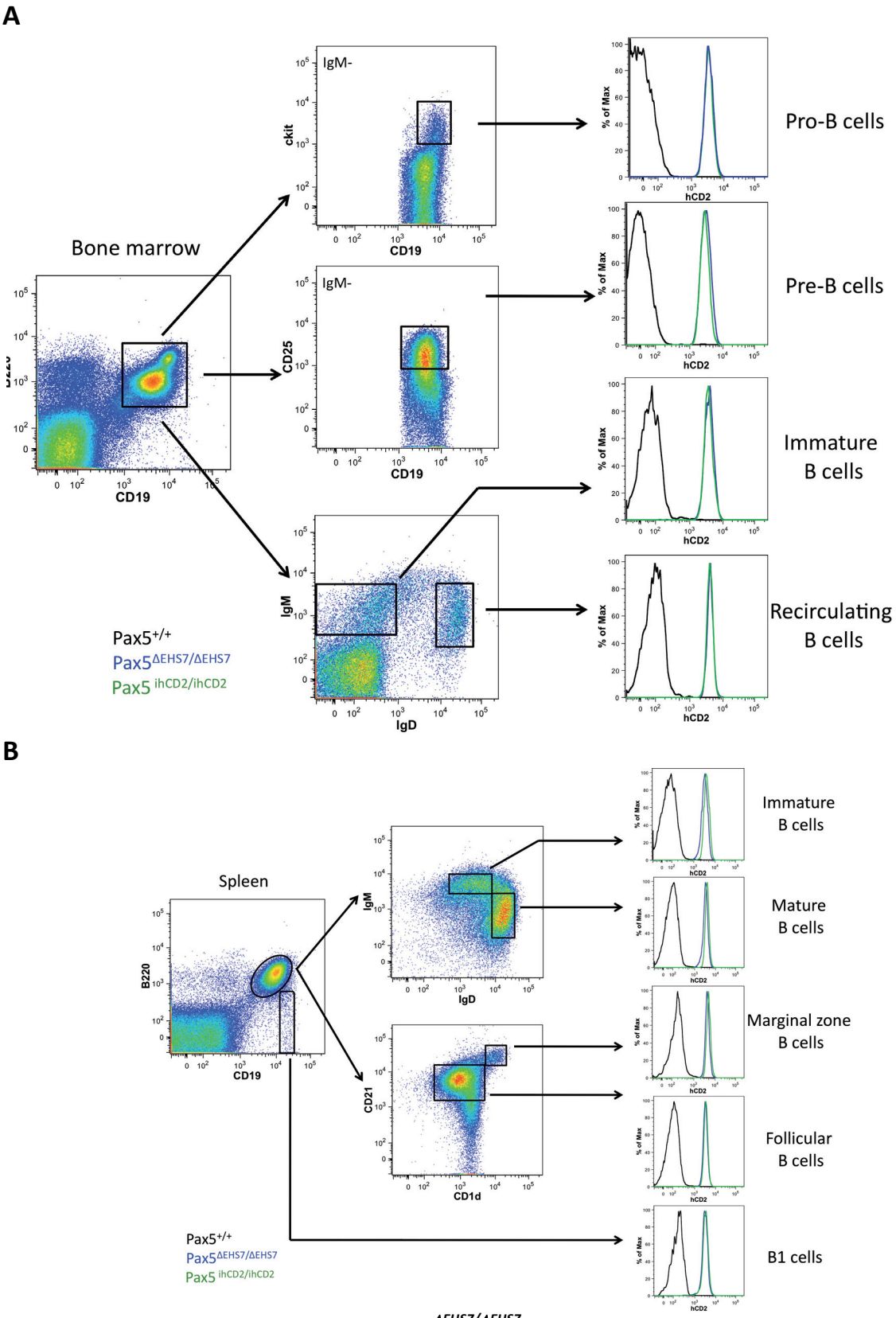


Figure 18. FACS Analysis of *Pax5$^{\Delta EHS7/\Delta EHS7}$* mice. (see explanation on next page)

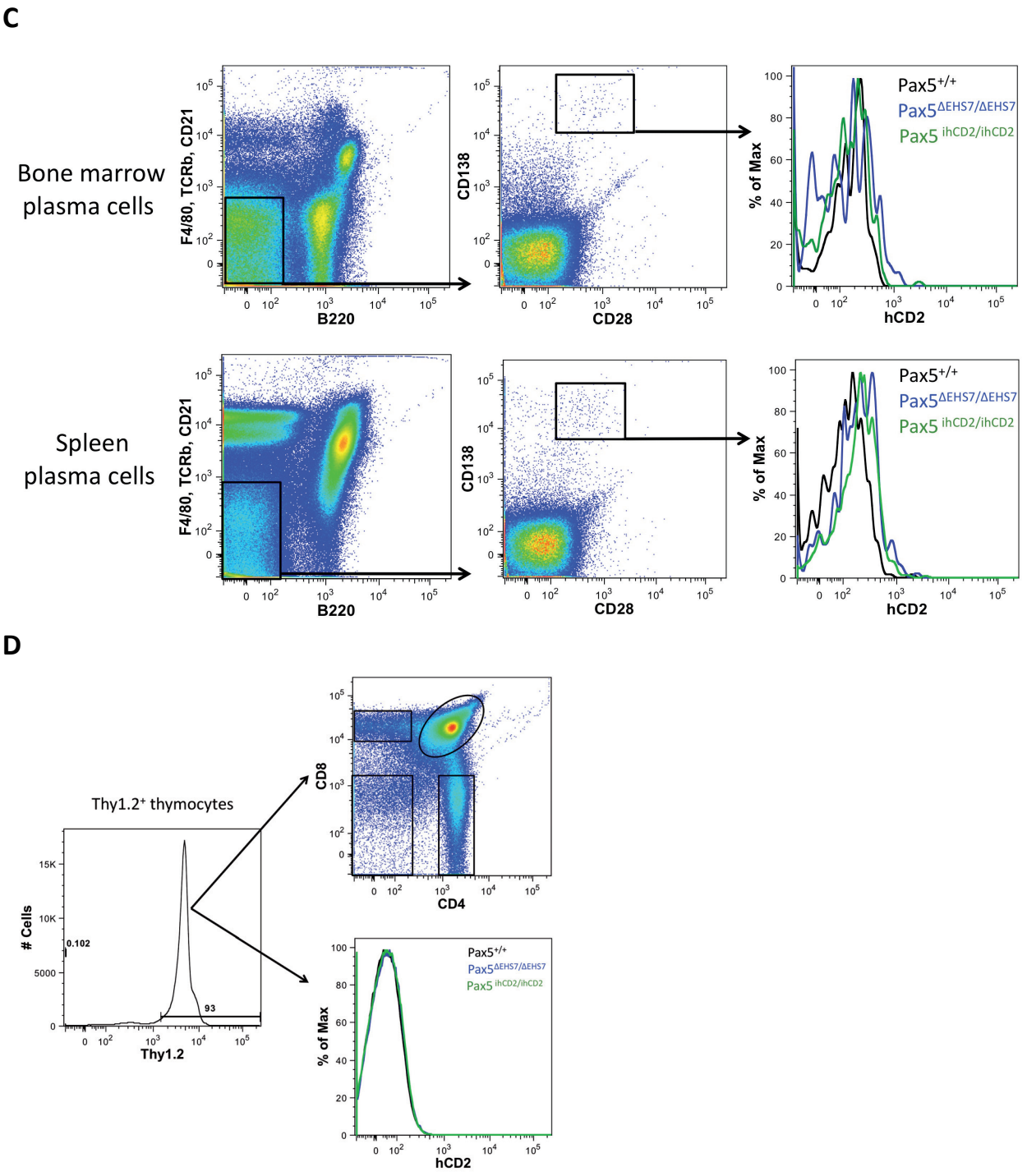


Figure 18. FACS Analysis of *Pax5*$^{\Delta EHS7/\Delta EHS7}$ mice.

Bone marrow **(A)**, spleen **(B)**, and thymus **(D)** of 8-week-old *Pax5*$^{\Delta EHS7/\Delta EHS7}$, *Pax5*$^{ihCD2/ihCD2}$, *Pax5*$^{+/+}$ were isolated and analyzed by FACS. Plasma cells in the bone marrow and spleen of non-immunized mice were also analyzed **(C)**. One illustrative gating example is shown, and expression levels of hCD2 are depicted for each gated population as overlap between the three genotypes of interest, as shown by arrows. The FACS phenotype of the different cell types was defined as in Figure 17.

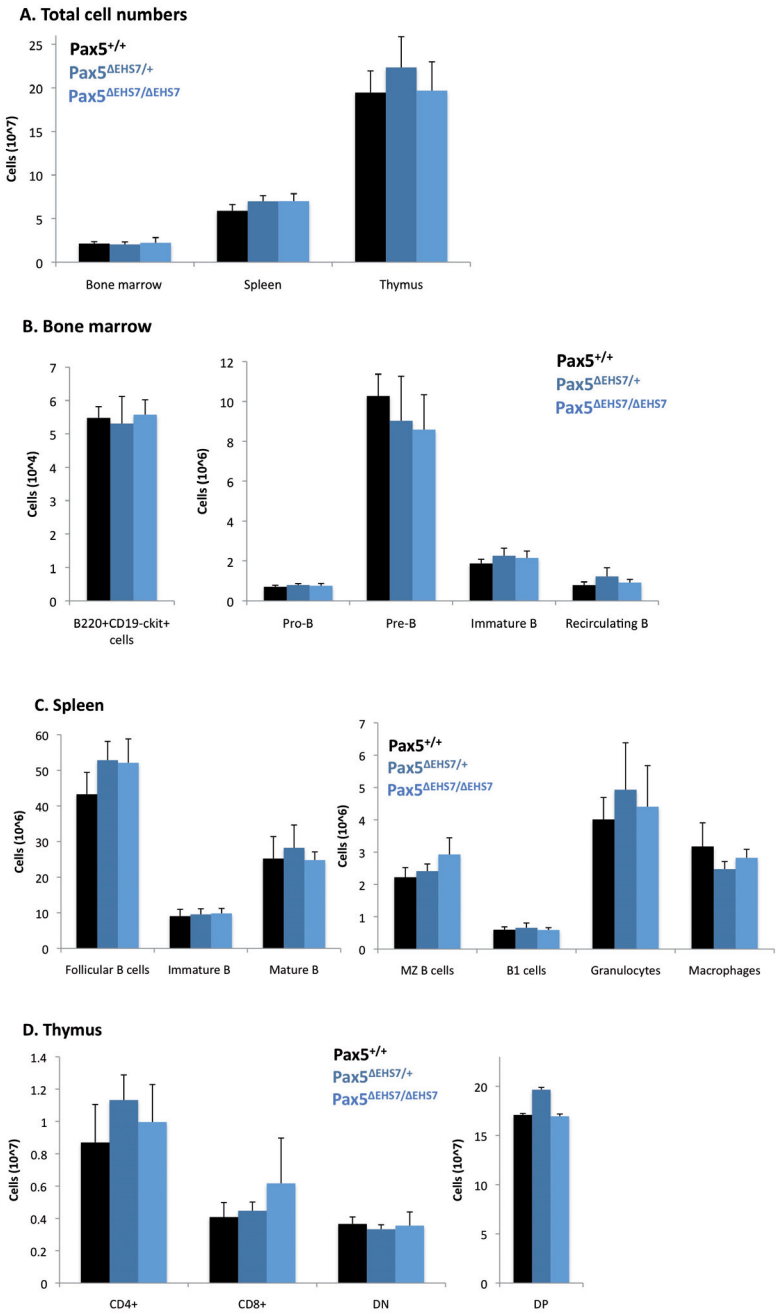


Figure 19. Evaluation of total cell numbers for *Pax5*$^{\Delta EHS7/+}$ and *Pax5*$^{\Delta EHS7/\Delta EHS7}$ mice.

Four mice of each genotype (7-week-old) were analyzed by FACS. Populations were defined as previously described. **(A)** Total cell numbers in the organ of interest. In the bone marrow and spleen B220^{+}CD19^{+} are analyzed as total cell numbers. In the thymus, Thy1.2^{+} T cells are quantified. Then individual populations were quantified in **B.** Bone marrow, **C.** Spleen, and **D.** Thymus. Standard error of the mean was calculated. In addition to already mentioned populations, macrophages were defined as B220^{-}CD19^{-}MCSF-R^{+}Mac1^{+} and granulocytes as B220^{-}CD19^{-}MCSF-R^{+}Gr1^{+}. In the bone marrow, B220^{+}CD19^{-}ckit^{+} progenitors were also quantified. In the thymus, single positive (CD4^{+} and CD8^{+}), double negative (DN) and double positive (DP) cells were identified based on CD4 and CD8 expression.

2.2 Chromatin immunoprecipitation (ChIP) analysis of *Pax5*$^{\Delta EHS7/\Delta EHS7}$ mice

Although the EBF1-binding site in HS7 of the *Pax5* locus has been replaced in the *Pax5*$^{\Delta EHS7/\Delta EHS7}$ mice, it is still necessary to prove that EBF1 cannot bind HS7 in vivo. ChIP experiments with an anti-EBF1 antibody followed by quantitative-PCR (qPCR)-specific amplification of HS7 were performed with isolated B220^{+} bone-marrow cells from *Pax5*$^{\Delta EHS7/\Delta EHS7}$ mice. The cells were in vitro expanded for five days. As positive control, bone marrow cells isolated from *Rag2*$^{-/-}$ mice were used. *Pax5*$^{\Delta EHS7/\Delta EHS7}$ cells were analyzed by FACS after 5 days in culture, in order to ensure that the majority are B220^{+}CD19^{+} (Figure 20A).

First, it was attempted to test two new sets of primers that detect HS7 (one set (EBF1 site in HS7) amplifies a region in HS7 immediately upstream of the EBF1-binding site; the other set (HS7) amplifies across the EBF1-binding site). This is necessary since the binding sequence for the previously reported primers that detect EBF1 binding at HS7 (Decker et al, 2009) was removed by the targeting strategy used, and thus was no longer present in the *Pax5*$^{\Delta EHS7/\Delta EHS7}$ context. As seen in Figure 20B, the two new primer sets are comparable to the previously reported primers used by Decker and colleagues in detection of HS7 sequences on DNA material precipitated with an anti-EBF1 antibody from *Rag2*$^{-/-}$ cells. As positive control, primers that amplify the region bound by EBF1 in the *mb-1 (CD79a)* promoter were used; negative control primers amplify a random intergenic region in chromosome 1. Furthermore, when the newly designed primer sets are used under the same experimental conditions, they do not detect enrichment of EBF1-bound sequences at HS7 in *Pax5*$^{\Delta EHS7/\Delta EHS7}$ cells, whereas EBF1 binding at HS6 can be observed in these cells (Figure 20C). Thus, removal of the EBF1-binding sequence at HS7 in the *Pax5* promoter abrogates EBF1 binding at HS7 in *Pax5*$^{\Delta EHS7/\Delta EHS7}$ mice.

In parallel, anti-E2A and anti-Pax5 antibody ChIP experiments were performed in *Pax5*$^{\Delta EHS7/\Delta EHS7}$ mice. E2A and Pax5 also bind at HS6 in the *Pax5* promoter to sequences immediately adjacent to the EBF1-binding site (Figure 10 and data not shown). As shown in Figure 21, the binding sites at HS6 for all three transcription factors of interest are unaffected by the removal of the EBF1-binding site in HS7, and thus the results obtained in the analysis of *Pax5*$^{\Delta EHS7/\Delta EHS7}$ mice represent solely the effect of abrogated EBF1-binding at HS7 in the *Pax5* promoter.

