FOG-1, a transcriptional regulator within the haematopoietic system

Inauguraldissertation

zur

Erlangung der Würde eines Doktors der Philosophie
vorgelegt der Philosophisch-Naturwissenschaftlichen
Fakultät der Universität Basel

von

Aude VERSAVEL

aus der Frankreich

Basel, 2013

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät auf Antrag von :

Prof. Dr. Patrick Matthias
Prof. Dr. Antonius Rolink

Basel, den 12. November 2013

Prof. Dr. Jörg Schibler (Dekan)

Summary

Friend of GATA-1 (FOG-1) is a member of the friend of GATA (FOG) family of proteins, which consists of large multitype zinc finger cofactors that bind to the amino zinc finger of GATA transcription factors and modulate their activity. FOG-1 also interacts with the C-terminal binding protein (CtBP), mainly known as a corepressor and the nucleosome remodelling and histone deacetylase repressive (NuRD) complex ; thus, integrating FOG-1 into the transcription factor and chromatin modifier networks. Remarkably, the protein activates or represses gene transcription by facilitating binding of GATA factors to DNA, recruiting chromatin remodelling complexes, or by stabilizing tissue-specific chromatin loops. Physical interaction between FOG and GATA proteins *in vivo* is essential for the development of a broad array of tissues, reflecting the overlapping expression patterns of these factors. Notably, within the haematopoietic system, FOG-1 is absent in most of the myeloid lineages ; it is expressed at high level in multipotent progenitors, erythroid and megakaryocytic cells, low level in lymphoid and haematopoietic stem cells. The cofactor is essential for differentiation of the erythroid and megakaryocytic lineages, notably by interacting with GATA-1. FOG-1 also plays a role in the T-lineage by repressing GATA-3 dependent induction of Th2 development. Interestingly, overexpression of FOG-1 in avian eosinophils, which do not normally express FOG-1, reprograms these differentiated cells into multipotent cells. To study FOG-1 in mammals, we used a novel transgenic mouse model strategy which we had designed to generate mice with conditional overexpression of FOG-1. Our work with enforced expression of FOG-1 in the whole murine haematopoietic system led to a reduction in the number of circulating eosinophils, confirming and extending to mammals the previously reported role of FOG-1 in repressing this lineage development. Strikingly, we have identified the expression of FOG-1 in early B lymphocytes, but not in late developmental stages such as mature B cells and plasma cells. Moreover, FOG-1 function had never been described in the B-lineage, where GATA factors are not expressed. Therefore, we were intrigued by both the regulated expression of FOG-1 during B cell development and its molecular mechanism of action in the absence of GATA factors. Thus, we generated transgenic mice in which FOG-1 expression was enforced at a physiologically relevant level in the B lymphoid system : in mature B cells and from early B cell stages. We found that sustained FOG-1 expression in mature and late B cells did not affect their development or function, contrary to our expected hypothesis. Although the mice overexpressing FOG-1 from early B cell lineages showed only a weak phenotype, we extensively studied FOG-1 partners in early B cell stages. Indeed, describing FOG-1 molecular mechanism of action in the absence of GATA factors is a question that warrant further investigation. We notably found FOG-1 in complex with Ikaros, a transcription factor well described as crucial for B cell development. The cofactor was also found associated to the CtBP and NuRD epigenetic complexes in B cell lines.

Table of contents

Chapter 1 : Introduction

The human organism is a complex set of numerous organs, which are themselves a set of tissues ; all composed of billions of cells in various stages of differentiation. Its homeostasis, of which any deregulation can lead to pathology, depends on appropriate cellular development and communication. Thus, understanding precisely how cell differentiation occurs, notably inside the haematopoietic system is a crucial question from both a fundamental and applied point of view.

Since many years, scientists have developed projects implying model organisms (for instance, *Mus musculus*) in order to study pathological situations. Especially, genetically modified mice with loss- or gain-of-function for a particular gene have been established. Besides, in human and mouse organisms, the haematopoietic tree constitutes a useful system for dissecting cell differentiation. Indeed, thanks to the expression of given cell surface markers, one can isolate step after step the differentiating cells by flow cytometry. The haematopoietic cell differentiation is under the control of various transcription and epigenetic factors ; whose expression can be genetically manipulated in an entire organism as precised before.

During the course of my thesis, I focused on the role of the transcriptional cofactor Friend of GATA-1 (FOG-1) in mammals. Different mice lines were generated and analyzed : some overexpressed FOG-1 in the whole haematopoietic system, which is presented at the beginning of the introduction. Others were knockdown for or overexpressed FOG-1 specifically in the B cell lineage. Inside the haematopoietic tree, T and B lineages share common characteristics which are exposed further in the introduction. Patrick Matthias' laboratory aims at understanding the molecular interplay of transcription and epigenetic factors during B cell development. These networks within B cells are described in a second part of the introduction. Finally, a statement on what is already known in the litterature on FOG-1 function in the haematopoietic system constitutes the third part of the introduction.

1. The haematopoietic system

The haematopoietic system or, in other words, the bodily system of organs and tissues involved in the formation and functioning of blood elements is the site of haematopoiesis. It includes the bone marrow, spleen, thymus, tonsils and lymph nodes.

Stem cells can divide either to produce more stem cells (self-renewal) or to replenish differentiated cells. The critical stem cell in haematopoiesis is the pluripotential haematopoietic stem cell, often referred to simply as the haematopoietic stem cell (HSC). All mature haematopoietic cells are generated from HSCs by stepwise differentiation into intermediate precursors with increasingly restricted differentiation potential (Fig. 1 (Orkin and Zon, 2008)).

Within the haematopoietic system, the myeloid system is defined in opposition to the lymphoid system or adaptive immune system. The following parts review the constituents of the haematopoietic system of the mouse,

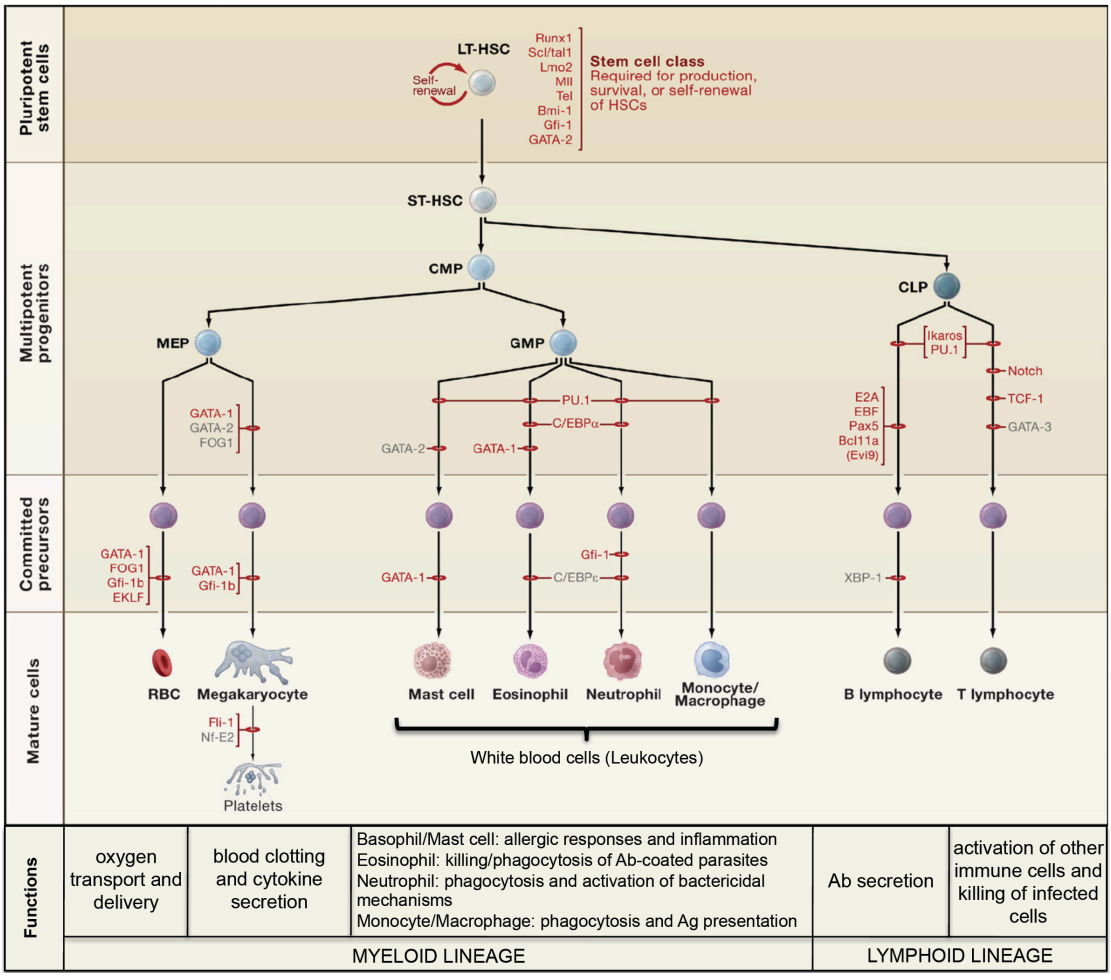


Figure 1 : Differentiation of haematopoietic stem cells (HSCs) to the different myeloid and lymphoid lineages. In definitive haematopoiesis, long-term HSCs (LT-HSCs) give rise to short-term HSCs (ST-HSCs). ST-HSCs produce common myeloid progenitors (CMPs) and common lymphoid progenitors (CLPs). CMPs give rise to megakaryocyte/erythroid progenitors (MEPs) and granulocyte/macrophage progenitors (GMPs), whereas CLPs are the source of committed precursors of B and T lymphocytes. GMPs give rise to the committed precursors of mast cells, eosinophils, neutrophils, and macrophages. The transcription factors that govern haematopoiesis in the mammals are depicted. The stages at which haematopoietic development is blocked in the absence of a given transcription factor, as determined through gene knockouts, are indicated by red loops. The factors depicted in red have been associated with oncogenesis. Factors in grey have not yet been found translocated or mutated in human/mouse haematologic malignancies. Adapted from Orkin *et al.*, 2008.

whose one of the physiological function is the defense against pathogens (antigens, Ag) invading the body.

1.1. The myeloid system

After bifurcation between the myeloid and lymphoid lineage, the common myeloid progenitor cell, or CMP produces the megakaryocyte/erythroid progenitor cell (MEP), which can generate either the red blood cell (RBC or erythrocyte) lineage or the platelet lineage. The CMP can also give rise to the granulocyte/monocyte progenitor cell (GMP), which generates the basophils/mast cells, eosinophils, neutrophils, and monocytes/macrophages. Eventually, MEPs and GMPs produce precursor cells that can divide but produce only one type of cell in addition to renewing themselves.

The erythroid precursor cell is a committed stem cell that can form only RBCs. Its immediate progeny is capable of responding to the hormone erythropoietin to produce the first recognizable differentiated member of the erythrocyte lineage, the proerythroblast, a RBC precursor. This cell begins to make globin (Krantz and Goldwasser, 1965a, b). As the proerythroblast matures, it becomes an erythroblast, synthesizing enormous amounts of hemoglobin. The mammalian erythroblast eventually expels its nucleus and becomes a reticulocyte. Although reticulocytes, lacking a nucleus, can no longer synthesize globin messenger RNA (mRNA), they can translate existing messengers into globins. The final stage of differentiation of the erythroid lineage is the erythrocyte, or mature RBC. In this cell, no division, RNA synthesis, or protein synthesis takes place. The DNA of the erythrocyte condenses and is not transcribed anymore. Amphibians, fishes and birds retain the functionless nucleus ; mammals extrude it from the cell. The erythrocyte then leaves the bone marrow and enter the circulation, where it delivers oxygen to the body tissues (Gilbert, 2010).

During megakaryocytopoiesis, a primitive megakaryocyte progenitor cell differentiate and produce colony-forming unit-megakaryocytes (CFU-MKs), which are then giving rise to megakaryocytes (MKs). These latter are multinucleate cells (containing up to 20 nuclei) from which platelets are derived. Platelets are small, colorless, irregularly shaped nonnucleated cells. The primary role of platelets is to mediate blood clotting, but activated platelets also secrete cytokines that influence the functions and migration of other leukocytes during an inflammatory response. The cytokines are produced from performed mRNAs that are translated upon platelet activation (Saunders, 2006).

Macrophages are resident in almost all tissues and are the mature form of monocytes, which circulate in the blood and continually migrate into tissues, where they differentiate. Together, monocytes and macrophages make up one of the three types of phagocytes in the immune system : the others are the granulocytes (the collective term for the white blood cells called neutrophils, eosinophils, and basophils) and the dentritic cells (DCs). Macrophages are relatively long-lived cells and perform several different functions throughout the innate immune response and the subsequent adaptive immune response. One is to engulf and kill invading microorganisms. In this phagocytic role, they represent an important first defense in innate

immunity and also dispose of pathogens and infected cells targeted by an adaptive immune response. Both monocytes and macrophages are phagocytic, but, as most infections occur in the tissues, it is primarily macrophages that perform this important protective function. An additional and crucial role of macrophages is to orchestrate immune responses : they help induce inflammation, which is a prerequisite to a successful immune response, and they secrete signalling proteins that activate other immune-system cells and recruit them into an immune response. As well as their specialized role in the immune system, macrophages act as general scavenger cells in the body, clearing dead cells and cell debris.

In addition to the committed precursor of monocytes and macrophages, GMPs give rise to the committed precursors of neutrophils, eosinophils, basophils and mast cells.

The granulocytes are so called because they have densely staining granules in their cytoplasm. There are three types of granulocytes –neutrophils, eosinophils and basophils-, which are distinguished by the different staining properties of the granules. In comparison with macrophages, they are all relatively short-lived, surviving for only a few days and are produced in increased numbers during immune responses. At that time, they leave the blood to migrate to sites of infection or inflammation. The phagocytic neutrophils are the most numerous and most important cells in innate immune responses : they take up a variety of microorganisms by phagocytosis and efficiently destroy them in intracellular vesicles using degradative enzymes and other antimicrobial substances stored in their cytoplasmic granules. The protective functions of eosinophils and basophils are less well understood. Their granules contain a variety of enzymes and toxic proteins, which are released when the cell is activated. Eosinophils and basophils are thought to be important chiefly in defense against parasites, which are too large to be ingested by macrophages or neutrophils, but their main medical importance is their involvement in allergic inflammatory reactions, in which their effects are damaging rather than protective.

Mast cells, whose blood-borne precursors are not well defined, differentiate in the tissues. Although best known for their role in orchestrating allergic responses, they are believed to play a part in protecting the internal surfaces of the body against pathogens and are involved in the response to parasitic worms. They have large granules in their cytoplasm that are released when the mast cell is activated ; these help induce inflammation (Kenneth M. Murphy, 2008a).

Apart from cells belonging to the myeloid lineage, able to carry out innate immune responses, cells of the lymphoid lineage are in charge of the adaptive immune responses. Cells considered to be of the lymphoid lineage include T and B lymphocytes, natural killer (NK) cells, and natural killer T (NKT) cells. T- and B-lymphopoiesis are described in the following section.

1.2. The lymphoid system

The major effector cells of the adaptive immune responses are the T and B lymphocytes that arise from a common lymphoid progenitor cell (CLP), detectable early in embryonic development, in the fetal liver and during

adulthood, in the bone marrow. Both T and B cells bear highly diverse receptors on their cell surface that are able to recognize a vast diversity of Ag. Both share an ordered rearrangement of their Ag receptor genes, obligatory expression of a surrogate, invariant component of their pre-receptors, and ligand-dependent positive and negative selection of their mature Ag receptor repertoires. Each cell is genetically programmed to encode unique cell surface receptors specific for a particular Ag. The T cell Ag receptor (TCR) is a membrane-bound molecule related to immunoglobulin (Ig), which recognizes a complex of a peptide fragment bound to a molecule specialized in Ag presentation, a major histocompatibility complex (MHC) molecule. The B cell Ag receptor (BCR) is the membrane-bound form of an antibody (Ab or Ig) that would be secreted upon activation of the cell. Each B cell presents only one type of BCR on its surface.

1.2.1. T-lymphopoiesis

1.2.1.1. Main stages of T cell development in the thymus

As mentioned before, T cells originate from HSCs and develop in the thymus, through a process that can be separated into three broad steps (Fig. 2 (Carpenter and Bosselut)). The first spans thymic colonization to T cell commitment ; the second proceeds to the divergence of $\alpha\beta$ and $\gamma\delta$ lineages. The third step sees $\alpha\beta$ and $\gamma\delta$ lineage cells complete their differentiation and acquire immunological properties and, in some cases, effector functions. In more details, HSCs can give rise to early T cell progenitors (ETP), that migrate to the thymus, the main site of the T cell development. The loss of multipotency that defines T commitment is a gradual process. It occurs in double negative (DN) thymocytes, which do not express neither CD4 nor CD8 – a subset that is itself separated into four sequential phenotypic stages on the basis of CD44 and CD25 expression (DN1 : $CD44^+CD25^-$, DN2 : $CD44^+CD25^+$, DN3 : $CD44^-CD25^+$ and DN4 : $CD44^-CD25^-$) – and is not complete before the DN2 stage. ETPs are part of a heterogeneous DN1 subset that includes both subsequent intermediates in the T cell differentiation pathway and cells belonging to DC, myeloid (M) or NK lineages. The DN2 and DN3 subsets are themselves divided into two stages, based on expression of the receptor c-Kit and of CD27, respectively. $\alpha\beta$ lineage cells emerge from β-selection as double-positive (DP) and are the precursors of conventional CD4 and CD8 cells. The developmental progression of these cells is notably marked by a positive selection consisting in the rescue from programmed cell death of DP thymocytes as a consequence of TCR$\alpha\beta$ interaction with self peptide-MHC complexes expressed by the thymic epithelium (or with other MHC or MHC-like molecules). Most $\alpha\beta$ lineage cells undergo a MHC-induced selection which results in the differentiation of ‘single-positive’ (SP) thymocytes that express either CD4 or CD8, two molecules that contribute to TCR recognition of MHC class II or MHC class I, respectively ; such SP thymocytes are the direct precursors of mature T cells. Whereas the vast majority of $\alpha\beta$ T cells are restricted by MHC class I or MHC class II molecules, as detailed before, small subsets are selected on other MHC or MHC-like molecules. The most abundant of these, invariant natural killer T (iNKT) cells

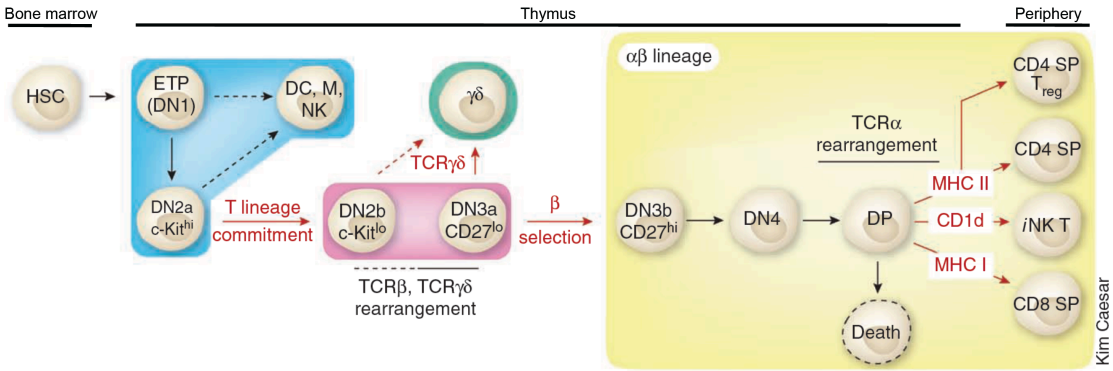


Figure 2 : Overview of T cell development, depicting thymic developmental stages. Expression of CD4 and CD8 separates $CD4^-CD8^-$ double-negative (DN), $CD4^+CD8^+$ double-positive (DP) and single-positive (SP) cells, whereas the expression of CD44 and CD25 defines four DN subsets : $CD44^+CD25^-$ (DN1), $CD44^+CD25^+$ (DN2), $CD44^-CD25^+$ (DN3) and $CD44^-CD25^-$ (DN4). The earliest precursors that enter the thymus derive from haematopoietic stem cells (HSCs) in the bone marrow and are known as early T cell progenitors (ETP). They are part of a heterogeneous DN1 subset that includes both subsequent intermediates in the T cell differentiation pathway and cells belonging to dendritic (DC), myeloid (M) or natural killer (NK) lineages. The DN2 and DN3 subsets are themselves divided into two stages, based on expression of the receptor c-Kit and of CD27, respectively. Critical checkpoints are shown in red. Cell subsets are grouped according to key developmental steps : early uncommitted progenitors (blue), T cell committed progenitors before the separation of αβ and γδ lineages (pink) and cells commited to the αβ (yellow) or γδ (green) lineages. Dashed lines indicate minor or speculative differentiation routes. Adapted from Carpenter and Bosselut, 2010.

recognize lipid-bound MHC class I-like CD1d molecules. In addition to the positive selection, a negative one consists in the elimination of self-reactive cells, which is the cause of extensive cell death in the thymus. The main stages of thymic T cell development, which takes place in the thymus are depicted in the figure 2 and reviewed extensively in Carpenter and Bosselut (Carpenter and Bosselut).

The thymus is structured in several lobules, each containing discrete cortical (outer) and medullary (central) regions. As shown in figure 3, thymus-seeding progenitors enter through blood vessels at the cortico-medullary junction (CMJ). Distinct thymocyte subsets occupy discrete regions of the thymus, indicating that differentiation is coupled with migration through the stroma in response to chemokine signals. Progenitors travel from the CMJ through the cortex to the subcapsular zone (SCZ) during the early DN stages. Therefore, the cortex consists mainly of immature thymocytes embedded in a network of branched cortical thymic epithelial cell (cTEC) and intermittent macrophages. In the outer cortical area, DN3 cells undergo TCR β rearrangement and β-selection, and the resulting immature single-positive (ISP) intermediate cells and then DP cells migrate back through the cortex. Positively selected SP thymocytes enter the medulla, where effective negative selection occurs. $\gamma\delta$ and $\alpha\beta$ T cell lineages diverge at the DN3 stage. Mature thymocytes surrounded by medullary thymic epithelial cell (mTEC), macrophages and medullary DC make up the medulla (Ciofani and Zuniga-Pflucker, 2007).

1.2.1.1.1. Classical pathway of $\alpha\beta$ T cell development

The majority of lymphoid progenitor cells present in the thymus follow the main pathway of T cell development (Fig. 3 (Ciofani and Zuniga-Pflucker, 2007)), resulting in the generation of a repertoire of T cells that carry a unique TCR, consisting of an alpha and a beta chain, on their surface, and a co-receptor molecule that can be either CD4 or CD8.

However, there are several branch points during early thymocyte development that can divert the cells into other lineages.

1.2.1.1.2. iNKT cells

The iNKT cells constitute a lymphocyte subpopulation that is abundant in the thymus, spleen, liver, and bone marrow and is also present in the lung. iNKT cells express surface markers that are characteristic of both natural killer cells (such as NK1.1 and CD122) and conventional T cells (such as TCRs). They have been implicated in the regulation of immune and autoimmune responses. The majority of iNKT cells utilizes a restricted TCR repertoire that recognizes glycolipids in association with the non-polymorphic MHC-like molecule CD1d. These cells carry TCRs made of a nearly invariant TCRα chain ($V_{\alpha}14J_{\alpha}18$ in mice) associated to a small set of TCRβ partners ($V_{\beta}8.2$, $V_{\beta}7$ or $V_{\beta}2$)(Carpenter and Bosselut). The development and maintenance of NKT cells are extensively reviewed in Godfrey *et al*. (Godfrey et al.).

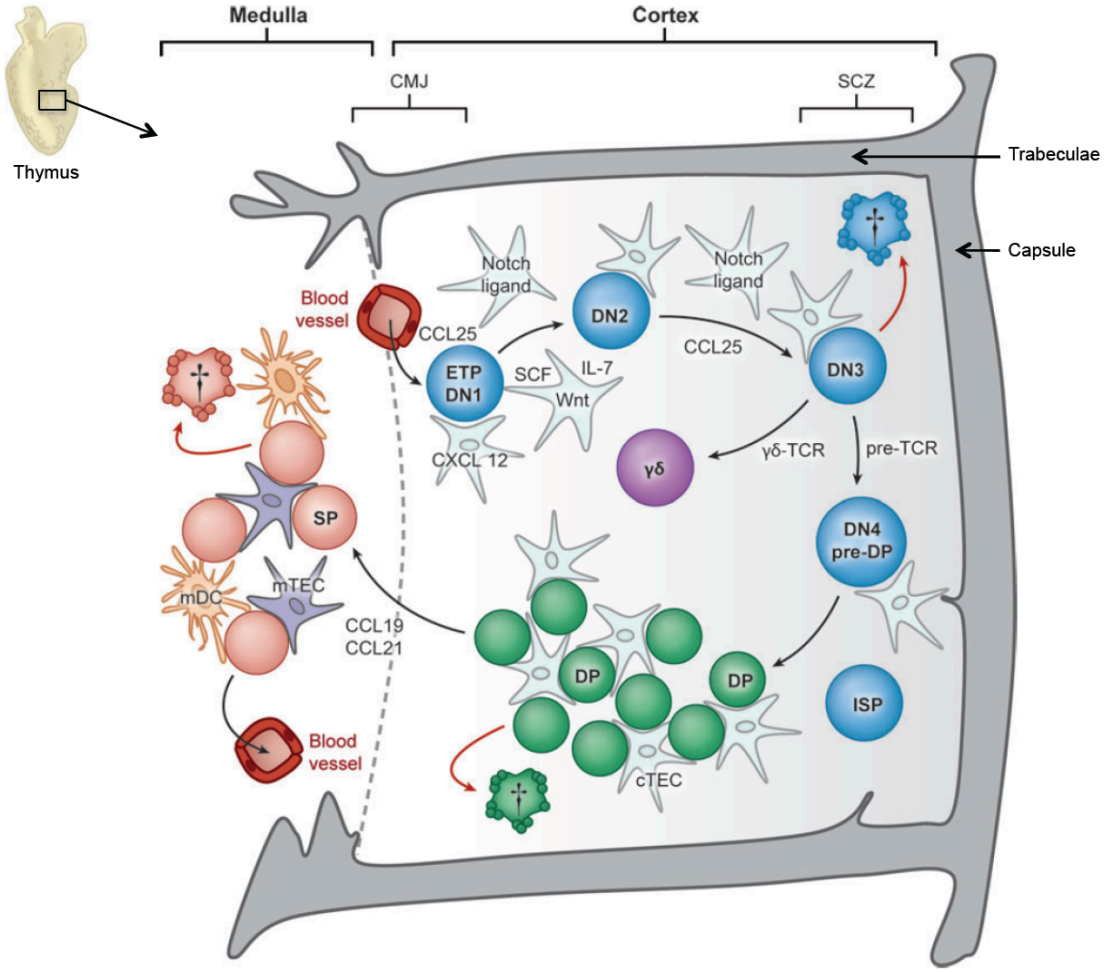


Figure 3 : Schematic view of T lymphopoiesis within the thymus. The thymus is divided into two histologically defined regions, the cortex and the medulla, each of which contains specialized thymic epithelial cell (TEC) subsets. Thymus-seeding progenitors enter through blood vessels at the cortico-medullary junction (CMJ) and undergo an ordered process of development. Distinct thymocyte subsets occupy discrete regions of the thymus, indicating that differentiation is coupled with migration through the stroma in response to chemokine signals. Progenitors travel from the CMJ through the cortex to the subcapsular zone (SCZ) during the early double-negative (DN) stages. Here, DN3 cells undergo T cell receptor (TCR) β rearrangement and β-selection, and the resulting immature single-positive (ISP) intermediate cells and then double-positive (DP) cells migrate back through the cortex. Positively selected single-positive (SP) thymocytes enter the medulla, where effective negative selection occurs. γδ and αβ T cell lineages diverge at the DN3 stage. Developmental stages are depicted by different cell colors, black arrows indicate developmental progression and/or migration, and red arrows indicate outcomes leading to cell death. Daggers mark cell populations undergoing apoptosis or death by neglect at critical developmental checkpoints. Thymus-derived factors vital to the early stages of T-lymphopoiesis are denoted on or near TEC representations. The thymus is a bilobed organ, composed of many nodular lobules separated by a thin connective septum, shown as a gray enclosing structure. cTEC, cortical TEC ; ETP, early thymic progenitor ; IL-7, interleukin-7 ; mDC, medullary dendritic cell ; mTEC, medullary TEC ; SCF, stem cell factor. Adapted from Ciofani and Zùniga-Pflücker, 2007.

1.2.1.1.3. γδ T cells

Another cell lineage, the γδ T cells, express TCRs γδ instead of TCRs αβ, and make up 1-5% of the T cells in the thymus. γδ T cells are also the first T cells that appear during embryonic mouse development. αβ T cells appear only a few days after the earliest γδ T cells and rapidly become the predominant thymocyte population (Havran and Allison, 1988).

Following the definitive commitment to the T cell lineage in the thymus, T cell precursors express the IL-2 receptor α chain (CD25) and begin to rearrange and express their TCRδ, γ, and β genes. Cells that successfully rearrange TCRγ and TCRδ genes express a TCRγδ and can proceed along the γδ lineage pathway. Similarly, cells that successfully rearrange their TCRβ gene express a pre-TCR (formed by association of TCRβ with the invariant pTα chain) and are able to differentiate along the αβ lineage (Kang and Raulet, 1997).

Current models of lineage commitment at the DN stage distinguish between two possible roles for the TCR : instructing commitment or reinforcing a prior commitment event (Hayday et al., 1999). Unlike in αβ cells, there is no 'pre-TCRγδ' ; the DN3 checkpoint assesses signalling by complete TCRγδ complexes. Moreover, γδ lineage differentiation does not depend on Notch signalling. A current view for αβ versus γδ lineage commitment is that strong signals promote γδ lineage choice and that γδ TCR signals are stronger than pre-TCR signals. Nevertheless, the αβ versus γδ cell type divergence has been a contentious issue (Carpenter and Bosselut).