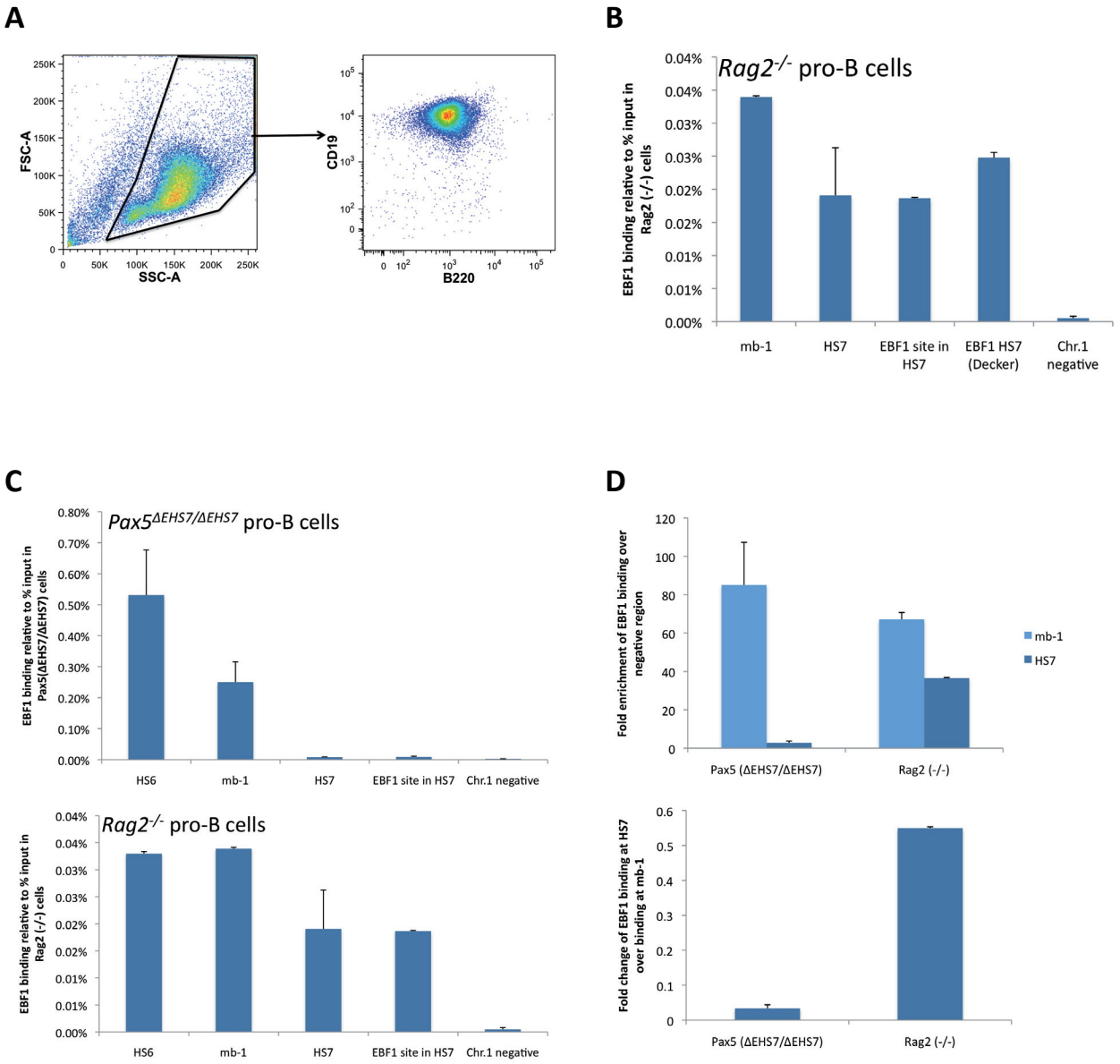


Figure 20. EBF1 binding is abrogated at HS7 in *Pax5*$^{\Delta EHS7/\Delta EHS7}$ mice.
(see explanation on next page)

Figure 20. EBF1 binding is abrogated at HS7 in *Pax5*$^{\Delta EHS7/\Delta EHS7}$ mice. (see previous page)
A. B220 MACS sorted bone marrow cells of *Pax5*$^{\Delta EHS7/\Delta EHS7}$ mice were expanded in culture for 5 days as previously reported (Nutt et al, 1997) and analyzed by FACS at the harvesting point for presence of B220 and CD19 markers.
B. *Rag2*$^{-/-}$ cells were grown in parallel with *Pax5*$^{\Delta EHS7/\Delta EHS7}$ cells, harvested after 5 days in culture, and the chromatin from these cells was precipitated with and anti-EBF1 antibody. Primers that amplify a positive control EBF1-bound region (*mb-1*; *Cd79a*) and a negative control region (chromosome 1) were used in qPCR amplifcation alongside with two new sets of primers (HS7, and EBF1 site at HS7) and previously reported primers for HS7 (Decker et al, 2009). The two new primer sets are as good as the previous primers in detecting enrichment of EBF1 binding at HS7.
C. Comparison of anti-EBF1 antibody ChIPs from *Pax5*$^{\Delta EHS7/\Delta EHS7}$ and *Rag2*$^{-/-}$ cells for enrichment of EBF1-binding sites; qPCRs were performed with previously mentioned primers; alongside, a primer set that amplifies the EBF1-binding site in HS6 of the *Pax5* promoter was used. No sequences enriched for EBF1 binding are amplified at HS7 in *Pax5*$^{\Delta EHS7/\Delta EHS7}$ cells.
D. Fold enrichment for binding of EBF1 relative to negative control region, as well as fold change of EBF1 binding relative to the binding at the positive region, were calculated for HS7 in both *Pax5*$^{\Delta EHS7/\Delta EHS7}$ and *Rag2*$^{-/-}$ cells in order to better illustrate abrogation of EBF1 binding only at HS7 in *Pax5*$^{\Delta EHS7/\Delta EHS7}$ cells.

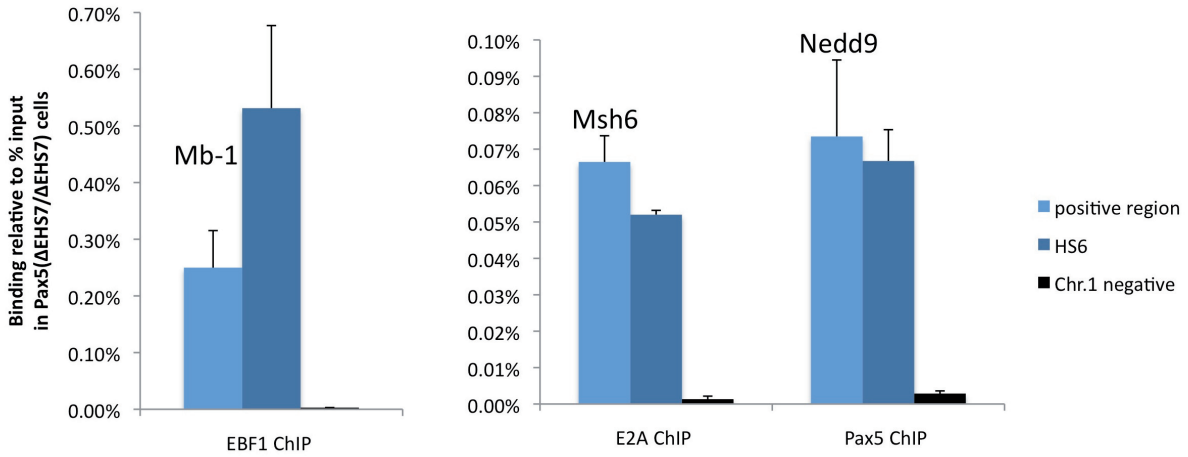


Figure 21. E2A, EBF1, and Pax5 binding is normal at HS6 in *Pax5*$^{\Delta EHS7/\Delta EHS7}$ mice.
B220 MACS sorted bone marrow cells of *Pax5*$^{\Delta EHS7/\Delta EHS7}$ mice were expanded in culture and harvested after 5 days; the chromatin from these cells was precipitated with anti-E2A, anti-EBF1, and anti-Pax5 antibodies individually. Primers that amplify a positive control region bound by the respective transcription factor (indicated) were used in qPCR amplifcation, alongside with random primers binding to a negative control region on chromosome 1 and primers that specifically amplify HS6 in the *Pax5* promoter region.

3. Analysis of *Pax5*$^{\Delta EHS6+7/+}$ and of *Pax5*$^{\Delta EHS6+7/\Delta EHS6+7}$ mice

3.1 FACS Analysis of *Pax5*$^{\Delta EHS6+7/+}$ mice

Heterozygous *Pax5*$^{\Delta EHS6+7/+}$ mice have both EBF1-binding sequences in the *Pax5* promoter replaced by restriction sites on one allele, and carry a second wild-type *Pax5* allele. To investigate the effect of these mutations on the expression of the *Pax5* gene, FACS experiments were conducted. *Pax5*$^{\Delta EHS6+7/+}$ mice were analyzed alongside with *Pax5*$^{ihCD2/+}$ controls, in the same experimental setting as previously described for *Pax5*$^{\Delta EHS7/+}$ mice. The bone marrow and spleen were analyzed for all B cell compartments of interest. At the early B cell stages (pro-B to immature B cells) in the bone marrow, the expression levels of *Pax5* (monitored by cell surface hCD2 expression) are two-fold reduced between *Pax5*$^{\Delta EHS6+7/+}$ and *Pax5*$^{ihCD2/+}$ mice. There is no difference between *Pax5*$^{\Delta EHS6+7/+}$ and *Pax5*$^{ihCD2/+}$ mice in *Pax5* expression levels at the later stages of B cell development and in the mature B cell subsets of the spleen (Figures 22A and 22B).

As expected, *Pax5* expression is downregulated in plasma cells in *Pax5*$^{\Delta EHS6+7/+}$ mice (Figure 22C). The thymus was used as a negative control to show that *Pax5* expression is not aberrantly initiated in other cells of the lymphoid system (Figure 22D).

Total cell numbers were not significantly different between *Pax5*$^{+/+}$ and *Pax5*$^{\Delta EHS6+7/+}$ littermates, as shown in Figure 23. All experiments were performed with mice that were only two or three times backcrossed into the C57BL/6 background, as colony establishment is still in progress. In all the analyzed mice, the neomycin cassette had been removed.

Thus, disruption of the two EBF1-binding sites in the *Pax5* promoter results in a two-fold decrease in expression of *Pax5* from pro-B to immature B cell stage in the bone marrow, in a heterozygous context in the presence of a second, wild-type *Pax5* allele.

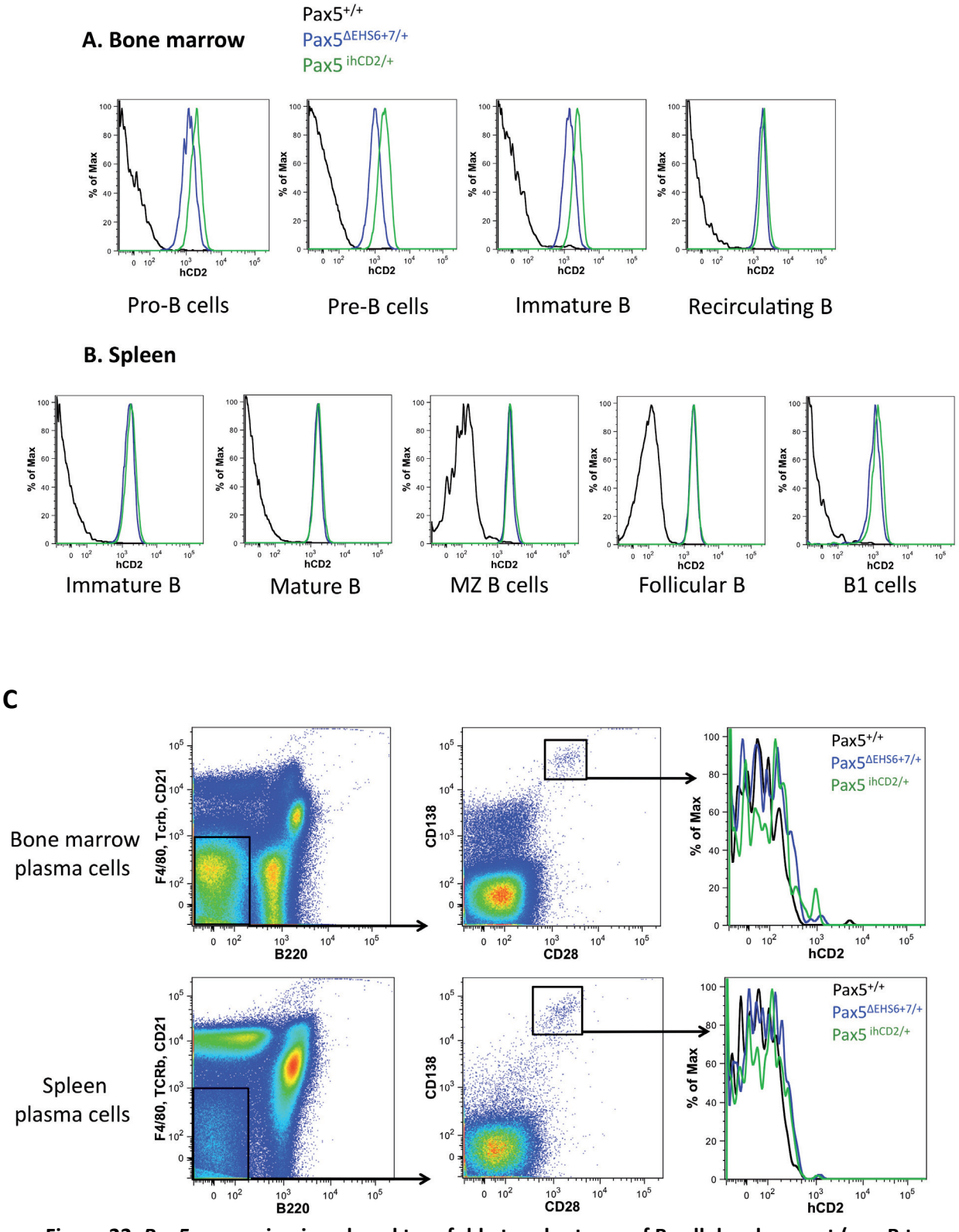


Figure 22. *Pax5* expression is reduced two-fold at early stages of B cell development (pro-B to immature B) in *Pax5*$^{\Delta EHSS6+7/+}$ mice. (see explanation on next page)

D

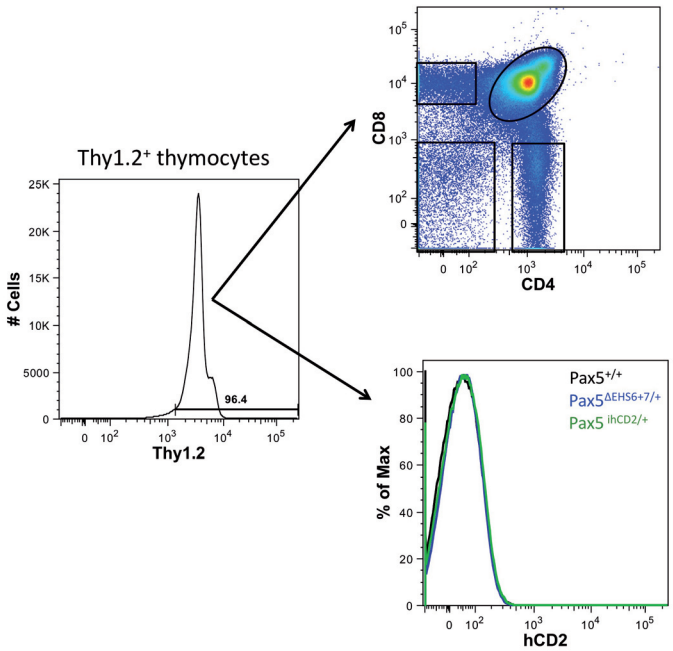


Figure 22. *Pax5* expression is reduced two-fold at early stages of B cell development (pro-B to immature B) in *Pax5*$^{\Delta EHSS6+7/+}$ mice. Bone marrow **(A)**, spleen **(B)**, and thymus **(D)** of 6-week-old *Pax5*$^{\Delta EHS6+7/+}$, *Pax5*$^{ihCD2/+}$, *Pax5*$^{+/+}$ were isolated and analyzed by FACS. Plasma cells in the bone marrow and spleen of non-immunized mice were also analyzed **(C)**. Expression levels of hCD2 are depicted for each gated population as overlap between the three genotypes of interest. The FACS phenotype of the different cell types was defined as in Figure 17.

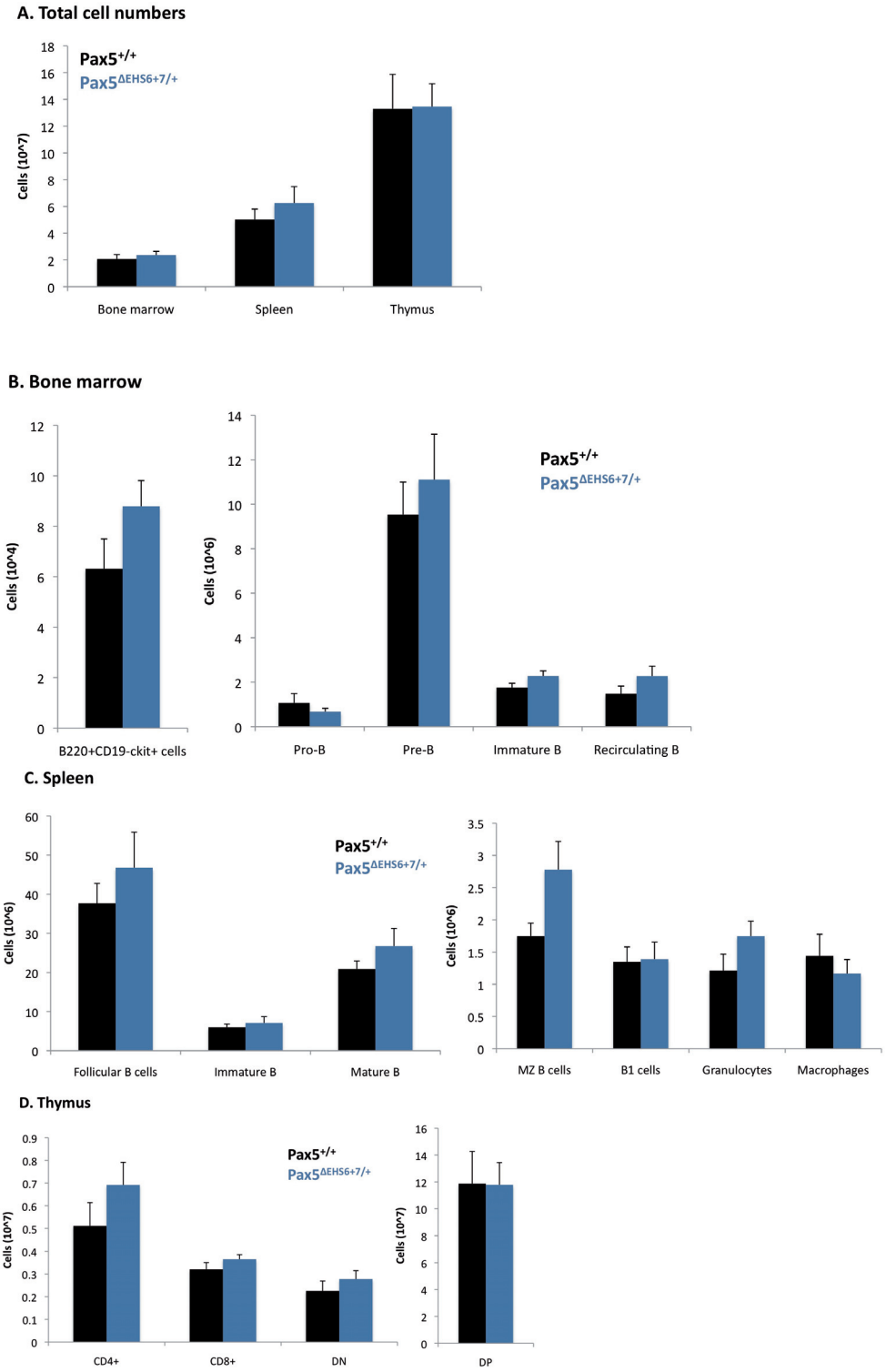


Figure 23. Total cell numbers evaluation for *Pax5$^{\Delta EHS6+7/+}$* mice.