γδ T cells are often considered to be a more primitive cell type, since they are emerging earlier than αβ T cells in phylogeny and ontogeny, and might therefore define an interface between the innate and adaptive immune systems (Pennington et al., 2003). At the functional level, they effectively have a role as a bridge between the innate and the adaptive immune systems, based on the interpretation that γδ T cells use their TCR as a pattern recognition receptor (Holtmeier and Kabelitz, 2005). In addition, γδ T cells rapidly produce a variety of cytokines and usually exert potent cytotoxic activity, also towards many tumor cells.

1.2.1.2. T cell function

T cells are able to distinguish cells that are harboring pathogens by recognizing peptide fragments of pathogen-derived proteins displayed by the cell in a complex with the MHC molecule. There are two different classes of MHC, class I and II (Fig. 4 (Klein and Sato, 2000)), that are specialized in the presentation of peptides derived from different cellular compartments. MHC class I molecules present peptides from the cytosol and are recognized by $CD8^+$ killer T cells, while MHC class II molecules display peptides generated in vesicles and are recognized by $CD4^+$ helper T cells (Fig. 5 (Klein and Sato, 2000)). Thus, the two different functional subsets of T cells can become activated and contribute to the host defense against different types of pathogens. On the one hand, proteins from viruses or other intracellular parasites are displayed on MHC class I molecules and the infected cell can be

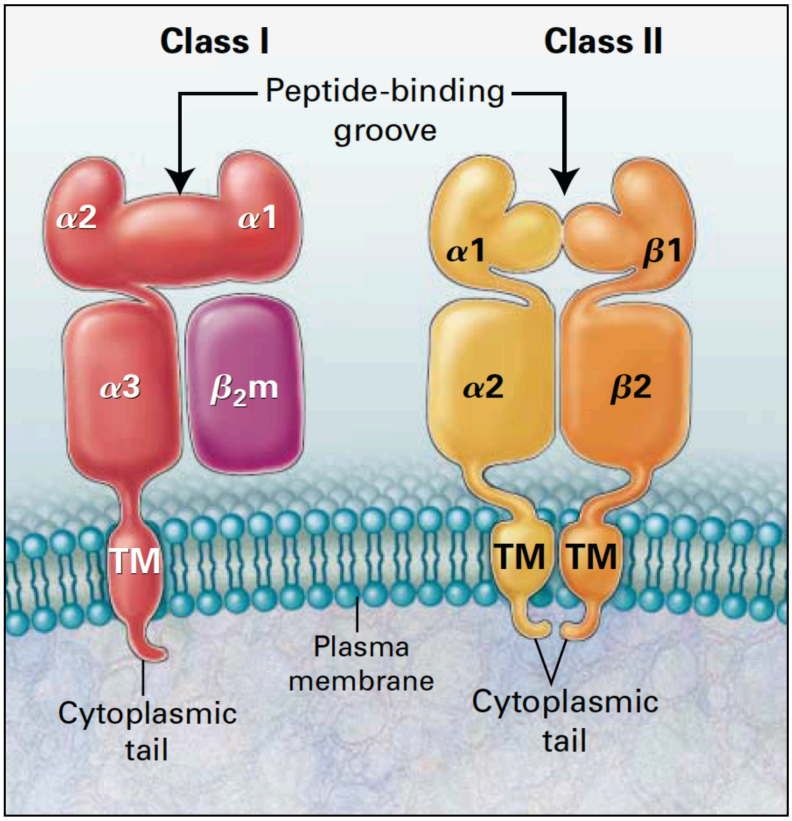


Figure 4 : Structure of HLA Class I and Class II molecules. Beta$_2$-microglobulin (β_2m) is the light chain of the class I molecule. The α chain of the class I molecule has two peptide-binding domains (α1 and α2), an immunoglobulin-like domain (α3), the transmembrane region (TM), and the cytoplasmic tail. Each of the class II α and β chains has four domains : the peptide-binding domain (α1 or β1), the immunoglobulin-like domain (α2 or β2), the transmembrane region, and the cytoplasmic tail. Taken from Klein and Sato, 2000.

killed by a killer T cell ($CD8^+$ T cell) that recognizes the Ag as non-self. On the other hand, extracellular pathogens and toxins taken up in vesicles, usually by phagocytic cells such as macrophages, are presented in MHC class II complexes to helper T cells ($CD4^+$ T cells) that activate B cells to secrete specific Abs that help to eliminate the pathogens (Kenneth M. Murphy, 2008b).

Each individual TCR is specific for a particular MHC-peptide complex or, in other words, MHC-restricted for Ag recognition. Thymocytes with TCRs that are able to recognize self-MHC molecules and function in self-MHC-restricted responses to foreign Ags are positively selected for survival in the thymus. The positive selection relies critically on cTEC that make close contacts with the T cells which express CD4 and CD8 co-receptors on their cell surface. This process depends on the engagement of both the Ag receptor and co-receptor with an MHC molecule, and determines the survival of SP cells that express only the appropriate co-receptor. Positively selected SP thymocytes are ready for export to the periphery. The function of the CD4 or CD8 co-receptors of the mature T cells lies in their distinct abilities to bind invariant sites on MHC class II or class I molecules, respectively.

In the thymus, negative selection occurs to cells that encounter their corresponding Ag (a self-peptide) on a self-MHC of an Ag presenting cell (APC). The most important types of APCs are the bone marrow-derived DCs and macrophages. These are professional APC types that also activate mature T cells in peripheral lymphoid tissues. Each individual thymocyte must undergo negative selection in the thymus to enable the immune system to respond to virtually any foreign Ags but avoid harmful responses to self in the periphery. The education process of thymocytes could, in principle, be accomplished either by recognizing one or a set of « common » peptides, which could represent all peripheral self-Ags in the thymus, or by « seing » the actual peripheral, tissue-restricted or age-dependent, self-peptides expressed in the thymus, or by both. Discovery of the autoimmune regulator (Aire) gene provided the first direct evidence for the latter hypothesis (Jiang and Chess, 2009).

Once they have completed their development in the thymus, T cells enter the bloodstream and are carried by the circulation. On reaching a peripheral lymphoid organ, they leave the blood to migrate through the lymphoid tissue, returning again to the bloodstream to recirculate between blood and peripheral lymphoid tissue until they encounter their specific Ag.

The activation of naive T cells in response to Ag, and their subsequent proliferation and differentiation, constitutes a primary immune response. At the same time as providing armed effector T cells, this response generates immunological memory, which gives protection from subsequent challenge by the same pathogen. The generation of memory long-lived T cells that give an accelerated response to Ag and the generation of effector T cells present similarities and differences (reviewed in Seder and Ahmed (Seder and Ahmed, 2003)). Memory T cells differ in several ways from naive T cells, but like naive T cells, they are quiescent and require activation by APCs with co-stimulatory activity in order to regenerate effector T cells.

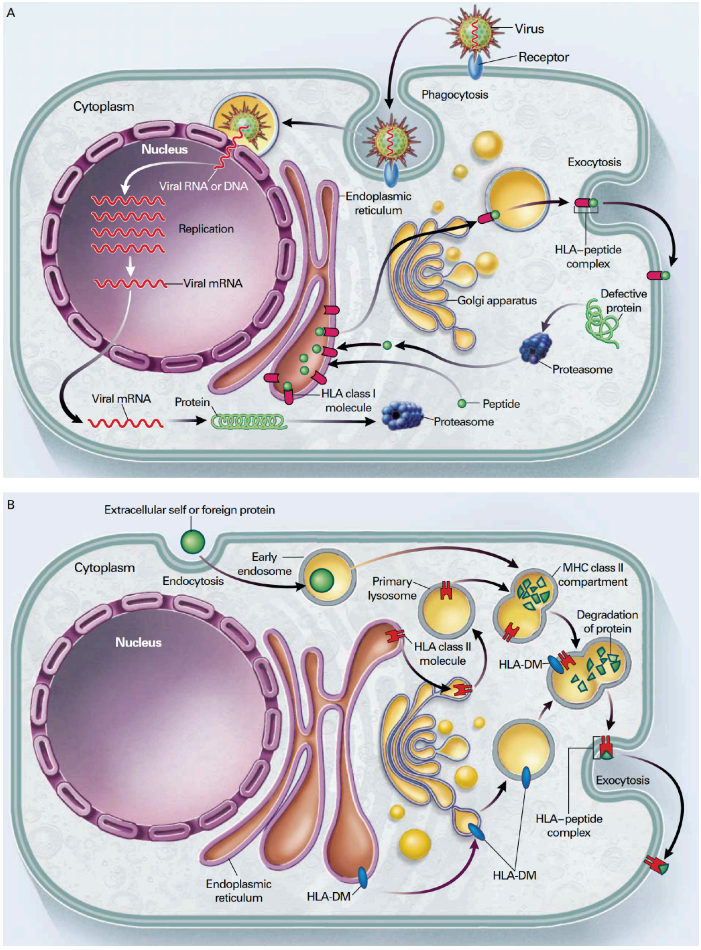


Figure 5 : Antigen processing. Panel A shows the principal pathways of generating peptides for loading onto HLA class I molecules. Worn-out or defective proteins in the cytosol are degraded into peptides in proteasomes. Selected peptides are then transported into the endoplasmic reticulum, where they are loaded onto newly synthesized class I molecules. The HLA-peptide complexes are exported by way of the Golgi apparatus to the surface of the cell. In tissues infected with a virus, viral particles are taken up by cells and uncoated. The viral DNA or RNA enters the nucleus and replicates within it. The viral messenger (mRNA) then enters the cytosol and is transcribed into proteins. Some of the proteins are subsequently degraded in proteasomes, and the peptides are delivered into the endoplasmic reticulum, where they are loaded onto class I molecules for export to the surface of the cell. Panel B shows the processing of extracellular proteins. Self or foreign proteins are taken up by endocytosis (or phagocytosis) and sequestered into endosomes. Class II molecules synthesized in the endoplasmic reticulum are delivered by way of the Golgi apparatus into primary lysosomes, which fuse with the early endosomes to form the major-histocompatibility-complex (MHC) class II compartment. Enzymes brought into this compartment by the lysosomes degrade the engulfed proteins into peptides. HLA-DM molecules synthesized in the endoplasmic reticulum and delivered into the MHC class II compartment by transport vesicles help load the peptides onto the class II molecules. The HLA-peptide complexes are then exported to the surface of the cell. Taken from Klein and Sato, 2000.

1.2.2. B-lymphopoiesis

1.2.2.1. Main stages of B cell development in the bone marrow and spleen

The bone marrow is the site of early events in B cell maturation. Indeed, in the bone marrow, B lymphocytes go through successive stages of differentiation : the Pro B cell (B cell Progenitor), the Pre B cell (B cell Precursor) and the immature B cell. They become mature B cells in the spleen.

The spleen is the major site of immune responses to blood-borne Ags. The lymphocyte-rich regions of the spleen, called the white pulp, are organized around branches of central arteries (Fig. 6 (Mebius and Kraal, 2005)). The architecture of the white pulp is analogous to the organization of lymph nodes, with segregated T and B cell zones. In mouse spleen, the central arteries are surrounded by cuffs of lymphocytes, most of which are T cells. Because of their anatomic location, morphologists call these T cell zones periarteriolar lymphoid sheaths (PALS). Several smaller branches of each central arteriole pass through the PALSs and drain into a vascular sinus. B cell-rich follicles occupy the space between the periarteriolar sheath and the marginal sinus. Outside the marginal sinus is a distinct region called the marginal zone (MZ), which forms the outer boundary of the white pulp and is populated by B cells and specialized macrophages. The B cells in the MZ are functionally distinct from follicular B cells and are known as marginal zone B (MZ B) cells. The segregation of T lymphocytes in the PALSs, and B cells in follicles and MZs is a highly regulated process, dependent on the production of different cytokines and chemokines by the stromal cells in these different areas.

Some of the arteriolar branches of the splenic artery ultimately end in extensive vascular sinusoids, distinct from the marginal sinuses, scattered among which are large numbers of erythrocytes, macrophages, DCs, sparse lymphocytes and plasma cells ; these constitute the red pulp. The plasma cells, generated in peripheral lymphoid tissues as a consequence of antigenic stimulation of B cells migrate to the bone marrow, where they may live and continue to produce Abs for many years (Abul K. Abbas, 2010).

1.2.2.1.1. Development of immature B cells in the bone marrow

Just like the aforementioned T lymphocytes, B cells arise from pluripotent progenitors in the fetal liver or in the bone marrow of adult mice. They start their maturation in these tissues and complete it in the secondary lymphoid organs, such as the spleen and the lymph nodes. Commitment to the B cell lineage is followed by rearrangement and expression of the B cell receptor genes : the Ig heavy and light chain genes. The process of B cell differentiation involves first the rearrangement of the Ig heavy chain genes (*Igh*), followed by a burst of proliferative clonal expansion (as a consequence of signalling through the Pre B Cell Receptor or Pre-BCR), and then, the rearrangement of the Ig light chain loci at the Pre B cell stage. The Pre-BCR is composed of the heavy chain and the surrogate light chains λ5 and V-preB.

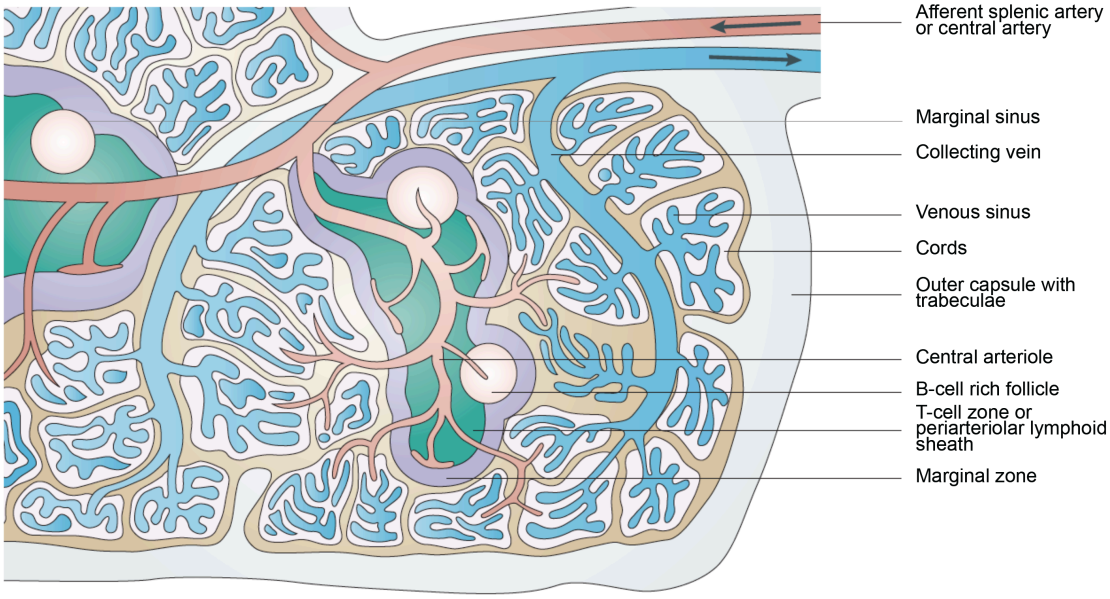


Figure 6 : Structure of the spleen. The afferent splenic artery branches into central arterioles, which are sheathed by white-pulp areas ; these white-pulp areas consist of the T-cell zone (also known as the periarteriolar lymphoid sheath, PALS), arterioles and B-cell follicles. The arterioles end in cords in the red pulp, from where the blood runs into venous sinuses, which collect into the efferent splenic vein. The larger arteries and veins run together in connective-tissue trabeculae, which are continuous with the capsule that surrounds the spleen. Adapted from Mebius and Kraal, 2005.

Successful pairing of one of the κ or λ light chains with the heavy chain results in surface expression of Immunoglobulin M (IgM), the hallmark of immature B cell stage, at which cells are checked for auto-reactivity. Upon encounter with autoantigens that are recognized by this sIgM, the cells can downregulate the IgM surface expression and induce a further rearrangement of the Ig light chain gene by a process termed receptor editing. Cells that remain autoreactive undergo programmed cell death (apoptosis), but those that lose the autoreactivity through receptor editing can proceed normally in their development. IgM-expressing cells that were not negatively selected for recognition of self molecules can subsequently exit the bone marrow and complete their maturation in the spleen. Discrete stages of B cell development have been defined based on :

- the cell cycle status and cell size
- the status of Ig heavy (H) and light (L) chain rearrangement
- the expression of several products of genes on the cell membrane or in the cytoplasm (especially those of the rearrangement machinery : Rag-1, Rag-2, TdT and the Pre-BCR and BCR complex : IgH chain, (surrogate) IgL chain, Igα and Igβ chains). Figure 7 depicts these successive stages of Ag independent B cell development. In addition, it indicates the stages at which various transcription factors have been found to be important, usually on the basis of knockout studies.

1.2.2.1.2. T cell dependent (TD) immune responses

Immature B cells enter the spleen from the blood at the level of the MZ where they can be further checked for self-reactivity in the periphery and consequently enter an anergic state ; alternatively, they can mature further into terminally differentiated plasma cells. B cells that have freshly entered the spleen can be identified by their cell surface reaction with the monoclonal Ab 493 (Rolink et al., 1998), which presumably recognizes the cell surface marker AA4.1, which is the complement component C1q like receptor C1qRp (Rolink et al., 2002). Subsequently, the (naive) B cells cross the marginal sinus and migrate *via* the PALS to the lymphoid follicles, which are areas specialized in Ag retention and presentation. Upon encounter with T cell dependent Ag, such mature IgM^+IgD^+ B cells become activated and proliferate with T cell help. The T helper cells (Th) required for this activation are $CD4^+CD8^-$ cells that can be divided into two subpopulations according to their cytokine profile : activated Th1 cells secrete mainly IL-2 and IFN-γ, while activated Th2 cells secrete IL-4 and IL-5. Th cell get activated by binding of Ag processed and displayed by APCs. Upon activation, Th cells upregulate the expression of a number of accessory receptors and ligands (e.g. CD40L and CD28) that are essential for B cell activation in TD immune responses. The activated mature B cells then differentiate predominantly into low affinity Ab-forming cells (AFCs), which are short-lived plasmocytes that initially secrete IgM, but subsequently switch to secondary isotypes and die by apoptosis within two weeks of immunization. Some cells from the expanded B cell pool do not become AFCs but migrate into the adjacent follicles, which are subsequently transformed into germinal centers (GCs). GC B cells undergo iterative cycles of proliferation, somatic hypermutation and apoptosis in cooperation with antigen-primed T cells and follicular DCs, a process

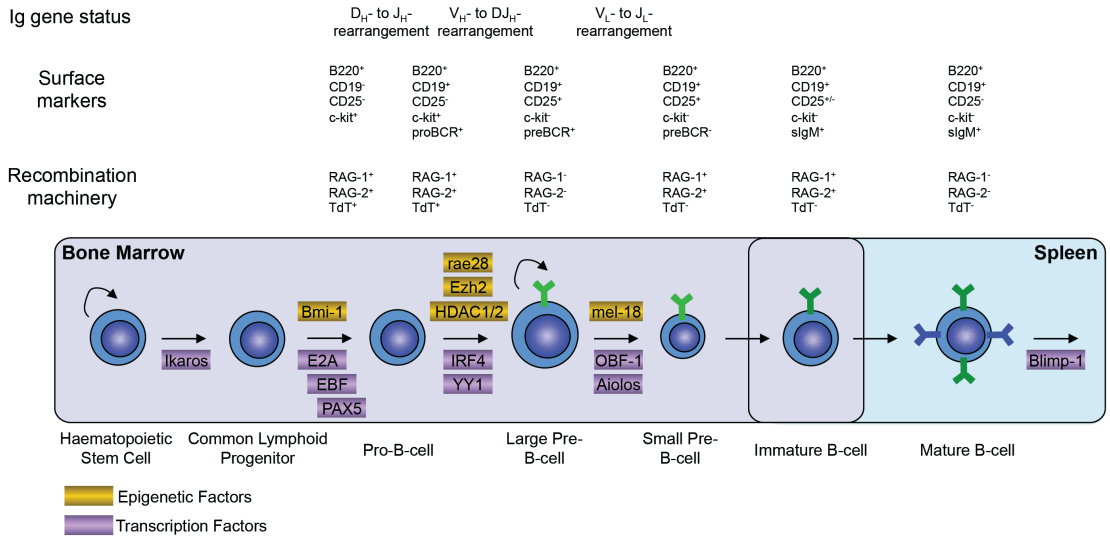


Figure 7 : Murine B cell development. Discrete stages of B cell development, based on cell size, on the rearrangement status of the immunoglobulin genes and the expression of genes on the cell membrane or in the cytoplasm, are shown. Immature B cells are depicted overlapping the bone marrow and the spleen because these cells are present in both organs. Some epigenetic and transcription factors identified to be crucial for B cell development are depicted. Adapted from Bartholdy and Matthias, 2004 (Bartholdy and Matthias, 2004).

resulting in the selection of B cell clones that make an Ab with high affinity for a cognate Ag. Additionally, Ig class switching occurs during the GC reaction. These steps expand, refine and diversify the repertoire of the early immune response to assure the long-term maintenance of protective immunity. Two types of terminally differentiated B cells eventually emerge from the GC reaction : nonsecreting, Ig surface-positive memory B cells and high affinity Ab-secreting plasmablasts that carry no Igs at their surface and that are the final mediators of the humoral immune response. Figure 8 (Oracki et al.) gives an overview over the B cell fate in the spleen.

1.2.2.1.3. T cell independent (TI) immune responses

As we have seen in the previous section, Ags recognized by the BCR usually require costimulatory signals from helper T cells to elicit a specific (TD) immune response. Some Ags however, including mostly bacterial cell wall components, such as LPS and certain sugars (e.g. Ficoll), can stimulate the B cells by themselves to proliferate and to secrete Abs independently of T cell help. These TI Ags do not lead to the formation of immunological memory, and the Abs that arise from a TI immune response usually do not undergo affinity maturation or class switching. The B cell subsets that are required for TI immune responses are preactivated splenic MZ B cells, as well as B-1 B cells which provide a bridge between the very early innate and the later appearing adaptive immune response by generating an initial wave of IgM producing plasmablasts during the first three days of a primary response to particulate bacterial Ags (Martin and Kearney, 2000; Martin et al., 2001).

1.2.2.1.3.1. B-1 B cells

B-1 cells constitute a subset of B cells that are dominant in the peritoneal and pleural cavities. Their precursors develop in the fetal liver and omentum. In adult mice, lymphoid organs contain precursors that can develop into B-1 cells when they express the appropriate BCR specificity and/or signalling threshold (Duan and Morel, 2006). B-1 cells recognize self components, as well as common bacterial Ags, and they secrete natural Abs important for innate immunity that tend to have low affinity and broad specificity - mostly self-reactive antibodies - (Mayer and Zaghouani, 1991; Shirai et al., 1991; Welner et al., 2008). B-1 cells make Ab responses mainly to polysaccharide Ags in a TI manner. They can be distinguished from the conventional B-2 B cells - which originate from bone marrow and are distributed to mucosal and systemic immune compartments for the continuous secretion of Abs with high affinity and fine specificity -, by their self-renewing capacity, which allows their size population to be kept constant and unique cell surface proteins. In contrast to other B cell populations, they are $B220^{lo}$, $IgM^{hi}IgD^{lo}CD23^{-}CD43^{+}MAC1^{+}$. The B-1a cell subset expresses CD5, but is otherwise almost indistinguishable from the B-1b cells (Kantor et al., 1992; Stall et al., 1992).

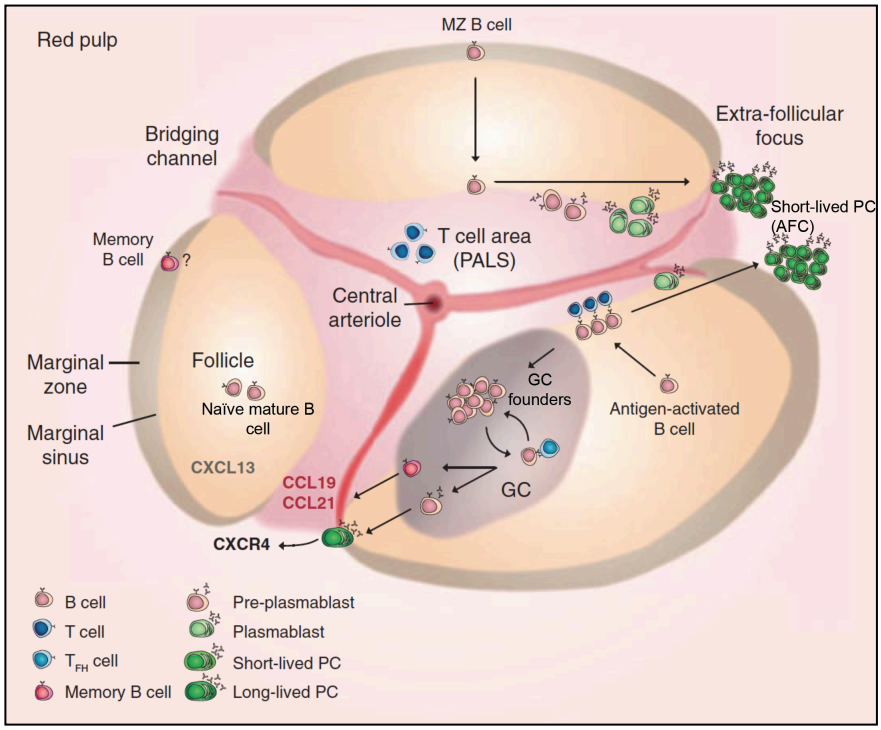


Figure 8 : Architecture of the splenic white pulp. The white pulp comprises a T cell area, called the periarteriolar lymphoid sheath (PALS), surrounding a central arteriole. Adjoining this are the follicles, which are home to the majority of naive B cells and are enveloped by the marginal zone (MZ). Partitioning of these areas is based on differential expression of the chemokines CXCL13 in the B cell follicle, CCL19 and CCL21 in the PALS, and CXCR4 in the red pulp. Antigen entering through blood can activate MZ B cells, causing their migration toward the PALS and subsequent migration, proliferation and differentiation to short-lived plasma cells in extra-follicular foci. Follicular B cells activated by antigen also traffic to the boundary of the T cell zone. From here, these cells may continue their migration and form an extra-follicular focus or re-enter follicles to form germinal centers. Taken from Oracki *et al.*, 2010.

1.2.2.1.3.2. MZ B cells

MZ B cells consist mainly of a large, mostly non-circulating subset of mature B cells that are localized in the MZ of the spleen. As mentioned before, they have an important function at the early stages of the immune response. This could be attributed to their observed lower activation threshold that triggers them more easily into proliferation or differentiation than immature or recirculating mature B cells.

A number of genes that are crucial for the generation or maintenance of the MZ B cell compartment have been identified by targeted disruption or overexpression studies. There are four broad categories of mutations in mice that result in defects in MZ B cells : mutations involving negative regulators of the BCR that result in enhanced BCR signalling ; mutations involving the BAFFR (B cell activating factor receptor) and the canonical NF-κB (Nuclear Factor-kappaB) pathway ; mutations in the Notch pathway ; and mutations involving signalling molecules that are linked to integrin or chemokine activation, which might be required for the post-migration retention of MZ B cells in the MZ.

Interestingly, in Aiolos-deficient mice, in which the strength of BCR signalling is increased, the numbers of splenic $IgM^{hi}IgD^{lo}CD23^{-}CD21^{hi}CD1d^{hi}$ MZ B cells were markedly decreased. The number of $IgM^{hi}IgD^{hi}CD23^{+}CD21^{hi}CD1d^{hi}$ B cells, which are now recognized as MZ precursor B cells, was also markedly decreased in these mice (Cariappa et al., 2001). This indicates that increased BCR signalling decreased the differentiation of maturing B cells into MZ B cells.

Otherwise, BAFF is a trimeric member of the tumour necrosis factor family, required for marginal B cell development (Batten et al., 2000; Schiemann et al., 2001). In addition, constitutively active IKKβ (Inhibitor of nuclear factor kappa-B kinase subunit beta, a component of the canonical NF-κB pathway) can correct the phenotype of BAFFR-deficient mice, resulting in the rescue of both follicular B cells and MZ B cells.

In addition, the requirement for Notch2 for the generation of both precursor and mature MZ B cells has led to some intriguing insights about the follicular versus MZ B cell transition. In the Notch signalling pathway, intracellular Notch, RBP-Jκ (recombination signal binding protein for immunoglobulin kappa J region) and Mastermind proteins form a ternary complex that activates Notch target genes. Knockout mice lacking Notch2 or RBP-Jκ in B cells, Mastermind-like 1 (MAML1) or the Notch ligand DL1 all have defects in MZ B cell development (Hozumi et al., 2004; Saito et al., 2003; Tanigaki et al., 2002; Wu et al., 2007).

Clearly, signals for commitment to a MZ B cell fate (through the BCR, the canonical NF-κB pathway and Notch2) dictate the acquisition of the ability of the committed cell to migrate, differentiate further, self-renew and be retained in the MZ. Several molecules contribute to the migration and/or retention of MZ B cells in the spleen. For instance, the sphingosine 1-phosphate receptor (S1PR1) participates in MZ B cell migration and retention *in vivo* and is crucial to overcome CXC-chemokine ligand 13 (CXCL13)-mediated attraction of B cells to the follicles (Cinamon et al., 2004). Otherwise, a mutation in the *Pik2* (also known as *Ptk2b*) gene, which encodes a tyrosine kinase that has been linked to integrin signalling, results in the

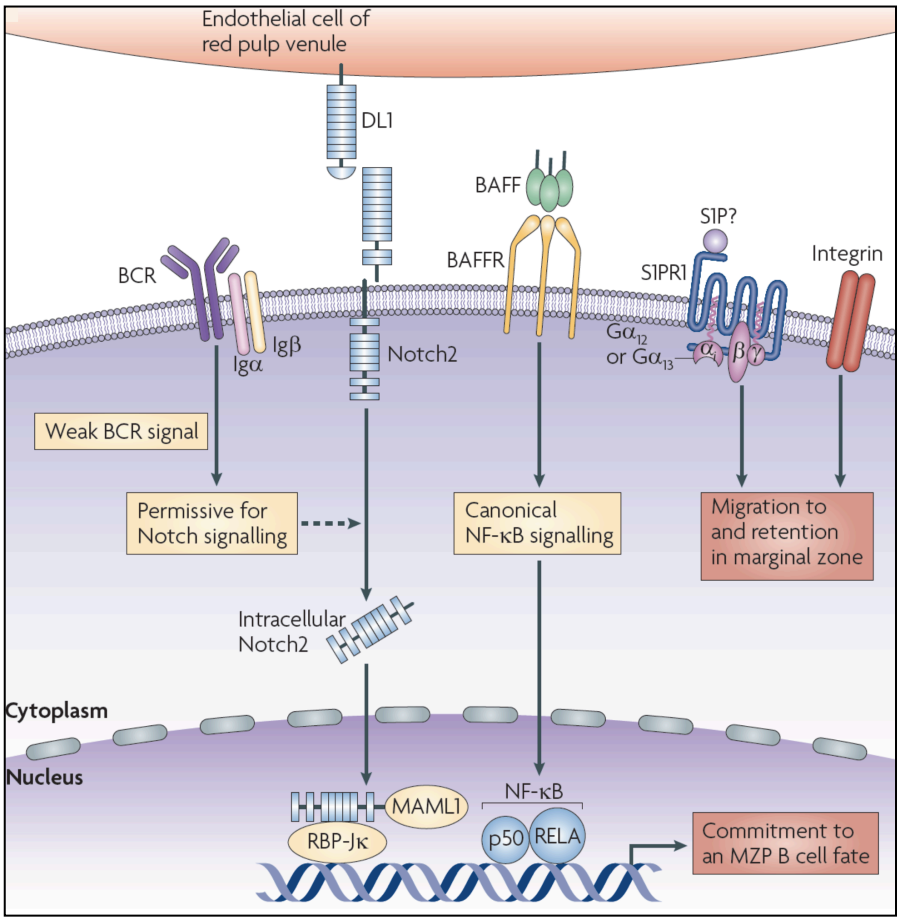


Figure 9 : Commitment to a marginal zone B cell fate. Although weak BCR signalling is the initial event driving commitment, acquisition of the marginal zone B cell phenotype also requires further signalling pathways. Signals resulting from the interaction of delta-like 1 (DL1) expressed by endothelial cells of red pulp venules with Notch2 on developing B cells lead to activation of the transcription factors Mastermind-like 1 (MAML1) and RBP-Jκ. Signals resulting from the interaction of BAFF with BAFFR activate NF-κB through the canonical pathway. A Gα12- or Gα13-coupled serpentine receptor, probably the sphingosine 1-phosphate receptor (S1PR1), and integrin signalling contribute to the processes of migration to and retention in the marginal zone, although the identity of the chemotactic factor involved remains to be definitively determined. Taken from Pillai and Cariappa, 2009.

absence of MZ B cells (Guinamard et al., 2000). In addition, there are certain mutations in genes also linked to integrin signalling that might have effects downstream of the BCR. The OBF-1 transcriptional coactivator lies downstream of the BCR but might also be required for the migration of B cells as the number of MZ B cells is decreased in the absence of OBF-1. However, the exact role of OBF-1 in B cell migration is not fully understood (Samardzic et al., 2002). Figure 9 summarizes the signals required for commitment to a marginal B cell fate (Pillai and Cariappa, 2009).

1.2.2.2. B cell function

The expression of Igs, either as cell surface receptors, or in secreted form as Ab molecules, is a unique feature of B cells. An Ig molecule consists of two identical set of heavy (H) and light (L) polypeptide chains held together by a combination of noncovalent bonds and covalent disulfide bonds (Fig. 10 (Kazatchkine and Kaveri, 2001)). Each polypeptide chain features one variable (V) and one constant (C) domain. The Ag is bound by both the V_H and V_L domains. The V region can be further divided into hypervariable regions, also called complementarity determining regions (CDRs), and more conserved framework regions (FR). The C region is characteristic for each distinct isotype group. There are five isotypes or main H chain classes, some of which have several subtypes : IgM, IgD, IgG, IgA and IgE. Abs of different isotypes operate in distinct functional activities. The corresponding H chains are denoted μ, δ, γ, α and ε, respectively. Initially, all immature B cells express the IgM isotype and may switch to a different isotype at later stages of B cell maturation.

1.2.2.2.1. Organization and transcription of the Ig loci

The functional gene segments that make up the Ig H and L chains are organized into three clusters, located on different chromosomes : the H chain genes, and the κ and λ L chain genes. Multiple copies of all of these gene segments exist in germline DNA (Fig. 11 (Schatz and Ji)). The two loci of L chain genes, κ and λ, are alternatively used, with species-specific preferences (κ to λ ratios of 2:1 in man and 20:1 in mouse), but apparently without functional differences.

In order to form the functional Ig molecules, the gene segments have to be assembled by somatic recombination. Two types of recombinatorial events take place at the Ig locus during B cell development : V(D)J recombination at both H and L loci, and class switch recombination (CSR) at the H locus.

1.2.2.2.1.1. V(D)J recombination, class switch recombination and recombination machinery

During the process of V(D)J recombination, the H chain rearrangement occurs first, starting with a rearrangement joining one of the D_H gene segments with one of the J_H gene segments. Subsequently, one V_H gene segment is joined to shape a VDJ-C_H gene. Rearrangement of the L chain follows, where one of the V_L gene segments is directly joined to a J_L gene segment, leading to a VJ-C_L κ or λ gene. In both cases, RNA splicing after

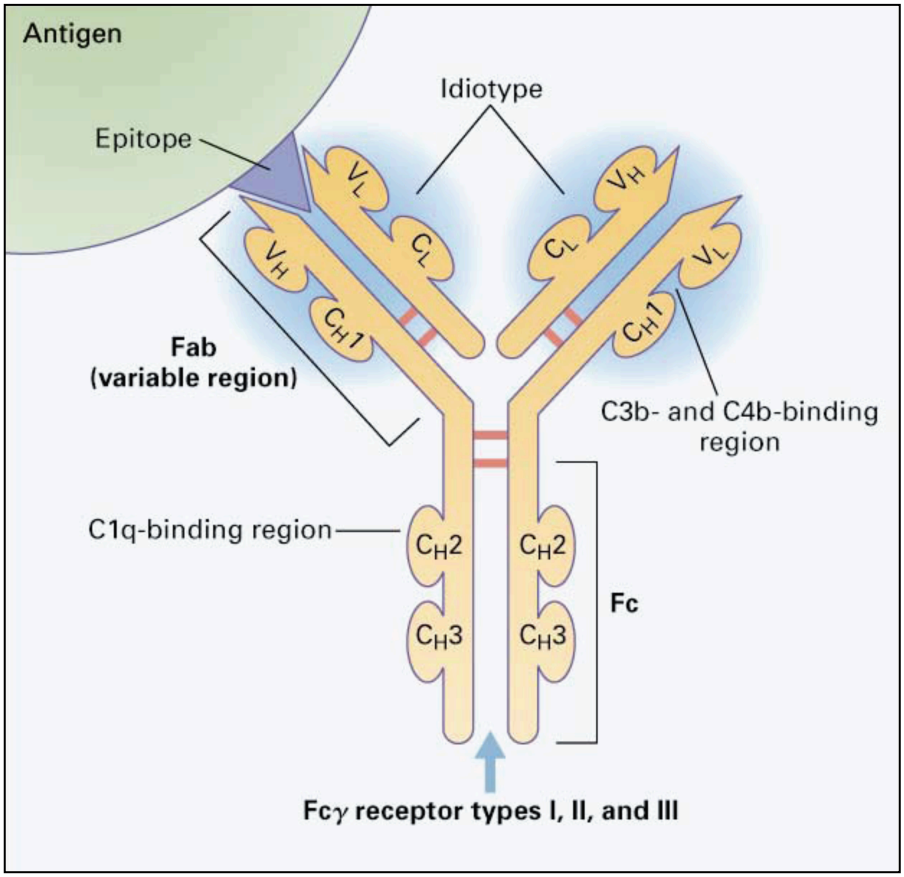


Figure 10 : The IgG molecule. The site of interactions between IgG and antigen (epitope) is shown, as are binding sites for C1q and activated C3b and C4b and sites of interaction between the heavy (H) chains of IgG and Fcγ receptor types I, II, and III. V denotes variable region, C constant region, and L light chain. Note that V_H and V_L indications have been inverted on the right of the scheme. Taken from Kazatchkine and Kaveri, 2001.

transcription joins the V(D)J segments to the closest downstream C region coding sequence. In immature B cells, this is Cμ or, if alternative splicing occurs during late phases of B cell development, Cδ. The C genes are organized in the following order in the mouse genome : 5'-V(D)J-Cμ-Cδ-Cγ3-Cγ1-Cγ2b-Cγ2a-Cε-Cα-3' (D'Eustachio et al., 1980).

The second type of recombinatorial event occuring at the Ig locus during B cell development is the CSR at the IgH locus. CSR is a deletional process that places a C gene other than Cμ or Cδ directly downstream of the rearranged VDJ segments, leading to the expression of other secondary Ig isotypes encoded by downstream C genes, such as IgG, IgA and IgE (Stavnezer, 1996). CSR occurs between highly repetitive G-rich switch (S) sequences with abundant palindromes that precede every C region, except for Cδ. This process is controlled by a serie of signals involving BCR and cytokine signalling, and in which B cell - T cell interactions play a critical role. The activation and targeting of CSR correlates with the capability of certain mitogens and cytokines to either induce or suppress germline transcription of specific C genes (Manis et al., 2002; Snapper et al., 1997).

Both recombinatorial events are thus intimately linked to transcription of the locus.

In both V(D)J rearrangements and CSR, chromosomal DNA double-strand breaks (DSB) are generated and differentially re-ligated, thereby deleting the interjacent genomic sequence. In the case of V(D)J rearrangements, this relies on highly conserved recombination signal sequences (RSS), short stretches of DNA present 3' of each V region, 5' and 3' of D regions and 5' of each J region. RSS always consist of a conserved heptamer and nonamer elements separated by a less well conserved spacer region of either 12 or 23 base pairs, which defines the 12RSS and 23RSS, respectively. V(D)J rearrangements occur only between gene segments located on the same chromosome and follow the 12-23 rule, a restriction due to a strong preference for recombination between a segment flanked by a 12mer-spaced RSS to one with a 23mer-spaced RSS. As both V_H and J_H segments are flanked by 23 bp spacers and the D_H segments are flanked by 12 bp spacers, only D_H to J_H segment joining and V_H to D_H segment joining are possible ; V_H to J_H segment joining is not possible (Eastman et al., 1996; Tonegawa, 1983).

V(D)J recombination is initiated by the RAG recombinase (referred to hereafter as RAG) – a protein complex consisting primarily of the proteins encoded by recombination activating gene 1 (*Rag-1*) and *Rag-2* -. RAG binds and cleaves the DNA at RSSs ; it catalyzes both the DSB and re-ligation reactions. DNA repair during V(D)J recombination is channelled into the non-homologous end joining (NHEJ) pathway by the RAG proteins. RAG1 is able to interact with the NHEJ factors Ku70 (also known as XRCC6) and Ku80 (also known as XRCC5) (Schatz and Ji). Ku70 and Ku80 form a complex with the DNA-dependent protein kinase catalytic subunit to form the full DNA-dependent protein kinase (DNA-PK).

Isotype switch recombination also requires DNA-PK, but can occur in the absence of *Rag-1* and *Rag-2* gene expression.

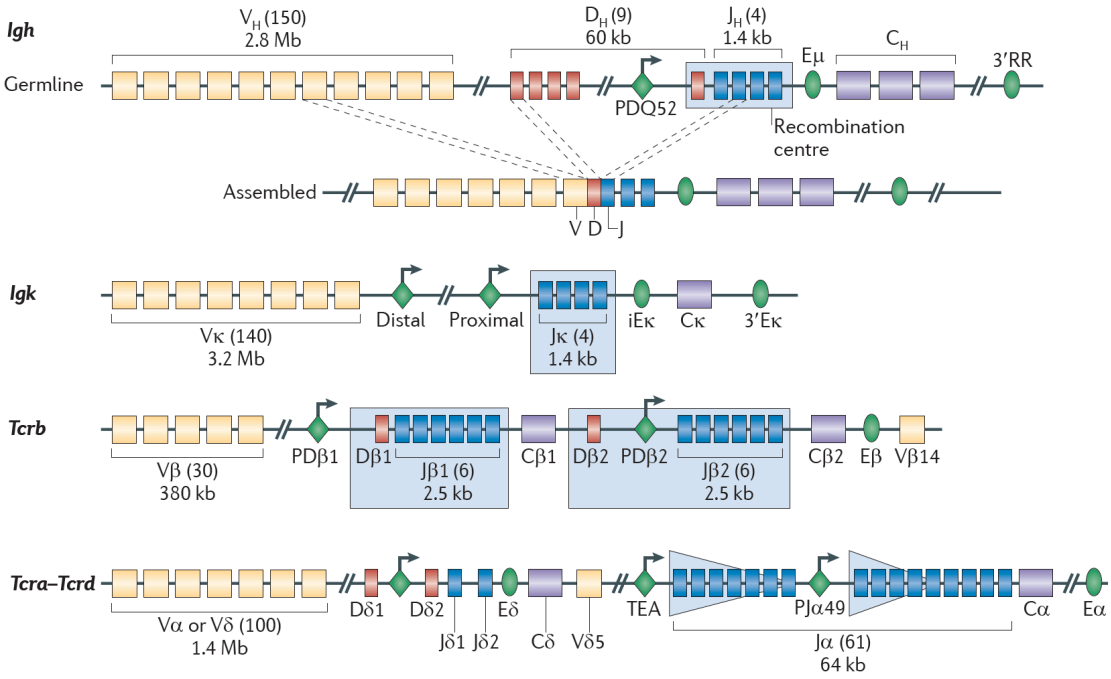


Figure 11 : The structure of antigen receptor genes. Schematic diagrams of the four mouse antigen receptor loci for which the binding patterns of the recombination activating gene (RAG) proteins have been reported. Variable (V), diversity (D) and joining (J) gene segments are represented as yellow, red and blue rectangles, respectively, constant (C) regions are shown as purple rectangles, enhancer elements as green ovals and germline promoters associated with recombination centres as green diamonds with arrows. The promoters associated with each V gene segment and the recombination signal sequences (RSSs) are not shown. The approximate sizes of the regions within each locus are indicated, together with the approximate number of gene segments in each region (shown in parentheses). The recombination centres (blue shaded areas) are the regions that are bound by RAG1 and RAG2 when the loci are in their germline configurations. From top to bottom, the immunoglobulin heavy chain (*Igh*) locus in its germline and assembled configurations, the germline *Igk* locus, the germline T cell receptor β-chain (*Tcrb*) locus, the germline *Tcra-Tcrd* locus. The *Tcra* recombination centre is represented by two triangles to depict the 5' to 3' gradient of RAG binding observed downstream of the T early α (TEA) element and the increase observed downstream of the Jα49 germline promoter. Taken from Schatz and Ji, 2011.

1.2.2.2.1.2. Sterile and productive transcripts

Two types of Ig transcripts are expressed during B cell development : sterile germline and productive transcripts. Sterile transcripts result from Ig gene transcription preceding gene rearrangement and class switching and are not translated. It is generally believed that sterile transcripts alter the accessibility of the Ig loci and thereby influence the regulation of gene rearrangements. Sterile transcripts have been characterized from most of the C_H genes, as well as for Cκ and Cλ L chain genes.

Productive transcripts originate from the transcription of rearranged Ig genes and are translated into functional proteins.

1.2.2.2.1.3. Affinity maturation and AID

The efficiency of Ag elimination is also enhanced by affinity maturation, which is accomplished by a high number of point mutations in the V-region gene, coupled with selection of high-affinity APCs by limited amounts of Ag. Point mutations are introduced by two types of molecular mechanisms : non-templated somatic hypermutation (SHM) and gene conversion (Reynaud et al., 2003). The process of SHM can be divided into three phases : targeting, DNA recognition and cleavage, and repair. The first phase of targeting a nuclease to the Ig locus necessitates transcription (Jacobs and Bross, 2001). While some sequence specificity was reported, it remains unclear what is recognized by this nuclease (Michael et al., 2002). The cleavage mediated by the nuclease is part of the second phase leading to a DNA DSB (Bross et al., 2000; Papavasiliou and Schatz, 2000), which might be preceded by a single strand lesion (Kong and Maizels, 2001). In the final phase, the DSB is probably repaired into a mutation by a subset of error-prone polymerases (Gearhart and Wood, 2001; Papavasiliou and Schatz, 2002).

The activation-induced deaminase (AID) is an enzyme specifically expressed in GC B cells, which is required for CSR, SHM and gene conversion. The analyses of several AID mutants indicate a requirement for class-switch-specific cofactors. Nevertheless, the exact function of AID is still a matter of considerable debate. Two mechanisms : RNA-editing and DNA-deamination have been suggested and extensively reviewed in Honjo *et al*. (Honjo et al., 2005). A model integrating AID-dependent mechanisms that diversify Ig genes in SHM, gene conversion and CSR pathways is shown in figure 12 (Papavasiliou and Schatz, 2002).

1.2.2.2.1.4. Allelic exclusion, receptor editing

For each Ig locus, there are two alleles that can undergo gene rearrangement, and this increases the chance of a successful event. After the first allele has successfully rearranged, the recombination machinery is quickly downregulated, so that generally only one correctly rearranged IgH and IgL locus is found in any B cell. This phenomenon is called allelic exclusion. The other allele remains in germline configuration, is partially rearranged or has completed an unproductive recombination.

Another rescue mechanism that prevents excessive unproductive rearrangements and also that avoids the recognition of self Ags with high

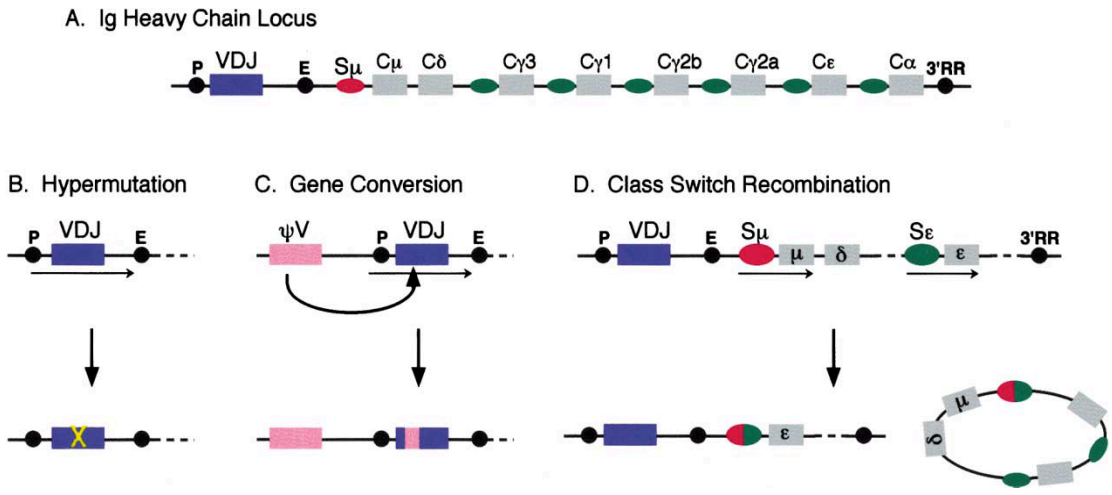


Figure 12 : AID-dependent mechanisms that diversify Ig genes. A. Schematic diagram of the Ig heavy chain locus, with the variable (VDJ) and constant (C) region exons represented as blue and gray rectangles, respectively, switch regions as red and green ovals, and promoter (P), enhancer (E), and 3' regulatory regions (3'RR) regions as black circles. Not to scale. B. Somatic hypermutation causes point mutations (x) in the vicinity of the V exon. Thin arrows represent transcription in B, C, and D. C. Gene conversion involves the transfer of sequence information from a pseudogene (ψV) into the variable region exon. D. Class switch recombination involves looping out and deletion of DNA between two switch regions (in this case Sμ and Sε), thereby swapping the constant region of the expressed heavy chain. Taken from Papavasiliou and Schatz, 2002.

avidity is the process of receptor editing. Unlike the H chain genes, repeated L chain gene rearrangements of unused V and J gene segments can occur, thereby leading to production of a new IgL chain and increasing the chance of the B cell precursors to generate progeny (the immature B cells) that bears intact IgM molecules.

The entire sequencing of the human genome has revealed that it was composed of a surprisingly small number of genes. Among them, only a limited number were encoding the V, D, and J regions ; which was not sufficient to explain the 10^9 clones of mature lymphocytes of the human organism. Indeed, several sophisticated mechanisms are existing in mammals so as to generate the diversity of the Ab repertoire. Many V region gene segments exist in the genome of an individual, providing a hereditable source of diversity. Random recombination of separate V, D, and J gene segments during V(D)J recombination provide additional diversity. Variability of the junctions between the gene segments is increased by the random insertion of P- and N-nucleotides and by deletion of nucleotides at the ends of some coding sequences. The association of different L and H chain V regions adds diversity. For increasing the diversity of the Ag-binding site, in the periphery, mature B lymphocytes undergo additional DNA alterations induced by Ag stimulation. In the periphery, activated B lymphocytes proliferate vigorously in GCs of lymphoid follicles, in which two different types of genetic alterations, SHM and CSR take place.

1.2.2.2.2. Generation of memory B and plasma cells

Iterative cycles of proliferation, SHM and apoptosis in the GC result in the selection of B cells that make an Ab with high affinity for cognate Ag. In addition, CSR occurs during the GC reaction which usually peaks ~ 10-12 days after immunization, giving rise to two types of B cells : non-secreting memory B cells that are Ig surface-positive and Ab-secreting plasmablasts. Plasmablasts leave the GC and develop into terminally differentiated plasma cells that secrete high affinity Abs (Calame, 2001).

As precised above, in haematopoiesis, more than ten lineages of mature blood cells are derived from HSCs. This process is tightly regulated by extrinsic environmental cues and an intrinsic genetic program. The establishment of the lineage-specific gene expression patterns underlying differentiation is believed to be driven by the interplay of intricate networks of transcription factors and chromatin regulators.

2. The transcriptional and epigenetic networks

2.1. The transcription

Eukaryotic gene transcription constitutes the basis of the most common and most immediate point of regulation of gene expression in a cell.

Transcription of the eukaryotic genes is performed by three RNA polymerases : RNA polymerase I (Pol I) synthesizes the large ribosomal RNA (rRNA), Pol II synthesizes mRNA and Pol III synthesizes tRNA and 5S rRNA.

We will focus here on Pol II transcription ; Pol I and Pol III transcription is reviewed in Paule and White (Paule and White, 2000).

2.1.1. Pol II transcription

2.1.1.1. The basal transcription machinery

Synthesis of mRNA by RNA polymerase II (Pol II) is governed by two distinct DNA elements : a core promoter and upstream – or downstream – enhancer sequences. The core promoter determines the transcription start site (TSS) and directs the assembly of the pre-initiation complex (PIC), which consists of Pol II and general transcription factors (GTFs), namely TFIIA, TFIIB, TFIID, TFIIE, TFIIF and TFIIH. TFIID itself is a multi-protein complex containing the TATA-binding protein (TBP) and a set of TBP-associated factors (TAF_ns), which have sequence-specific DNA-binding activity and are involved in promoter selectivity (Burley and Roeder, 1996).

The core promoter often contains a TATA box, an AT-rich motif located 25-30 bp upstream of the TSS which is bound by TBP. Another common element is the initiator (Inr), which encompasses the TSS and which serves as a binding site for additional regulatory factors. The sequence elements of a typical core promoter and a model of the Pol II transcription initiation machinery are depicted in figure 13 (Baumann et al.; Hochheimer and Tjian, 2003).

The largest subunit of Pol II contains a highly conserved C-terminal domain (CTD) consisting of tandem repeats of the heptapeptide YSPTSPS, which can be differentially phosphorylated during the different stages of transcription. During the initiation of transcription, the CTD is unphosphorylated and binds to the large SRB/Mediator complex. At the transition to transcription elongation, the CTD becomes phosphorylated and Pol II dissociates from the SRB/Mediator complex and recruits the elongator complex as well as several pre-mRNA processing factors. The cell cycle-specific regulation of Pol II activity also depends on the phosphorylation of its CTD.

2.1.1.2. Transcription factors

To achieve a specific and controlled regulation of gene transcription, the basal transcription machinery interacts with additional proteins so-called transcription factors that bind to the promoter or enhancer regions. They are specialized nuclear proteins with the ability to bind to DNA *via* a sequence-specific DNA-binding domain ; they are additionally defined by at least an independent regulatory domain that can either transactivate or corepress. Factors that do not feature a DNA-binding domain but are recruited by transcription factors to the active site of transcription and thereby modulate transcriptional activity are termed transcriptional cofactors (coactivators or corepressors).

There are various ways in which transcription factors regulate transcription. First, they can interact with components of the basal transcription machinery and thereby modulate the assembly of the PIC (Ptashne and Gann, 1997). Other processes that can be targeted by

transcription factors are transcription elongation and re-initiation. They can also recruit chromatin modifying activities that modulate the acetylation state of the local chromatin and thereby can facilitate or inhibit transcription.

Moreover, transcription factors could be involved in targeting active promoters to regions in the nucleus with high transcriptional activities. Alternatively, supporting the model of nucleus organizers, they could selectively recruit repressed genes to silent domains (Brown et al., 1997).

2.1.2. Transcriptional regulation of B cell development

The development, maturation and selection of mammalian B cells is a complex and strictly regulated process. As we have described previously, the successive stages of B cell differentiation can be characterized by sequential rearrangement of the Ig genes, and by the ordered appearance or disappearance of proteins at the cell surface, or within the cell (Rolink and Melchers, 1993). The entire program of differentiation involves not only the expression of B cell lineage-specific genes but, also and importantly, the silencing of lineage-inappropriate genes. The correct temporal establishment of this lineage-specific expression pattern is notably determined by the action of lineage-restricted transcriptional regulators. Experiments using targeted gene disruptions have shown that the loss of some transcription factors, such as PU.1 and Ikaros, affects multiple haematopoietic lineages (Nichogiannopoulou et al., 1998; Scott et al., 1994), whereas the loss of other transcriptional regulators specifically affects early B cell development. Remarkably, three transcription factors E2A, EBF (early B cell factor) and PAX-5 (paired box protein 5) have been identified to be crucial for early B cell development (Bain et al., 1994; Lin and Grosschedl, 1995; Nutt et al., 1999; Sun, 1994; Zhuang et al., 1994) by regulating notably the B cell specific gene that encodes the Igα subunit (*mb-1*) of the BCR (Sigvardsson et al., 2002). Yin Yang 1 (YY1), which is a zinc-finger protein that functions as a transcriptional activator, repressor or transcription-initiator element-binding protein, depending on the promoter context, is another critical regulator of B cell development : its specific deletion in mouse B cells causes a defect in *Igh* rearrangement and a block at the Pro B to Pre B cell transition (Liu et al., 2007). Besides, the zinc-finger transcription factors Ikaros and Aiolos can activate or repress target genes by recruiting chromatin remodeling complexes (nucleosome remodeling and deacetylase complex, NuRD) or corepressors (C-terminal binding protein, CtBP) and are essential for the development of CLPs and mature B cells, respectively (Cariappa et al., 2001; Kim et al., 1999; Koipally et al., 1999). Recently, the involvement of CtBP in Bcl-6 autoregulation has been demonstrated (Mendez et al., 2008). This is of particular interest since Bcl-6 has homology with zinc finger transcription factors and is required for the formation of GCs –a microenvironment where TD Ag responses occur-.

2.1.3. Immunoglobulin and surrogate light chain genes transcription

Like the globin genes, the Ig genes were among the first genes studied in depth in mammalian cells. For this reason, they are some of the best-

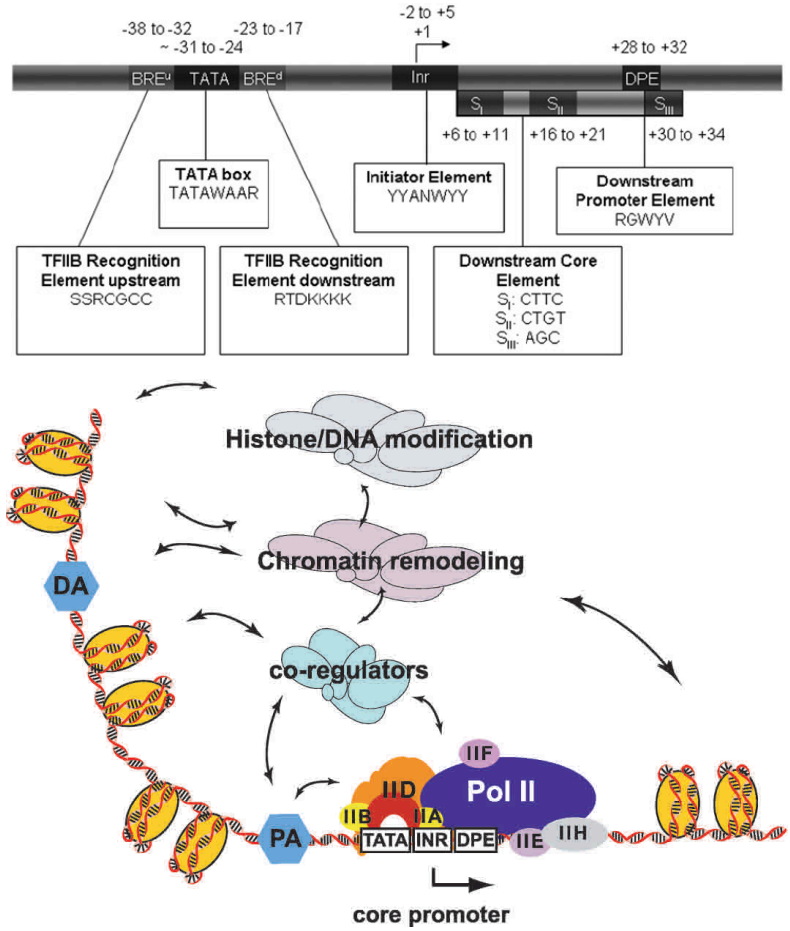


Figure 13 : Top panel. Core promoter elements. This diagram shows common core promoter elements with their consensus sequences (degenerate nucleotides according to IUPAC code) and relative position to the transcription start site (+1). The core promoter can show considerable variability as there are no universal elements. Each of the motifs is found in only a fraction of core promoters with different combinations. The TATA box, Initiator (Inr), downstream promoter element (DPE), and downstream core element (DCE) are recognition sites for binding of TFIID, whereas TFIIB recognition element upstream, BRE^u and downstream BRE^d interact with TFIIB. Bottom panel. The transcription apparatus, a multilayered ensemble of multisubunit complexes. The transcription machinery is responsible for the decoding and expression of genes in a regulatory network composed of various co-regulatory complexes that are interconnected to control RNA synthesis from a given promoter. This includes covalent modifications of histone tails and DNA that can influence whether a chromatin template is programmed to be transcriptionally active or silent. Chromatin remodeling activities catalyze the ATP-dependent deposition, removal, or sliding of nucleosomes to create a DNA template that is accessible to transcription factors. Activated transcription is controlled by co-regulatory co-activator complexes that mediate activation of transcription by activators located in a promoter proximal position (PA) or many kilobases away in a promoter distal position (DA). Also, the composition and sequence of the core promoter directs the formation of the preinitiation complex (PIC) and defines the start site of transcription. This involves the interplay of basal transcription factors (TFIIA, TFIIB, TFIID, TFIIE, TFIIF, and TFIIH) and RNA polymerase with distinct core promoter elements like TATA box (TATA), Inr, and DPE to position the PIC. Taken from Baumann *et al*., 2010 and Hochheimer and Tijan, 2003.

described transcriptional units in this model and, therefore, constitute a very useful tool to analyze the regulation of gene transcription in cells. Ig gene transcription is a highly regulated process driven by a number of ubiquitous and tissue-specific transcription factors that bind to regulatory regions in the Ig locus. These control regions include the V gene proximal promoters and promoters used for sterile transcription of the H and L chain loci, the intronic enhancers in the J-C region of the H and κ chain loci (Eμ, iEκ), the silencers located near the intronic enhancers, and the 3' enhancers of the Cα, κ and λ C region genes (3'EH, 3'Eκ and 3'Eλ). Additional regulatory elements are the matrix attachment regions (MARs), flanking the Eμ and iEκ intronic enhancers. Most of the control elements are depicted in figure 11 (Schatz and Ji).