Three mice of each genotype (6-week-old) were analyzed by FACS. Populations were defined as previously described. **(A)** Total cell numbers in the organ of interest. In the bone marrow and spleen B220^{+}CD19^{+} are compared as total cell numbers. In the thymus, Thy1.2^{+} T cells are quantified. Then individual populations were quantified in **B.** Bone marrow, **C.** Spleen, and **D.** Thymus. Standard error of the mean was calculated.

3.2 FACS Analysis of *Pax5*$^{\Delta EHS6+7/\Delta EHS6+7}$ mice

Pax5$^{\Delta EHS6+7/\Delta EHS6+7}$ mice were available only in limited numbers, since the colony is still being established. At this time point, all mice available had only been crossed three times into the C57BL/6 background. However, one 4-week-old mouse was analyzed alongside with *Pax5*$^{ihCD2/ihCD2}$ and *Pax5*$^{+/+}$ controls, in the same experimental setting as for analysis of *Pax5*$^{\Delta EHS6+7/+}$ mice. As seen in Figure 24, the expression levels of *Pax5* (monitored by cell surface hCD2 expression) are two-fold reduced, from the pro-B to the immature B cell stages in the bone marrow, between *Pax5*$^{\Delta EHS6+7/\Delta EHS6+7}$ and *Pax5*$^{ihCD2/ihCD2}$ mice. No difference in *Pax5* expression levels can be detected between *Pax5*$^{\Delta EHS6+7/\Delta EHS6+7}$ and *Pax5*$^{ihCD2/ihCD2}$ mice at the later stages of B cell development and in the mature B cell subsets of the spleen. Therefore, the phenotype of *Pax5*$^{\Delta EHS6+7/\Delta EHS6+7}$ mice mirrors the findings obtained for *Pax5*$^{\Delta EHS6+7/+}$ mice. As negative control, the thymus was analyzed in this experiment (not shown). Macrophages and granulocytes were also stained for hCD2 expression. No aberrant expression of *Pax5* is initiated in any of these lineages (not shown).

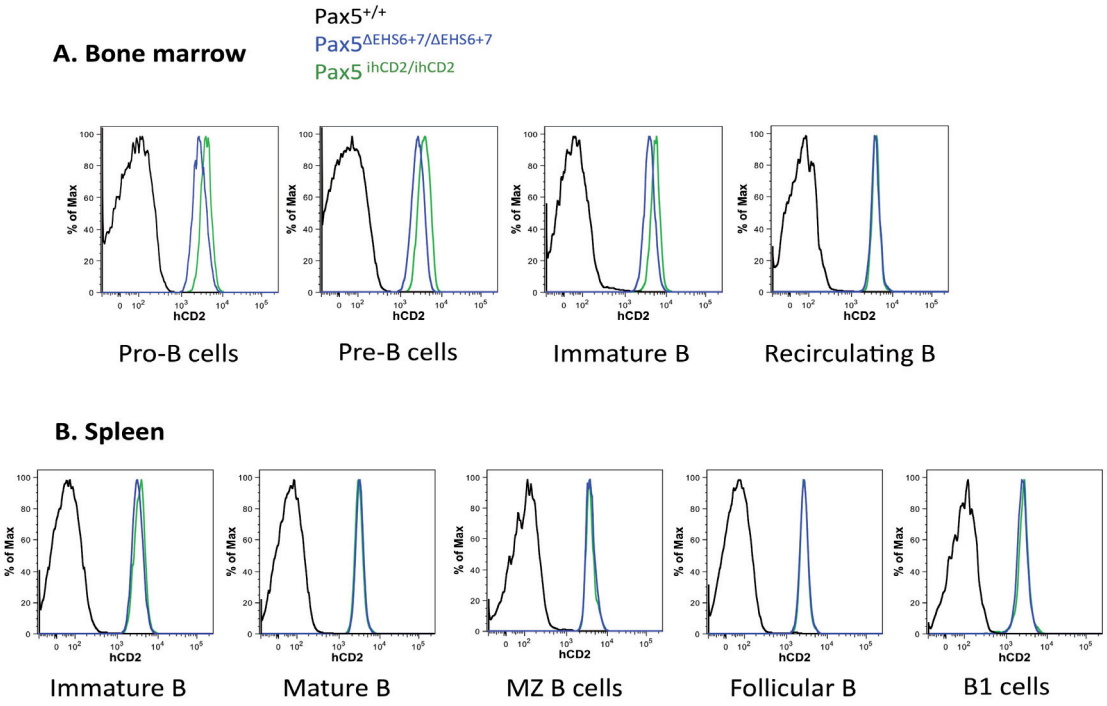


Figure 24. *Pax5* expression is reduced two-fold at early stages of B cell development (pro-B to immature B) in *Pax5*$^{\Delta EHSS6+7/\Delta EHSS6+7}$ mice. Bone marrow **(A)** and spleen **(B)** of 4-week-old *Pax5*$^{\Delta EHS6+7/\Delta EHS6+7}$, *Pax5*$^{ihCD2/ihCD2}$, *Pax5*$^{+/+}$ were isolated and analyzed by FACS. Expression levels of hCD2 are depicted for each gated population as overlap between the three genotypes of interest. The phenotype of the different cell types was defined as in Figure 17.

DISCUSSION

In this study I have generated two genetically modified mice, in which the EBF1-binding sites in the *Pax5* promoter are inactivated. One mouse line $Pax5^{\Delta EHS7/+}$ lacks the EBF1-binding site in DNaseI hypersensitive site HS7 in the *Pax5* promoter, whereas the second mouse line $Pax5^{\Delta EHS6+7/+}$ lacks both EBF1-binding sites in HS6 and HS7. While disruption of EBF1 binding at HS7 does not alter the levels of *Pax5* expression, disruption of both EBF1-binding sites in the *Pax5* promoter results in a two-fold reduction in expression from pro-B to immature B cell stages in the bone marrow. These results strongly indicate that the direct action of EBF1 in the regulation of *Pax5* expression, mediated by binding at HS6 and HS7, is responsible for the fine-tuning of the correct *Pax5* expression levels in early B cell development. The EBF1-bound HS6 and HS7 do not seem to be the sites from which *Pax5* expression is initiated.

Interestingly, a conundrum arises since it is clear from previous work (Decker et al, 2009) that in the absence of EBF1, *Pax5* is not transcribed and the promoter is in an inactive state, characterized by the presence of repressive chromatin marks and the absence of DNaseI hypersensitive sites. Therefore, additional EBF1-dependent mechanisms have to be in place to regulate *Pax5* expression that do not necessarily require binding of EBF1 at HS6 and HS7 in the *Pax5* promoter. Speculation about these mechanisms prompts interesting future experiments, as discussed below.

1. Two-fold reduction of *Pax5* levels in $Pax5^{\Delta EHS6+7/\Delta EHS6+7}$ mice from pro-B to immature B cell stages in the bone marrow

It was not surprising that deletion of the two EBF1-binding sites in the *Pax5* promoter has no effect on *Pax5* expression at later stages of B cell development. Vilagos et al (2012) have clearly shown that the Pax5 protein is expressed at wild-type levels in EBF1-deficient follicular B cells of the spleen. The data in this study further support the fact that EBF1 is not essential for the regulation of *Pax5* in late B cell development.

In contrast, a plethora of experiments clearly showed that the initiation of *Pax5* expression is EBF1-dependent (Decker et al. 2009; Treiber et al, 2010; Zandi et al, 2008). To this end, it was unexpected to detect only a two-fold reduction in *Pax5* levels upon deletion of the two EBF1-binding sites in the gene promoter. This two-fold reduction in *Pax5* levels also makes the search for EBF1 partners in the regulation of *Pax5* expression much more difficult. B cell development is normal in

Pax5$^{+/-}$ mice suggesting that even the expression of *Pax5* from only one allele is sufficient (Nutt et al, 1999b), thus mild reductions in the *Pax5* gene dosage can easily go unnoticed in analysis of phenotypes resulted from inactivation of individual transcription factors regulating *Pax5* expression. Deletion of EBF1-binding sites in HS6 and HS7 mildly reduces *Pax5* expression without significant perturbation of the B cell development thereby providing an excellent system to study contribution of additional transcription factors that are proposed to regulate *Pax5* expression.

It is very possible that the two *Pax5* promoters are differentially regulated and that EBF1 binding at HS6 and HS7 influences only the *Pax5*-1B promoter. Both promoters are used in B cells (Busslinger et al, 1996) and differential regulation of promoter elements has been reported for other transcription factor genes in this hierarchy, including *Ebf1* (Nutt and Kee, 2007). Therefore, the two-fold reduction in *Pax5* expression in *Pax5*$^{\Delta EHS6+7/\Delta EHS6+7}$ mice could be explained by a failure to initiate transcription from the *Pax5*-1B promoter in absence of EBF1 binding at HS6 and HS7. The expression levels detected for *Pax5* could be the result of transcription from the *Pax5*-1A promoter, which could be initiated by different EBF1-dependent mechanisms (for example, indirect effects discussed in Section 3). Therefore, the absence of EBF1 binding at the *Pax5* promoter could influence the local chromatin architecture. A mechanistic analysis of changes in promoter usage, alongside with investigations of histone tail modifications, nucleosome positioning, and position of DNaseI hypersensitive regions at HS6 and HS7 upon removal of EBF1-binding sites, could provide noteworthy information and highlight future research directions.

There are two further possibilities to rationalize the findings of my study and integrate them into the larger picture of the transcription factor network that drives early B cell development (discussed in the following paragraphs).

Firstly, it is possible that additional EBF1-binding sites are present in the *Pax5* locus, and/or that EBF1 binding acts cooperatively with other transcription factors to induce changes in chromatin architecture at the *Pax5* promoter. Genome-wide ChIP-sequencing experiments performed in the laboratory provide a handle to answer this question, and will be addressed below (Section 2).

However, further experimental endeavors can also shed more light into the matter. Initially, a ChIP experiment with anti-EBF1 antibody should be performed as soon as homozygous *Pax5*$^{\Delta EHS6+7/\Delta EHS6+7}$ mice become available, to prove that EBF1 does not interact with the mutated sites at HS6 and HS7 (NdeI and ClaI, respectively). A very straightforward explanation for the observed phenotype would be available if EBF1 binding was only rendered inefficient, but not abrogated by the mutations introduced in the *Pax5* promoter in *Pax5*$^{\Delta EHS6+7/\Delta EHS6+7}$ mice; however, I regard this hypothesis as unlikely. Another interesting

question in this direction is whether EBF1 is part of a protein complex assembled on the *Pax5* promoter at HS6 and HS7. EBF1 has been shown to assemble complexes on the *Cd79a* promoter with RUNX1, CBFβ, and E2A, to promote DNA demethylation (Maier et al, 2004; Hagman et al, 2012), and to recruit the SWI/SNF chromatin remodeler complex (Gao et al, 2009). Cooperativity between EBF1 and other factors is further supported by reported effects on B cell development of *Ebf1/E2a* and *Ebf1/Runx1* haploinsufficiencies (O'Riordan and Grosschedl, 1999; Lukin et al, 2010). This study's analysis of EBF1 function in tuning the correct expression levels of the *Pax5* gene could be a good platform to investigate such potential cooperative effects. If this was the case, it is possible that some of the EBF1 cooperation partners are still able to bind the *Pax5* promoter or change their binding pattern in the absence of EBF1 binding at HS6 and HS7. This could give rise to the expression of Pax5-IRES-hCD2 seen for $Pax5^{\Delta EHS6+7/\Delta EHS6+7}$ mice. Several experiments could be conducted to identify interesting changes in binding patterns, like in vivo DMS footprinting in the wild-type and $Pax5^{\Delta EHS6+7/\Delta EHS6+7}$ contexts, or electrophoretic mobility shift assays (EMSA) with wild-type and EBF1-sites mutated oligo probes (Dey et al, 2012).

Furthermore, it might prove useful to repeat the FACS analysis performed in Figures 22 and 24 with a statistically relevant number of experimental and control mice under identical experimental conditions (the results presented for the heterozygous mice were representative of three different experiments, conducted with different antibody staining patterns to ensure validity of the observed phenotype; one mouse was analyzed for the homozygous situation). This would allow quantifying and comparing the mean fluorescence intensity of hCD2 cell surface expression between $Pax5^{\Delta EHS6+7/\Delta EHS6+7}$ and $Pax5^{ihCD2/ihCD2}$ mice at each stage of B cell development. The exact quantification of the differences in expression resulted from removal of EBF1-binding sites could also indicate the stage in B cell development at which this difference is most pronounced. This would narrow the search for EBF1 cooperation partners to factors specific/important for that particular cell stage.

Secondly, EBF1 could have both a direct influence on the *Pax5* promoter mediated through binding at HS6 and HS7, and an indirect effect by activating other transcription factors which in turn influence *Pax5* expression. Candidate transcription factors in this respect are discussed in Section 3.

2. E2A, EBF1, and Pax5 binding patterns at the *Pax5* locus; results of genome-wide ChIP-seq experiments. Cooperative interactions

The presence of the two EBF1-binding sites at HS6 and HS7 in the *Pax5* promoter was confirmed by several approaches, including ChIP-Seq, EMSA (HS7), and in vivo DMS footprinting (Decker et al, 2009; Hiromi Tagoh, personal communication; Busslinger laboratory, unpublished data). These two binding sites are the only prominent ones at the *Pax5* locus, from the $Pax5^{-/-}$ pre-pro-B cell stage to the mature B cell stage (Busslinger laboratory, unpublished observations). Interestingly, the EBF1-binding site in HS6 is in close proximity with an E2A and a Pax5 peak in pro-B cells (Figure 10, Figure 20C, Figure 21). E2A is placed genetically upstream of EBF1 (Lin and Groschedl, 1995) and induces *Ebf1* (Kee and Murre, 1998; Ikawa et al, 2004). Compound haploinsufficiency of EBF1 and E2A has been shown to dramatically reduce *Pax5* levels, more than *Ebf1* haploinsufficiency alone does (O'Riordan and Grosschedl, 1999). It is therefore tempting to speculate about the role of E2A in the regulation of *Pax5* expression. Whereas EBF1 and E2A synergistically regulate a large proportion of important B cell-specific targets (Lin et al, 2010; Treiber et al, 2010), ample evidence suggests that, at the *Pax5* locus, E2A binding is secondary in importance to EBF1 action. Firstly, in $Ebf1^{-/-}$ progenitor cells, in which E2A is already expressed, the *Pax5* promoter is repressed by H3K27me3 (Decker et al, 2009) and is most probably still in a bivalent state characterized by the co-presence of lower amounts of H3K4me2 and H3K4me3 together with the H3K27me3 mark (Treiber et al, 2010). De-repression (loss of H3K27me3) and acquisition of H3K4me3 and H3K9Ac only happens at the pro-B cell stage upon EBF1 expression (Decker et al, 2009; Treiber et al, 2010). Secondly, DNaseI hypersensitive sites are not formed in $Ebf1^{-/-}$ progenitor cells (Decker et al, 2009) and E2A binding in HS6 is not detected (Busslinger laboratory, unpublished data). E2A binds at the *Pax5* promoter only upon EBF1 expression in both $Pax5^{-/-}$ and wild-type pro-B cells (Figure 25A). This line of evidence thus excludes an important role for E2A at the *Pax5* promoter prior to EBF1 expression. Concerted action of the two transcription factors in driving maximum *Pax5* expression at the pro-B cell stage is however very likely and should be further addressed in the $Pax5^{\Delta EHS6+7/\Delta EHS6+7}$ context (by inactivating the E2A-binding site at HS6 in the $Pax5^{\Delta EHS6+7/\Delta EHS6+7}$ genetic background).

Upon expression in pro-B cells, Pax5 binds its own locus at several locations (Figure 25B), including the enhancer and at HS6 in the promoter. Thus, it is highly possible that once activated Pax5 is able to auto-regulate its own expression. Observations so far lead to the idea that Pax5 downregulates its own expression three-fold (Revilla-i-Domingo et al, 2012). This probably becomes important at later stages

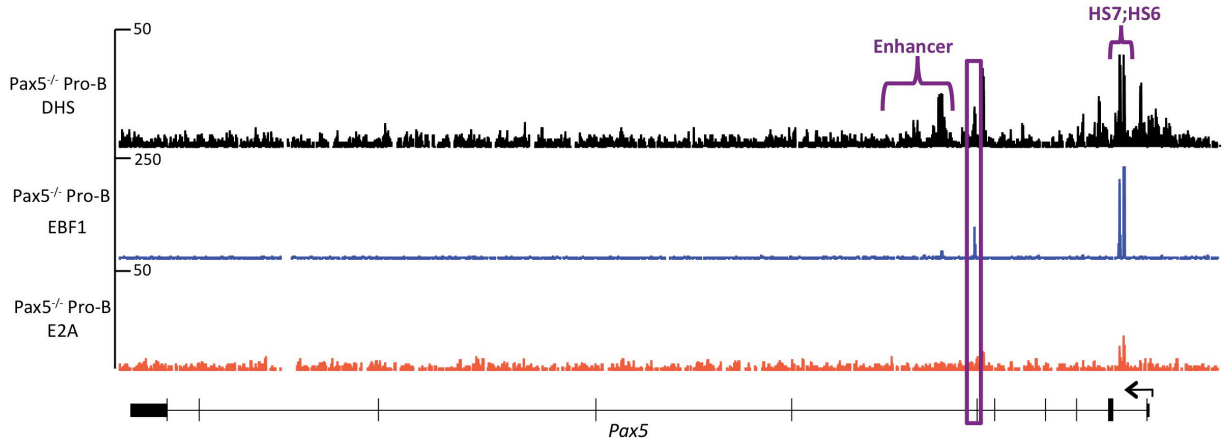
of development when EBF1 becomes dispensable (Vilagos et al, 2012) and E2A does not bind the *Pax5* locus any more (Busslinger laboratory, unpublished observations).