The Ig surrogate light chain genes *Igll1* and *VpreB*, while not required for early B cell differentiation, are also expressed in a lineage-specific manner to form a Pre-BCR signalling complex with μ heavy chain. Thus, the *Igll1* and *VpreB* loci are a model system for studying gene regulation during early B cell development.

2.1.3.1. V region promoters of Ig genes

The V region promoters contain several control elements, the most highly conserved elements being a TATA box about 25 bp upstream of the transcription start site and an octamer element (ATGCAAAT) or its reverse complement. This element is also conserved in some Ig enhancers and in the promoters of other B cell specific genes. Other consensus elements in many V promoters and enhancers include CAAT enhancer binding protein (C/EBP) binding sites and μE3 elements, which are binding sites for E box proteins (Beckmann et al., 1990).

2.1.3.2. IgH intronic enhancer (Eμ enhancer)

The intronic enhancer is located between the J regions and Cμ and spans approximately 220 bp sequence featuring octamer and C/EBP binding sites as well as a number of E box motifs (CANNTG), such as μE1, μE2, μE3, μE4, μE5, μA (Arulampalam et al., 1997; Engel and Murre, 2001; Ernst and Smale, 1995). The E box binding-proteins (E-proteins) like E2A belong to a subclass of basic helix-loop-helix (bHLH) factors that act as transcriptional activators and can homo- or heterodimerize on the E box motifs μE2, μE4 and μE5 (Hu et al., 1992; Murre et al., 1989). Additional motifs termed μA and μB were also implicated in the activation of IgH gene expression. These binding sites are recognized by the transcription factors of the Ets family, including Ets-1, Ets-2 and Elf-1, which can bind to the μA motif, and PU.1, which binds to the μB site (Nelsen et al., 1993; Rivera et al., 1993). The Ets transcription factors are characterized by a conserved DNA-binding domain (the ETS domain) that forms a HLH structure, which binds specific purine-rich DNA sequences with a GGA core (Crepieux et al., 1994; Sharrocks, 2001). They are generally weak activators of transcription by themselves and act mainly through cooperative binding with other transcription factors.

Of great interest, the intronic enhancer is flanked by two MARs, consisting of A/T-rich sequences that associate with the nuclear matrix (Cockerill et al., 1987). They have been associated with the control of chromatin structure of entire gene loci and thus overlap functionally with locus control regions (LCRs). The MARs can contribute positively to Eμ function by enhancing transcription from a transgene promoter (Forrester et al., 1994; Jenuwein et al., 1997). In addition, MARs have been implicated in constituting physical boundaries between genes by forming chromosome loops (Cockerill and Garrard, 1986; Cockerill et al., 1987).

2.1.3.3. 3' IgH enhancer

The 3' IgH enhancer is located downstream of the IgH locus, 3' of the Cα region. It spans over 30 kb and encompasses four DNase I hypersensitive (HS) sites with enhancer activity termed HS1 to HS4. The initially identified major enhancer region, whose activity is inducible in activated B cells and plasma cells contains HS1 and HS2 sites. It was discovered early that during B cell activation, an octamer site present in several of the 3' enhancer elements is crucial (Yuan et al., 1995) as well as the interaction of Oct-2 and its coactivator OBF-1 with the 3' IgH (Tang and Sharp, 1999).

It is generally believed that the 3' IgH enhancer plays an important role in the high expression of switched Ig loci in plasma cells compared to unswitched loci, presumably because it is then brought into the vicinity of the V_H promoter and interacts with it.

2.1.3.4. *Igll1-VpreB* gene control elements

The regulatory elements required for correct levels of expression and controlling time-specificity of *Igll1* and *VpreB* are located within a 19 kb fragment containing both genes. This fragment was shown to contain a LCR that is able to promote efficient and stage-specific expression of both genes. This LCR, having strong enhancer activity, includes a cluster of five DNase I HS sites located 3' of the *Igll1* gene. There is also evidence that additional elements located elsewhere in the locus can contribute to the LCR effect : sequences that can enhance transcription in pre B cells have been identified in the *Igll1* promoter, in two HS sites (HS7 and HS8) in the intergenic region and in the HS1 site. This is an interesting picture of regulatory elements distributed across the locus instead of being concentrated into a discrete unit of multiple HS site like it is the case for the β-globin LCR (Sabbattini and Dillon, 2005).

The observation that the promoters of *Igll1* and *VpreB* genes contain multiple binding sites for EBF and E2A in addition to *in vivo* study (Sigvardsson et al., 1997) demonstrated the involvement of these two transcription factors in activating *Igll1* and *VpreB* expression. Moreover, two adjacent binding sites for Ikaros are present in the *Igll1* promoter and one of these sites overlaps with an EBF binding site. The fact that Ikaros binds DNA more efficiently as a dimer raised the possibility that co-operative binding of Ikaros to the two sites could repress *Igll1* transcription by preventing binding of EBF to the overlapping site. Ikaros is indeed directly involved in stage-

specific silencing of the *Igll1* gene in pre BI and mature B cells (Sabbattini et al., 2001; Thompson et al., 2007).

2.1.4. Transcription factors regulating Ig and surrogate light chain gene transcription

2.1.4.1. Oct factors and their coactivator OBF-1

The best studied transcription factors that interact specifically with the octamer motif in B cells are the POU homeodomain proteins Oct-1 and Oct-2. While Oct-1 is ubiquitous, Oct-2 is predominantly expressed in B cells, as well as in activated T cells and in the nervous system. Oct-1 and Oct-2 are members of the POU family of transcription factors (POU stands for Pit-1, Oct-1 and nematode Unc-86 ; for review see Herr and Cleary, Latchman and Philips and Luisi (Herr and Cleary, 1995; Latchman, 1999; Phillips and Luisi, 2000)). The common characteristic of the POU family is the 150-160 amino-acid bipartite DNA-binding domain (POU domain), consisting of two structurally independent subdomains : the homeodomain POU_H (which is related to the DNA binding domain of homeobox proteins) and the specific domain POU_S (which is unique to the POU factors). Oct-1 and Oct-2 show identical DNA-binding activities and share 90% sequence identity in their DNA-binding POU domain, but differ in other protein regions, mostly in the C-terminal activation domain (Tanaka et al., 1992). The analysis of Oct-2 deficient mice showed that this factor is required for B cell maturation and also, surprisingly, for postnatal survival (Corcoran and Karvelas, 1994; Corcoran et al., 1993). Oct-1 appears to rely on co-operative binding with other factors to achieve its full activation potential. Ternary complex formation with tissue-restricted coactivators, such as OBF-1 (Ford et al., 1998) can increase the specificity of the transcriptional activation mediated by Oct factors. While it was known early that Oct-1 and Oct-2 can bind to DNA elements as monomers (Staudt et al., 1986), it was discovered only recently that they can form homo- and heterodimers on specific octamer-related DNA motifs : the PORE (Palindromic Oct-factor Recognition Element : ATTTGAAATGCAAAT) and the MORE (More PORE : ATGCATATGCAT) (Remenyi et al., 2001; Tomilin et al., 2000). This can confer additional selectivity.

OBF-1 (OCA-B, Bob-1) is a proline-rich lymphoid-specific transcriptional coactivator protein that interacts with the POU domains of Oct-1 or Oct-2 and with the conserved octamer element on the DNA (Gstaiger et al., 1996; Luo and Roeder, 1995; Strubin et al., 1995). OBF-1 strongly potentiates transcription from octamer-containing promoters such as Igκ light chain promoter in transcription assays and *in vitro* (Luo and Roeder, 1995; Schubart et al., 1996). OBF-1 seems to play a role for cell survival at early B cell stages, since a significant reduction of transitional (immature 493^+) B cells in the spleen was observed by FACS analyses, indicating a malfunction in the B cell homing from bone marrow to the spleen in the absence of OBF-1 (Schubart et al., 2000) or an impaired production of transitional B cells in the bone marrow of OBF-1 deficient mice (Hess et al., 2001). Importantly, histochemical studies have shown that OBF-1 function is crucial for the formation of GCs since they are absent in secondary lymphoid organs of

OBF-1 deficient mice. As a consequence, these mice have defects in the Ag-dependent B cell development and show a strongly impaired immune response to T cell-dependent Ags (Kim et al., 1996; Schubart et al., 1996). Interestingly, structural and functional data have demonstrated that, at least on some of the Ig heavy chain promoters, Oct-1 binds to the octamer site in a dimeric conformation (MORE) that precludes recruitment of OBF-1 to the DNA (Tomilin et al., 2000). In B cells, the *B29* and *mb-1* promoters are directly targeted by OBF-1 (Malone and Wall, 2002). The severe reduction of the number of immature B cells observed in the spleen of OBF-1 deficient mice is even more dramatic in OBF-1/Oct-2 (Schubart et al., 2001) and OBF-1/btk (Bruton's tyrosine kinase) compound mutant mice (Schubart et al., 2000). In line with this, mice lacking simultaneously OBF-1 and the chromatin regulator Aiolos show a severe developmental block in the bone marrow, at the transition between Pre B and immature B cells (Sugai et al., 2003; Sun et al., 2003). This phenotype is all the more surprising given the fact that the individual knockouts of these two factors show defects only in late stages of B cell development. However, OBF-1 is critical for the activity of a subset of IgL promoters (Casellas et al., 2002).

Apart from OBF-1, investigation in mice revealed that Aiolos is a Ikaros family factor crucial for the function of mature B cells in the periphery.

2.1.4.2. Ikaros and Aiolos

Ikaros is the founding member of a growing family of krüppel-like zinc-finger transcription factors. Gene targeting studies in mice have shown that inactivation of Ikaros family proteins leads to a complete absence of B, T, NKs and DCs, whereas a reduction of Ikaros activity induce hyperproliferation and lymphomas (Georgopoulos et al., 1997; Takanashi et al., 2002; Winandy et al., 1995). Like the Drosophila protein hunchback, Ikaros proteins are characterized by a highly conserved C_2H_2 zinc-finger DNA-binding domain near the N terminus and a C_2H_2 zinc-finger protein-protein interaction domain near the C terminus. The latter mediates the formation of homo- or heterodimers and multimers between Ikaros family members, which is essential for the regulation of their functions (Georgopoulos et al., 1992; Hahm et al., 1994).

Upon alternative splicing, several isoforms of Ikaros are expressed (Fig.14 (Rebollo and Schmitt, 2003)). Long isoforms (Ik1 to Ik3) with at least three zinc fingers can efficiently bind to DNA, while shorter versions (Ik4 to Ik8) with less than three zinc fingers behave, upon dimerization, as dominant negative isoforms. The role of Ikaros in lymphoid development has been studied following Ikaros gene inactivation in mice, by deletion of exons 3 to 5 leading to overexpression of a dominant negative Ikaros protein (Ikaros $DN^{-/-}$) ; these mice exhibited a complete lack of lymphoid cells (B, T, NKs and DCs) (Georgopoulos et al., 1994). A less drastic phenotype is observed in Ikaros null mutant resulting from targeting of exon 7, required for stable protein expression. In these mice, T cells still appears few weeks after birth, suggesting some complementation by a related protein (Wang et al., 1996).

A protein related to Ikaros was identified by degenerate PCR with primers complementary to sequences encoding the carboxy-terminal zinc fingers. This protein, Aiolos, exhibits considerable homology to the largest

Ikaros isoform, with strong similarities in the DNA-binding, activation and dimerization domains. It interacts with Ikaros through its carboxy terminal fingers (Morgan et al., 1997). Aiolos is first detected at low levels in early B and T cell precursors and its levels increase as they progress to more definitive stages of differentiation (pre B cells and double positive T cells). Aiolos expression increases further in peripheral B cells and peaks in recirculating B cell populations. Disruption of Aiolos leads to phenotypes that are consistent with its pattern of expression : an increase in pre B and immature B cells and a severe reduction in recirculating B cells (Wang et al., 1998).

The interplay between Ikaros and Aiolos in the regulation of gene expression is further complicated by additional Ikaros and Aiolos isoforms (Fig.14 (Rebollo and Schmitt, 2003)) or by other proteins that can sequester either member in transcriptionally inert complexes. Variations in the expression levels of these factors are important for cell progression through the lymphoid lineage. Normal progression through B and T lineages requires the sequential expression of Ikaros/Ikaros, Aiolos/Ikaros and Aiolos/Aiolos complexes (Morgan et al., 1997), suggesting that both members act in concert to regulate lymphoid differentiation and function.

Ikaros and Aiolos can activate transcription by binding to given promoters *via* their common consensus DNA-binding site (C/TGGGAA/T with GGGA described as the core motif) (Molnar and Georgopoulos, 1994; Sun et al., 1996). Although Ikaros and Aiolos share similarity in the activation domain, Aiolos is a stronger activator in mammalian cells. Moreover, a study revealed that Ikaros may not be a typical transcriptional activator since it was found to localize primarily to centromeric heterochromatin in B cells (Brown et al., 1997; Klug et al., 1998). Therefore, Ikaros may recruit genes that are destined for inactivation to centromeric foci. In the same line, the Ikaros-Helios complexes localize to the centromeric region of T cell nuclei. However, only a fraction of Ikaros exhibited a centromeric localization, presumably the fraction associated to Helios, suggesting that Helios is a limiting regulatory subunit for Ikaros within T cell centromeric heterochromatin (Hahm et al., 1998). Helios-Ikaros complexes might bind to DNA sequence elements within the promoters, enhancers, or silencers of genes that require repression ; especially, centromeric foci are characterized by γ-satellites (Hahm et al., 1998). Of great interest, while Ikaros is expressed throughout the ontogeny of lymphoid cells, a diminution of its activity by means of overexpression of a dominant negative isoform in T lymphocytes leads to hyperproliferation, loss of heterozygosity and malignant T cell development (Wang et al., 1996). The restructuring of chromatin is obviously a key aspect of Ikaros function in lymphocyte development. Not only Ikaros proteins may be involved in gene silencing but, also and importantly, by colocalizing with DNA replication foci, they may control the stable propagation of genetic material (Avitahl et al., 1999; Wang et al., 1998).

Among key developmental genes for the B lineage, the *Igll1* gene, which encodes the pre-BCR component λ5 is tightly regulated by Ikaros and Aiolos. In pre BI cells, Ikaros modulates λ5 expression in competition with the transcriptional activator EBF. At the transition from the pre BI to the pre BII stage, the expression of Aiolos is upregulated and required for the efficient silencing of *Igll1* (Thompson et al., 2007). Silencing of *Igll1* is not

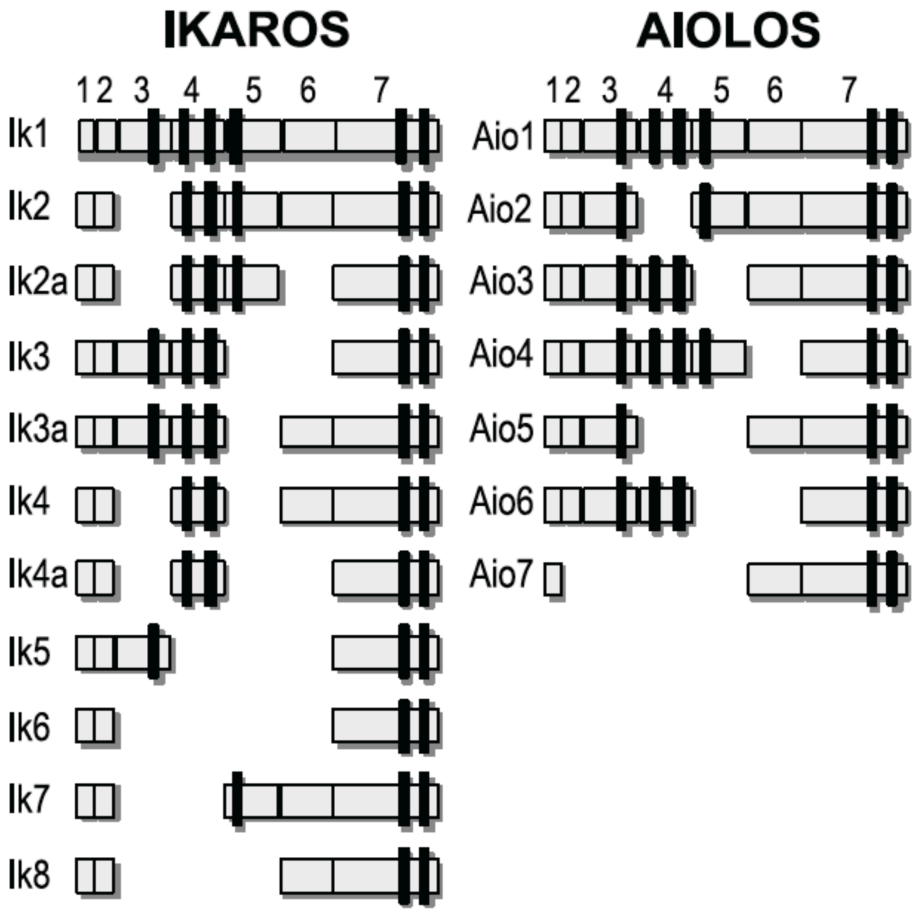


Figure 14 : Ikaros and Aiolos isoforms. Schematic representation of the various isoforms described for Ikaros and Aiolos. Ikaros isoform Ikx described by Payne *et al*. is here named Ik3a. Taken from Rebollo and Schmitt, 2003.

accompanied by DNA (CpG) methylation, delayed replication in S phase (Azuara et al., 2003), or a drastic reduction in chromatin accessibility (Sabbattini et al., 2001) but with repositioning to the vicinity of pericentromeric heterochromatin (Brown et al., 1997). Ikaros proteins may drive this by binding to high-affinity sites in pericentromeric γ-satellite repeats (Cobb et al., 2000); also, silencing involves increased Aiolos occupancy of the *Igll1* promoter.

In erythroid cells, Ikaros is not detectable as free protein but only found associated to chromatin-remodeling complexes. Alternatively, Ikaros may also be in complex with a similar Ikaros-like protein and/or perhaps with erythroid-specific factors such as GATA-1 (O'Neill et al., 2000). An Ikaros-GATA-1 interaction has been hypothesized but never identified. Most likely because biochemical purification of Ikaros-associated complexes has been generally carried out in lymphoid cells and most importantly, because beside Ikaros interaction with other Ikaros family proteins, it is known that Ikaros forms low stability complexes with other proteins *in vivo*, which might not resist high stringency washes (Bottardi et al., 2009).

2.1.4.3. Other key transcription factors in B cell development

Other Ikaros family members such as Helios, Eos and Pegasus are also expressed in B cells (Perdomo et al., 2000). Their function in this lineage remains poorly reported if at all. Ets factors, such as PU.1 and Spi-B, E2A, EBF, PAX-5, SOX4 (sex-determining region Y box 4), Lef-1 (lymphoid enhancer binding factor 1), IRFs (interferon-regulatory factor), the NF-κB family, BCL-6 (B cell lymphoma 6), XBP-1 (X-box-binding protein 1) and Blimp-1 (B lymphocyte induced maturation protein 1) are further discussed.

PU.1 – a transcription factor of the ETS family – is a protein crucial in the early stages of haematopoietic-cell lineage specification. *Pu.1*$^{-/-}$ mice die during embryonic development at day 18.5 (E18.5) ; they lack B and T cells, monocytes, and granulocytes (McKercher et al., 1996; Scott et al., 1994).

The highly related ETS protein SPI-B, expressed at all stages of B cell development (similarly to PU.1) cannot compensate for the absence of PU.1. Indeed, expression of SPI-B from the *Pu.1* locus in mice was found to rescue myeloid differentiation but not defect in lymphopoiesis, arguing for specific roles of these two proteins in haematopoietic-cell development (Dahl et al., 2002).

The gene that encodes E2A is a founding member of the bHLH gene family ; it encodes the broadly expressed E12 and E47 proteins (collectively known as E2A), generated by alternative splicing. Because these proteins bind the E box – CANNTG (where N denotes any nucleotide), a motif initially identified in the mouse IgH enhancer-, they are known as E-box factors or E proteins. Although the *E2a* gene is widely expressed, its mutation in mice leads to defects that are most marked in the B cell lineage (Bain et al., 1994; Kee et al., 2000), but other cell lineages (such as T cells) are also affected (Bain et al., 1997).

EBF is a transcription factor expressed specifically by pro B, pre B and mature B cells ; it has been implicated in the regulation of the B cell specific gene *mb-1* that encodes the Igα subunit of the BCR. Ablation of EBF in mice

leads to a very early block in B cell differentiation, similar to the block observed in the absence of E2A (Lin and Grosschedl, 1995).

PAX-5 is a paired homeodomain protein expressed by B cells and in the nervous system at the midbrain-hindbrain boundary. Mutation of the *Pax5* gene leads to a complete block in B cell differentiation downstream of the block seen in the absence of E2A or EBF (Urbanek et al., 1994).

The transcription factors SOX4 and Lef-1 are members of the high mobility group (HMG)-box family and have a crucial role at an early stage of B cell development. Mice lacking SOX4 die at E14 due to a defect in cardiac formation. However, when *Sox4*$^{-/-}$ fetal liver cells were used to reconstitute the haematopoietic system of lethally irradiated mice, a very strong block was observed at the pro B to pre B cell transition (Schilham et al., 1996). In *Lef-1*$^{-/-}$ mice, the size of the B cell compartment is reduced (Reya et al., 2000).

IRF4 and IRF8 are highly related factors that are expressed in both lymphoid- and myeloid-lineages. Mice lacking IRF4 have normal B cell development but greatly impaired B and T cell activation, resulting in a strongly impaired immune response (Rettig et al., 1997). By contrast, mice lacking IRF8 show defects in macrophage development and have a chronic-myeloid-leukaemia-like phenotype (Holtschke et al., 1996).

NF-κB is an inducible transcription factor comprising related proteins (Ghosh et al., 1998). Ablation of various NF-κB subunits in mice revealed that early B cell development and immunoglobulin-gene transcription do not require this factor. Instead, several aspects of late B cell differentiation and/or maturation are affected by the absence of NF-κB (Sha et al., 1995).

IRF4, BCL-6 and XBP-1 are involved in the terminal differentiation of B cells into plasma cells. In B cells deficient in Blimp-1, plasma cell differentiation is impaired (Shapiro-Shelef and Calame, 2005).

2.2. Epigenetics

2.2.1. Definitions

When Conrad Waddington coined the word ‘epigenetics’ (literally ‘over’ or ‘upon’ genetics) in the early 1940s, the term was used to explain why genetic variations sometimes did not lead to phenotypic variations and how genes might interact with their environment to yield a phenotype. However, the word currently refers specifically to the study of heritable changes in gene expression that occur during mitosis or meiosis without changes in the DNA sequence (Berger et al., 2009). The disruption of such changes underlies a wide variety of pathologies, including cancers (Esteller, 2008). Epigenetic regulation includes heritable modifications of the DNA –such as cytosine methylation- and aspects of chromatin structure –including histone modifications-.

2.2.2. The CtBP complex

The vertebrates have two *CtBP* genes. Genetic studies indicate that the *CtBP* genes play pivotal roles in animal development. The vertebrate C-terminal binding proteins (CtBP1 and CtBP2) are highly related and are

functionally redundant for certain developmental processes and non redundant for others. CtBP proteins are evolutionarily-conserved, non DNA-binding co-regulator phosphoproteins, interacting with sequence-specific DNA binding transcriptional regulators. The corepressor complex of CtBP1 contains enzymatic constituents that mediate coordinated histone modification by deacetylation (histone deacetylases, HDACs) and methylation and demethylation (CoREST and LSD1) (Chinnadurai, 2007). Notably, CtBP is a BCL-6 corepressor ; during the GC stage of B cell development, CtBP is required for BCL-6 autoregulation and subsequently for the control of the expression of other BCL-6 target genes (Mendez et al., 2008). Well-known as a repressor complex, the CtBP complex can nevertheless mediate activation of some targets (Fang et al., 2006).

2.2.3. The NuRD complex

The NuRD complex was first purified about a decade ago in cells from different species, and it contains six core subunits. In mammals, the subfamily II members CHD3 and CHD4 are subunits of NuRD complexes. The latter also contain HDACs and function generally as transcriptional repressors. Mammalian NuRD complexes achieve diversity in regulatory function through combinatorial assembly. The core ATPase is CHD3 or CHD4. HDAC1 and HDAC2 catalyse protein deacetylation. There are three main accessory subunits, which are encoded by gene families : MTA (metastasis-associated), MBD (Methyl-CpG-binding domain) and RbBP (retinoblastoma-associated-binding protein). Each complex contains one MTA protein : MTA1, MTA2 or MTA3. These are mutually exclusive and nucleate complexes with markedly different, and sometimes opposite, functions. Each complex also contains MBD2 or MBD3, which are functionally distinct and contribute to different forms of the complex, and RbBP4 and/or RbBP7. The RbBP4 and RbBP7 subunits are thought to be structural components of the NuRD complex and have been shown to directly associate with histone tails. The composition of the NuRD complexes varies with cell type and in response to signals within a tissue, giving rise to a diversity of complexes with distinct functions (Ho and Crabtree). For instance, the transcriptional repressor BCL-6 regulates B lymphocyte cell fate during the GC reaction by preventing terminal differentiation of B lymphocytes into plasma cells until appropriate signals are received. The cofactor MTA3, a cell type-specific subunit of the corepressor complex NuRD is required for BCL-6-dependent cell fate determination (Fujita et al., 2004). The NuRD complex is involved in the transcriptional regulation of key genes that promote the progression of T and B lymphocyte development. Transcriptional repression of multiple lineage-specific genes by the NuRD complex during haematopoiesis is mediated through Friend Of GATA-1 (FOG-1), which binds MTA family proteins and recruit NuRD to GATA family transcription factors. Other haematopoietic lineage-specific transcription factors that are also associated with the NuRD complex include Ikaros (Lai and Wade).

2.2.4. Other key epigenetic regulator complexes

Sin3 and PRC2 (Polycomb Repressive Complex 2) complexes are two other epigenetic complexes containing RbBP4 and RbBP7.

Besides, SWI/SNF complexes are crucial for the proper development of all organisms. The second family of SWI-like ATP-dependent chromatin-remodeling complexes is the ISWI family. The ISWI ATPase is the core component of three types of ISWI complex : NuRF (nucleosome remodeling factor), ACF (chromatin-assembly factor) and CHRAC (chromatin accessibility complex) complexes. The third family of SWI-like ATP-dependent chromatin-remodeling complexes is the CHD family. Theses complexes contain members of the CHD family of ATPases, which comprises nine chromodomain-containing members. CHD proteins are broadly classified into three subfamilies based on their constituent domains : subfamily I (CHD1 and CHD2), subfamily II (CHD3 and CHD4) and subfamily III (CHD5, CHD6, CHD7, CHD8 and CHD9). The last family of SWI-like ATP-dependent chromatin-remodeling complexes is the INO80. These complexes contain INO80 ATPases, which in mammals include INO80 and SWR1 and are characterized by the presence of a conserved split ATPase domain. The INO80 and SWR1 complexes are large multisubunit machines with *in vitro* nucleosome-remodeling activity, which might contribute to their *in vivo* roles in transcriptional regulation (Ho and Crabtree).

3. FOG-1, a master developmental cofactor within the haematopoietic system

Haematopoiesis is a highly regulated multi-step process controlled by transcription factors and epigenetic modifiers. From a very small population of self-renewing, pluripotent HSCs, all specialized and distinct blood cell types arise. During the proliferation, differentiation and maturation of pluripotent HSCs, cellular specification and commitment are initiated by primary lineage determinants, such as the cofactor FOG-1.

Within the haematopoietic cells, FOG-1 expression is restricted : it is detected at high level in early multipotent progenitors, erythroid and megakaryocytic cells, low level in HSCs and lymphoid cells, and it is absent from myeloid lineages (Cantor and Orkin, 2005).

Zfpm1 encodes the developmental regulator FOG-1. It is a multitype zinc finger protein of 995 and 1006 amino acids in the mouse and human, respectively. FOG-1 interacts with the DNA-binding transcription factor GATA-1 (Tsang et al., 1997). Aside from GATA-1, three main interacting partners of FOG-1 have been defined : CtBP and the NuRD complex, which both link FOG-1 to chromatin regulation, and transforming acidic coiled-coil 3 (TACC3), which sequesters FOG-1 in the cytosol (Fox et al., 1999; Garriga-Canut and Orkin, 2004; Hong et al., 2005).

This review summarizes genetic and biochemical evidence underlying the key role of FOG-1 as a primary determinant of haematopoietic cell fate. Current facts showed its substantial role in the cell fate control of HSC and progenitor populations rather than in the maturation of terminally differentiated cells. FOG-1 functions are intrinsically linked to its nuclear interactors, which put this protein at the crossroads of transcription and chromatin.

3.1. FOG-1 functions in embryonic stem cells

The transcriptional cofactor FOG-1, like the transcription factors GATA-1, GATA-2 and Ikaros, is stably expressed in $Lin^{-}Sca^{+}$ HSC over a limited number of time points of culture in IL-3, IL-6, IL-11, and steel factor (Dooner et al., 2008).

The analysis of the functions of FOG-1 in mouse embryonic stem (ES) cells was mainly performed by Tanaka *et al*. (Tanaka et al., 2004). Using a TET-Off conditional gene expression system, they overexpressed FOG-1 in ES cell line in the presence of tetracycline and could show that FOG-1 misexpression in ES cells induced cell cycle arrest at the G0/G1 phase. Similar enforced expression of FOG-1 in GATA-1 null ES cells lead to an inhibitory proliferation effect of FOG-1, suggesting a GATA-1 independent role of FOG-1 in this cellular setting. On the contrary and interestingly, the overexpression of FOG-1 mutant unable to bind CtBP did not affect ES cell proliferation and morphology. As a result, FOG-1/CtBP interaction may well be crucial for regulating the proliferation of ES cells (Fig. 15 (Tanaka et al., 2004)).

3.2. FOG-1 functions in the myeloid compartment

3.2.1. FOG-1, a master regulator of erythroid and megakaryocytic differentiation

Zfpm1$^{-/-}$ mice die between E10.5 and E12.5 of gestation from severe anaemia with a developmental block during erythropoiesis and a failure of megakaryopoiesis (Tsang et al., 1998). In both *Zfpm1*$^{-/-}$ and GATA-1^{-} mice, the block in erythroid maturation was observed at the proerythroblast stage (Fujiwara et al., 1996). This was providing genetic evidence that FOG-1 and GATA-1 functions were coordinately linked in erythroid development ; which was expected according to biochemical data demonstrating that FOG-1 acts as a cofactor for GATA-1 at least in erythroid differentiation (Tsang et al., 1997). Nevertheless, contrary to GATA-1^{-} mice, *Zfpm1*$^{-/-}$ mice exhibited a complete failure of megakaryopoiesis. FOG-1 might therefore directs megakaryocyte development independently of GATA-1.

So far, in the erythroid lineage, FOG-1 molecular mechanism of action was mainly associated to GATA-1 ; especially in the chromatin context, FOG-1/GATA-1 interaction was well-described to trigger RBC differentiation by tighly orchestrating gene regulation. For instance, Ikaros in complex with GATA-1, FOG-1 and NuRD complex components is responsible for human β-globin gene regulation in mouse transgenic erythroid cells (Bottardi et al., 2009). Moreover, in the nucleus of GATA-1 erythroid committed cell line G1E, FOG-1 mediates activation or silencing of genes by facilitating chromatin occupancy by GATA-1 at sites bound by GATA-2 (Fig. 16 (Pal et al., 2004)), by recruiting epigenetic modifiers or by stabilizing tissue-specific chromatin loops (Vakoc et al., 2005). FOG-1 has also been shown to be required for reconfiguring higher order chromatin organization *via* GATA factor exchange (Jing et al., 2008).

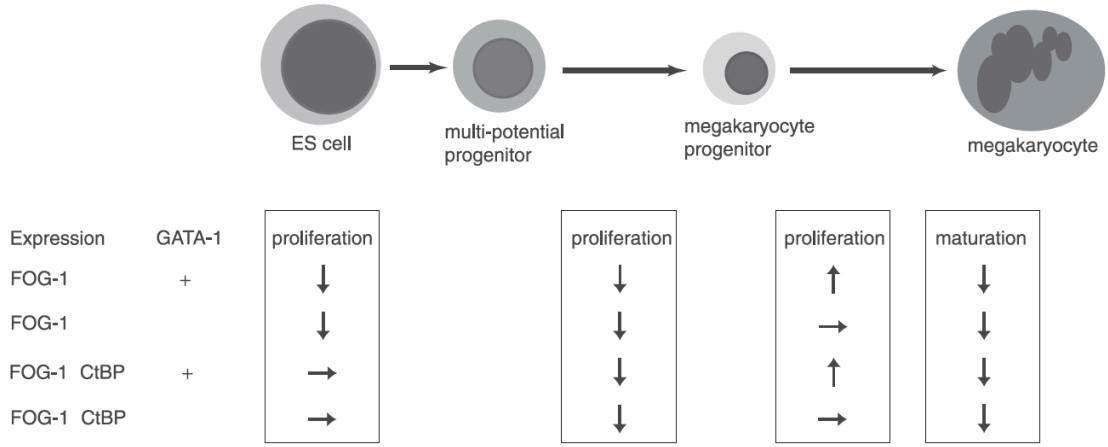


Figure 15 : Schematic diagram of the effects of overexpressing FOG-1 and FOG-1ΔCtBP in the wild-type and GATA-1-null conditions. Taken from Tanaka *et al*., 2004.

In addition, recent genetic studies showed that FOG-1 mediated recruitment of NuRD was required for re-enforcing the engagement towards erythrocyte and megakaryocyte differentiation during haematopoiesis (Gao et al.). This was confirming *in vitro* findings from Hong *et al*. : they isolated the NuRD repressor complex in a GST pull-down assay using the N-terminal domain of FOG-1 as prey in MEL cell ; they additionally demonstrated that in G1E cells, the ability of GATA-1 to repress transcription was impaired while overexpressing a FOG-1 disrupted in the NuRD-binding site (Hong et al., 2005). According to Rodriguez *et al*. work, overlapping functions of FOG-1 and GATA-1 in erythropoiesis occur in the context not only of the GATA-1/FOG-1 complex but also of the GATA-1/FOG-1/MeCP1 (Rodriguez et al., 2005). The association with MeCP1 provided an additional molecular basis for the well-documented repressive properties of FOG-1 and GATA-1 interactions (Crispino et al., 1999; Fox et al., 1999; Letting et al., 2004; Pal et al., 2004).

In megakaryocytes, FOG-1 molecular mechanism of action is less well-known. However, mice in which FOG-1/NuRD interaction is disrupted display defects similar to germline mutations in the *Zfpm1* and *Gata1* genes, including anaemia and macrothrombocytopenia. Miccio *et al*. provided clear evidence that NuRD is involved in several key functions of FOG-1 in the erythroid and megakaryocyte lineages (Gregory et al.; Miccio et al.).

3.2.2. FOG-1, a cofactor balancing differentiation in the granulocyte/monocyte progenitor derived cells

FOG-1 expression in myeloid (neutrophil, granulocyte and macrophage) progenitors inhibits mast cell differentiation and enhances development of neutrophils. Especially, cells with overexpression of FOG-1 expressed *Gr-1* and *Mac-1* genes and displayed phagocytic activity (neutrophilic characteristics). Otherwise, in mature mast cells, FOG-1 reprogrammed the fate of differentiated cells into multipotent progenitors. From forced expression studies in both cell lines and primary cells, it is evident that GATA-1 and PU.1, two FOG-1 interactors are able to specify erythroid and myeloid cell fates (Heyworth et al., 2002). FOG-1 is not expressed in mast cells. Nevertheless, mast cell differentiation was clearly driven by the molecular and functional interactions between GATA-1, PU.1 and FOG-1. On the contrary, the loss of the mast cell phenotype was not arising from disruption of the complex between the GATA factors and PU.1 but would rather be caused by the down-regulation of *MITF*, a gene encoding an essential mast cell regulator (Sugiyama et al., 2008). In addition, FOG-1 reduced the mRNA of transcription factors and mast cell-specific genes such as *PU.1*, *c-Kit*, *mast cell protease 4* (*p4*), *mast cell protease 5* (*p5*), *mast cell protease 6* (*p6*), *granzyme B* (*granB*) and *tryptophan hydroxylase* (*tph*). Confirming FOG-1 involvement in the correct establishment of mast cell lineage specific gene expression, Maeda *et al*. demonstrated that FOG-1 expression level regulated the GATA-1 dependent *FcepsilonRI beta-chain* promoter (Maeda et al., 2006). Notably, FcεRIβ-chain or mouse high-affinity IgE receptor is involved in IgE-dependent allergic reaction, and, therefore linked to IgE-dependent allergic disease. Moreover, results occuring from cells cultured from male GATA-1^{-} or GATA-1^{V205G} (non-FOG-1-binding GATA-1 mutant) demonstrated that GATA functions completely independent of FOG

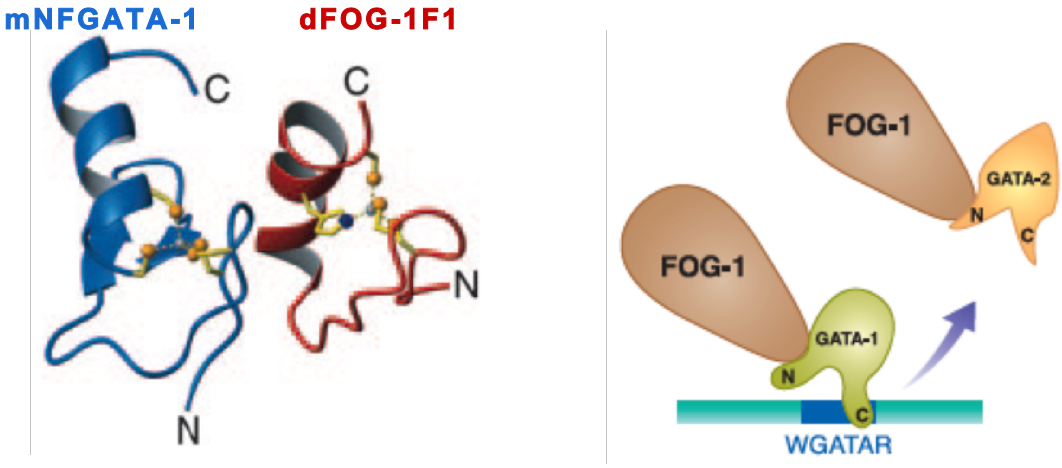


Figure 16 : Left panel. Structure of the GATA/FOG complex. Ribbon diagram of the lowest energy structure of the complex. Zinc ion are shown in gray and zinc-ligating side chains are shown in gold and blue. Right panel. Model of chromatin occupancy of GATA factors and FOG-1. GATA-1-FOG-1 and GATA-2-FOG-1 complexes switch at the chromatin template is depicted. Adapted from Liew *et al*., 2005 and taken from Pal *et al*., 2004, respectively.

proteins in terminal mast cell development (Cantor et al., 2008).

Also, the finding that FOG-1 was rather required for the development of the thrombocytic lineage at a very early stage (Tsang et al., 1998) led to examine the role of FOG-1 in the generation of MEPs and their differentiation. MEPs, but not eosinophils or myeloid cells express high levels of FOG-1. This could partially explain the role for FOG-1 in the GATA-1 mediated conversion of myeloid cells to a multipotent phenotype ; GATA-1 was indeed able to reprogram HD50M avian myelomonocytic cell line into MEPs and eosinophils (Kulessa et al., 1995). Moreover, expression of FOG-1 in avian eosinophils leads to the acquisition of a multipotent phenotype (Querfurth et al., 2000). Remarkably, myeloid cell types are characterized by the expression of PU.1 and C/EBP transcription factors, which regulate most myeloid-specific promoters (Tenen et al., 1997). Down-regulation of FOG-1 notably coincides with and is a prerequesite for the C/EBPβ-mediated up-regulation of eosinophils specific gene expression in MEPs. Concomitant with FOG-1 down-regulation, up-regulation of *mim-1* expression, which is an eosinophil/myeloid-specific marker was observed. Besides, some of the processes that are negatively affected by loss or misexpression of FOG-1 in the myeloid compartment are reminiscent of the B cell reprogramming into macrophages following ectopic expression of C/EBPα and C/EBPβ (Xie et al., 2004).

3.3. FOG-1 functions in the lymphoid compartment

FOG-1 plays a role in the T lineage by repressing GATA-3-dependent induction of Th2 development. Indeed, in reporter assays, FOG-1 can repress GATA-3-dependent activation of several cytokine promoters in T cells and notably the IL-5 promoter. Also, FOG-1 overexpression during primary activation of naive T cells, which express significant levels of FOG-1 mRNA inhibited Th2 development in $CD4^+$ T cells. In fact, FOG-1 mRNA is rapidly down-regulated upon commitment to both Th1 and Th2 lineages. Moreover, retroviral expression of FOG-1 in developing Th2 cells suppressed both IL-4 and IL-5 and allowed for IFN-γ production, which was accompanied by a significant level of T-bet mRNA expression. FOG-1 fully repressed GATA-3-dependent Th2 development and GATA-3 autoactivation, but not Stat6-dependent induction of GATA-3. Although FOG-1 overexpression repressed development of Th2 cells from naive T cells, it did not reverse the phenotype of fully committed Th2 cells (Zhou et al., 2001). Using the GATA-3 Nf as a bait in a yeast two-hybrid screening, Kurata *et al*. identified FOG-1 in the Th2 cell-specific library. Serial deletion mutation analysis indicated that the N-terminal region, but not the consensus CtBP-binding motif of FOG-1 is critical for the effects. The results clearly indicate that first of all FOG-1 is a repressor of GATA-3 in naive T cells and second the down-regulation of FOG-1 induces Th2 differentiation by releasing GATA-3 from its repression (Kurata et al., 2002).

To conclude, appropriate expression of FOG-1 within the «haematopoietic tree» is crucial for normal cell commitment (Fig. 17).

The role of FOG-1 in the B lineage is presented in the article entitled « FOG-1 is a novel co-factor for Ikaros in B-cells. » and in the discussion.

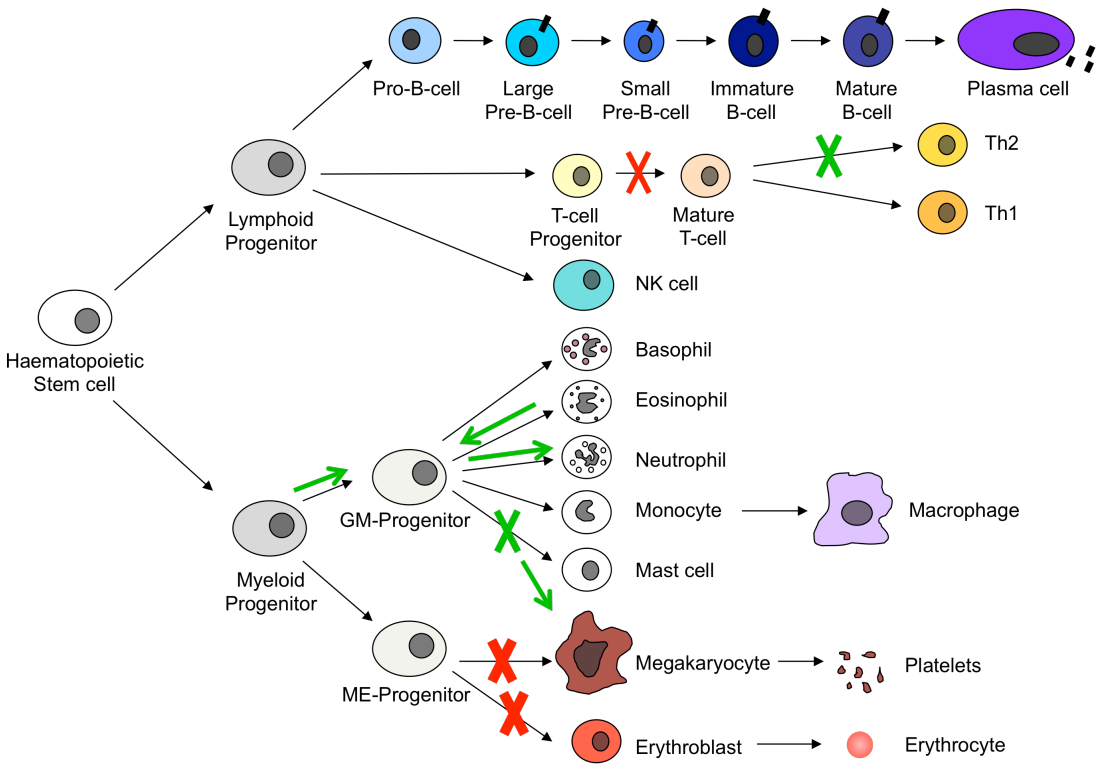


Figure 17 : FOG-1 requirement and impact during haematopoietic development. Processes that are negatively affected by loss or misexpression of FOG-1 are indicated by red and green crosses, while those that are enhanced by FOG-1 overexpression are depicted by green arrows.

Chapter 2 : Research publications

1. Inappropriate hematopoietic expression of FOG-1 does not affect B cells but reduces the number of circulating eosinophils

2. FOG-1 is a novel co-factor for Ikaros in B-cells

Inappropriate hematopoietic expression of FOG-1 does not affect B cells but reduces the number of circulating eosinophils

Camille Du Roure[1,3], Aude Versavel[1], Thierry Doll[2], Markus Kaller[1,4], Jean-François Spetz[1], Patrick Kopp[1], Chun Cao[1], Gabriele Matthias[1], Matthias Müller[2] and Patrick Matthias[1]

[1] Friedrich Miescher Institute for Biomedical Research, Maulbeerstrasse 66, CH-4058 Basel, Switzerland. [2] Novartis Institutes for Biomedical Research, CH-4056 Basel, Switzerland. [3] Present address: Actelion Pharmaceuticals Ltd, Gewerbestrasse 16, CH-4123 Allschwil, Switzerland. [4] Present address: Institute of Pathology, Ludwig-Maximilians-University Munich, Thalkirchnerstrasse 36, 80337 München, Germany.

Keywords: FOG-1, Gata factors, Hematopoiesis, B cells, Gene targeting, ES cells, *Rosa26* locus

ABSTRACT

We have identified expression of the gene encoding the transcriptional coactivator FOG-1 (Friend of GATA-1; *Zfpm1*, *Zinc finger protein multitype 1)* in B lymphocytes. We found that FOG-1 expression is directly or indirectly dependent on the B cell-specific coactivator OBF-1 and that it is modulated during B cell development: expression is observed in early but not in late stages of B cell development. To directly test *in vivo* the role of FOG-1 in B lymphocytes, we developed a novel embryonic stem cell recombination system. For this, we combined homologous recombination with the FLP recombinase activity to rapidly generate embryonic stem cell lines carrying a Cre-inducible transgene at the *Rosa26* locus. Using this system, we successfully generated transgenic mice where FOG-1 is conditionally overexpressed in mature B cells or in the entire hematopoietic system. While overexpression of FOG-1 in B cells did not significantly affect B cell development or function, we found that enforced expression of FOG-1 throughout all hematopoietic lineages led to a reduction in the number of circulating eosinophils, confirming and extending to mammals the known function of FOG-1 in this lineage.

INTRODUCTION

The development of specialized hematopoietic cells from self-renewing hematopoietic stem cells proceeds through a number of precursor stages with progressively restricted differentiation potential and requires a complex interplay of transcription factors and epigenetic modifiers. These regulators are responsible for orchestrating the establishment of lineage-specific gene expression patterns that underlie cellular differentiation (reviewed in [1,2]). While many factors involved in this process are already known, a complete molecular understanding is still missing. Friend of GATA-1 (FOG-1), which is encoded by the *Zfpm1* (*Zinc finger protein multitype 1*) gene, was previously thought to be expressed primarily in cells of the erythroid and megakaryocytic lineages, where it is essential for differentiation [3,4]. FOG-1 is a zinc finger protein initially identified as an interacting partner of GATA factors which contributes to activation or repression of their target genes [3,5,6]. FOG-1 also interacts with the C-terminal binding protein (CtBP), mainly described as a corepressor and the nucleosome remodelling and histone deacetylase repressive (NuRD) complex; and, thus makes a link between transcription factors and chromatin modifiers. In addition, FOG-1 activates or represses gene transcription by facilitating binding of GATA factors to DNA [7], by recruiting chromatin remodelling complexes [5,8], or by stabilizing tissue-specific chromatin loops [9]. FOG-1 is expressed at high level in multipotent progenitors, erythroid and megakaryocytic cells, low level in lymphoid and hematopoietic stem cells; it is absent in myeloid lineages [3]. *Zfpm1*-deficient mice lack megakaryocytes and show severe defects in erythropoiesis, leading to embryonic lethality [4]. FOG-1 also plays a role in the T-lineage by repressing GATA-3-dependent induction of Th2 development [10,11]. Interestingly, overexpression of FOG-1 in avian eosinophils, which do not normally express FOG-1, reprograms these differentiated cells into multipotent cells [12], reminiscent of the reprogramming of B cells into macrophages following ectopic expression of C/EBPalpha and C/EBPbeta [13,14]. Thus, FOG-1 is essential for specific branches of the hematopoietic system, and its inappropriate expression leads to abnormal cell differentiation.

Strikingly, we have identified the expression of FOG-1 in early B lymphocytes, but not in late developmental stages such as mature B cells and plasma cells. In analogy to some of the systems mentioned above, we were intrigued by the regulated expression of FOG-1 during B cell development and hypothesized that the downregulation of FOG-1 might be a necessary step for proper differentiation and function of mature B cells. We therefore set out to test this hypothesis and for this made use of a novel transgenic mouse model strategy which we had designed to generate mice with conditional overexpression of any cDNA. Using this system, we generated transgenic mice in which FOG-1 expression was enforced at a physiologically relevant level in mature B cells or in the entire hematopoietic system. We found that sustained FOG-1 expression in mature and late B cells did not affect their development or function, contrary to our hypothesis. In contrast, overexpressing FOG-1 in the whole hematopoietic system led to a reduction in the number of circulating eosinophils, confirming and extending to mammals the previously reported role of FOG-1 in repressing avian eosinophil development.

MATERIALS AND METHODS

Ethics statement

Animal experimentation was carried out according to regulations effective in the canton of Basel-Stadt, Switzerland. All experimental procedures were approved by the Animal Committee of the Friedrich Miescher Institute for Biomedical Research and the Veterinary office of the Kanton Basel Stadt.

Generation of the targeting pR26-SA-FRT-Hygro^r vector

The backbone of a pre-existing targeting vector pR26-STOP-FRT-Hygro [15] was adapted to allow transcription from the endogenous *Rosa26* promoter. After removal of the unwanted sequences, the backbone vector contained two homology arms and the hygromycin B resistance gene flanked by an FRT3 and an FRTwt sites in 5' and 3', respectively. The 5' and 3' homology arms correspond to Chr6: 113 024 284-113 026 000 and to Chr6: 113 021 493-113 024 090 using the mm9 assembly on the UCSC genome browser. The integration site maps about 2 kb downstream of the insertion point obtained with the targeting vector pROSA26-1 [16]. The splice acceptor sequence (SA) of the STOP-eGFP-Rosa26TV vector (Adgene plasmid 11739) was PCR amplified and cloned upstream of the FRT3 site [17].

Generation of the control and FOG-1 donor vectors

The backbone of the donor vector, containing the FRT3 site, a polyA sequence and the FRTwt site, was derived from the FRT3-CAG-lox-stop-lox-Enpp1-tkNeo-FRTwt [15] vector and was further modified as follows. The loxP-Neo-STOP-loxP cassette was PCR amplified from the STOP-eGFP-Rosa26TV vector (Adgene plasmid 11739) with 5'-tagccctaggcttcgcggtctttccagtggt-3' and 3'-atgcaccggtcttcggtaccgaattgatcg-5' containing an AvrII and an AgeI site, respectively. This fragment was cloned downstream of the FRT3 site using SpeI and AgeI sites. The IRES-hCD2t fragment was obtained from the pBS-IRES-hCD2t vector (kindly provided by M. Busslinger, Vienna, [18]) and cloned downstream of the loxP-Neo-STOP-loxP cassette. The resulting control donor vector FRT3-loxP-Neo-STOP-loxP-IRES-hCD2t-FRTwt harbors a unique NotI site in between the second loxP site and the IRES sequence. The FlagFOG-1

cDNA was obtained from a pcDNA3-FlagFOG-1 vector (kindly provided by M. Crossley, Sydney) and cloned in the NotI site of the control donor vector, resulting in the 10.9 kb FOG-1 donor vector: FRT3-loxP-Neo-STOP-loxP-FlagFOG-1-IRES-hCD2t-FRTwt.

Targeting of ES cells

The SacI-linearized targeting pR26-SA-FRT-Hygro^r vector was electroporated into 129 sv jae ES cells. Electroporated ES cells were then selected with 0.1 mg/ml hygromycin B. 480 hygromycin-resistant clones were collected, and five potentially successfully recombined clones were identified by PCR screening using the following primer pair (fwd: ctactggaaagaccgcgaag, rev: tacctttctgggagttctctgc, 2 kb product).

Southern blot analysis

Verification of ES cell targeting: 10 μg of genomic DNA was analyzed by standard Southern blotting. Genomic DNA was restricted with BamHI, PstI or PvuII to confirm the 5', 3' and single integration of the targeting vector, respectively. The 5' and 3' probes were PCR amplified from genomic DNA with the following primers: 5' fwd: cgcctaaagaagaggctgtg and 5' rev: gactggagttgcagatcacg; 3' fwd: agccatctgggccttttaac and 3' rev: aagggcacagacaatccttc. The 5' probe highlighted a 5.8 kb wild-type and a 4.9 kb targeted bands. The 3' probe detected a 6.5 kb wild-type and a 7.5 kb targeted bands. The internal hygromycin probe was obtained from the targeting vector and highlighted an 8 kb band.

Verification of the RMCE targeting: 20 μg of genomic DNA was analyzed by standard Southern blotting. Genomic DNA was digested with PvuII or BglI to confirm the 5', internal and 3' single integration of the donor vector, respectively. The 5' and 3' probes were PCR amplified from genomic DNA with the following primers: 5' fwd: cgcctaaagaagaggctgtg and 5' rev: gactggagttgcagatcacg and 3' fwd: ggacaggacagtgcttgtttaagg and 3' rev: acaccacaaatgaacagtgccaag. The 5' probe highlighted a 5.8 kb wild-type and a 6.3 kb targeted bands. The 3' probe detected a 6.5 kb wild-type and a 6.2 kb targeted bands. The internal neomycin probe was obtained by PCR from the control donor vector using the following primers: Neo fwd: gaactcgtcaagaaggcgatagaag and Neo rev: gaacaagatggattgcacgcagg. It

highlighted a 3.2 kb and a 2.4 kb bands in the control clone and FOG-1 clone, respectively in addition to a 6.3 kb targeted band detected in both clones.

Recombinase-Mediated Cassette Exchange: FLP-mediated recombination of the control and FOG-1 donor vectors into the pre-targeted ES cells

The pre-targeted $R26^{Hygro}$ ES cells were thawed and cultured for 2 days on feeders in ES cell culture medium without hygromycin. 0.8 x 10^6 cells in a 6 cm dish were then transfected with a FLP-expressing vector together with the FOG-1 donor vector (FRT3-loxP-Neo-STOP-loxP-FlagFOG-1-IRES-hCD2t-FRTwt), or the control donor vector (FRT3-loxP-Neo-STOP-loxP-IRES-hCD2t-FRTwt) using Effectene reagent according to the manufacturer's instructions. One day later, the medium was replaced by ES cell medium containing 0.2 mg/ml geneticin. Ten days later, 48 or 94 colonies were picked and single-cell suspensions made by trypsin treatment. Each colony was then plated in medium containing geneticin only or geneticin and hygromycin (0.1 mg/ml). Five days later, colonies that were geneticin resistant and hygromycin sensitive were picked and seeded in 35 mm dishes for further expansion. Clones of interest were checked by PCR for correct insertion (see Supporting Figure 2A for a scheme of the strategy). The 5' insertion was verified by primer pair 1 (1F: 5'-aactcttcgcggtctttcc-3' and 1R: 3'-tctggattcatcgactgtgg-5') and the 3' insertion by primer pair 2 (2F: 5'-gccttcttgacgagttcttctgag-3' and 2R: 3'-gaaggacggtacaccagagaac-5'). A primer pair 3 (3F: 5'-aactcttcgcggtctttcc-3' and 3R: 3'-gactttccacacctggttgc-5') and a primer pair 4 (4F: 5'-agagcttggcgtaatcatgg-3' and 4R: 3'-cgtaagggattactcggtga-5') amplifying only products in the unrecombined allele were used as negative controls. For checking the integration of the neomycin cassette by PCR, the same primers than those used for the synthesis of the internal neomycin probe and specified above were used. The integration of the hCD2t sequence was checked by PCR with the following primers: fwd: tctgaagaccgatgatcagg and rev: tcattacctcacaggtcagg. The primer pair 2' (2F': 5'-aacagatggtccccagatgc-3' and 2R': 3'-agtggctcattagggaatgc-5') was used for genotyping the $R26^{FOG-1}$ mice.

Cre-expression in the recombined ES cells

10^6 $R26^{FOG-1}$ ES cells were electroporated with a Cre-expressing vector (pCAGS-nlsCre-PGK-Puro, kindly provided by D. Schübeler, Basel). Neomycin-selection was performed at a concentration of 0.2 mg/ml geneticin for 48 hrs in absence of feeders.

Mouse work

Animal experimentation was carried out according to regulations effective in the canton of Basel-Stadt, Switzerland. All experimental procedures were approved by the Animal Committee of FMI and the Veterinary office of the Kanton Basel Stadt. The mice were housed in groups of one to six animals at 25°C with a 12-h light-dark cycle (12 h light, 12 h dark) and were fed a standard laboratory diet containing 0.8% phosphorus and 1.1% calcium (NAFAG 890; Kliba, Basel, Switzerland). Food and water were provided *ad libitum*.

$R26^{FOG-1}$ ES cells were used to generate chimeric mice which were then crossed with C57Bl/6 mice to generate transgenic animals. The $R26^{FOG-1}$ mice were subsequently crossed with different Cre-expressing mouse lines to obtain overexpression of FOG-1 in specific B cell subpopulations or throughout hematopoietic lineages. In particular, they were crossed with Cd23-Cre [18], Vav-iCre [19] or mb1-Cre [20] mice to obtain $R26^{FOG-1}$:Cd23-Cre, $R26^{FOG-1}$:Vav-iCre or $R26^{FOG-1}$:mb1-Cre mice, respectively.