Interestingly, a newly identified EBF1-binding site of lesser intensity than the two peaks at HS6 and HS7 is present in a hypersensitive region that does not belong to the enhancer, in intron 5 of *Pax5* locus (Figure 25A). This peak is present only in *Pax5*$^{-/-}$ cells (Figure 25 and Busslinger laboratory database). Removal of intron 5 abrogates *Pax5* expression (Decker et al, 2009), which led to the identification of the B cell-specific enhancer at this location. These findings did however not exclude the presence of additional regulatory elements in intron 5. One example could be the hypersensitive site bound by EBF1 in *Pax5*$^{-/-}$ cells, which was later identified by more sensitive experimental approaches such as genome-wide sequencing experiments.

The *Pax5* enhancer is already active in *Ebf1*$^{-/-}$ progenitor – characterized by hypersensitivity to DNaseI digestion, presence of H3K9Ac, and binding of PU.1 (Decker et al, 2009). One interesting study suggested two separate stages for the location of a λ5 transgene outside of heterochromatin and its activation – formation of a hypersensitive region that is bound by E2A, and pre-B cell-specific EBF1-dependent modulation of higher-order chromatin structure and activation of transcription (Lundgren et al, 2000). By analogy, it is possible that EBF1 binding in intron 5 of the *Pax5* gene represents an intermediate stage between activation of the enhancer and binding of EBF1 at the gene promoter, followed by transcriptional activation. EBF1 binding requires a permissive chromatin context (Treiber et al, 2010), and thus it could be that EBF1 binding in the proximity of the enhancer is more favorable initially. Activation of the *Pax5* promoter by EBF1 binding at HS6 and HS7 could subsequently be facilitated by CTCF mediated looping (Gillen and Harris, 2011; Weth and Renkawitz, 2011). CTCF-binding sites are identified at the enhancer and promoter of *Pax5* (Busslinger laboratory, unpublished observations).

Further experiments are necessary to investigate the function of EBF1 binding in intron 5 in *Pax5*$^{-/-}$ cells in order to elucidate the steps that lead to the activation of the *Pax5* gene in early B lympopoiesis.

A



B

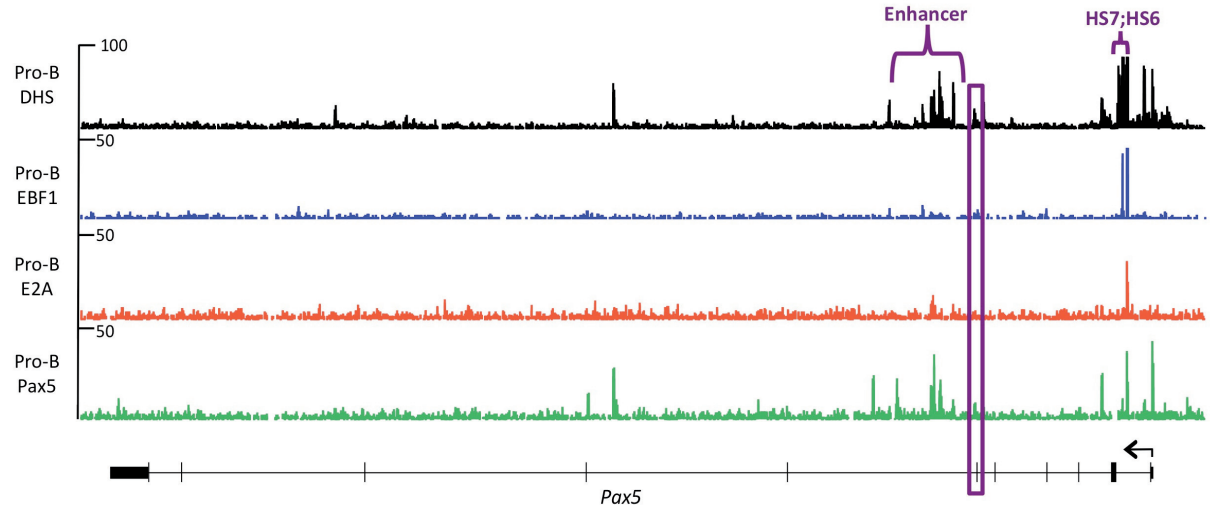


Figure 25. E2A, EBF1, and Pax5 binding at the *Pax5* locus.
E2A, EBF1, and Pax5 binding sites as well as DNaseI hypersensitive sites (DHS) were mapped at the *Pax5* locus from ChIP-sequencing experiments in **A.** *Pax5*$^{-/-}$ pro-B cells and **B.** *Rag2*$^{-/-}$ pro-B cells (Busslinger laboratory database). Binding patterns are presented for the *Pax5* locus in the respective cell types. For our purposes, *Rag2*$^{-/-}$ cells are wild-type pro-B cells, the deficiency in the *Rag* gene being used to arrest the cells at the pro-B cell stage (Shinkai et al, 1992). The positions for the enhancer and HS6 and HS7 at the promoter are indicated. Note that the EBF1 binding peak in intron 5 of *Pax5* present in *Pax5*$^{-/-}$ cells is not present in Rag2$^{-/-}$ cells (marked by the purple column).
All data presented comes from the Busslinger laboratory database.

Additionally, factors with functions at earlier progenitor stages could extent their action, via cooperative interactions, to the regulation of B cell commitment by playing a role in the activation of the *Pax5* gene. One such factor that could act in cooperation with EBF1 is RUNX1, which has been reported to rely on its interaction partners to achieve tissue-specific target gene regulation (Lam and Zhang, 2012). Unless an unknown EBF1-dependent factor recruits RUNX1 specifically to the *Pax5* locus, it is unlikely that the action of RUNX1 alone at the *Pax5* locus can explain the discrepancies outlined in this study, since RUNX1 is expressed very early in development (Lam and Zhang, 2012).

Therefore, RUNX1 can only be regarded as a potentially important EBF1 cooperation partner, the contribution of which needs to be elucidated. Particularly, it has been integrated in the transcription factors network that promotes B cell specification and commitment since it directly regulates *Ebf1*, and *Ebf1, E2a*, and *Pax5* levels are decreased in its absence (Seo et al, 2012; Kuo et al, 2008). The deficiencies in the B cell compartment were restored by EBF1 expression in these studies. Interestingly, RUNX1 has been shown to cooperatively bind the *Cd79a* promoter together with EBF1, and reactivate transcription in the context of hypermethylated DNA (Maier et al, 2004). The RUNX1 motif is among the overrepresented binding motifs in regions bound by EBF1 (Treiber et al, 2010) and effects on B cell development of compound *Ebf1* and *Runx1* haploinsufficiencies have been reported (Lukin et al, 2010).

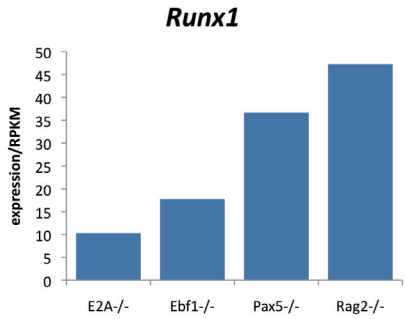
In pro-B cells two RUNX1-binding peaks have been identified in the promoter region of *Pax5* in the same region like the EBF1 binding peaks at HS6 and HS7 (Figure 26B). The RUNX1-binding motif is present in the DNA sequence of HS6 and HS7 (not shown), however further experiments are needed to confirm these as bona-fide RUNX1 binding sites. Also, a RUNX1-binding site in a DNaseI hypersensitive site in intron 4 of the *Pax5* gene has been identified in pro-B cells (Figure 26B). However, removal of intron 4 alone did not influence *Pax5* expression in a transgenic reporter assay of *Pax5* expression (Decker et al, 2009). It could be that the local chromatin context is important in such a setting, which cannot be recapitulated in a transgene experiment unless the transgene is an independent chromatin domain (Martin and Whitelaw, 1996).

Site-directed ChIP experiments as well as in vivo DMS footprinting and EMSAs could be done to investigate whether RUNX1 binding is confirmed at the three sites (in intron 4, at HS6 and HS7), and more importantly whether cooperative binding of EBF1 and RUNX1 at HS6 and/or HS7 is detected. Although *Runx1* expression is initiated very early in hematopoietic development, it progressively increases as commitment to the B cell lineage progresses (Figure 26A). This could be indicative of a continuous requirement for this transcription factor. Furthermore, analysis of the conditional ablation of *Runx1* with B cell-specific Cre lines is in incipient form (Growney et al, 2005; Seo et al, 2012). Recently, a

block in B cell development before emergence of IgM^{+} cells was reported in *Runx1$^{fl/fl}$mb-1Cre* mice (Seo et al, 2012).

In what the *Pax5* expression is concerned, it is moderately reduced in *Ebf1$^{+/-}$* mice and this reduction is not exacerbated by in *Ebf1$^{+/-}$Runx1$^{+/-}$* mice (Lukin et al, 2010). However, in the same study expression of *Cd79a* was not affected at all by the transcription factor haploinsufficiencies analyzed (both alone, and in combination). Thus it is highly possible, as the authors of the study also claim, that at genes where the two factors bind in a highly cooperative manner, this interaction is able to overcome the effects of their reduced dosage. Therefore, an alternative experimental approach is necessary to investigate the possibility of cooperative regulation of *Pax5* by EBF1 and RUNX1.

A



B

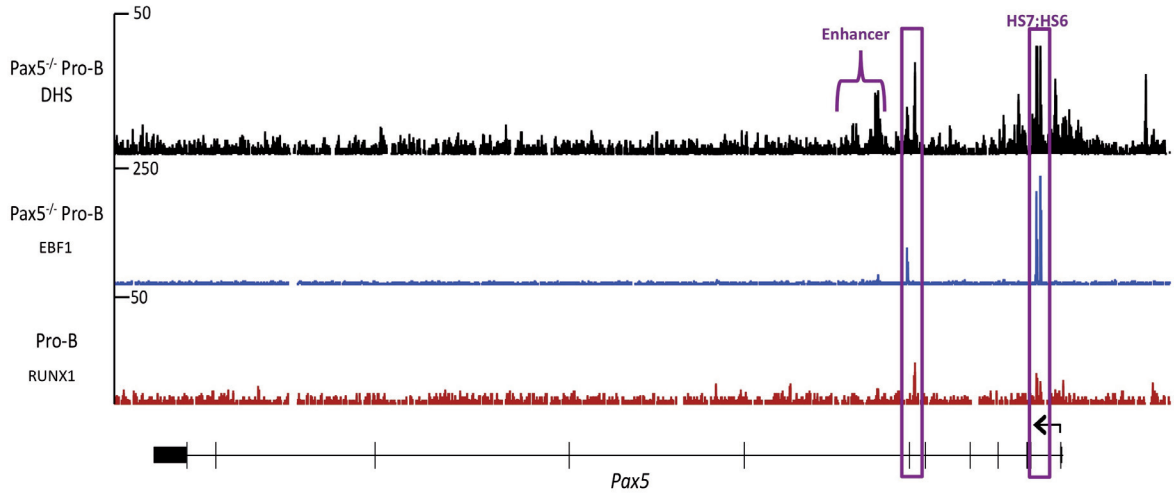


Figure 26. **RUNX1 as a potential co-regulator of the *Pax5* locus.**
A. The expression levels of *Runx1* resulting from mRNA sequencing experiments in the respective cell types. **B.** The RUNX1 binding profile in pro-B cells (Rag2$^{-/-}$) was aligned with the profile for EBF1 binding and DHS mapping in Pax5$^{-/-}$ cells. The positions for the enhancer and HS6 and HS7 at the promoter are indicated. The RUNX1 binding sites of interest are marked by the purple columns.
All the data presented comes from the Busslinger laboratory database.

3. Indirect regulation of *Pax5* expression by EBF1 - candidate mediators

The most plausible scenario that would reconcile the fact that EBF1-binding sites deletion in the *Pax5* promoter results in only a two-fold reduction in gene expression, with previous reports of strict dependence of *Pax5* expression on EBF1 (Decker et al, 2009; Treiber et al, 2010; Zandi et al, 2008), suggests that EBF1 has both a direct (this study) and an indirect effect on the *Pax5* gene. To this end, transcription factors that are EBF1-activated targets that could subsequently influence *Pax5* expression should be identified. For this search, transcription factors that change their expression level in response to EBF1 expression are immediate candidates. Such factors have already been identified both by the Busslinger laboratory (Bojan Vilagos, personal communication) and by others (Treiber et al, 2010). Among these OBF1 (OCA-B, Pou2af1) and Foxo1 particularly caught the attention of this study, and they will be discussed in detail below. The thought process that led to highlighting these two particular transcription factors was by no means exhaustive, and thus implication of other EBF1-activated genes in the regulation of *Pax5* expression cannot be ruled out.

Furthermore, it becomes clear from this study that additive effects of more transcription factors are most likely responsible for the activation of the *Pax5* gene. Thus the search for new candidate players should be focused on identification and quantification of their direct effects on the *Pax5* gene, followed by integration into a network that is just being unraveled. To this end, no transcription factor-binding site can be pointed out as the site through which an all or nothing effect on *Pax5* expression is mediated.

3.1 OBF1/Pou2af1/OCA-B

OBF1 is a B cell-specific transcriptional co-activator also known under the names OCA-B, Pou2af1, and Bob1. It fulfills functions at various stages of B cell development, but its most striking phenotype is the complete lack of germinal centers in *OBF1 (Pou2af1)* deficient mice (Teitell 2003). The most studied function of this transcription factor is at the immunoglobulin heavy-chain locus (*IgH*) where it binds a highly conserved octamer motif present in promoters of V_H genes and enhancers, together with Oct-1 and Oct-2, thus stimulating transcription (Matthias 1998). In this respect, a study showed that OBF1 interacts with the promoter-bound TFII-I, thus relieving histone deacetylase 3 (HDAC3) repression and facilitating promoter-enhancer interactions at the *IgH* locus (Ren et al, 2011).

OBF1 was identified as a direct EBF1 target that is upregulated in early lymphoid compartments by several studies (Zhang et al, 2008; Treiber et al, 2010; Inlay et al, 2009). This is in agreement with previous findings that showed that *OBF1 (Pou2af1)* B cell-specific expression is a result of the action of cell-type specific elements that regulate its gene locus (Massa et al, 2003). Furthermore, mice that overexpress *OBF1* under the control of an *Igh* promoter and E_{μ} enhancer display among other defects, impaired commitment to the B cell lineage. In these mice, OBF1 overexpression results in increased expression of *Ebf1*, *Pax5*, and the E2A antagonists *Id2* and *Id3* (Bordon et al, 2008).

Although direct binding of OBF1 to the *Pax5* locus has never been reported, its specific octamer motif is found several times across the locus (not shown); whether or not OBF1 binds these sites remains an open question.

As mentioned, levels of *OBF1* transcripts increase in response to *Ebf1* expression in *Pax5*$^{-/-}$ cells (Figure 27A). Once Pax5 is activated, it also directly regulates OBF1 (Revilla-i-Domingo et al, 2012) leading to a further increase in *OBF1* expression (Figure 27A). As seen in Figure 27A, the same pattern of expression is valid for the direct OBF1 target Spi-B (Bartholdy et al, 2006) and the OBF1 binding partner Oct-2/Pou2f2 (Teitell, 2003; Matthias 1998). Similar to OBF1, Spi-B is also a direct target of Pax5 in pro-B cells (Revilla-i-Domingo et al, 2012).

It is therefore possible that OBF1/Oct-2 upregulation is necessary at the *Pax5*$^{-/-}$ stage to contribute to the activation of *Pax5* expression by EBF1. Higher levels of OBF1 are detrimental for B cell development prior to commitment (Bordon et al, 2008) and are only detected in response to *Pax5* expression, consistent with the reported functions of OBF1 at later stages of B cell development (Teitell, 2003). This tight regulation of transcript dosage could also indicate distinct functions at the different stages. However, early B cell development is normal in OBF1-deficient mice (Schubart et al, 1996), which again hints at the possibility of cumulative effects in cooperation with other factors on *Pax5* expression.

An additional possibility is that OBF1 indirectly affects *Pax5* transcription, through its direct target Spi-B (Bartholdy et al, 2006). The Ets trancription factor Spi-B shares about 70% Ets-domain amino acid homology with PU.1 (encoded by *Sfpi1* gene) and they have indistinguishable DNA-binding activities (Wei et al, 2010). Spi-B has been shown to partially replace PU.1 in B cell differentiation (Schweitzer et al, 2006). A recent study showed that mice deficient in both PU.1 and Spi-B (*CD19*$^{+/Cre}$*Sfpi1*$^{lox/lox}$*Spib*$^{-/-}$) develop an aberrant pre-B cell compartment and subsequent acute lymphoblastic leukemia (ALL); in these leukemic cells, *Pax5* expression was two-fold reduced (Sokalski et al, 2011). A PU.1-binding site has already been identified in the *Pax5* enhancer

(Decker et al, 2009). It would be interesting to know whether at any stage of B cell development, PU.1 is replaced by Spi-B at this site, and whether additional Spi-B binding sites are located in the *Pax5* gene.

3.2 Foxo1

Perhaps the highest ranked candidate to fulfill the function of mediating the indirect effects of EBF1 on *Pax5* transcription is Foxo1. It belongs to the highly conserved family of forkhead box transcription factors, which counts over 40 members with varied functions such as transcriptional regulation and modification of chromatin structure (Lalmansingh et al, 2012; Monsalve and Olmos, 2011). Foxo1-deficient mice show a block at the pro-B cell stage; conditional inactivation of the gene also points out important functions for Foxo1 at later B cell stages (Dengler et al, 2008). Recently, Foxo1 has been implicated in the transcription factor network that controls B cell commitment (Lin et al, 2010).

Of note for the purposes of this study is the mode in which Foxo1 participates in the activation of transcription of early B lineage factors. At the ALP stage *Foxo1* is induced by E2A in concert with HEB (Welinder et al, 2011); subsequently, expression of *Ebf1* is initiated in response to E2A, Foxo1, and other factors (Lin et al, 2010; Nutt and Kee, 2007). The transcripts of *Foxo1* are further upregulated in response to *Ebf1* expression (Figure 27B) and EBF1 binding at the *Foxo1* locus is detected in $Pax5^{-/-}$ pro-B cells (Busslinger laboratory, data not shown). *Foxo1* has also been identified as a direct EBF1 target in earlier studies (Zandi et al, 2008).