Mice or targeted ES cells will be made available upon request.

Cell culture and RT-PCR analysis

Cells were cultured in a humidified tissue culture incubator set up at 5% CO_2 and 37°C. Abelson lines (OBF-1 wt or OBF-$1^{-/-}$), B3, A20J, X63Ag8 and J558L cell lines were cultured in RPMI 1640 completed with 10% heat-inactivated FCS, 1% penicillin-streptomycin and 4mM L-glutamine. Total cellular RNA was extracted using RNeasy Mini Kit (Qiagen), DNaseI treated and reverse transcribed using oligo(dT) and SuperScript II RT (Invitrogen) kit according to standard procedures. Subsequent quantitative real-time PCRs were performed with the MESA GREEN qPCR MasterMix Plus for SYBR® (Eurogentec) on an ABI prism 7000 instrument. The FOG-1 primers were 5'-ccaactgtgaacgccatctc-3' and 3'-gatctcacccttggagcctg-5'. The primers specific

for transgene-derived FOG-1 (FlagFOG-1) were 5'-atggactacaaggacgacg-3' and 3'-tccatggccttggcttcttc-5'. The RNA polymerase II (RPII) primers were 5'-aggagcgccaaatgccgataa-3' and 3'-aggagcgccaaatgccgataa-5'. The GAPDH primers were 5'-TGCACCACCAACTGCTTAG-3' and 3'-TGGAAGAGTGGGAGTTGCTG-5'.

Western blot analysis

Proteins were separated on a 7-8% SDS–PAGE, transferred to a PROTRAN® Nitrocellulose Transfer Membrane (Whatman®) or to an Immobilon-P Membrane PVDF (Millipore) and immunoblotted with the appropriate primary and secondary antibodies. The following antibodies were used: anti-FOG-1 (Santa Cruz, sc-9362), anti-β-tubulin (Sigma, T4026), anti-actin (NeoMarkers, MS-1295-P1), anti-goat IgG, HRP (Abcam, ab7125), anti-mouse IgG, HRP (GE Healthcare, NA931V), anti-goat 680 and anti-mouse 680 (Molecular Probes). Signals were detected either with Amersham HyperFilm ECL (GE Healthcare) or quantified using a LI-COR Odissey instrument and the Odissey 2.1 software (Biosciences).

Flow cytometry analysis

Cells were stained in PBS-3%FCS for 30 min on ice with the following antibodies: anti-B220 (BD 553092), anti-CD25 (BD 553050), anti-IgM (Southern biotechnology 1140-02), anti-TER119 (BD 557915), anti-CD71 (BD 553267), anti-CD4 (BD 553729), anti-CD8 (BD 553032), and anti-hCD2t (R&D FAB 18561P). For sorting, the following antibody combinations were used: Pro B cells, B220+, cKit+, CD25-, IgM-; large or small Pre B, B220+, cKit-, CD25+, IgM-; immature B, B220+, IgM^{low}; splenic mature B, B220+, IgM^{high}. Cells were analyzed on a Becton Dickinson FACSCalibur or sorted on a Cytomation MoFlo instrument.

Mature B cell activation

10^6 splenic mature B cells purified using CD43-magnetic beads (Miltenyi Biotec) were cultured for 4 days in the presence of IL4 (10 ng/ml) and/or LPS (5 μg/ml) and/or anti-CD40 antibody (1 μg/ml; BD 553721).

ELISA Assay

IgM or IgG1 antibody titers in the mature B cell activation cultures were determined by standard ELISA protocol.

Full blood analysis

Tail blood samples were collected in EDTA-coated tubes and analyzed with a Sysmex XT-2000iV blood analyzer.

RESULTS

Identification of FOG-1 expression in B cells

Expression profiles of Pre B cells lacking the transcriptional coactivator OBF-1 [21,22] and transgenic thymocytes overexpressing OBF-1 [23] identified several genes that were downregulated in null cells and upregulated in overexpressing cells. Among these genes, we identified *Zfpm1*, which encodes the developmental regulator FOG-1 (Figure 1A, 1B and data not shown). *Zfpm1* had also been identified as a gene consistently activated or repressed in *Ebf1* (early B-cell factor-1) gain- and loss-of-function experiments, respectively and ChIP-Seq data demonstrated that Ebf1 binds to the promoter of *Zfpm1* within 10 kb of the transcription start site [24].

In agreement with previous results [3], we detected a lower expression level of FOG-1 in total bone marrow B cells than in red blood cells (Figure 1C). Microarray and quantitative RT-PCR analysis demonstrated that FOG-1 was expressed from Pro B cell to immature B cell stages at a relatively high level and was downregulated in mature B cells and plasma cells. This specific expression pattern was observed in primary cells and also in cultured cell lines representative of different B cell developmental stages (Figure 1D, 1E and 1F).

Together these results show that FOG-1 is expressed in a regulated manner during B cell development and suggest that this factor may play a role not previously appreciated in this lineage. To examine this in greater detail, we wished to test the effect of overexpressing FOG-1 in B cells, in particular in late stages, hypothesizing that this might affect their differentiation or function. For this, we made use of a system that we had designed to generate mice overexpressing a gene of interest in a conditional manner.

Pre-targeting of the *Rosa26* locus

Our strategy has been to generate ES cells pre-targeted at the *Rosa26* (R26) locus, so that appropriate expression constructs can rapidly be inserted by recombination mediated cassette exchange (RMCE, see Figure 2). We chose to use the *Rosa26* locus, which was first described in a gene trap experiment and was shown to be expressed in the whole mouse [25]. This locus is believed to encode two non-coding transcripts and an antisense transcript of unknown function and can be used to drive the expression of any cDNA [25]. In the original gene trap experiment, the insertion of a splice acceptor and a promoter-less cDNA in intron 1 of the gene led to expression of this cDNA from the endogenous *Rosa26* promoter. For these reasons, we decided to target the *Rosa26* locus and we used 5' and 3' homology arms located in intron 1. The targeting vector containing a splice acceptor sequence to allow expression from the endogenous *Rosa26* promoter and an hygromycin B resistance gene flanked by FRT3 and FRTwt sites was linearized and introduced by electroporation into 129 sv jae ES cells to generate the pre-targeted $R26^{Hygro}$ allele (Figure 2, step 1 and Supporting Figure 1A). 480 hygromycin-resistant ES cell clones were first screened by PCR using a forward primer located upstream of the 5' homology arm and a reverse primer located in the hygromycin cassette. Four clones (1-4, Supporting Figure 1B) showed the expected 2kb band for successful homologous recombination, as compared to the aberrant product obtained for clone 5. To further confirm the correct homologous recombination, we performed Southern blot analysis using 5' and 3' probes with BamHI or PstI digested genomic DNA blots. Clones 1-4 showed the expected 4.9 kb band for correct integration of the selection cassette in 5' (Supporting Figure 1C), as well as the 7.5 kb band for correct integration at the 3' end (Supporting Figure 1D). As a control, clone 5 only showed the wild-type bands. To ensure that a single copy was integrated in the genome, an additional Southern blot analysis was performed with a hygromycin probe and PvuII digested genomic DNA blots. Clones 2-4 showed a single band at the expected size, whilst clone 1 showed an additional smaller band, suggesting multiple insertions of the transgene (Supporting Figure 1E). Thus, our targeting vector was successfully homologously recombined into the *Rosa26* locus to generate the pre-

targeted R26^{Hygro} allele. Clone 4 was selected as our R26^{Hygro} ES cell clone for further use.

Efficient recombination of the FOG-1 cassette in the pre-targeted R26^{Hygro} locus

The combined use of the heterospecific FRT3/FRTwt sites allows replacement of a target DNA by an incoming plasmid donor cassette upon transient FLP expression [26]. Hence, any cassette of interest flanked by FRT3/FRTwt sites can readily be introduced into pre-targeted R26^{Hygro} ES cells upon transient expression of the FLP recombinase (Figure 2, step 2). Cassettes of interest contain a loxP-Neo-STOP-loxP sequence upstream of the gene of interest to allow neomycin selection of the recombined clones and Cre-dependent expression of the transgene from the endogenous *Rosa26* promoter. The cassettes also contain an internal ribosome entry site (IRES) sequence derived from the Encephalomyocarditis virus placed downstream of the cDNA of interest to allow the concomitant expression of a "reporter" gene (here a truncated version of the human cell surface marker CD2, hCD2t), and selective monitoring of the recombined cells. In order to study the effect of enforced expression of FOG-1 in transgenic mice, two donor vectors were generated: (i) a control donor vector with loxP-Neo-STOP-loxP and the IRES-hCD2t sequences, but no gene of interest (Supporting Figure 2A), (ii) the FOG-1 donor vector which contained a cDNA encoding Flag-tagged FOG-1 downstream of the loxP-Neo-STOP-loxP sequence and upstream of the IRES-hCD2t sequence (Figure 3A and Supporting Figure 2A).

To test the efficiency of our system, RMCE experiments were performed in R26^{Hygro} ES cell clone 4: cells were transfected with either the FOG-1 donor vector or the control donor vector, together with an expression vector encoding the FLP recombinase. The R26^{Hygro} ES cells are resistant to hygromycin B and, upon successful RMCE, become sensitive to this antibiotic, while acquiring neomycin-resistance (see Figure 2). In two independent RMCE experiments, a total of 142 Neo^{R} colonies were picked for each vector and were tested for hygromycin B resistance, as described in the Materials and Methods section. We found that at this step 48.5% of the FOG-1 colonies and 62.5% of the control colonies were both neomycin-resistant and hygromycin-sensitive, indicative of successful RMCE. We next selected 12

$Neo^R/Hygro^S$ clones of each kind (control or FOG-1 vector) for an extensive PCR analysis which demonstrated that all the clones analyzed were properly recombined (Supporting Figure 2A-B). Successful integration of the neomycine cassette and of the hCD2t cassette in the genome of these ES cell clones was also tested and all clones were positive for these PCRs as well (Supporting Figure 2C-D). Thus, using the system described here, we efficiently recombined a large cassette of 5.0 kb or 8.0 kb at the $R26^{Hygro}$ allele of the pre-targeted ES cells to generate ES cell clones carrying the $R26^{Control}$ allele or $R26^{FOG\text{-}1}$ allele, respectively. Based on the extensive PCR analysis presented here, 100% of the $Neo^R/Hygro^S$ clones appear to be correctly recombined upon RMCE. We next chose one control clone and one FOG-1 clone for verifying the successful RMCE by Southern blot analysis and used genomic digests and probes allowing us to interrogate the 5' and the 3' boundaries, as well as the copy number. This analysis as well confirmed the correct recombination of the targeting vectors into the *Rosa26* locus (Figure 3A-B).

Before generating mice, we tested the Cre-inducible expression of FOG-1 and hCD2t in our recombined FOG-1 clone. To this end, a Cre-expressing vector was transiently transfected in the targeted ES cells and expression of hCD2t and FOG-1 was tested two days later by flow cytometry and western blot analysis, respectively. As expected, the recombined $R26^{FOG\text{-}1}$ ES cells expressed hCD2t and FOG-1 only when Cre was expressed (Figure 4A-B). The partial expression of hCD2t observed by flow cytometry is due to the experimental settings. Here, transfected ES cells were only selected for a short period of time to avoid FOG-1-induced cell cycle arrest in ES cells [27]. As a result of this, not all ES cells were expressing Cre, leading to a heterogeneous cell population.

In conclusion, we efficiently recombined our FOG-1 donor vector in the pre-targeted $R26^{Hygro}$ locus and demonstrated that, upon Cre expression, both FOG-1 and hCD2t were expressed in these ES cells. We therefore used the $R26^{FOG\text{-}1}$ ES cell clone to generate transgenic $R26^{FOG\text{-}1}$ mice.

Enforced expression of FOG-1 does not affect B cell differentiation or function

To investigate the potential effect of sustained level of FOG-1 in mature B cells, we first crossed the $R26^{FOG\text{-}1}$ mice with Cd23-Cre mice which express Cre specifically in mature B cells [18]. Flow cytometry analysis of hCD2t expression in B cells derived from $R26^{FOG\text{-}1}$:Cd23-Cre mice showed that in the early developmental stages (Pre B cells), no hCD2t could be detected, as expected (Figure 5A, second panel). In contrast, more than 85% of the mature B cells derived from these animals were found to express hCD2t (Figure 5A, fourth panel). Importantly, the expression of hCD2t was never detected in control animals which carry the $R26^{FOG\text{-}1}$ allele but do not express Cre. Next, quantitative RT-PCR and western blot analysis were performed to estimate the level of FOG-1 overexpression in mature B cells derived from $R26^{FOG\text{-}1}$:Cd23-Cre animals. As shown in Figure 5B, FOG-1 mRNA was upregulated slightly more than three fold in the overexpressing cells, as compared to control cells. At the protein level, the increase was even larger: careful quantification of the FOG-1 signal in relation to the expression of actin showed that the protein was upregulated about 6-fold (Figure 5C). Overall this analysis demonstrated that our system allows reliable overexpression of FOG-1 *in vivo*.

To investigate the potential biological effect of this elevated level of FOG-1 on plasma cell development, mature B cells derived from $R26^{FOG\text{-}1}$:Cd23-Cre mice were activated *in vitro* with different stimuli and the antibody titers in the culture supernatants were determined by ELISA. Irrespective of the stimulus used, no difference in the level of IgM or IgG1 was observed between overexpressing and control cultures, indicating that enforced expression of FOG-1 in mature B cells did not impair or alter their ability to differentiate into antibody-secreting cells in vitro (Figure 5D). In addition, we also performed immunization experiments to test whether FOG-1 overexpression in mature B cells might have an impact on the immune response *in vivo*. For this, $R26^{FOG\text{-}1}$:Cd23-Cre and control mice were immunized with DNP-KLH and the serum titers of antigen-specific immunoglobulins were measured by ELISA 8 and 15 days later. However, also in this case, no significant difference was found (data not shown). Thus, although the expression of FOG-1 is normally downregulated during B cell development, enforcing expression of this factor at late B cell stages did not reveal any detrimental effect, contrary to our

hypothesis. Finally, $R26^{FOG-1}$ mice were also crossed with mb1-Cre mice to induce FOG-1 overexpression from the earliest stages of B cell development. However, this did not impact B cell development as examined by flow cytometric analysis (data not shown).

Reduction of eosinophil numbers upon enforced expression of FOG-1

To further analyze the consequences of an elevated level of FOG-1 in the hematopoietic system, we next crossed the $R26^{FOG-1}$ mice with Vav-iCre mice [19] to overexpress FOG-1 in all hematopoietic lineages. Remarkably, all B lymphocytes, myeloid cells, and erythrocytes derived from the bone marrow of $R26^{FOG-1}$:Vav-iCre animals expressed hCD2t (Figure 6A, 6B and 6C). Similarly, all thymocytes and splenic B lymphocytes as well as all splenic erythrocytes expressed hCD2t (Figure 6D, 6E and 6F), thus demonstrating the usefulness of our reporter system. Expression of transgene-derived FOG-1 mRNA was analyzed by quantitative RT-PCR and showed that FlagFOG-1 mRNA is produced at a similar level in bone marrow, spleen, and thymus of $R26^{FOG-1}$:Vav-iCre mice (Supporting Figure 3). Altogether these data indicated that in $R26^{FOG-1}$:Vav-iCre animals virtually all hematopoietic cells express the transgene at a roughly similar level.

The total numbers of bone marrow cells, splenocytes, and thymocytes in overexpressing mice were comparable to the numbers obtained in control animals (Figure 7A). Using a panel of antibodies against lineage-specific surface markers we analyzed by flow cytometry the major hematopoietic cell populations. $R26^{FOG-1}$:Vav-iCre mice showed normal B cell development (Figure 7B-C), normal T cell development (Figure 7D-E), as well as a normal myeloid population (Figure 7G). Developing erythrocytes can be subdivided into early and late erythroblasts based on the expression of the cell surface markers TER119 and CD71 [28]. Using this method, we found a largely normal erythroid development, although a slight increase of the early erythroblast population (CD71+ TER119+) was seen in the spleen of some (≤50%) of the $R26^{FOG-1}$:Vav-iCre animals (exemplified in Figure 7H). However, no difference was observed in the bone marrow (Figure 7F). More detailed analysis will be required to understand this phenotype and the reason for its partial penetrance. Full blood count analysis of $R26^{FOG-1}$:Vav-iCre females and males

revealed no major difference in red blood cell count, hemoglobin content, hematocrit and platelet counts (Table 1). Similarly, no significant changes were observed in the total numbers of white blood cells, lymphocytes, monocytes, neutrophils and basophils in the blood of these mice (Table 1). In contrast, a consistent decrease in the total number of circulating eosinophils was observed for $R26^{FOG\text{-}1}$:Vav-iCre females and males (Table 1). A highly reproducible and significant difference of more than 3-fold was found when comparing the average numbers of circulating eosinophils in control and $R26^{FOG\text{-}1}$:Vav-iCre animals (Figure 8).

DISCUSSION

We have shown here that FOG-1 expression is regulated during B cell development, being high in early stages (e.g. pre B cells) and low or absent in late stages such as mature B cells and plasma cells. Based on these observations, we hypothesized that this downregulation is important for effective B cell development and that artificially maintaining FOG-1 expression at late stages might have an impact on normal B cell development or function. However, we did not observe any major phenotypes when FOG-1 was selectively overexpressed in mature B cells. Flow cytometric analysis of the major B cell populations failed to reveal alterations in $R26^{FOG\text{-}1}$:Cd23-Cre mice, demonstrating no obvious developmental impact. Furthermore, *in vivo* immune response to a T cell-dependent antigen or antibody secretion following *in vitro* stimulation were also not affected, thus suggesting that B cell function was not impaired. In addition, mice overexpressing FOG-1 from the pre B cell stage onwards also did not show any remarkable phenotype. Nevertheless, it remains possible that FOG-1 affects a B cell subset that was not examined here and additional experiments will be required to determine this. Furthermore, since B cells do not express GATA factors (our unpublished data), other B cell ancillary factors must be postulated for a role for FOG-1 in these and potentially other cells (Versavel *et al*., in preparation).

In contrast, $R26^{FOG\text{-}1}$:Vav-iCre mice with enforced expression of FOG-1 in the entire hematopoietic system showed two main phenotypes. First, consistent with the known role of FOG-1 in red blood cell development [4,6,29,30], we observed a moderately altered erythropoiesis in the spleen of some (≤50%) $R26^{FOG\text{-}1}$:Vav-iCre

animals. The reason for the partial penetrance of this phenotype is unclear and additional work will be required. Second, we found a striking and highly significant reduction of the total number of circulating eosinophils. This is of great interest since a previous study concluded that FOG-1 is a repressor of the eosinophil lineage in avian cells [12]. Our results, therefore, extend this finding to mammals. In the future, the $R26^{FOG\text{-}1}$:Vav-iCre mice will be analyzed further to draw a more complete view of FOG-1 functions in hematopoiesis.

The Recombinase-Mediated Cassette Exchange technology relies on the exquisite selectivity of recombinases such as Cre or FLP and has been developed to facilitate the generation of transgenic ES cell lines [26,31] [31-37] [38,39]. To allow inducible or lineage-specific expression, the RMCE technology has been combined with tetracycline-inducible systems [40-42], or with the Cre recombinase activity [43,44]. Recent studies made use of a pre-targeted locus to generate shRNA-expressing mice [45,46] and Hitz *et al*. combined the C31 integrase-mediated recombination at a pre-targeted locus with Cre-dependent expression to establish shRNA-expressing mice [47]. Other systems that allow RMCE into pre-modified ES cells [48,49] or in human ES cells pre-targeted at the *HPRT* locus [50] have also been developed.

To alter the expression pattern of FOG-1 *in vivo*, we developed a novel and rapid transgenic system that is also based on the RMCE technology and that allows rapid insertion of expression cassettes into the *Rosa26* locus. In this system, expression of the transgene from the endogenous *Rosa26* promoter is dependent on Cre recombinase-mediated excision of a STOP sequence, allowing cell- or temporal-specific control of expression. A hCD2t cDNA is also included as a reporter to track transgene-expressing cells *in vitro* and *in vivo*.

Our strategy is broadly similar to the one described by Hitz and colleagues, where C31 integrase-mediated recombination is used instead of FLP-mediated recombination to insert a transgene at a modified *Rosa26* locus and where the loxP/Cre system is used for conditional expression of shRNAs [47]; the RMCE efficiency of the two systems is also comparable. A particularly useful feature of our system is the concomitant expression of a "marker" gene, hCD2t, together with our transgene (i.e. FOG-1); this allows to monitor Cre-recombined cells *ex vivo* or *in vivo*.

Expression of hCD2t can conveniently be detected by flow cytometry and is therefore particularly well suited for analyses in the hematopoietic system. Furthermore, since the antibodies recognizing hCD2t are species-specific, expression of the endogenous mouse protein (e.g. by T cells) does not interfere. This is especially useful in situations where Cre expression, and therefore transgene expression, is only partial or mosaic and leads to a mixture of recombined and non-recombined cells; having a marker gene such as hCD2t then allows to selectively examine –and potentially isolate- the transgene-expressing cells. Finally, hCD2t can also be detected by immunohistochemistry, further extending the range of cells that can be selectively examined.

Using this system, we generated mice with moderate overexpression of FOG-1 either in mature B cells ($R26^{FOG-1}$:Cd23-Cre mice), throughout B cell development ($R26^{FOG-1}$:mb1-Cre mice), or in all hematopoietic lineages ($R26^{FOG-1}$:Vav-iCre mice). Remarkably, flow cytometry analysis of bone marrows and spleens derived from $R26^{FOG-1}$:Vav-iCre mice revealed that all hematopoietic cells expressed hCD2t. In contrast and as expected, in $R26^{FOG-1}$:Cd23-Cre mice, where Cre starts being expressed just before the mature B cell stage, hCD2t was only detected in mature B cells. Importantly, cells derived from control $R26^{FOG-1}$ mice lacking Cre expression did not express any hCD2t. These results demonstrate that conditional expression of a transgene using our system is tightly regulated and underscore the utility of having a marker gene.

We found that the *Rosa26* promoter drives moderate FOG-1 expression in the hematopoietic system. Interestingly, this expression level was sufficient to marginally alter splenic erythropoiesis and to significantly reduce the number of circulating eosinophils in $R26^{FOG-1}$:Vav-iCre mice, demonstrating its physiological relevance. Such moderate expression level is an advantage to unravel the roles of proteins in physiologic and pathologic situations, as it avoids phenotypes that may be partly caused by too strong overexpression. Since targeted *Rosa26* homozygous mice are viable and appear normal, the expression level of the transgene could be doubled by generating $R26^{FOG-1}$/$R26^{FOG-1}$ homozygous mice [16] or by inserting additional regulatory elements in the expression cassette. Following this idea, Tchorz

et al. generated a modified RMCE-compatible *Rosa26* locus for the expression of transgenes and characterized several promoters with different strengths [15].

In conclusion, despite our finding that the expression of FOG-1 is tightly regulated throughout B cell differentiation and is dependent on the B cell-specific coactivator OBF-1, we could not demonstrate a role for FOG-1 in B cell differentiation or function. However, we could confirm the finding that FOG-1 is a negative regulator of eosinophil development and extend it to mammals. Further work will be required to better understand this important function of FOG-1 in the mouse.

ACKNOWLEDGEMENTS

We thank M. Busslinger for providing the Cd23-Cre mice and the pBS-IRES-hCD2t vector, M. Crossley for providing the pcDNA3-FlagFOG-1 vector, D. Kioussis for providing the Vav-iCre mice. We also thank N. Reichert, F. Brellier, T. Rolink, R. G. Clerc and F. Cubizolles for useful discussions and critical reading of the manuscript.

AUTHOR CONTRIBUTIONS

Conceived and designed the experiments: CDR, MK, AV, MM, PM. Performed the experiments: CDR, AV, MK, CC, GM, TD, PK, JFS, MM, PM. Analyzed the data: CDR, AV, MM, PM. Contributed reagents/materials/analysis tools: MM. Wrote the paper: CDR, PM, AV.

REFERENCES

1. Matthias P, Rolink AG (2005) Transcriptional networks in developing and mature B cells. Nature reviews Immunology 5: 497-508.
2. Orkin SH, Zon LI (2008) Hematopoiesis: an evolving paradigm for stem cell biology. Cell 132: 631-644.
3. Cantor AB, Orkin SH (2005) Coregulation of GATA factors by the Friend of GATA (FOG) family of multitype zinc finger proteins. Seminars in cell & developmental biology 16: 117-128.
4. Tsang AP, Fujiwara Y, Hom DB, Orkin SH (1998) Failure of megakaryopoiesis and arrested erythropoiesis in mice lacking the GATA-1 transcriptional cofactor FOG. Genes Dev 12: 1176-1188.
5. Fox AH, Liew C, Holmes M, Kowalski K, Mackay J, et al. (1999) Transcriptional cofactors of the FOG family interact with GATA proteins by means of multiple zinc fingers. The EMBO journal 18: 2812-2822.
6. Tsang AP, Visvader JE, Turner CA, Fujiwara Y, Yu C, et al. (1997) FOG, a multitype zinc finger protein, acts as a cofactor for transcription factor GATA-1 in erythroid and megakaryocytic differentiation. Cell 90: 109-119.
7. Pal S, Cantor AB, Johnson KD, Moran TB, Boyer ME, et al. (2004) Coregulator-dependent facilitation of chromatin occupancy by GATA-1. Proceedings of the National Academy of Sciences of the United States of America 101: 980-985.
8. Hong W, Nakazawa M, Chen YY, Kori R, Vakoc CR, et al. (2005) FOG-1 recruits the NuRD repressor complex to mediate transcriptional repression by GATA-1. The EMBO journal 24: 2367-2378.
9. Vakoc CR, Letting DL, Gheldof N, Sawado T, Bender MA, et al. (2005) Proximity among distant regulatory elements at the beta-globin locus requires GATA-1 and FOG-1. Molecular cell 17: 453-462.
10. Kurata H, Lee HJ, McClanahan T, Coffman RL, O'Garra A, et al. (2002) Friend of GATA is expressed in naive Th cells and functions as a repressor of GATA-3-mediated Th2 cell development. J Immunol 168: 4538-4545.
11. Zhou M, Ouyang W, Gong Q, Katz SG, White JM, et al. (2001) Friend of GATA-1 represses GATA-3-dependent activity in CD4+ T cells. J Exp Med 194: 1461-1471.
12. Querfurth E, Schuster M, Kulessa H, Crispino JD, Doderlein G, et al. (2000) Antagonism between C/EBPbeta and FOG in eosinophil lineage commitment of multipotent hematopoietic progenitors. Genes Dev 14: 2515-2525.
13. Rapino F, Robles EF, Richter-Larrea JA, Kallin EM, Martinez-Climent JA, et al. (2013) C/EBPalpha induces highly efficient macrophage transdifferentiation of B lymphoma and leukemia cell lines and impairs their tumorigenicity. Cell reports 3: 1153-1163.
14. Xie H, Ye M, Feng R, Graf T (2004) Stepwise reprogramming of B cells into macrophages. Cell 117: 663-676.
15. Tchorz JS, Suply T, Ksiazek I, Giachino C, Cloetta D, et al. (2012) A modified RMCE-compatible Rosa26 locus for the expression of transgenes from exogenous promoters. PloS one 7: e30011.
16. Soriano P (1999) Generalized lacZ expression with the ROSA26 Cre reporter strain. Nat Genet 21: 70-71.
17. Sasaki Y, Derudder E, Hobeika E, Pelanda R, Reth M, et al. (2006) Canonical NF-kappaB activity, dispensable for B cell development, replaces BAFF-

receptor signals and promotes B cell proliferation upon activation. Immunity 24: 729-739.
18. Kwon K, Hutter C, Sun Q, Bilic I, Cobaleda C, et al. (2008) Instructive role of the transcription factor E2A in early B lymphopoiesis and germinal center B cell development. Immunity 28: 751-762.
19. de Boer J, Williams A, Skavdis G, Harker N, Coles M, et al. (2003) Transgenic mice with hematopoietic and lymphoid specific expression of Cre. Eur J Immunol 33: 314-325.
20. Hobeika E, Thiemann S, Storch B, Jumaa H, Nielsen PJ, et al. (2006) Testing gene function early in the B cell lineage in mb1-cre mice. Proc Natl Acad Sci U S A 103: 13789-13794.
21. Bordon A, Bosco N, Du Roure C, Bartholdy B, Kohler H, et al. (2008) Enforced expression of the transcriptional coactivator OBF1 impairs B cell differentiation at the earliest stage of development. PloS one 3: e4007.
22. 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.
23. Bartholdy B, Du Roure C, Bordon A, Emslie D, Corcoran LM, et al. (2006) The Ets factor Spi-B is a direct critical target of the coactivator OBF-1. Proceedings of the National Academy of Sciences of the United States of America 103: 11665-11670.
24. Treiber T, Mandel EM, Pott S, Gyory I, Firner S, et al. (2010) Early B cell factor 1 regulates B cell gene networks by activation, repression, and transcription-independent poising of chromatin. Immunity 32: 714-725.
25. Zambrowicz BP, Imamoto A, Fiering S, Herzenberg LA, Kerr WG, et al. (1997) Disruption of overlapping transcripts in the ROSA beta geo 26 gene trap strain leads to widespread expression of beta-galactosidase in mouse embryos and hematopoietic cells. Proc Natl Acad Sci U S A 94: 3789-3794.
26. Schlake T, Bode J (1994) Use of mutated FLP recognition target (FRT) sites for the exchange of expression cassettes at defined chromosomal loci. Biochemistry 33: 12746-12751.
27. Tanaka M, Zheng J, Kitajima K, Kita K, Yoshikawa H, et al. (2004) Differentiation status dependent function of FOG-1. Genes Cells 9: 1213-1226.
28. Du Roure C, Takacs K, Maxwell PH, Roberts I, Dazzi F, et al. (2006) Correction of severe anaemia using immuno-regulated gene therapy is achieved by restoring the early erythroblast compartment. Br J Haematol 132: 608-614.
29. Harju-Baker S, Costa FC, Fedosyuk H, Neades R, Peterson KR (2008) Silencing of Agamma-globin gene expression during adult definitive erythropoiesis mediated by GATA-1-FOG-1-Mi2 complex binding at the -566 GATA site. Mol Cell Biol 28: 3101-3113.
30. Katz SG, Cantor AB, Orkin SH (2002) Interaction between FOG-1 and the corepressor C-terminal binding protein is dispensable for normal erythropoiesis in vivo. Mol Cell Biol 22: 3121-3128.
31. Sauer B, Henderson N (1990) Targeted insertion of exogenous DNA into the eukaryotic genome by the Cre recombinase. New Biol 2: 441-449.
32. Dymecki SM (1996) Flp recombinase promotes site-specific DNA recombination in embryonic stem cells and transgenic mice. Proc Natl Acad Sci U S A 93: 6191-6196.

33. Dymecki SM, Tomasiewicz H (1998) Using Flp-recombinase to characterize expansion of Wnt1-expressing neural progenitors in the mouse. Dev Biol 201: 57-65.
34. O'Gorman S, Fox DT, Wahl GM (1991) Recombinase-mediated gene activation and site-specific integration in mammalian cells. Science 251: 1351-1355.
35. Rodriguez CI, Buchholz F, Galloway J, Sequerra R, Kasper J, et al. (2000) High-efficiency deleter mice show that FLPe is an alternative to Cre-loxP. Nat Genet 25: 139-140.
36. Seibler J, Bode J (1997) Double-reciprocal crossover mediated by FLP-recombinase: a concept and an assay. Biochemistry 36: 1740-1747.
37. Soukharev S, Miller JL, Sauer B (1999) Segmental genomic replacement in embryonic stem cells by double lox targeting. Nucleic Acids Res 27: e21.
38. Branda CS, Dymecki SM (2004) Talking about a revolution: The impact of site-specific recombinases on genetic analyses in mice. Dev Cell 6: 7-28.
39. Schnutgen F, Stewart AF, von Melchner H, Anastassiadis K (2006) Engineering embryonic stem cells with recombinase systems. Methods Enzymol 420: 100-136.
40. Beard C, Hochedlinger K, Plath K, Wutz A, Jaenisch R (2006) Efficient method to generate single-copy transgenic mice by site-specific integration in embryonic stem cells. Genesis 44: 23-28.
41. Masui S, Shimosato D, Toyooka Y, Yagi R, Takahashi K, et al. (2005) An efficient system to establish multiple embryonic stem cell lines carrying an inducible expression unit. Nucleic Acids Res 33: e43.
42. Wutz A, Rasmussen TP, Jaenisch R (2002) Chromosomal silencing and localization are mediated by different domains of Xist RNA. Nat Genet 30: 167-174.
43. Hohenstein P, Slight J, Ozdemir DD, Burn SF, Berry R, et al. (2008) High-efficiency Rosa26 knock-in vector construction for Cre-regulated overexpression and RNAi. Pathogenetics 1: 3.
44. Nyabi O, Naessens M, Haigh K, Gembarska A, Goossens S, et al. (2009) Efficient mouse transgenesis using Gateway-compatible ROSA26 locus targeting vectors and F1 hybrid ES cells. Nucleic Acids Res 37: e55.
45. Seibler J, Kuter-Luks B, Kern H, Streu S, Plum L, et al. (2005) Single copy shRNA configuration for ubiquitous gene knockdown in mice. Nucleic Acids Res 33: e67.
46. Seibler J, Kleinridders A, Kuter-Luks B, Niehaves S, Bruning JC, et al. (2007) Reversible gene knockdown in mice using a tight, inducible shRNA expression system. Nucleic Acids Res 35: e54.
47. Hitz C, Wurst W, Kuhn R (2007) Conditional brain-specific knockdown of MAPK using Cre/loxP regulated RNA interference. Nucleic Acids Res 35: e90.
48. Schebelle L, Wolf C, Stribl C, Javaheri T, Schnutgen F, et al. (2010) Efficient conditional and promoter-specific in vivo expression of cDNAs of choice by taking advantage of recombinase-mediated cassette exchange using FlEx gene traps. Nucleic acids research 38: e106.
49. Sandhu U, Cebula M, Behme S, Riemer P, Wodarczyk C, et al. (2011) Strict control of transgene expression in a mouse model for sensitive biological applications based on RMCE compatible ES cells. Nucleic acids research 39: e1.

50. Sakurai K, Shimoji M, Tahimic CG, Aiba K, Kawase E, et al. (2010) Efficient integration of transgenes into a defined locus in human embryonic stem cells. Nucleic acids research 38: e96.

FIGURE LEGENDS

Figure 1: FOG-1 is expressed in a regulated manner during B cell development.

A. Gene expression level of FOG-1 determined by Affymetrix microarray in wild-type or *OBF-1*$^{-/-}$ Abl Pro B cells. **B**. Reverse transcriptase (RT)-PCR analysis of FOG-1 mRNA level in wild-type or *OBF-1*$^{-/-}$ Abl Pro B cells. GAPDH mRNA amount was used as a control. +, - refer to cDNA synthesis reactions performed in the presence or absence of reverse transcriptase (RT), respectively. **C**. mRNA expression level of FOG-1 in bone marrow B cells (B220^{+}) and in red blood cells (TER 119^{+}) determined by semi-quantitative RT-PCR. Actin mRNA amount was used as a reference. **D**. mRNA expression level of FOG-1 during B cell development determined by semi-quantitative RT-PCR. The analysis was performed on primary Pro B, large Pre B, small Pre B, immature B and splenic mature B cells. Actin mRNA amount was used as a reference. **E**. mRNA expression level of FOG-1 measured by real-time RT-PCR during the B cell development. The analysis was performed on primary B cells as in (**D**). RNA Polymerase II mRNA was used as a reference. **F**. mRNA expression level of FOG-1 in B cell lines representing different stages of B cell development, determined by RT-PCR. From early to late stages, the analysis was performed on the five following cell lines : wild-type Abelson (Abl+/+), B3, A20J, X63 and J558. GAPDH mRNA was used as a control.

Figure 2: Strategy for the rapid generation of conditionally overexpressing ES cells

Schematic representation of the wild-type *Rosa26* (R26) locus, the pre-targeted R26^{Hygro} and the recombined R26 alleles. The *Rosa26* locus was first pre-targeted by homologous recombination with a cassette containing a splice acceptor and a hygromycin B resistance gene flanked by a FRT3 site in 5' and a FRTwt site in 3' to generate the pre-targeted R26^{Hygro} allele (step 1). The hygromycin B cassette is then replaced by the cassette of interest via the FRT sites using transient expression of the FLP recombinase (step 2). Our cassette of interest contains a loxP-Neo-STOP-loxP cassette for neomycin selection of ES cells and conditional expression of the transgene, and an IRES-hCD2t sequence to monitor expressing cells. Notably, Cre-mediated recombination will excise the neomycin gene from the targeted allele in

the transgenic mice. Neo: neomycin resistance gene. R26 pro: promoter of the R26 gene.

Figure 3: Correct recombination at the pre-targeted $R26^{Hygro}$ allele by RMCE

A. Schematic representation of the wild-type R26, pre-targeted $R26^{Hygro}$ and recombined $R26^{FOG\text{-}1}$ alleles. In the FOG-1 cassette, the neomycin resistance gene (Neo, black) and the STOP sequence (red) are flanked by LoxP sites (white triangles) to allow conditional expression. The cDNA encoding FlagFOG-1 (pink) followed by an IRES sequence and the sequence coding for hCD2t (pale blue) were inserted downstream. A polyA signal was placed at the 3' end of the hCD2t coding sequence (purple oval). The probes (green bars) used for Southern blot analysis are shown. **B**. Correct insertion of the vectors was confirmed by Southern blotting. To test the 5' boundary, PvuII digested ES cell genomic DNA was hybridized with the radioactively labeled 5' probe to detect the wild-type (5.8 kb) and the targeted (6.3 kb) bands (Left panel). To verify single-copy insertion, PvuII digested DNA was hybridized with a radioactively labeled internal probe to detect the 3.2 kb or 2.4 kb targeted bands in control or FOG-1 clones, respectively, in addition to the 6.3 kb targeted band (Middle panel). To test the 3' boundary, BglI digested DNA was hybridized with the radioactively labeled 3' probe to detect the 6.5 kb wild-type and the 6.2 kb targeted bands (Right panel).

Figure 4: Cre-dependent expression of hCD2t and FOG-1 in recombined ES cells

$R26^{FOG\text{-}1}$ ES cells were electroporated with a Cre expressing vector. Cre-induced expression of hCD2t and FOG-1 was detected by flow cytometry (**A**) and western blot (**B**) analysis, respectively.

Figure 5: Overexpression of FOG-1 in mature B cells

A. In $R26^{FOG\text{-}1}$:Cd23-Cre mice, hCD2t expression is restricted to mature B cells. Cell surface expression of hCD2t was analyzed by flow cytometry in Pre B cells (B220+, CD25+) and in mature B cells (B220+, IgM^{high}) of control ($R26^{FOG\text{-}1}$ without Cre) and $R26^{FOG\text{-}1}$:Cd23-Cre mice. Representative results of at least 3 independent experiments are shown. **B.** FOG-1 mRNA is increased 3-fold in $R26^{FOG\text{-}1}$:Cd23-Cre

mature B cells. RNA extracted from 3 control ($R26^{FOG-1}$ without Cre) and 3 $R26^{FOG-1}$:Cd23-Cre mice was reverse transcribed and subjected to quantitative PCR to detect FOG-1 and RNA Polymerase II (RPII, for normalization) transcripts. Standard error of the mean and p-value are shown. **C.** FOG-1 protein is up-regulated ca. 6-fold in mature B cells derived from $R26^{FOG-1}$:Cd23-Cre mice. FOG-1 and actin proteins were detected by western blotting in mature B cells derived from 3 control and 3 $R26^{FOG-1}$:Cd23-Cre mice (upper panel). The band intensities were quantified by LiCor Odyssey scanning and normalized to expression of actin (lower panel). Standard error of the mean and p-value are shown. **D**. FOG-1-overexpressing mature B cells respond normally to *in vitro* stimulation. Splenic resting mature B cells isolated from 3 to 6 control (black dots) or $R26^{FOG-1}$:Cd23-Cre mice (grey dots) were activated *in vitro* by LPS, LPS+IL4 or anti-CD40+IL4 for 4 days. Titers of IgM (left panel) or IgG1 (right panel) in the culture supernatants were determined by ELISA; means are shown (red bar).

Figure 6: hCD2t is expressed in all hematopoietic cells of $R26^{FOG-1}$:Vav-iCre mice.

A-F. Flow cytometry analysis of hCD2t expression in $R26^{FOG-1}$:Vav-iCre (red line) and control ($R26^{FOG-1}$ without Cre, blue line) mice. **A.** Bone marrow B lymphocytes (B220+ cells). **B**. Bone marrow granular cells (based on Forward and Side Scatters). **C**. Bone marrow erythrocytes (TER119+ cells). **D**. Thymocytes (CD4+, CD8+ cells). **E**. Splenic B lymphocytes (B220+ cells). **F**. Splenic erythrocytes (TER119+ cells). Data for one representative animal of each genotype are shown (n = 5).

Figure 7: Normal B cell, T cell and granular cell populations in $R26^{FOG-1}$:Vav-iCre mice

A. Cells of the bone marrow (BM), spleen (Spl) and thymus (Thy) of 3 $R26^{FOG-1}$ (black bars) and 3 $R26^{FOG-1}$:Vav-iCre (grey bars) mice were enumerated. Standard error of the mean is shown. **B**. Bone marrow cells were stained with anti-B220 and anti-IgM antibodies to analyze B cell development. **C**. Splenocytes were stained with anti-B220 and anti-IgM antibodies to identify B cells. **D**. Thymocytes were stained with anti-CD4 and anti-CD8 antibodies to analyze T cell development. **E**. Splenocytes were stained with anti-CD4 and anti-CD8 antibodies to identify mature T cells. **F**. Bone

marrow cells were stained with anti-TER119 and anti-CD71 antibodies to analyze erythropoiesis. **G**. Bone marrow cells were stained with anti-Gr1 and anti-CD11b antibodies to identify Gr1+ CD11b+ myeloid cells. **H**. Splenocytes were stained with anti-TER119 and anti-CD71 antibodies to analyze splenic erythropoiesis.

Cells were analyzed by flow cytometry in $R26^{FOG\text{-}1}$ (control) and $R26^{FOG\text{-}1}$:Vav-iCre animals; data for one representative animal are shown, n = 5. Percentages of the populations are shown next to the gates. A diagram representing the developmental pathway of the different lineages from pale (progenitors) to dark grey (differentiated cells) is shown next to the pseudo-dotplots **B**, **D**, **F** and **H**.

Figure 8: Altered eosinophil numbers in $R26^{FOG\text{-}1}$:Vav-iCre mice

Reduction of circulating eosinophils. The numbers of eosinophils obtained in full blood count analysis of 8 control ($R26^{FOG\text{-}1}$ without Cre) and 8 $R26^{FOG\text{-}1}$:Vav-iCre female mice were averaged. Standard error of the mean is shown.

Table 1: Full blood count of R26^{FOG-1}:Vav-iCre mice

Blood samples from 4 control (R26^{FOG-1} without Cre) and 4 R26^{FOG-1}:Vav-iCre mice were examined with a mouse blood analyzer. Individual values are shown, as well as the corresponding averages (in red). Note the reduction of circulating eosinophils in FOG-1 overexpressing animals (highlighted in pink). RBC: red blood cells; HGB: hemoglobin; HCT: hematocrit; PLT: platelets; WBC: white blood cells; LYMPH: lymphocytes; MONO: monocytes; NEUT: neutrophils; BASO: basophils; EO: eosinophils.

	Control					R26^{FOG-1}:Vav-iCre				
	Males		Females		Mean	Males		Females		Mean
RBC (10^4/uL)	1054	1032	979	1132	1049	1063	1136	1100	1067	1091
HGB (g/L)	177	166	169	172	171	160	172	183	166	170
HCT (10^(-1)%)	534	505	508	516	515	492	525	532	507	514
PLT (10^3/uL)	1124	1352	914	1399	1197	1292	1241	1172	831	1134
WBC (10/uL)	855	901	1595	1305	1164	653	573	1434	1314	993
LYMPH (10/uL)	602	657	1205	970	858	485	412	1081	978	739
MONO (10/uL)	150	130	255	207	185	75	106	230	226	159
NEUT (10/uL)	79	83	87	92	85	82	49	107	103	85
BASO (10/uL)	0	0	1	0	0	0	0	0	2	0.5
EO (10/uL)	24	31	47	36	34	11	6	16	5	9

LEGENDS FOR SUPPORTING FIGURES

Figure S1: Pre-targeting of ES cells with the pR26-SA-FRT-Hygro^{R} vector

A. Schematic representation of the wild-type and pre-targeted R26^{Hygro} alleles. The splice acceptor (yellow dot) and the hygromycin B resistance cassette (Hygro, orange) flanked by FRT sites were inserted using the homology arms (thick grey) between exons 1 and 2 (blue). The PCR primers (arrows) as well as the restriction sites and probes (green bars) used for Southern blotting are shown. For clarity, only the relevant PstI sites are shown. **B**. PCR screening of the ES cell clones. Clones 1-4 are positive, clone 5 shows an aberrant product. **C-E**. Correct insertion of the transgene confirmed by Southern blotting. Clone 5 was included as a negative control. To test the 5' insertion, BamHI digested ES cell genomic DNA was hybridized with the radioactively labeled 5' probe to detect the wild-type (WT, 5.8 kb) and the targeted (Targ, 4.9 kb) bands (**C**). To test the 3' insertion, PstI digested DNA was hybridized with the radioactively labeled 3' probe to detect the 6.5 kb WT and the 7.5 kb targeted bands (**D**). To verify single-copy insertion, PvuII digested DNA was hybridized with a radioactively labeled internal probe to detect the 8 kb targeted band (arrow). Note that clone 1 shows an aberrant extra band, indicating multiple insertions in this clone (**E**).

Figure S2: Efficiency of RMCE at the pre-targeted R26^{Hygro} allele

A. Schematic representation of the different alleles, from top to bottom: wild-type R26, R26^{Hygro}, R26^{Control} and R26$^{FOG\text{-}1}$. The different primer pairs used for PCR analysis of the ES clones are depicted by arrows. **B.** PCR analysis of Neo^{R}/Hygro^{S} ES cell clones for testing RMCE recombination at the 5' (FRT3) and at the 3' (FRTwt) sites. Lanes 1-12: control clones, cells derived from RMCE with the control donor vector; lanes 1'-12': FOG-1 clones, cells derived from RMCE with the FOG-1 donor vector; + Ctl, positive control. From top to bottom: PCR screening with primer pairs 1F/1R and 3F/3R at the 5' end junction of the recombined cassette. PCR screening with primer pairs 2F/2R and 4F/4R at the 3' end junction of the recombined cassette. Appropriate positive controls were chosen for each PCR set up. **C.** PCR analysis of Neo^{R}/Hygro^{S} ES cell clones for presence of the Neomycin resistance gene; - Ctl,

negative control. **D.** PCR analysis of Neo^R/$Hygro^S$ ES cell clones for presence of the human CD2t gene. Lanes labeling as in (**B**) above. As shown, all clones analyzed are positive for both the Neomycin and the hCD2t gene.

Figure S3: Expression of transgene-derived FOG-1 in $R26^{FOG\text{-}1}$:Vav-iCre animals

Total RNA from bone marrow (BM), spleen (Spl), and thymus (Thy) of 3 $R26^{FOG\text{-}1}$:Vav-iCre animals was extracted, reverse transcribed and subjected to quantitative PCR to specifically detect transgene-derived FlagFOG-1 mRNA. Values are relative to RNA Polymerase II (RPII) expression. Standard error of the mean is shown. FlagFOG-1/RPII relative expression in bone marrow was arbitrarily set to 1.

Figure 1 Du Roure *et al.*

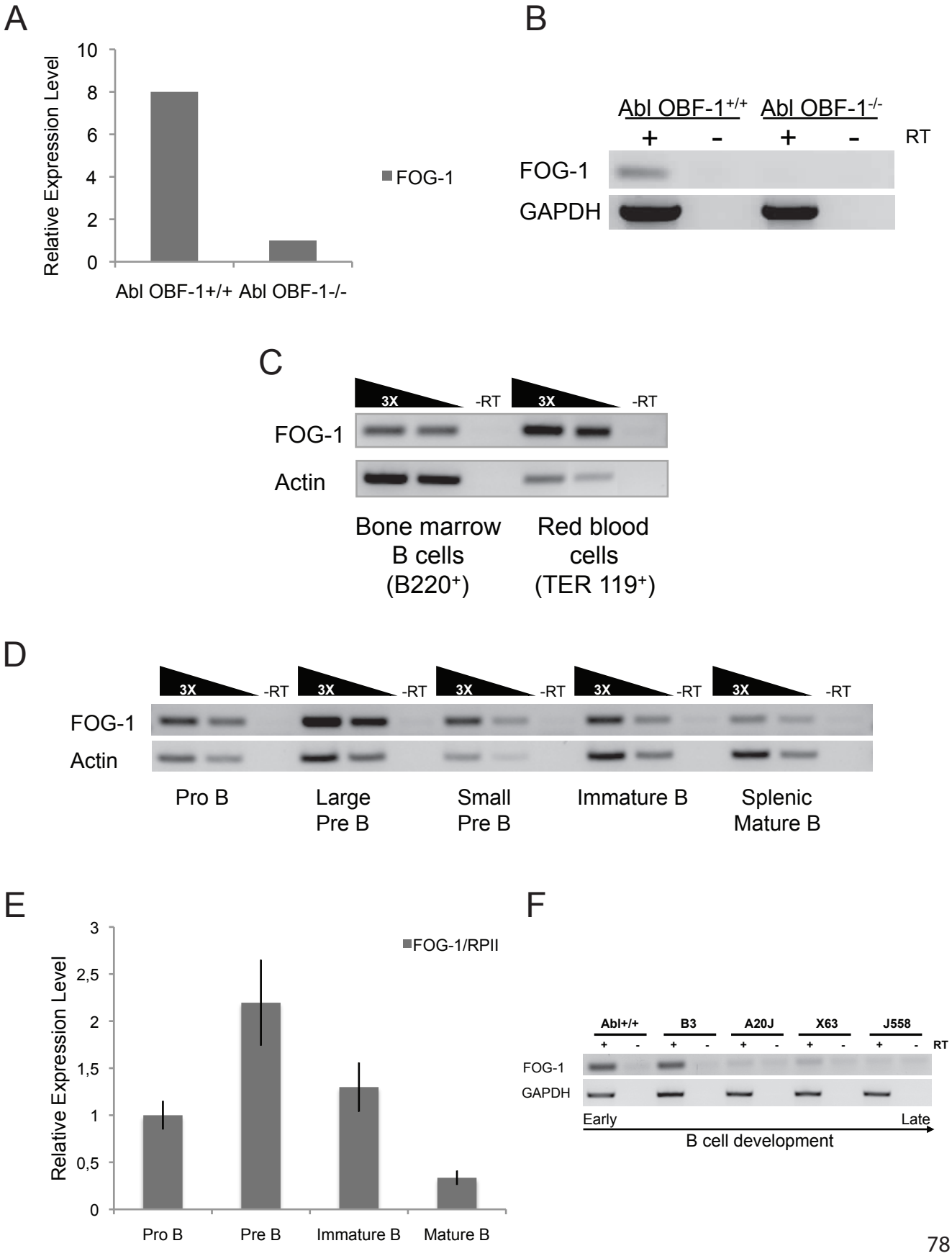


Figure 2 Du Roure *et al.*

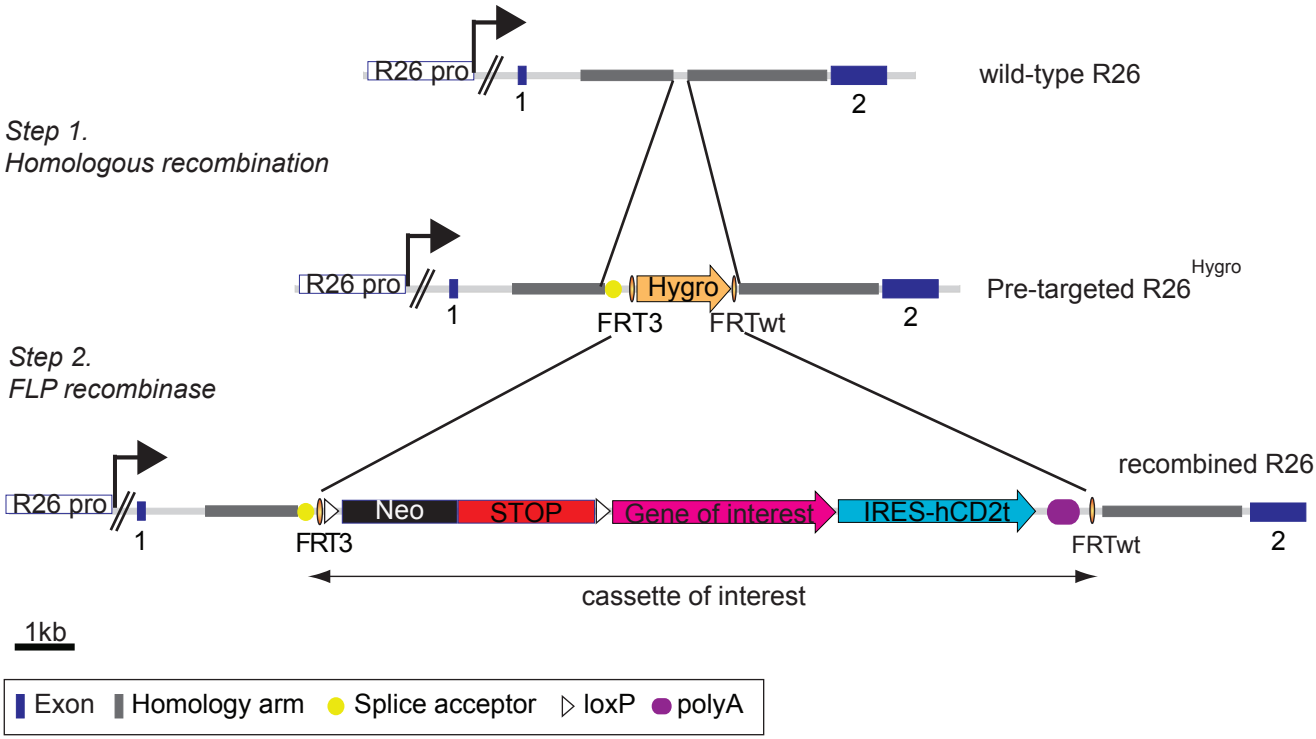
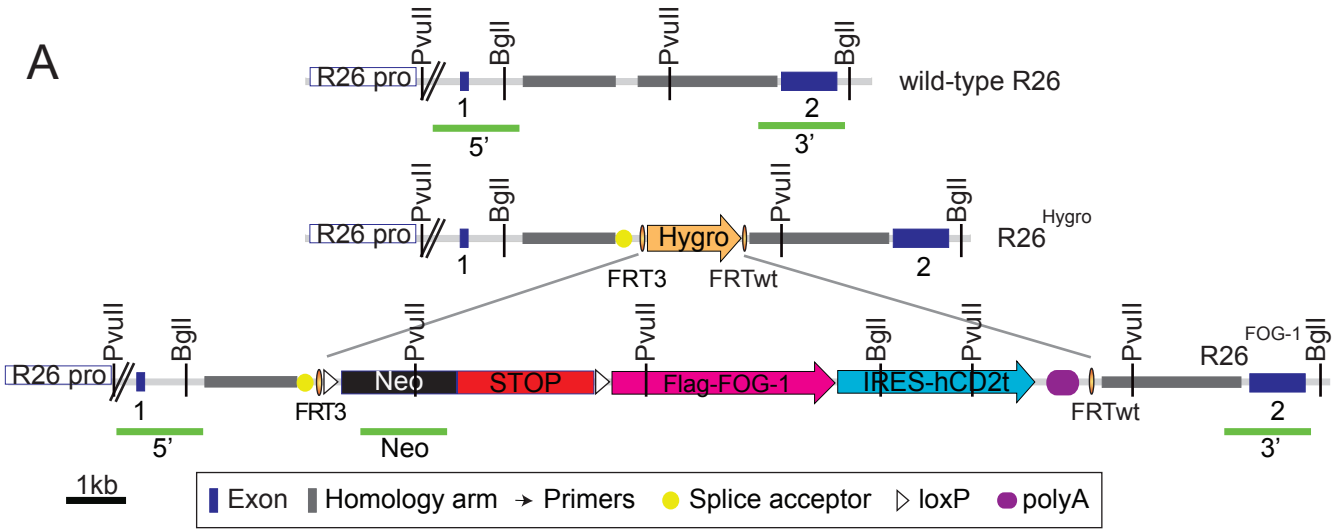


Figure 3 Du Roure *et al.*



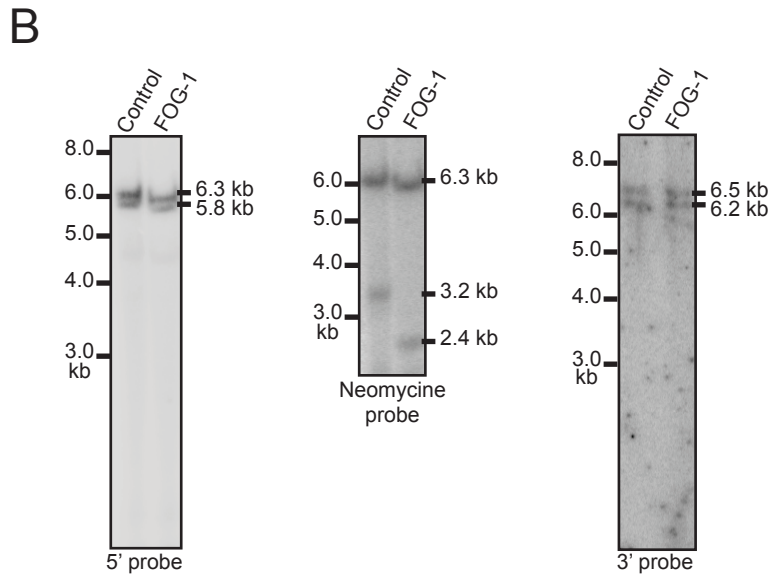


Figure 4 Du Roure *et al.*

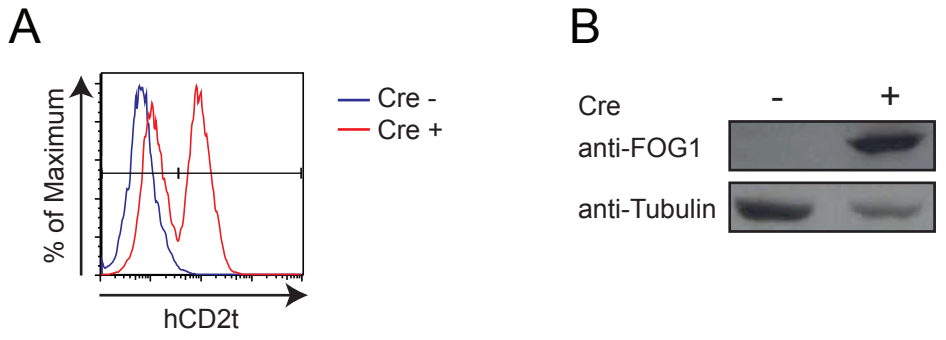
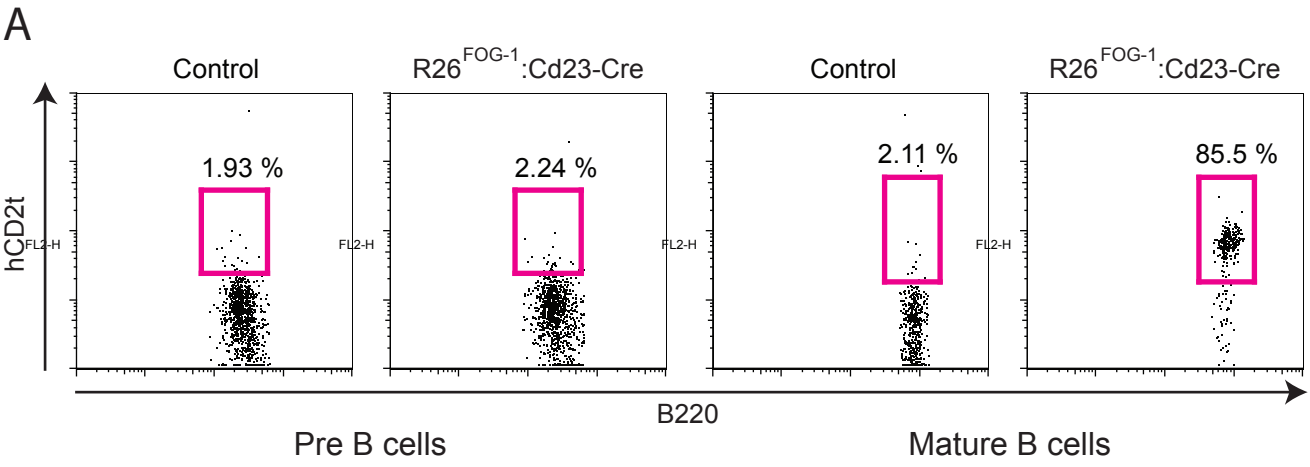
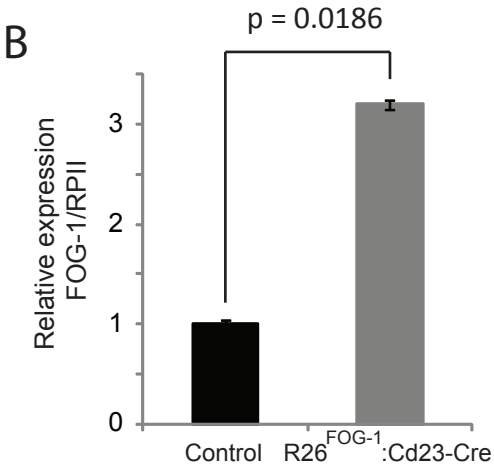
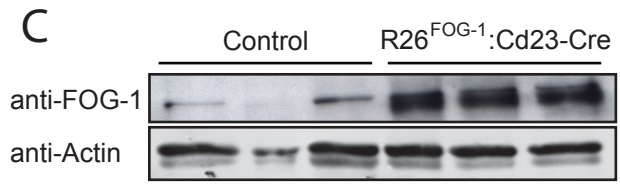