Particularly relevant is the fact that Foxo1 binding is detected in pro-B cells at the *Pax5* enhancer, as well as in other regions of the *Pax5* locus (Lin et al, 2010 and Figure 27C). Thus it is very likely that concerted activation of *Pax5* expression by EBF1 and Foxo1 is the next step in the transcription factors relay outlined above. To round up the picture, once expressed *Pax5* directly regulates *Foxo1* at the pro-B cell stage (Revilla-i-Domingo et al, 2012).

Foxo1 binding at the *Pax5* locus should thus be studied in the $Pax5^{\Delta EHS6+7/\Delta EHS6+7}$ context, as well as in the context of stepwise B cell development to investigate the exact time points at which its binding sites are first detected.

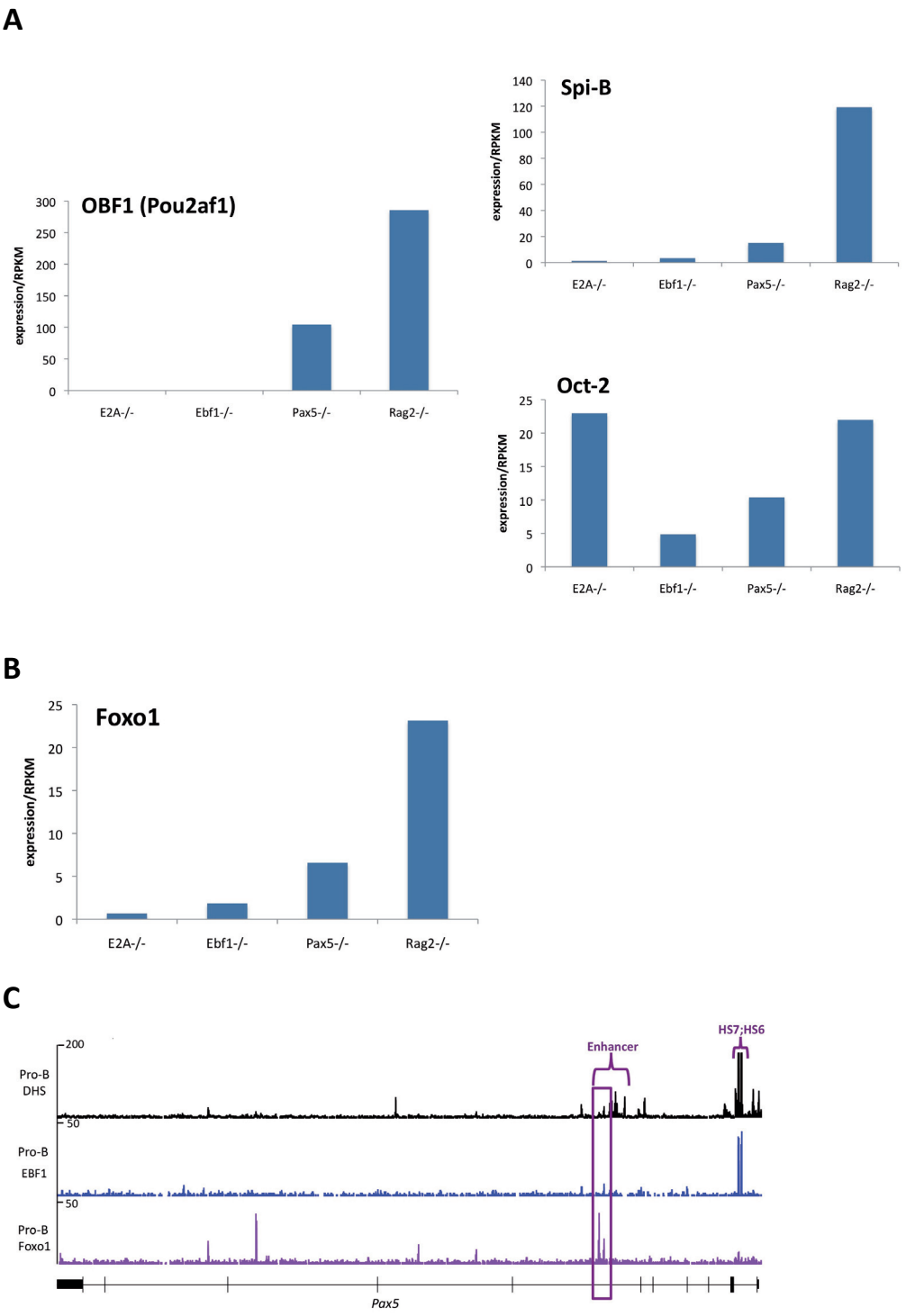


Figure 27. Expression profiles and binding patterns at *Pax5* locus for potential mediators of indirect EBF1 effects on *Pax5* expression. The expression levels resulting from mRNA sequencing experiments in the respective cell types for **A.** *Pou2af1, Spi-B, Pou2f2 (Oct-2)* **B.** Foxo1.
C. The Foxo1 binding profile (from Lin et al, 2010) in pro-B cells ($Rag2^{-/-}$) was aligned with the profile for EBF1 binding and DHS mapping in the same cell type from the Busslinger laboratory database. The positions for the enhancer and HS6 and HS7 at the promoter are indicated. The Foxo1-binding sites in the enhancer are marked by the purple column.

4. Perspective

Whereas genetic gain- and loss-of-function experiments have so far led to the accumulation of a large body of knowledge on the transcription factors that regulate B cell development (Nutt and Kee, 2007; Busslinger 2004), the arrival of genome-wide sequencing (MacQuarrie et al, 2011; Park 2009) poses a further challenge for the assembly of gene-regulatory networks important in this context (Lin et al, 2010). B cell development has been long regarded as a model pathway for the study of general principles of gene regulation (Bryder and Sigvardsson, 2010), which is why it becomes increasingly more important and challenging to integrate available information at the level of the regulation of single genes. Such approaches should eventually lead to the emergence of a global picture of how transcription factor relays and networks assemble regulatory modules at gene promoters/enhancers and drive cell-type-specific gene expression.

For the B cell lineage, understanding how the commitment factor Pax5 is regulated represents an ideal starting point (Decker et al, 2009). Implications for such findings extent beyond principles of gene regulation, to treatment of leukemia and manipulation of antibody-mediated responses (Cobaleda et al, 2007a; Medvedovic et al, 2011). To this end, a first step was the identification of the regulatory elements of the Pax5 locus (Decker et al, 2009). This knowledge was complemented and expanded by recent findings from genome-wide sequencing experiments (Busslinger laboratory, unpublished observations). The present study constitutes a logical next step in this undertaking, by investigating the direct effects of EBF1 at the *Pax5* promoter. The finding that EBF1 binding at HS6 and HS7 modulates, rather than initiates, *Pax5* expression opens the door to interesting and challenging research avenues. Some of them have been discussed in detail above. Others, like an extensive search for binding sites of potential EBF1 partners and EBF1-regulated factors at the *Pax5* promoter, remain to be addressed.

The mouse lines generated in this study represent a genetic tool that can be used as a starting point for the bottom-up assembly of the gene regulatory network that drives *Pax5* expression and B cell development.

MATERIALS AND METHODS

1. The *Pax5*$^{\Delta EHS7/+}$ and *Pax5*$^{\Delta EHS6+7/+}$ mice - cloning, targeting; Southern screening

1.1 Targeting vector (Pax5$^{\Delta EHS7}$)

The targeting vector was assembled by recombineering (Liu et al, 2003) in the pSP65 plasmid, which contains HSV-tk and DT-A genes for negative selection. In short, a minitargeting vector was firstly assembled in the pL451-E plasmid that already contained the following sequence of elements: Frt-site; PGK promoter; EM7 promoter; Neomycin gene; polyA; Frt-site; lox-P site. Upstream of these elements, a homology box (boxD) (amplified from the mPax5 BAC RP-23-182C16 with primers listed in Table 1) was cloned into SalI and EcoRI sites. Downstream, a homology box (boxE) (amplified from the mpax5 BAC RP-23-182C16 with primers listed in Table 1) was cloned into BamHI and NotI sites. Two retrieval vectors were assembled in parallel in the pSP65 vector, by cloning of homology boxes A (NotI and HindIII) and B (XbaI and SpeI), or A (NotI and HindIII) and C (XbaI and SpeI) (amplified from the mpax5 BAC RP-23-182C16 with primers listed in Table 1) into restriction sites present in the plasmid. Firstly, the targeting cassette in the pL451-E was excised and purified (from boxD to boxE, containing the sequence of elements between them), and 1-2 μl of it were electroporated into EL350 cells (recombineering active bacteria strain, Liu et al 2003) containing the mPax5 BAC RP-23-182C16 (bacteria were previously induced at 42°C for 15 minutes and made electro-competent), in order to introduce by recombination the desired modification at the *Pax5* promoter (see Figure 11A). Clones that integrated the modification cassette were selected on chloramphenicol/kanamycin resistant plates and partially sequenced; correct colonies were grown, induced at 42°C, made electro-competent and electroporated with the linearized retrieval vector (either AB, or AC – named according to which homology boxes were used). This step yielded the final Pax5$^{\Delta EHS7}$AB and Pax5$^{\Delta EHS7}$AC targeting vectors that were isolated, sequenced completely, and re-introduced into DH5α individually.

For ES cell electroporation experiments the Pax5$^{\Delta EHS7}$AC (containing a longer arm of homology that spans up to box C, downstream of *Pax5* exon 1B) was used; bacteria were grown to allow plasmid isolation by Qiagen maxiprep; the vector isolated was re-sequenced, and then 100 μg of the vector were linearized by digestion with PvuI, purified, kept under sterile conditions, and used in all necessary targeting experiments.

OLIGO NAME	SEQUENCE	USED FOR AMPLIFICATION OF:
MB 9675	CCATGCGGCCGCACTACCCAGTGCTAGTATCCA	Box A
MB 9676	GCGAAGCTTCCAAACTTGAACATCTCCAGATT	Box A
MB 9716	GCGGTCGACAAGAACTCTAGTGAGGAAGGTCAG	Box D
MB 9678	GCGGAATTCGGTGATGCTCTTAGCTCAAAGCT	Box D
MB 9679	GCGGGATCCTATCAGCTCCTAGGCTAAGACA	Box E
MB 9680	GGTTGCGGCCGCTCTATCATTCTTCTATCCACAG	Box E
MB 9681	GCGCTCGAGAGGTGTTTTCTGATCTCTGCT	Box B
MB 9682	GCGACTAGTGAGAGTGAGAGAAGCGAGCT	Box B
MB 9683	GCGCTCGAGTCCTATGAGACAAGCGCTCTTC	Box C
MB 9718	TACGACTAGTACCCGAGTGCAGTCTGCCAAA	Box C

Table 1. Oligos used for the PCR amplification of homology boxes for cloning of Pax5$^{\Delta EHS7}$.

1.2 Targeting vector (Pax5$^{\Delta EHS6+7}$)

The Pax5$^{\Delta EHS6+7}$ targeting vector (described in Figure 14) was cloned starting from the Pax5$^{\Delta EHS7}$AB construct (that spans up to the end of *Pax5* exon 1B). Firstly, the old targeting cassette was digested with enzymes that cut upstream of HS6 in the long arm and downstream the mutated EBF1-binding site in HS7. The deleted sequences were replaced with a linker that contains restriction sites necessary for the introduction of the new targeting cassette. This cassette was assembled in the previously cloned plasmid pL451E-iloxP that contained the necessary elements in the following order: Frt-site; PGK promoter; EM7 promoter; Neomycin gene; polyA; lox-P site; Frt-site. To introduce the desired mismatches-containing short arm (see Figure 14B), a linker was cloned into the pL451E-iloxP plasmid that contained restriction sites for the introduction of PCR fragments spanning the regions of interest in *Pax5* intron 1 (Table 2); the cloning strategy first introduced the NdeI mutation at HS6 by cut and paste of the two PCR fragments into the pL451E-iloxP linker. Then, the entire targeting cassette was cut from the pL451E-iloxP plasmid and pasted into the previously digested Pax5$^{\Delta EHS7}$AB construct (pSP65 backbone as mentioned) by using previously introduced ClaI sites, and thus the mutation of the EBF1 site at HS7 was introduced. This step yielded the final Pax5$^{\Delta EHS6+7}$ targeting vector that was isolated, sequenced completely, and re-introduced into DH5α.

For ES cell electroporation experiments bacteria containing the Pax5$^{\Delta EHS6+7}$ targeting vector were grown to allow plasmid isolation by Qiagen maxiprep; the vector isolated was re-sequenced, and then 100 μg of the vector were linearized by digestion with PvuI, purified, kept under sterile conditions, and used in all necessary targeting experiments.

OLIGO NAME	SEQUENCE	USED FOR:
MB 10986	CTTCAGGATCCGCAGCCTCTTGACATCCATAG	PCR amplification 1
MB 10988	CTGGCATATGAAGACTCAGAAGGCTGTGGTGTG	PCR amplification 1
MB 10987	GTCTTCATATGCCAGCCTTGGCTGTTTTCCTG	PCR amplification 2
MB 10989	GGTGGCGGCCGCATCGATAGAGCTGGTGATGCTCTTAGCTCAAAGC	PCR amplification 2
MB 10995	GATCCGTCTCATATGCCTCGC	Linker in pL451E-iloxP
MB 10996	GGCCGCGAGGCATATGAGACG	Linker in pL451E-iloxP
MB 10990	CCGGAATCGATTCCTATCAGCTCCTAGGCTAAGACAACAGAAAGGATCTATGCCTCTG	Linker in digested Pax5$^{\Delta EHS7}$AB
MB 10992	TACTTGAGGCAACCTGGTCCAAGCTTTGGAGACTCCCCAAATTTCTG	Linker in digested Pax5$^{\Delta EHS7}$AB
MB 10993	GTTACCAGAAATTTGGGGAGTCTCCAAAGCTTGGACCAGGTTGCCTCAAGTACAGAGGCAT	Linker in digested Pax5$^{\Delta EHS7}$AB
MB 10994	AGATCCTTTCTGTTGTCTTAGCCTAGGAGCTGATAGGAATCGATT	Linker in digested Pax5$^{\Delta EHS7}$AB

Table 2. Oligos used for the cloning of Pax5$^{\Delta EHS6+7}$.

1.3 Targeting experiments in AP5 ES cells

AP5 ES cells were grown on a confluent layer of γ-irradiated mouse embryonic fibroblast (MEF) feeder cells, which were isolated from E14 DR4 embryos, expanded in culture for five passages, irradiated and frozen. DR4 MEFs are resistant to four different antibiotics, including G418 (Tucker et al, 1997). ES cells were kept in high glucose DMEM medium supplemented with 15% FCS (Sigma, lot 055K3396, ES cell culture tested); 2mM Glutamine, 100U/ml Penicillin, 0.1mg/ml Streptomycin, 1% non-essential amino acids, 0.1mM β-mercaptoethanol were added additionally (all from Sigma) and 1000U/ml recombinant LIF (ESGRO) was added freshly. The medium was exchanged daily and cells were generally passaged every second day.

For electroporation, one confluent 15 cm plate was used; cells were harvested, resuspended in 900μl PBS and were added to a tube containing 20-30μg purified, linearized DNA (amount depended on the degree of confluence of the cells). The mixture was electroporated at 0.24kV and 500μF in a 4mm cuvette (Biorad), incubated five minutes at room temperature, an then distributed to four 10 cm plates with feeder cells. Medium was exchanged 12 hours after electroporation, and 24 hours after medium exchange selection was started by addition of 0.3μg/ml G418 (Geneticin from Gibco Life Technologies) to the medium. Selection medium was exchanged daily for 4-5 days, until resistant colonies became visible and were picked. At this point, the medium was exchanged with PBS and clones were picked in a volume of 20μl into a 96-well plate and trypsinized in 40μl. The entire suspension was immediately transferred to a feeder coated 96-well plate with fresh medium and expanded in this format. When most picked clones reached splitting density, the medium was aspirated, cells were washed once with 50μl PBS, and 20μl 1x trypsin solution was added; 100μl of medium was added after 4 minutes to stop

the reaction; 20μl of the mixture were transferred to a new feeder coated 96-well plate with fresh medium and expanded further. To the remaining 100μl, freezing medium was added to a total volume of 200μl (final concentrations, taking into account the volume of medium previously added: 50% medium, 40% FCS, 10% DMSO). Plates destined for freezing were immediately transferred to -80°C. At subsequent thawing, they would be left to defrost at 37°C and as soon as liquid, the entire mixture from one well would be transferred to one well of a 12-well plate which contains new feeders and excess fresh medium.

1.4 Southern blotting screening in 96-well format

For Southern analysis, 96-well plates with picked ES cell clones were expanded until fully confluent (about 3-4 days in culture, with medium exchange once or twice per day to allow rapid growth). When confluent, the cells were washed once with PBS, and incubated overnight in 50μl lysis buffer/well (10mM Tris pH7.5, 10mM EDTA pH8, 10mM NaCl, 0.5% Sarcosyl, and 1mg/ml Proteinase K) in a humidified chamber in a water bath at 60°C. After overnight incubation, plates could be stored indefinitely at -20°C. When needed, they were thaw at room temperature and 100μl/well of a freshly made solution of 75mM NaCl in cold ethanol was added. The plates were incubated at room temperature for up to one hour, until the precipitated DNA was clearly visible; they were then spun at maximum speed for 5 minutes for the precipitated DNA to stick to the bottom of the well, and washed once with 70% ethanol. The precipitated DNA was air-dried at room temperature for 4-6 hours, and subsequently 30μl/well of digestion mix (1x restriction buffer specified for the enzyme being used, 1mM spermidine, 100μg/ml bovine serum albumin, 10-15Units of enzyme) was added. The plates for digestion were incubated overnight in a humidified chamber at the temperature specified for the enzyme of interest. The next day, 5μl/well of DNA loading buffer were added and the contents of the wells was loaded onto an agarose gel and electophoresed. After the run was complete, the DNA in the gel was depurinated (0.25M HCl) for 15 minutes, denatured (0.2M NaOH / 0.6M NaCl) for 30 minutes, and neutralized (0.5M Tris pH7.5 / 1.5M NaCl) for 30 minutes. The DNA was transferred to an Immobilion charged nylon membrane from Millipore (according to the manufacturer's instructions) by capillary transfer overnight in 20x SSC buffer. The next day, the DNA probe was radioactively labeled (Prime-It RmT Random Primer Labeling Kit from Stratagene) with [α-32P]-dCTP according to the manufacturer's instructions. The membrane was pre-hybridized in PerfectHyb™ Plus Hybridization Buffer (from Sigma) for 1 hour at 65°C and incubated with the labeled probe overnight in the same

conditions. Subsequently, the membrane was washed 1x 20 minutes washing buffer 1 (2x SSC, 0.1% SDS) and 1x 20 minutes washing buffer 2 (0.5x SSC, 0.1% SDS). The bands where visualized by exposure on the phosphoimager.

1.5 General Southern blot protocol

For Southern blotting experiments that involved a restricted number of ES cell clones, the cells were expanded in 10 cm dishes until confluent, harvested and treated as described below. For genotyping of mice, tail DNA was prepared as follows (same procedure for harvested cells). The tissue/cells were digested with proteinase K at a concentration of 1mg/ml overnight in proteinase K buffer (50mM Tris pH 8, 20mM EDTA, 100mM NaCl, 1% SDS). The next day, the DNA was recovered by phenol-chloroform extraction and isopropanol precipitation, washed with 70% ethanol and resuspended in a suitable volume of TE. The concentration was measured and digestion with the enzyme of interest was performed according to the specific enzyme requirements at a ratio of 5 units of enzyme per 1μg of DNA; 7-10μg of DNA were used per digestion, depending on availability. The samples were electrophoresed on an agarose gel and Southern blotting was further performed as described in the previous section.

1.6 Southern probes

Throughout this study, three Southern blotting probes were used. Since the targeting vectors constructed were created in a similar fashion, it was possible to use the same 5' probe and 3' probe (named after their position relative to the targeting vector), but adjust the digestion scheme (enzyme used) accordingly. In addition, a third probe internal to the Pax5$^{\Delta EHS7}$ vector was used (this probe cannot be used with Pax5$^{\Delta EHS6+7}$ since it covers HS6). The position of the probes used was indicated in the schematic diagram of each Southern blot performed. The primers used to amplify them by PCR are listed in Table 3. Following PCR amplification, the probes were run on an agarose gel, excised, and purified (Qiagen gel extraction kit). The integrity of the purified probe was rechecked by running an aliquot on an agarose gel; the concentration of the final stock was measured, and 100ng of probe were used for one individual radioactive labeling reaction.

OLIGO NAME	SEQUENCE	USED FOR AMPLIFICATION OF:
MB 9943	GCGGGATCCAGCTTTTGATCTCCAGGCGCA	5' probe
MB 9944	GCGGTCGACGGTGGAAAAAAAGCGTCCGAAG	5' probe
MB 9945	GCGGAATTCGCTCGAGTCTCTGTGGGTTGC	3' probe
MB 9946	GCGGTCGACGAAGAATCCTAGTACACCTGTCACC	3' probe
MB 11076	GCGGAATTCTTGTAAGCTCAAGGAGCAGGGTT	internal probe
MB 9678	GCGGAATTCGGTGATGCTCTTAGCTCAAAGCT	internal probe

Table 3. PCR primers used for the amplification reactions to obtain the Southern blot probes used in this study. The mPax5 BAC RP-23-182C16 was used as a template.

2. Genotyping of *Pax5*$^{\Delta EHS7/+}$ mice

For the genotyping of the *Pax5*$^{\Delta EHS7/+}$ and *Pax5*$^{\Delta EHS7/\Delta EHS7}$ mice, primers previously reported were used to detect the presence of IRES-hCD2 (Fuxa and Busslinger, 2007). Additional primers were designed to amplify across the mutated EBF1 site in HS7 (Table 4).

FORWARD PRIMER	REVERSE PRIMER	ALLELE DETECTED/ PRODUCT SIZE
MB 3441 GAGACAAGAGCCCACAGAGTA	MB 3445 ACAGAGGAGCAAGACGGTATT	*Pax5*^{ihCD2} 327 bp
MB 3446 GCATTGGGCATAGTGTAGAGA		*Pax5*$^{+}$ 916 bp
MB 10991 TGGCTTCCGGTGCTCTCATTA	MB 10999 ACGGCCGTCTCCTCTGTATCC	*Pax5*$^{\Delta EHS7}$ 418 bp
		Pax5^{+} 312 bp

Table 4. Primers used to genotype the *Pax5*$^{\Delta EHS7/+}$ and *Pax5*$^{\Delta EHS7/\Delta EHS7}$ mice.

3. Analysis of *Pax5*$^{\Delta EHS6+7/+}$ mice

3.1 AdenoCre virus infection of ES cells

Before infection, ES cell clones were expanded and maintained in culture for one-two passages, in order to ensure good growing speed. When cells reached splitting confluence, they were washed with PBS and trypsinized, in order to ensure homogenization. Cells were plated in feeder-free (gelatinized) dishes with high glucose DMEM containing 2% FCS (no additional reagents added); the medium contained 40μl AdenoCre virus and 10μl polybrene/ 15 cm dish – the ratio was adjusted accordingly when cell culture dishes smaller than 15 cm dishes were used (AdenoCre virus and polybrene were a kind gift from Peters Lab, IMP). After 2 hours incubation at 37°C, the infection medium was diluted 1:2 with ES cell medium (previously described; concentration of reagents were adjusted according to the volume). After another 2-4 hours incubation 37°C (once the majority of cells has adhered to the plate), the medium was exchanged with fresh ES cell medium, and the cells were expanded to confluence and harvested; DNA was prepared as described for Southern blotting protocol. PCR reactions were carried out on 1:10 dilution of stock DNA with primers summarized in Table 5.

FORWARD PRIMER	REVERSE PRIMER	PRODUCT SIZE
MB 11423 CTCCCCCGTGCCTTCCTTGAC	MB 11424 GACCTTGCATTCCTTTGGCGAGAG	516 bp

Table 5. Primers used to monitor AdenoCre mediated deletion in Figure 15B.

3.2 PCR amplification across HS6 and HS7; genotyping

The primers used to amplify across HS6 and HS7 in Figures 15F and 16F are found in Table 6. For genotyping, primers that amplify upstream of the sites of interest, across the Frt site that remains after removal of Neomycin resistance (see Figure 14A) were designed and used. Alongside, previously mentioned primers that distinguish the allele carrying the IRES-hCD2 reporter from the wild-type allele were used (Table 4).

FORWARD PRIMER	REVERSE PRIMER	USED FOR AMPLIFICATION ACROSS:
MB 10991 TGGCTTCCGGTGCTCTCATTA	MB 10999 ACGGCCGTCTCCTCTGTATCC	HS7; followed by ClaI digest; on ES cell DNA
MB 11993 CACTCCCACTGTCCTTTCCTAATA	MB 11994 TCTTCTTGCTCCACTTCTAACATC	HS6; followed by NdeI digest; on ES cell DNA
MB 11995 TCGCTGGGGACTACGGGATAAAG	MB 11996 TTCAGGCATAGTAGAGGGAGGACA	HS6; followed by NdeI digest; on tail DNA
MB 11997 AAGCCTGTTCATTTTTGTCCTAA	MB 11998 AACGGCCGTCTCCTCTGTATCCT	HS7; followed by ClaI digest; on tail DNA
MB 11730 GCAGGGTTGGGAGCTTGAGGTA	MB 11994 TCTTCTTGCTCCACTTCTAACATC	Frt site; for genotyping
MB 11817 AGCATCACCAGCTCTATCGAT	MB 10999 ACGGCCGTCTCCTCTGTATCC	ClaI detection at HS7 only; genotyping
MB 11814 AGTCTTCATATGCCAGCCTTGG	MB 11994 TCTTCTTGCTCCACTTCTAACATC	NdeI detection at HS6 only; genotyping

Table 6. Primers used to genotype/analyze *Pax5*$^{\Delta EHS6+7/+}$ mice.

4. Flow cytometric analysis (FACS)

The bone marrow was isolated by crushing of the femur and tibia of the two hind legs. Spleen and thymus were also harvested. Single cell suspensions of the organs were generated and incubated for 20 minutes with γFc block (from BD biosciences) to prevent unspecific binding of antibodies. Cells were counted and stained at a density of $5x10^7$/ml with an antibody cocktail in FACS buffer (PBS, 2% FCS, 2mM EDTA) on ice. Cells were acquired on a FACS Canto machine (from Becton Dickinson) and the results were analyzed using the FlowJo software. Populations were defined as mentioned in the respective figure legends.

5. Chromatin immunoprecipitation experiments

Bone marrow was isolated like for FACS analysis, and the $B220^+$ cells were MACS sorted according to the manufacturer's instructions (Miltenyi biotec). Cells were cultured on OP9 stromal cells in IMDM medium containing IL-7, as described (Nutt et al, 1997). After five days in culture, cells were harvested and fixed with 1% formaldehyde in culture medium for 10 minutes followed by quenching with 0.125M glycine for 5 minutes (both steps at room temperature). The cells were then sequentially washed 20 min on ice in each of the following buffers: buffer A (10mM HEPES pH8, 10mM EDTA, 0.5mM EGTA, 0.25% Triton X-100) and buffer B (10mM HEPES pH8, 200mM NaCl, 1mM EDTA, 0.5mM EGTA, 0.01% Triton X-100) and lysed overnight at 4°C in lysis buffer (50mM Tris-HCl pH8, 10mM EDTA, 0.25% SDS). The chromatin was sonicated the following day to a size range of 200-1000 bp and then diluted to decrease the SDS concentration (dilution buffer: 250mM NaCl, 1.67% Triton X-100). About 350 μg of chromatin were used for one immunoprecipitation experiment, and were incubated with 3.5 μg of the antibody of interest. The chromatin was pre-cleared two hours at 4°C with protein A sepharose beads, and then incubated with the antibody overnight (4°C). The next day, immune complexes were harvested with 30μl protein A sepharose beads; they were washed (4°C each wash 5-10 minutes) in sequential washing steps with the following buffers (in this order): RIPA buffer (50mM Tris pH8.0, 150mM NaCl, 0.1% SDS, 1% NP-40, 0.5% sodium deoxycholate), high salt buffer (50mM Tris pH8.0, 500mM NaCl, 0.1% SDS, 1% NP-40), LiCl buffer (50mM Tris pH8.0, 250mM LiCl, 1% NP-40, 0.5% Na-deoxycholate) and twice with 1x TE. The complexes of DNA and protein were eluted two times with 450μl elution buffer (1% SDS, 100mM $NaHCO_3$, 0.5 mg/ml proteinase K). Cross-linking was reversed by overnight incubation at 65°C. Immunoprecipitated DNA was extracted with phenol chloroform, followed by isopropanol precipitation. Input controls were purified alongside (1/100 of IP material).

The primers used for qPCR amplification of the regions of interest are summarized in Table 7.

PRIMER 1	PRIMER 2	DETECTION OF:
MB 9225 CCACGCACTAGAGAGAGACTCAA	MB 9226 CCGCCTCACTTCCTGTTCAGCCG	*mb-1* promoter
MB 8121 AGCCCATGACTCCTGTGTTC	MB 8122 ACAGGGGAATGAACAGGATG	*Nedd9* enhancer
MB 9967 ACCCCAGCCAGGAGACTATT	MB 9968 CGGCTTTTCAGTAGCCAGAC	*Msh6* gene exon9/10
MB 7401 TGAGGAAGGTCAGGTTTTCC	MB 7402 GGGAGCTGGTGATGCTCTTA	*Pax5* promoter HS7 T. Decker et al, 2009
MB 11544 GACAAGCGGGGAAAGTCTC	MB 11545 TCACACACCACAGCCTTCTG	*Pax5* promoter HS6
MB 11550 CGGCATCCAGAGCTAATGA	MB 11551 GGCCATCCTGATGCTACACT	*Pax5* promoter HS7
MB 12002 GCTTTGAGCTAAGAGCATCACC	MB 12003 GGACCAGGTTGCCTCAAGTA	across HS7 in *Pax5* promoter, also in ΔEHS7 context
MB 10558 CATAGATGAAGCTGCCACATAGGT	MB 10559 GTGGGCAAGGACAAAGCATTA	Intergenic region in Chromosome 1

Table 7. Primers used for qPCR amplification following ChIP experiments.

6. Additional specifications

Experiments relevant for the project but not covered in the Results section were performed, and are briefly summarized here. Firstly, the AdenoCre infection experiment in ES cell lines was conducted for the $Pax5^{\Delta EHS7}$ targeting as well; these infections were done in collaboration with the Penninger lab, at IMBA. However, at the time the results were evaluated wrongly and correct analysis was performed only after the F1 generation of the chimeric mice already provided the answer to the scientific question of interest. DNA isolated from the $Pax5^{\Delta EHS7}$ AdenoCre infected ES cell clone 5E was used as a positive control in the optimization of the PCR detection of the band of interest for the $Pax5^{\Delta EHS6+7}$ targeting.
The targeting vector to obtain $Pax5^{\Delta EHS6/+}$ mice only was also generated, and targeting experiments in the AP5 ES cell line were conducted. Again, the targeting frequency was very high (approaching 100%) and thus clones were randomly selected and screened by AdenoCre virus infection. Unfortunately, this screen identified no clone that would be of the genotype $Pax5^{\Delta EHS6/+}$ and targeted on the Pax5-IRES-hCD2 allele.

REFERENCES

Adams B., Dörfler P., Aguzzi A., Kozmik Z., Urbánek P., Maurer-Fogy I., Busslinger M. (1992) Pax-5 encodes the transcription factor BSAP and is expressed in B lymphocytes, the developing CNS, and adult testis. Genes and Development 6: 1589-1607

Adli M., Zhu J., Bernstein BE. (2010) Genome-wide chromatin maps derived from limited numbers of hematopoietic progenitors. Nature Methods 7: 615-618

Adolfsson J., Borge OJ., Bryder D., Theilgaard-Mönch K., Astrand-Grundström I., Sitnicka E., Sasaki Y., Jacobsen SE. (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

Adolfsson J., Månsson R., Buza-Vidas N., Hultquist A., Liuba K., Jensen CT., Bryder D., Yang L., Borge OJ., Thoren LA., Anderson K., Sitnicka E., Sasaki Y., Sigvardsson M., Jacobsen SE. (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 EC., Izon DJ., Amsen D., Kruisbeek AM., Weintraub BC., Krop I., Schlissel MS., Feeney AJ., 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

Barberis A., Widenhorn K., Vitelli L., Busslinger M. (1990) A novel B-cell lineage-specific transcription factor present at early but not late stages of differentiation. Genes and Development 4: 849-859

Barlev NA., Emelyanov AV., Castagnino P., Zegerman P., Bannister AJ., Sepulveda MA., Robert F., Tora L., Kouzarides T., Birshtein BK., Berger SL. (2003) A novel human Ada2 homologue functions with Gcn5 or Brg1 to coactivate transcription. Molecular Cell Biology 23: 6944-6957

Bartholdy B., Du Roure C., Bordon A., Emslie D., Corcoran LM., Matthias P. (2006) The Ets factor Spi-B is a direct critical target of the coactivator OBF-1. PNAS 103: 11665-11670

Baumgarth N. (2011) The double life of a B-1 cell: self-reactivity selects for protective effector functions. Nature Reviews Immunology 11: 34-46

Beck K., Peak MM., Ota T., Nemazee D., Murre C. (2009) Distinct roles for E12 and E47 in B cell specification and the sequential rearrangement of immunoglobulin light chain loci. Journal of Experimental Medicine 206: 2271-2284

Bernstein BE., Mikkelsen TS., Xie X., Kamal M., Huebert DJ., Cuff J., Fry B., Meissner A., Wernig M., Plath K., Jaenisch R., Wagschal A., Feil R., Schreiber SL., Lander ES. (2006) A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell 125: 315-326

Boisset JC., Robin C. (2012) On the origin of hematopoietic stem cells: Progress and controversy. Stem Cell Research 8: 1-13

Bonasio R., Shengjiang T., Reinberg D. (2010) Molecular Signals of Epigenetic States. Science 330: 612-616

Bordon A., Bosco N., Du Roure C., Bartholdy B., Kohler H., Matthias G., Rolink AG., Matthias P. (2008) Enforced expression of the transcriptional coactivator OBF1 impairs B cell differentiation at the earliest stage of development. PLoS One 3: e4007

Borge OJ., Adolfsson J., Jacobsen AM. (1999) Lymphoid-restricted development from multipotent candidate murine stem cells: distinct and complimentary functions of the c-kit and flt3-ligands. Blood 94: 3781-3790

Brandeis M., Frank D., Keshet I., Siegfried Z., Mendelsohn M., Nemes A., Temper V., Razin A., Cedar H. (1994) Sp1 elements protect a CpG island from de novo methylation. Nature 371: 435-438

Bryder D, Sigvardsson M. (2010) Shaping up a lineage--lessons from B lymphopoesis. Currrent Opinion Immunology 22: 148-153

Busslinger M. (2004) Transcriptional control of early B cell development. Annual Reviews Immunology 22: 55-79

Busslinger M., Klix N., Pfeffer P., Graninger PG., Kozmik Z. (1996) Deregulation of PAX-5 by translocation of the Emu enhancer of the IgH locus adjacent to two alternative PAX-5 promoters in a diffuse large-cell lymphoma. PNAS 93: 6129-6134

Campos EI, Reinberg D. (2009) Histones: annotating chromatin. Annual Reviews Genetics 43: 559-599

Carotta S., Wu L., Nutt SL. (2010) Surprising new roles for PU.1 in the adaptive immune response. Immunological Reviews 238: 63-75

Carsetti R., Rosado MM., Wardmann H. (2004) Peripheral development of B cells in mouse and man. Immunological Reviews 197: 179-191

Cedar H., Bergman Y. (2011) Epigenetics of haematopoietic cell development. Nature Reviews Immunology 11: 478-488

Cobaleda C., Schebesta A., Delogu A., Busslinger M. (2007a) Pax5: the guardian of B cell identity and function. Nature Immunology 8: 463-470

Cobaleda C., Jochum W., Busslinger M. (2007b) Conversion of mature B cells into T cells by dedifferentiation to uncommitted progenitors. Nature 449: 473-477

Cotta CV., Zhang Z., Kim HG., Klug CA. (2003) Pax5 determines B- versus T-cell fate and does not block early myeloid-lineage development. Blood 101: 4342-4346

Coyaud E., Struski S., Prade N., Familiades J., Eichner R., Quelen C., Bousquet M., Mugneret F., Talmant P., Pages MP., Lefebvre C., Penther D., Lippert E., Nadal N., Taviaux S., Poppe B., Luquet I., Baranger L., Eclache V., Radford I., Barin C., Mozziconacci MJ., Lafage-Pochitaloff M., Antoine-Poirel H.,

Charrin C., Perot C., Terre C., Brousset P., Dastugue N., Broccardo C. (2010) Wide diversity of PAX5 alterations in B-ALL: a Groupe Francophone de Cytogenetique Hematologique study. Blood 115: 3089-3097

Cui K., Zang C., Roh TY., Schones DE., Childs RW., Peng W., Zhao K. (2009) Chromatin signatures in multipotent human hematopoietic stem cells indicate the fate of bivalent genes during differentiation. Cell Stem Cell 4: 80-93

Czerny T., Schaffner G., Busslinger M. (1993) DNA sequence recognition by Pax proteins: bipartite structure of the paired domain and its binding site. Genes and Development 7: 2048-2061

Dahl E., Koseki H., Balling R. (1997) Pax genes and organogenesis. Bioessays 19: 755-765

Decker T., Pasca di Magliano M., McManus S., Sun Q., Bonifer C., Tagoh H., Busslinger M. (2009) Stepwise activation of enhancer and promoter regions of the B cell commitment gene Pax5 in early lymphopoiesis. Immunity 30: 508-520

Delogu A., Schebesta A., Sun Q., Aschenbrenner K., Perlot T., Busslinger M. (2006) Gene repression by Pax5 in B cells is essential for blood cell homeostasis and is reversed in plasma cells. Immunity 24: 269-281

Dengler HS., Baracho GV., Omori SA., Bruckner S., Arden KC., Castrillon DH., DePinho RA., Rickert RC. (2008) Distinct functions for the transcription factor Foxo1 at various stages of B cell differentiation. Nature Immunology 9: 1388-1398

Dey B., Thukral S., Krishnan S., Chakrobarty M., Gupta S., Manghani C., Rani V. (2012) DNA-protein interactions: methods for detection and analysis. Molecular and Cellular Biochemistry 365: 279-299

Dzierzak E., Speck NA. (2008) Of lineage and legacy: the development of mammalian hematopoietic stem cells. Nature Immunology 9 (2): 129-135

Dörfler P., Busslinger M. (1996) C-terminal activating and inhibitory domains determine the transactivation potential of BSAP (Pax-5), Pax-2 and Pax-8. EMBO Journal 15: 1971-1982

Eberhard D., Busslinger M. (1999) The partial homeodomain of the transcription factor Pax-5 (BSAP) is an interaction motif for the retinoblastoma and TATA-binding proteins. Cancer Research 59: 1716-1725s

Eberhard D., Jiménez G., Heavey B., Busslinger M. (2000) Transcriptional repression by Pax5 (BSAP) through interaction with corepressors of the Groucho family. EMBO Journal 19:2292-2303

Emelyanov AV., Kovac CR., Sepulveda MA., Birshtein BK. (2002) The interaction of Pax5 (BSAP) with Daxx can result in transcriptional activation in B cells. Journal of Biological Chemistry 277: 11156-11164

Fitzsimmons D., Hodsdon W., Wheat W., Maira SM., Wasylyk B., Hagman J. (1996) Pax-5 (BSAP) recruits Ets proto-oncogene family proteins to form functional ternary complexes on a B-cell-specific promoter. Genes and Development 10: 2198-2211

Fujisaki J., Wu J., Carlson AL., Silberstein L., Putheti P., Larocca R., Gao W., Saito TI., Lo Celso C., Tsuyuzaki H., Sato T., Côté D., Sykes M., Strom TB., Scadden DT., Lin CP. (2011) In vivo imaging of Treg cells providing immune privilege to the haematopoietic stem-cell niche. Nature 474: 216-219

Fuxa M, Busslinger M. (2007) Reporter gene insertions reveal a strictly B lymphoid-specific expression pattern of Pax5 in support of its B cell identity function. Journal of Immunology 178: 8222-8228

Gao H., Lukin K., Ramírez J., Fields S., Lopez D., Hagman J. (2009) Opposing effects of SWI/SNF and Mi-2/NuRD chromatin remodeling complexes on epigenetic reprogramming by EBF and Pax5. PNAS 106: 11258-11263

Gillen AE., Harris A. (2011) The role of CTCF in coordinating the expression of single gene loci. Biochemistry and Cell Biology, 89: 484-494

Goetz CA., Harmon IR., O'Neil JJ., Burchill MA., Johanns TM., Farrar MA. (2005) Restricted STAT5 activation dictates appropriate thymic B versus T cell lineage commitment. Journal of Immunology 174: 7753-7763

Growney JD., Shigematsu H., Li Z., Lee BH., Adelsperger J., Rowan R., Curley DP., Kutok JL., Akashi K., Williams IR., Speck NA., Gilliland DG. (2005) Loss of Runx1 perturbs adult hematopoiesis and is associated with a myeloproliferative phenotype. Blood 106: 494-504

Hagman J., Lukin K. (2005) Early B-cell factor 'pioneers' the way for B-cell development. Trends in Immunology 26: 455-461

Hagman J., Ramírez J., Lukin K. (2012) B lymphocyte lineage specification, commitment and epigenetic control of transcription by early B cell factor 1. Current Topics in Microbiology and Immunology 356: 17-38

He T., Hong SY., Huang L., Xue W., Yu Z., Kwon H., Kirk M., Ding SJ., Su K., Zhang Z. (2011) Histone acetyltransferase p300 acetylates Pax5 and strongly enhances Pax5-mediated transcriptional activity. Journal of Biological Chemistry 286: 14137-14145

Heinz S., Benner C., Spann N., Bertolino E., Lin YC., Laslo P., Cheng JX., Murre C., Singh H., Glass CK. (2010) Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Molecular Cell 38: 576-589

Hirokawa S., Sato H., Kato I., Kudo A. (2003) EBF-regulating Pax5 transcription is enhanced by STAT5 in the early stage of B cells. European Journal of Immunology 33: 1824-1829

Höflinger S., Kesavan K., Fuxa M., Hutter C., Heavey B., Radtke F., Busslinger M. (2004) Analysis of Notch1 function by in vitro T cell differentiation of Pax5 mutant lymphoid progenitors. Journal of Immunology 173: 3935-3944

Holmes ML., Pridans C., Nutt SL. (2008) The regulation of the B cell gene expression programme by Pax5. Immunology and Cell Biology 86: 47-53

Horcher M., Souabni A., Busslinger M. (2001) Pax5/BSAP maintains the identity of B cells in late B lymphopoiesis. Immunity 14: 779-790

Ikawa T., Kawamoto H., Wright LY., Murre C. (2004) Long-term cultured E2A-deficient hematopoietic progenitor cells are pluripotent. Immunity 20: 349-360

Inlay MA., Bhattacharya D., Sahoo D., Serwold T., Seita J., Karsunky H., Plevritis SK., Dill DL., Weissman IL. (2009) Ly6d marks the earliest stage of B cell specification and identifies the branchpoint between B cell and T cell development. Genes and Development 23: 2376-2381

Ji H., Ehrlich LI., Seita J., Murakami P., Doi A., Lindau P., Lee H., Aryee MJ., Irizarry RA., Kim K., Rossi DJ., Inlay MA., Serwold T., Karsunky H., Ho L., Daley GQ., Weissman IL., Feinberg AP. (2010) Comprehensive methylome map of lineage commitment from haematopoietic progenitors. Nature: 467: 338-342

Jenuwein T. and Allis D. (2001) Translating the Histone Code. Science 293: 1074-1080

Kallies A., Hasbold J., Fairfax K., Pridans C., Emslie D., McKenzie BS., Lew AM., Corcoran LM., Hodgkin PD., Tarlinton DM., Nutt SL. (2007) Initiation of plasma-cell differentiation is independent of the transcription factor Blimp-1. Immunity 26: 555-566

Kee BL., Murre C. (1998) Induction of early B cell factor (EBF) and multiple B lineage genes by the basic helix-loop-helix transcription factor E12. Journal of Experimental Medicine 188: 699-713

Kim TK., Hemberg M., Gray JM., Costa AM., Bear DM., Wu J., Harmin DA., Laptewicz M., Barbara-Haley K., Kuersten S., Markenscoff-Papadimitriou E., Kuhl D., Bito H., Worley PF., Kreiman G., Greenberg ME. (2010) Widespread transcription at neuronal activity-regulated enhancers. Nature 465: 182-187

Kim K., Doi A., Wen B., Ng K., Zhao R., Cahan P., Kim J., Aryee MJ., Ji H., Ehrlich LI., Yabuuchi A., Takeuchi A., Cunniff KC., Hongguang H., McKinney-Freeman S., Naveiras O., Yoon TJ., Irizarry RA., Jung N., Seita J., Hanna J., Murakami P., Jaenisch R., Weissleder R., Orkin SH., Weissman IL., Feinberg AP., Daley GQ. (2010b) Epigenetic memory in induced pluripotent stem cells. Nature 467: 285-290

Kondo M. (2010) Lymphoid and myeloid lineage commitment in multipotent hematopoietic progenitors. Immunological Reviews 238: 37-46

Kuo YH., Gerstein RM., Castilla LH. (2008) Cbfbeta-SMMHC impairs differentiation of common lymphoid progenitors and reveals an essential role for RUNX in early B-cell development. Blood 111: 1543-1551

Kwon K., Hutter C., Sun Q., Bilic I., Cobaleda C., Malin S., Busslinger M. (2008) Instructive role of the transcription factor E2A in early B lymphopoiesis and germinal center B cell development. Immunity 28: 751-762

Lalmansingh AS., Karmakar S., Jin Y., Nagaich AK. (2012) Multiple modes of chromatin remodeling by Forkhead box proteins. Biochemica et Biophysica Acta in print

Lam K., Zhang DE. (2012) RUNX1 and RUNX1-ETO: roles in hematopoiesis and leukemogenesis. Frontiers in Biosciences 17: 1120-1139

Lin H., Grosschedl R. (1995) Failure of B-cell differentiation in mice lacking the transcription factor EBF. Nature 376: 263-267

Lin YC., Jhunjhunwala S., Benner C., Heinz S., Welinder E., Mansson R., Sigvardsson M., Hagman J., Espinoza CA., Dutkowski J., Ideker T., Glass CK., Murre C. (2010) A global network of transcription factors, involving E2A, EBF1 and Foxo1, that orchestrates B cell fate. Nature Immunology 11: 635-643

Liu P., Keller JR., Ortiz M., Tessarollo L., Rachel RA., Nakamura T., Jenkins NA., Copeland NG. (2003) Bcl11a is essential for normal lymphoid development. Nature Immunology 4: 525-532

Liu P., Jenkins NA., Copeland NG. (2003) A Highly Efficient Recombineering-Based Method for Generating Conditional Knockout Mutations. Genome Research 13: 476-484

Lukin K., Fields S., Lopez D., Cherrier M., Ternyak K., Ramírez J., Feeney AJ., Hagman J. (2010) Compound haploinsufficiencies of Ebf1 and Runx1 genes impede B cell lineage progression. PNAS 107: 7869-7874

Lukin K., Fields S., Guerrettaz L., Straign D., Rodriguez V., Zandi S., Månsson R., Cambier JC., Sigvardsson M., Hagman J. (2011) A dose-dependent role for EBF1 in repressing non-B-cell-specific genes. European Journal of Immunology 41: 1787-1793

Lundgren M., Chow CM., Sabbattini P., Georgiou A., Minaee S., Dillon N. (2000) Transcription factor dosage affects changes in higher order chromatin structure associated with activation of a heterochromatic gene. Cell 103: 733-743

MacQuarrie KL., Fong AP., Morse RH., Tapscott SJ. (2011) Genome-wide transcription factor binding: beyond direct target regulation. Trends in Genetics 27: 141-148

Maier H., Colbert J., Fitzsimmons D., Clark DR., Hagman J. (2003) Activation of the early B-cell-specific mb-1 (Ig-alpha) gene by Pax-5 is dependent on an unmethylated Ets binding site. Molecular Cell Biology 23: 1946-1960

Maier H., Ostraat R., Gao H., Fields S., Shinton SA., Medina KL., Ikawa T., Murre C., Singh H., Hardy RR., Hagman J. (2004) Early B cell factor cooperates with Runx1 and mediates epigenetic changes associated with mb-1 transcription. Nature Immunology 5: 1069-1077

Malin S., McManus S., Cobaleda C., Novatchkova M., Delogu A., Bouillet P., Strasser A., Busslinger M. (2010) Role of STAT5 in controlling cell survival and immunoglobulin gene recombination during pro-B cell development. Nature Immunology 11: 171-179

Månsson R., Hultquist A,. Luc S., Yang L., Anderson K., Kharazi S., Al-Hashmi S., Liuba K., Thorén L., Adolfsson J., Buza-Vidas N., Qian H., Soneji S., Enver T., Sigvardsson M., Jacobsen SE. (2007) Molecular evidence for hierarchical transcriptional lineage priming in fetal and adult stem cells and multipotent progenitors. Immunity 4: 407-419

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

Martin DI., Whitelaw E. (1996) The vagaries of variegating transgenes. Bioessays 18: 919-923

Massa S., Junker S., Schubart K., Matthias G., Matthias P. (2003) The OBF-1 gene locus confers B cell-specific transcription by restricting the ubiquitous activity of its promoter. European Journal of Immunology 33: 2864-2874

Matthias P. (1998) Lymphoid-specific transcription mediated by the conserved octamer site: who is doing what? Seminars in Immunology 10: 155-163

McManus S., Ebert A., Salvagiotto G., Medvedovic J., Sun Q., Tamir I., Jaritz M., Tagoh H., Busslinger M. (2011) The transcription factor Pax5 regulates its target genes by recruiting chromatin-modifying proteins in committed B cells. EMBO Journal 30: 2388-2404

Medina KL., Singh H. (2005) Genetic networks that regulate B lymphopoiesis. Current Opinion in Hematology 12: 203-209

Medvedovic J., Ebert A., Tagoh H., Busslinger M. (2011) Pax5: a master regulator of B cell development and leukemogenesis. Advances in Immunology 111: 179-206

Mercer EM., Lin YC., Murre C. (2011) Factors and networks that underpin early hematopoiesis. Seminars in Immunology 23: 317-325

Mercier FE., Ragu C., Scadden DT. (2012) The bone marrow at the crossroads of blood and immunity. Nature Reviews Immunology 12: 49-60

Mikkola HKA., Orkin SH. (2006) The journey of developing hematopoietic stem cells. Development 133: 3733-3744

Mikkola I., Heavey B., Horcher M., Busslinger M. (2002) Reversion of B cell commitment upon loss of Pax5 expression. Science 297:110-113

Miyamoto T., Iwasaki H., Reizis B., Ye M., Graf T., Weissman IL., Akashi K. (2002) Myeloid or lymphoid promiscuity as a critical step in hematopoietic lineage commitment. Developmental Cell 3: 137-147

Monsalve M., Olmos Y. (2011) The complex biology of FOXO. Current Drug Targets 12: 1322-1350

Morrison AM., Jäger U., Chott A., Schebesta M., Haas OA., Busslinger M. (1998) Deregulated PAX-5 transcription from a translocated IgH promoter in marginal zone lymphoma. Blood 92: 3865-3878

Müller U. (1999) Ten years of gene targeting: targeted mouse mutants, from vector design to phenotype analysis. Mechanisms of Development 82: 3-21

Mullighan CG., Goorha S., Radtke I., Miller CB., Coustan-Smith E., Dalton JD., Girtman K., Mathew S., Ma J., Pounds SB., Su X., Pui CH., Relling MV., Evans WE., Shurtleff SA., Downing JR. (2007) Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. Nature 446: 758-764

Murre C. (2007) Regulation and function of the E2A proteins in B cell development. Advances in Experimental Medicine and Biology 596: 1-7

Nagy A. (2000) Cre Recombinase: The Universal Reagent for Genome Tailoring. Genesis 26: 99-109

Nichogiannopoulou A., Trevisan M., Neben S., Friedrich C., Georgopoulos K. (1999) Defects in hemopoietic stem cell activity in Ikaros mutant mice. Journal of Experimental Medicine 190: 1201-1214

Nutt SL., Urbánek P., Rolink A., 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. Genes and Development 11: 476-491

Nutt SL., Morrison AM., Dörfler P., Rolink A., Busslinger M. (1998) Identification of BSAP (Pax-5) target genes in early B-cell development by loss- and gain-of-function experiments. EMBO Journal 17:2319-2333

Nutt SL., Heavey B., Rolink AG., Busslinger M. (1999a) Commitment to the B-lymphoid lineage depends on the transcription factor Pax5. Nature 401: 556-562

Nutt SL., Busslinger M. (1999b) Monoallelic expression of Pax5: a paradigm for the haploinsufficiency of mammalian Pax genes? Biological Chemistry 380: 601-611

Nutt SL., Kee BL. (2007) The transcriptional regulation of B cell lineage commitment. Immunity 26: 715-725

Oguro H., Yuan J., Ichikawa H., Ikawa T., Yamazaki S., Kawamoto H., Nakauchi H., Iwama A. (2010) Poised lineage specification in multipotential hematopoietic stem and progenitor cells by the polycomb protein Bmi1. Cell Stem Cell 6: 279-286

Orford K., Kharchenko P., Lai W., Dao MC., Worhunsky DJ., Ferro A., Janzen V., Park PJ., Scadden DT. (2008) Differential H3K4 methylation identifies developmentally poised hematopoietic genes. Developmental Cell 14: 798-809

O'Riordan M., Grosschedl R. (1999) Coordinate regulation of B cell differentiation by the transcription factors EBF and E2A. Immunity 11: 21-31

Orkin SH., Zon LI. (2002) Hematopoiesis and stem cells: plasticity versus developmental heterogeneity. Nature Immunology 3(4): 323-328

Orkin SH., Zon LI. (2008) Hematopoiesis: An Evolving Paradigm for Stem Cell Biology. Cell 132: 631-644

Park PJ. (2009) ChIP-seq: advantages and challenges of a maturing technology. Nature Reviews Genetics 10: 669-680

Parra M. (2009) Epigenetic events during B lymphocyte development. Epigenetics 4: 462-468

Pfeffer PL., Bouchard M., Busslinger M. (2000) Pax2 and homeodomain proteins cooperatively regulate a 435 bp enhancer of the mouse Pax5 gene at the midbrain-hindbrain boundary. Development 127: 1017-1028

Pillai S., Cariappa A., Moran ST. (2005) Marginal zone B cells. Annual Reviews Immunology 23: 161-196

Pongubala JM., Northrup DL., Lancki DW., Medina KL., Treiber T., Bertolino E., Thomas M., Grosschedl R., Allman D., Singh H. (2008) Transcription factor EBF restricts alternative lineage options and promotes B cell fate commitment independently of Pax5. Nature Immunology 9: 203-215

Poppe B., De Paepe P., Michaux L., Dastugue N., Bastard C., Herens C., Moreau E., Cavazzini F., Yigit N., Van Limbergen H., De Paepe A., Praet M., De Wolf-Peeters C., Wlodarska I., Speleman F. PAX5/IGH rearrangement is a recurrent finding in a subset of aggressive B-NHL with complex chromosomal rearrangements. Genes Chromosomes Cancer 44:218-223

Ramírez J., Hagman J. (2009) The Mi-2/NuRD complex: a critical epigenetic regulator of hematopoietic development, differentiation and cancer. Epigenetics 4: 532-536

Ramírez J., Lukin K., Hagman J. (2010) From hematopoietic progenitors to B cells: mechanisms of lineage restriction and commitment. Current Opinion in Immunology 22: 177-184

Ren X., Siegel R., Kim U., Roeder RG. (2011) Direct interactions of OCA-B and TFII-I regulate immunoglobulin heavy-chain gene transcription by facilitating enhancer-promoter communication. Molecular Cell 42: 342-355

Revilla-i-Domingo R., Bilic I., Vilagos B., Tagoh H., Ebert A., Tamir IM., Smeenk L., Trupke J., Sommer A., Jaritz M., Busslinger M. (2012) The B cell identity factor Pax5 regulates distinct transcriptional programs in early and late B lymphopoiesis. EMBO Journal in press

Ringrose L., Paro R. (2004) Epigenetic regulation of cellular memory by the Polycomb and Trithorax group proteins. Annual Reviews Genetics 38: 413-443

Robson EJD., He SJ., Eccles MR. (2006) A PANorama of PAX genes in cancer and development. Nature Reviews Cancer 6: 52-62

Rodríguez CI., Buchholz F., Galloway J., Sequerra R., Kasper J., Ayala R., Stewart AF., Dymecki SM. (2000) High-efficiency deleter mice show that FLPe is an alternative to Cre-loxP. Nature Genetics 25: 139-140

Roessler S., Györy I., Imhof S., Spivakov M., Williams RR., Busslinger M., Fisher AG., Grosschedl R. (2007) Distinct promoters mediate the regulation of Ebf1 gene expression by interleukin-7 and Pax5. Molecular Cell Biology 27: 579-594

Rolink AG., Nutt SL., Melchers F., Busslinger M. (1999) Long-term in vivo reconstitution of T-cell development by Pax5-deficient B-cell progenitors. Nature 401: 603-606

Samitas K., Lötvall J., Bossios A. (2010) B cells: from early development to regulating allergic diseases. Archivum Immunologiae et Therapiae Experimentalis 58: 209-225

Schaniel C., Gottar M., Roosnek E., Melchers F., Rolink AG. (2002a) Extensive in vivo self-renewal, long-term reconstitution capacity, and hematopoietic multipotency of Pax5-deficient precursor B-cell clones. Blood 99: 2760-2766

Schaniel C., Bruno L., Melchers F., Rolink AG. (2002b) Multiple hematopoietic cell lineages develop in vivo from transplanted Pax5-deficient pre-B I-cell clones. Blood 99: 472-478

Schebesta A., McManus S., Salvagiotto G., Delogu A., Busslinger GA., Busslinger M. (2007) Transcription factor Pax5 activates the chromatin of key genes involved in B cell signaling, adhesion, migration, and immune function. Immunity 27: 49-63

Schubart DB., Rolink A., Kosco-Vilbois MH., Botteri F., Matthias P. (1996) B-cell-specific coactivator OBF-1/OCA-B/Bob1 required for immune response and germinal centre formation. Nature 383: 538-542

Schweitzer BL., Huang KJ., Kamath MB., Emelyanov AV., Birshtein BK., DeKoter RP. (2006) Spi-C has opposing effects to PU.1 on gene expression in progenitor B cells. Journal of Immunology 177: 2195-2207

Seo W., Ikawa T., Kawamoto H., Taniuchi I. (2012) Runx1-Cbfβ facilitates early B lymphocyte development by regulating expression of Ebf1. Journal of Experimental Medicine, in print

Shinkai Y., Rathbun G., Lam KP., Oltz EM., Stewart V., Mendelsohn M., Charron J., Datta M., Young F., Stall AM., et al. (1992) RAG-2-deficient mice lack mature lymphocytes owing to inability to initiate V(D)J rearrangement. Cell 68: 855-867

Shaffer AL., Lin KI., Kuo TC., Yu X., Hurt EM., Rosenwald A., Giltnane JM., Yang L., Zhao H., Calame K., Staudt LM. (2002) Blimp-1 orchestrates plasma cell differentiation by extinguishing the mature B cell gene expression program. Immunity 17: 51-62

Sims RJ., Reinberg D. (2008) Is there a code embedded in proteins that is based on post-translational modifications? Nature Reviews Molecular Cell Biology 9: 1- 6

Smale ST. (2010) Pioneer factors in embryonic stem cells and differentiation. Current Opinion in Genetics and Development, 20: 519-526

Smith EM., Gisler R., Sigvardsson M. (2002) Cloning and characterization of a promoter flanking the early B cell factor (EBF) gene indicates roles for E-proteins and autoregulation in the control of EBF expression. Journal of Immunology 169: 261-270

Sokalski KM., Li SK., Welch I., Cadieux-Pitre HA., Gruca MR., DeKoter RP. (2011) Deletion of genes encoding PU.1 and Spi-B in B cells impairs differentiation and induces pre-B cell acute lymphoblastic leukemia. Blood 118: 2801-2808

Souabni A., Cobaleda C., Schebesta M., Busslinger M. (2002) Pax5 promotes B lymphopoiesis and blocks T cell development by repressing Notch1. Immunity 17: 781-793

Spivakov M., Fisher AG. (2007) Epigenetic signatures of stem-cell identity. Nature Reviews Genetics 8: 263-271

Stephan RP., Sanders VM., Witte PL. (1996) Stage-specific alterations in murine B lymphopoiesis with age. International Immunology 8: 509-518

Sugiyama T., Kohara H. Noda M., Nagasawa T. (2006) Maintenance of the Hematopoietic Stem Cell Pool by CXCL12-CXCR4 Chemokine Signaling in Bone Marrow Stromal Cell Niches. Immunity 25: 977-988

Sun XH., Baltimore D. (1991) An inhibitory domain of E12 transcription factor prevents DNA binding in E12 homodimers but not in E12 heterodimers. Cell 64: 459-470

Tarlinton D. (2006) B cell memory: are subsets necessary? Nature Reviews Immunology 6: 785-790

Teitell MA. (2003) OCA-B regulation of B-cell development and function.
Trends in Immunology 24: 546-553

Thal MA., Carvalho TL., He T., Kim HG., Gao H., Hagman J., Klug CA. (2009) Ebf1-mediated down-regulation of Id2 and Id3 is essential for specification of the B cell lineage. PNAS 106: 552-55

Treiber T., Mandel EM., Pott S., György I., Firner S., Liu ET., Grosschedl R. (2010) Early B cell factor 1 regulates B cell gene networks by activation, repression, and transcription- independent poising of chromatin. Immunity 32: 714-725

Tucker KL., Wang Y., Dausman J., Jaenisch R. (1997) A transgenic mouse strain expressing four drug-selectable marker genes. Nucleic Acids Research 25: 3745-3746

Turner BM. (2007) Defining an epigenetic code. Nature Cell Biology 9: 2-6

Urbánek P., Wang ZQ., Fetka I., Wagner EF., Busslinger M. (1994) Complete block of early B cell differentiation and altered patterning of the posterior midbrain in mice lacking Pax5/BSAP. Cell 79: 901-912

Vilagos B., Hoffmann M., Souabni A., Sun Q., Werner B., Medvedovic J., Bilic I., Minnich M., Axelsson E., Jaritz M., Busslinger M. (2012) Essential role of EBF1 in the generation and function of distinct mature B cell types. Journal of Experimental Medicine 209: 775-792

Visel A., Rubin EM., Pennacchio LA. (2009a) Genomic views of distant-acting enhancers. Nature 461:199-205

Visel A., Blow MJ., Li Z., Zhang T., Akiyama JA., Holt A., Plajzer-Frick I., Shoukry M., Wright C., Chen F., Afzal V., Ren B., Rubin EM., Pennacchio LA. (2009b) ChIP-seq accurately predicts tissue-specific activity of enhancers. Nature 457: 854-858

Walter K., Bonifer C., Tagoh H. (2008) Stem cell-specific epigenetic priming and B cell-specific transcriptional activation at the mouse Cd19 locus. Blood 112: 1673-1682

Wei GH., Badis G., Berger MF., Kivioja T., Palin K., Enge M., Bonke M., Jolma A., Varjosalo M., Gehrke AR., Yan J., Talukder S., Turunen M., Taipale M., Stunnenberg HG., Ukkonen E., Hughes TR., Bulyk ML., Taipale J. (2010) Genome-wide analysis of ETS-family DNA-binding in vitro and in vivo. EMBO Journal 29: 2147-2160

Weishaupt H., Sigvardsson M., Attema JL. (2010) Epigenetic chromatin states uniquely define the developmental plasticity of murine hematopoietic stem cells. Blood 115: 247-256

Weissman IL. (2000) Stem Cells: Units of Development, Units of Regeneration, and Units in Evolution. Cell 100: 157-168

Welinder E., Mansson R., Mercer EM., Bryder D., Sigvardsson M., Murre C. (2011) The transcription factors E2A and HEB act in concert to induce the expression of FOXO1 in the common lymphoid progenitor. PNAS 108: 17402-17407

Weth O., Renkawitz R. (2011) CTCF function is modulated by neighboring DNA binding factors. Biochemistry and Cell Biology, 89: 459-468

Wilson A. and Trumpp A. (2006) Bone marrow haematopoietic-stem-cell niches. Nature Reviews Immunology 6: 93-106

Xu HE., Rould MA., Xu W., Epstein JA., Maas RL., Pabo CO. (1999) Crystal structure of the human Pax6 paired domain-DNA complex reveals specific roles for the linker region and carboxy-terminal subdomain in DNA binding. Genes and Development 13: 1263-1275

Ye W., Bouchard M., Stone D., Liu X., Vella F., Lee J., Nakamura H., Ang SL., Busslinger M., Rosenthal A. (2001) Distinct regulators control the expression of the mid-hindbrain organizer signal FGF8. Nature Neuroscience 4: 1175-1181

Zandi S., Bryder D., Sigvardsson M. (2010) Load and lock: the molecular mechanisms of B-lymphocyte commitment. Immunological Reviews 238: 47-62

Zandi S., Mansson R., Tsapogas P., Zetterblad J., Bryder D., Sigvardsson M. (2008) EBF1 is essential for B-lineage priming and establishment of a transcription factor network in common lymphoid progenitors. Journal of Immunology 181: 3364-3372

Zhou VW., Goren A., Bernstein BE. (2011) Charting histone modifications and the functional organization of mammalian genomes. Nature Reviews Genetics 12: 7-18

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

CONTRIBUTIONS

The AP5 ES cell line described and used in this study was generated by Dr. Anton Wutz and maintained by Abdallah Souabni, at the IMP.

The genome-wide data presented in this study comes from the Busslinger laboratory database, unless stated otherwise.

All other experiments presented were performed by Anca-Sarmiza Tigan.

"Ich habe mich bemüht, sämtliche Inhaber der Bildrechte ausfindig zu machen und ihre Zustimmung zur Verwendung der Bilder in dieser Arbeit eingeholt. Sollte dennoch eine Urheberrechtsverletzung bekannt werden, ersuche ich um Meldung bei mir".

ACKNOWLEDGEMENTS

This thesis, like everything I have done in life, is dedicated to my parents.

The work presented in this thesis has been carried out at the Institute of Molecular Pathology (IMP) in Vienna.
First, I am grateful to Meinrad Busslinger for scientific guidance and discussions, support for the completion of this study, and for everything that I have learned during the past four years while working on my PhD project in his laboratory.

Throughout my PhD I received guidance and advice from two persons, without whose constant support I could not have succeeded: Leonie Ringrose and Ivan Bilic. Thank you for believing in me even when I did not, and for teaching me about the kind of person I want to be!

I am also grateful to Thomas Decker for his advice during committee meetings; to Robert Heinen (Ringrose lab), and Martin Leeb (and the Wutz laboratory) for sharing Southern expertise and knowledge; to Georg Busslinger (Peters lab) and Daniel Schramek (Penninger lab) for sharing reagents and knowledge for AdenoCre infection experiments; to Christian Theussl and Jacek Wojciechowski for ES cell injections; and to the entire IMP/IMBA Services.

I would also like to thank all members of the Busslinger laboratory, past and present, for helpful discussions, advice, and support at various time points during my stay in the lab: Abdallah, Alex, Anja, Babsi, Bojan, Erika, Heike, Hiromi, Ivan, Jasna, Leonie, Mareike, Markus, Martina, Qiong, Roger, Shane, Sinan, Stephen, Tanja, and Thorsten. I learned a lot from all of you!

Last but not least, a huge thank you to Sinan who has stayed by my side, encouraged and supported me all the way through.

And finally, thank you to all my friends (especially Mark for scientific discussions) for still being my friends at the end of these four years.

Curriculum Vitae

Personal Details

Name: Anca-Sarmiza Tigan, M.Sc.

Date of birth: October 16th, 1984

Place of birth: Cluj-Napoca, Romania

Nationality: Romanian

email address: tigan@imp.ac.at

Education:

Since 07/2008	**International PhD program at the Vienna Biocenter** Research group of Prof. Dr. Meinrad Busslinger, IMP
09/2007 – 07/2008	**"Aurel Vlaicu" University – Arad, Romania** Post-graduate level specialization in Pedagogy
09/2006 – 05/2008	**Jacobs University Bremen - Bremen, Germany** Master of Science degree in Biological Recognition *Thesis*: Peptide-induced conformational change of MHC class I
09/2003 – 05/2006	**Jacobs University Bremen – Bremen, Germany** Bachelor of Science degrees in Biochemistry & Cell Biology and Chemistry
09/1999 – 06/2003	"**Moise Nicoara" National College – Arad, Romania** Secondary school (top of generation, rank 1 out of 135)

Internships:

07/2007 – 08/2007	**Institut Curie, U653 Immunité et Cancer – Paris, France** Internship on antigen crosspresentation with Dr. Clotilde Thery
06/2005 – 08/2005	**Cancer Sciences Division, Univ. of Southampton UK** Internship with Prof. Dr. Tim Elliott on MHC class I presentation
01/2005	**Institut für Virologie, Uni Bremen – Bremen, Germany** Internship with Dr. Andreas Dotzauer on Hepatitis A infection
06/2004 – 07/2004	**S.C. "FARMEC" S.A. – Cluj-Napoca, Romania** Internship in cosmetics industry with Ing. Chem. Sofia Irimie

Employment:

09/2004 – 05/2008 **Jacobs University Bremen – Bremen, Germany**
Teaching assistant for Biochemistry and Molecular Biology,
Biological Chemistry, Immunology - theoretical and practical courses

09/2003 – 05/2004 **Jacobs University Bremen – Bremen, Germany**
Laboratory Assistant in the research group of Prof. Dr. Ulrich Kortz

Publications:

Cobalt Containing Silicotungstate Sandwich Dimer
[{Co3(B-β-SiW9O33(OH))(B-β-SiW8O29(OH)2}2]22
Bassem S. Bassil, Ulrich Kortz, **Anca S. Tigan**, Juan M. Clemente-Juan, Bineta Keita, Pedro de Oliveira, and Louis Nadjo
Inorganic Chemistry (2005); 44 (25): 9360-9368

Posters and Conferences:

11/2009 **Vienna BioCenter PhD Symposium – Vienna, Austria**
"Android and Eve – Bridging Biology, Medicine, and Technology"
Member of the Organizing Committee

05/2009 **7th B Cell Forum – Salzburg, Austria** (PhD student participant)

02/2007 **Bremen Life Sciences Meeting at Schloss Etelsen, Germany**
Poster: Thermal Stability Studies of Suboptimal MHC I Complexes

09/2006 **Ph.D. Student Symposium Horizons in Molecular Biology - Goettingen, Germany**
Participation stipend for the poster: The Carboxyl Terminus of the Peptide Determines ER Exit of MHC Class I Molecules

06/2006 **Lindau Meeting of Nobel Laureates and Students - Germany**

Awards:

03/2006 – 07/2008 **Studienstiftung des Deutschen Volkes**

09/2004 – 05/2006 **Member of President's List, Jacobs University Bremen**

09/2003 – 05/2006 **Merit-based scholarship from Jacobs University Bremen**

06/2003 **Honor Diploma from "Moise Nicoara" Foundation - Arad, Romania**

02/2003; 02/2002 **Third prize at county phase of Romanian National Olympiad of Chemistry**