Figure 5A-C Du Roure *et al.*







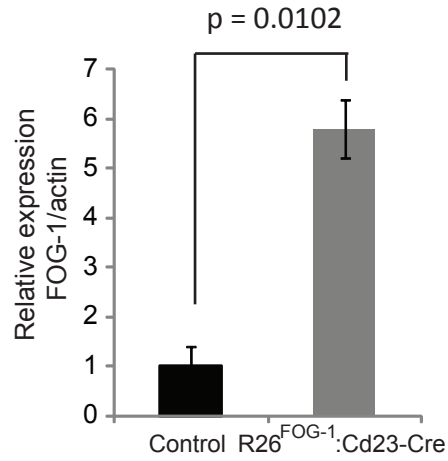


Figure 5D Du Roure *et al.*

D

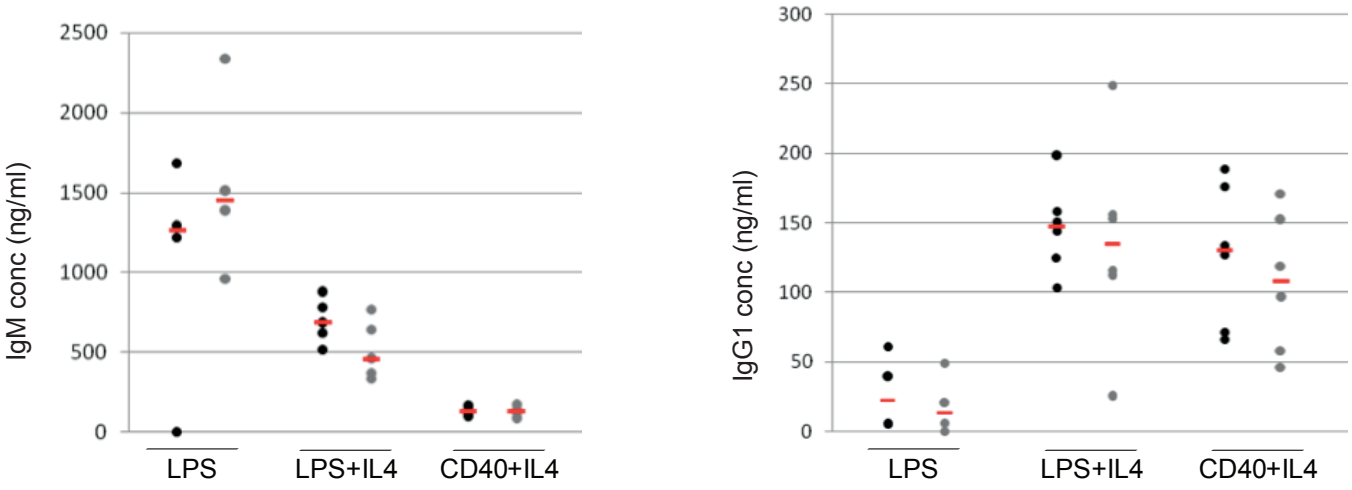


Figure 6 Du Roure *et al.*

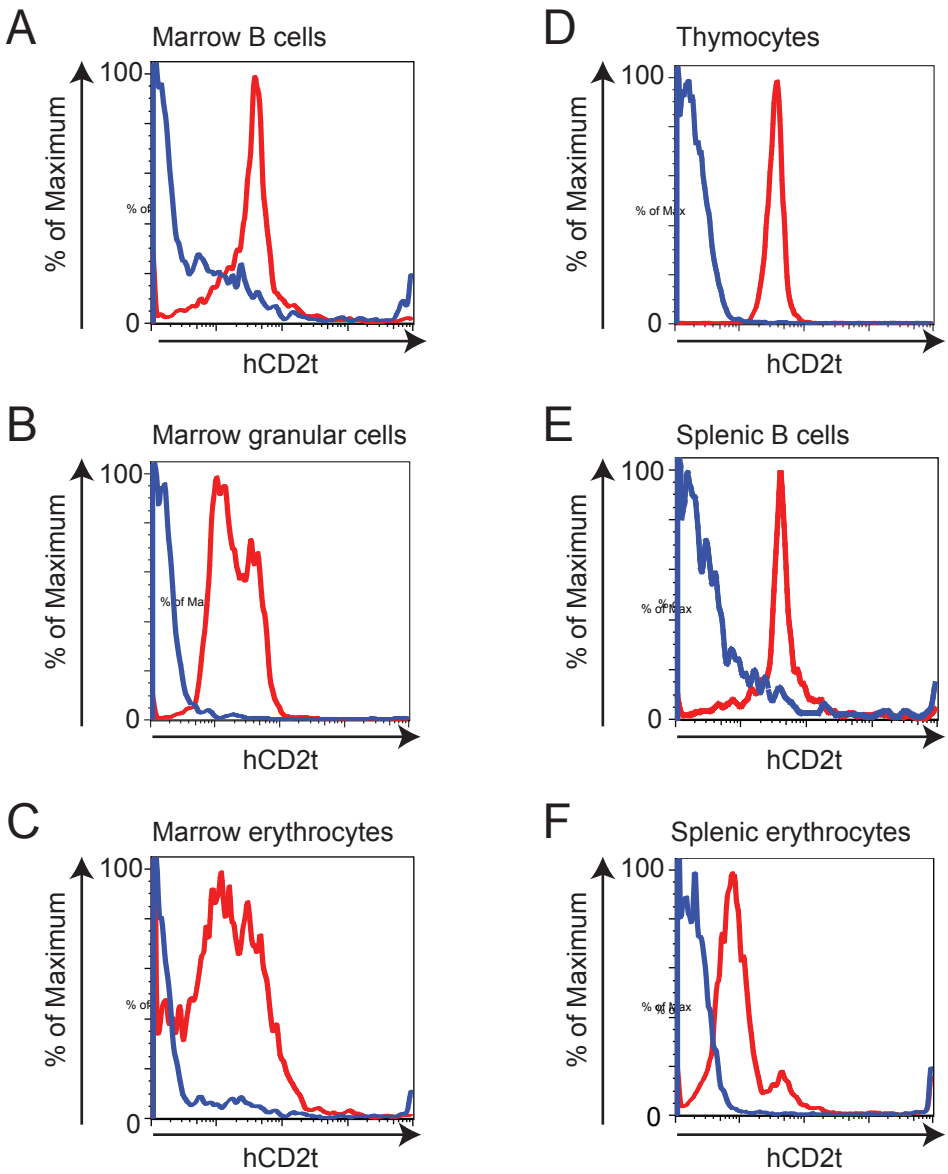


Figure 7 Du Roure *et al.*

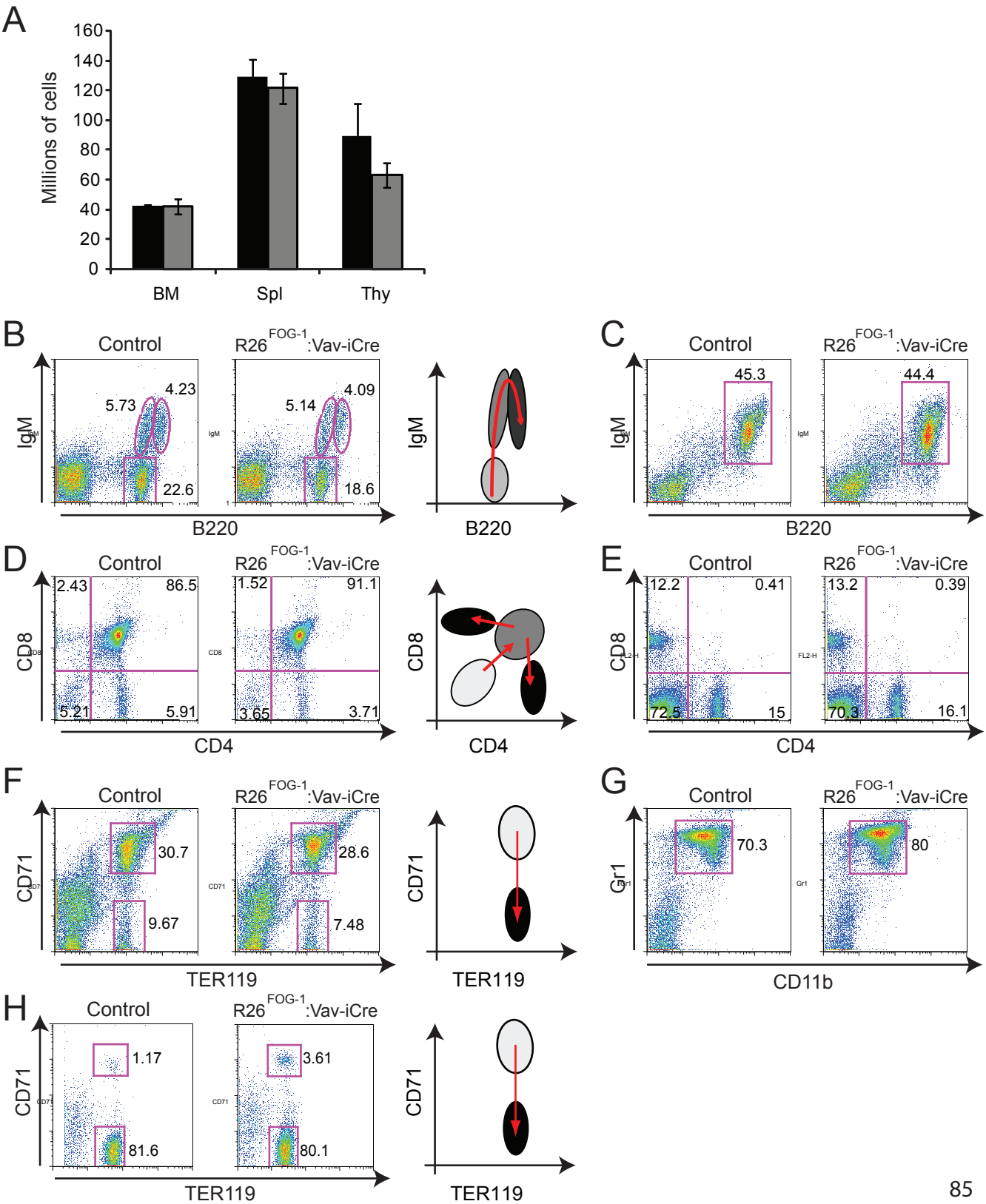
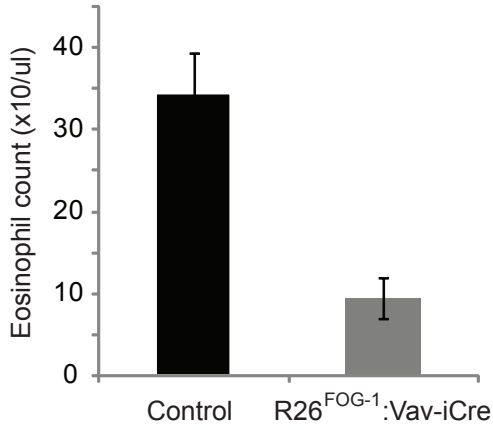
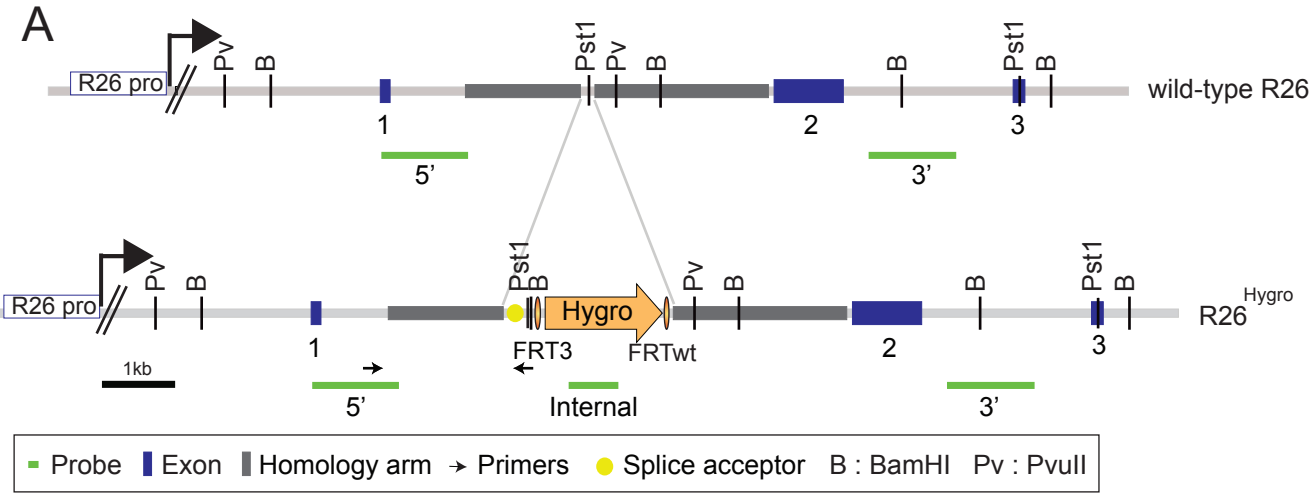
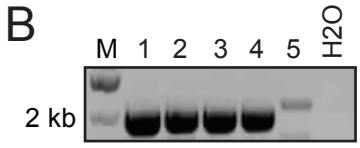


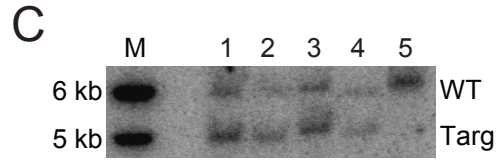
Figure 8 Du Roure *et al.*

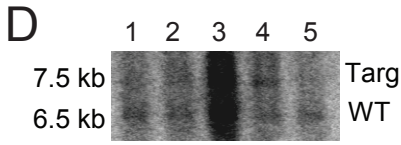


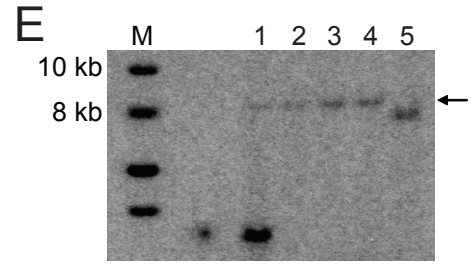
Supporting Figure 1 Du Roure *et al.*



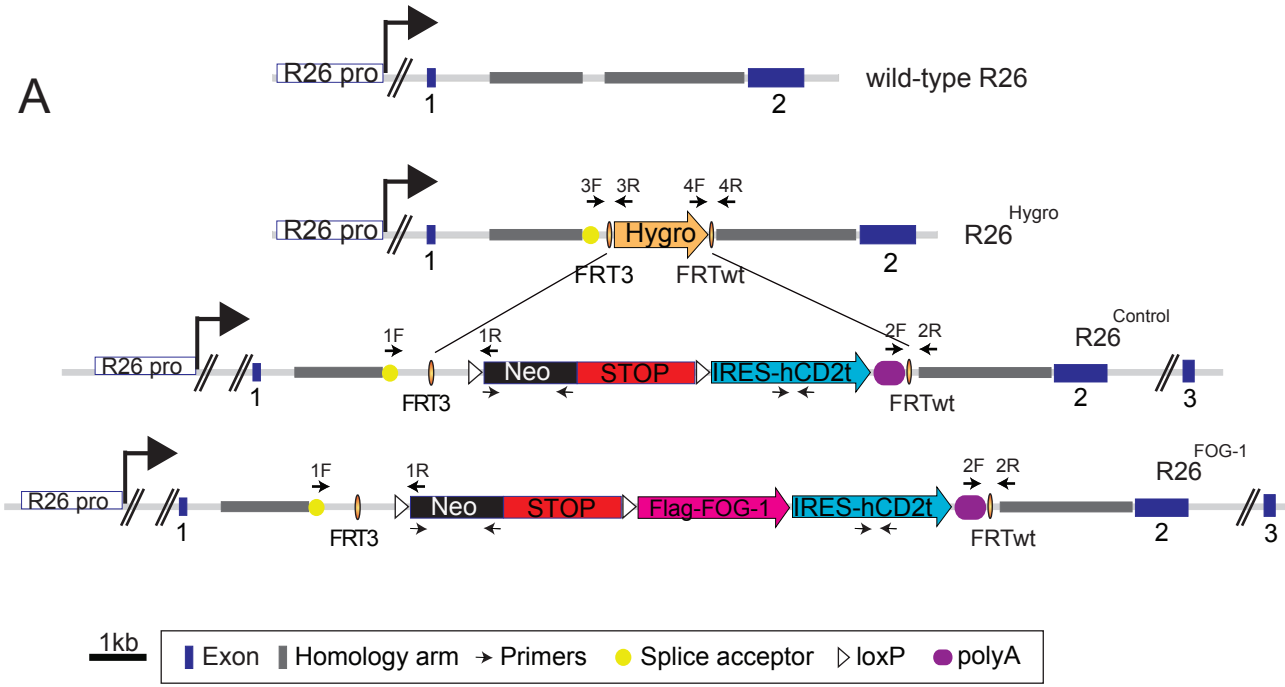


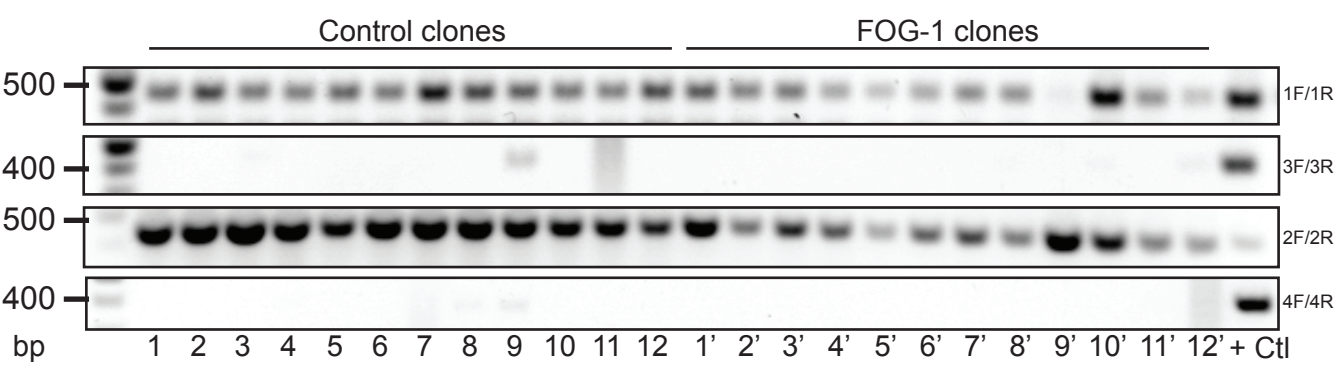


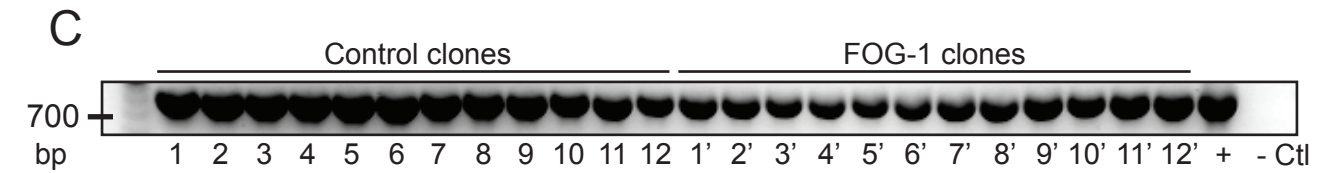


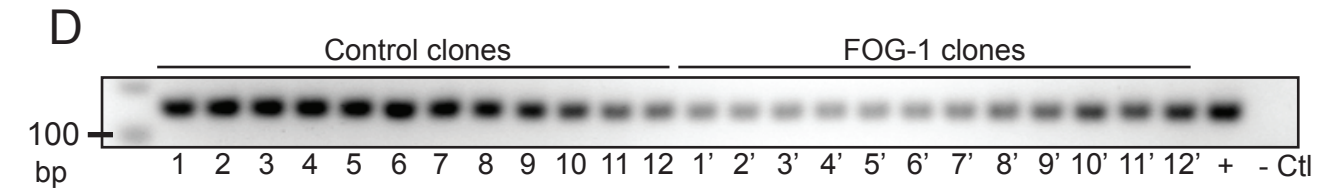


Supporting Figure 2 Du Roure *et al.*

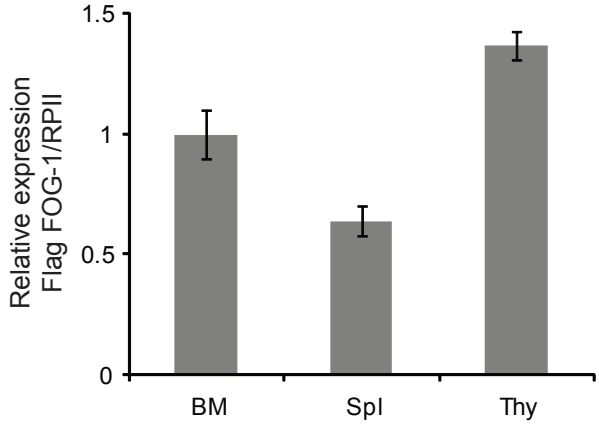








Supporting Figure 3 Du Roure *et al.*



FOG-1 is a novel co-factor for Ikaros in B-cells

Aude Versavel, Camille Du Roure[§], Ragna Sack, Hubertus Kohler, Daniel Hess and Patrick Matthias*

Friedrich Miescher Institute for Biomedical Research, Novartis Research Foundation, PO Box 2543, Maulbeerstrasse 66, 4058 Basel, Switzerland

§ Present address: Actelion Pharmaceuticals Ltd, Gewerbestrasse 16, CH-4123 Allschwil, Switzerland

* To whom correspondence should be addressed. Tel +41-61-697 66 61; Fax +41-61-697 39 76; E-mail: patrick.matthias@fmi.ch

Running title: FOG-1 is a co-factor of Ikaros

Keywords: B-cell development, transcription factors, FOG-1, Ikaros, chromatin

ABSTRACT

A number of transcriptional and epigenetic regulators involved in B-cell development have been well-described. Nevertheless, a detailed understanding of this process at the molecular level is not yet available. Of great interest, we have identified expression of the gene encoding FOG-1 (Friend of GATA-1; *Zfpm1*, *Zinc finger protein multitype 1*) in early B-lymphocytes, but not in mature B-cells and plasma cells. FOG-1 was hitherto thought to be expressed only in cells of the erythroid and megakaryocytic lineage, for whose differentiation it is essential as a transcriptional regulator. We found that in pre-B-cells, FOG-1 expression depends, directly or indirectly, on the B-cell restricted co-factor OBF-1 (OCT-Binding Factor 1). In addition, deregulation of FOG-1 expression in several systems has been shown to lead to alterations in cell-fate patterns. We showed that the regulated expression of FOG-1 in B-cells is functionally important, as enforced expression of FOG-1 in early B-cell stages leads to abnormal proliferation. These observations, in analogy to what has been seen for example in eosinophils, suggest that the specific expression pattern of FOG-1 in B-cells is required for the maintenance of early B-cell identity. Moreover, GATA factors, the main known partners of FOG-1 are not expressed in B-cells. Therefore, we hypothesized that FOG-1, as a co-factor may interact with other DNA-binding factors in this lineage. To address these questions, we used a combination of *in vivo* and *in vitro* approaches. Remarkably, we identified a gene called *epcam* (*epithelial cell adhesion molecule*) synergistically regulated by FOG-1 and Ikaros and used this locus as a readout for studying FOG-1 molecular mechanism of action in B-cells.

Strikingly and unexpectedly, our experiments showed that FOG-1 interacts with Ikaros. In addition to direct erythropoiesis and megakaryopoiesis, FOG-1 is a transcriptional regulator of B-cell development. This is reminiscent of recent findings on the function of Ikaros, long considered as only playing a role in lymphopoiesis, and, now known as being also crucial for red blood cell development.

INTRODUCTION

The development of B-cells from self-renewing haematopoietic stem cells proceeds through a number of precursor stages with progressively restricted differentiation potential and requires a complex interplay of transcription factors and epigenetic modifiers. These regulators are responsible for orchestrating the establishment of lineage-specific gene expression patterns that underly cellular differentiation (reviewed in (Orkin 1995; Matthias and Rolink 2005)). While many factors involved in this process are already known, a complete molecular scheme of the control of B-lymphocyte development is not yet available. Strikingly, we have identified the expression of Friend of GATA-1 (FOG-1) in early B-lymphocytes, but not in late developmental stages such as mature B-cells and plasma cells. FOG-1, which is encoded by the *Zfpm1* (*Zinc finger protein multitype 1*) gene, was previously thought to be expressed primarily in cells of the erythroid and megakaryocytic lineages, where it is essential for differentiation (Tsang et al. 1998; Cantor and Orkin 2005). FOG-1 is a zinc finger protein initially identified as an interacting partner of GATA factors (Tsang et al. 1997; Fox et al. 1999; Cantor and Orkin 2005). FOG-1 also interacts with the C-terminal binding protein (CtBP), mainly described as a corepressor and the nucleosome remodelling and histone deacetylase repressive (NuRD) complex; this makes a link between transcription factors and chromatin modifiers. FOG-1 also activates or represses gene transcription by facilitating binding of GATA factors to DNA (Pal et al. 2004), recruiting chromatin remodelling complexes (Fox et al. 1999; Hong et al. 2005), or by stabilizing tissue-specific chromatin loops (Vakoc et al. 2005). FOG-1 is expressed at high level in multipotent progenitors, erythroid and megakaryocytic cells, low level in lymphoid and haematopoietic stem cells; it is absent in myeloid lineages (Cantor and Orkin 2005). *Zfpm1*-deficient mice lack megakaryocytes and show severe defects in erythropoiesis, leading to embryonic lethality (Tsang et al. 1998). FOG-1 also plays a role in the T-lineage by repressing GATA-3-dependent induction of Th2 development (Zhou et al. 2001; Kurata et al. 2002). Interestingly, overexpression of FOG-1 in avian eosinophils, which do not express FOG-1, reprograms these differentiated cells into multipotent cells (Querfurth et al. 2000), reminiscent of the reprogramming of B-cells into macrophages following ectopic expression of C/EBPalpha and C/EBPbeta (Xie et al.

2004). Thus, FOG-1 is essential for specific branches of the haematopoietic system, and its inappropriate expression leads to abnormal cell differentiation.

We previously found that, in Pre-B-cells, FOG-1 expression depends on the B-cell restricted co-factor OBF-1 (OCT-Binding Factor 1) (Bartholdy et al. 2006). Besides, GATA factors, the main factors interacting with FOG-1 and recruiting it to the DNA, are not expressed in B-lymphocytes (our unpublished observations). In addition, the ability of FOG-1 to bind DNA on its own has not been demonstrated yet. Therefore, in the absence of detectable expression levels of GATA factors in B-lymphocytes, we hypothesized that FOG-1, as a co-factor, has other partners in these cells; we were particularly interested in DNA-binding transcription factors. Such interactors may also be relevant in other cellular settings, such as megakaryocytes, T-cells and non-haematopoietic cells, where a GATA-1-independent function of FOG-1 has been suggested (Tanaka et al. 2004; Amigo et al. 2009). Together these observations suggest that FOG-1 may play a role in B-cells and that in this environment its recruitment to the DNA may be achieved by a novel mechanism.

Using a FLAG-tagging/proteomics approach, we isolated FOG-1-containing complexes from cultured 230-238 Pre-B-cells. We identified distinct FOG-1 interactions. Consistent with a function of FOG-1 at the crossroad of transcription and chromatin, we described its binding to the essential haematopoietic factors Ikaros and Aiolos, in addition to the epigenetic regulator CtBP and the NuRD complex.

Ikaros and Aiolos encode DNA-binding zinc finger transcription factors, which have been found to be crucial for lymphoid development. Indeed, gene inactivation by targeted deletion of exons 3-5 of Ikaros generates a dominant negative isoform of Ikaros, which leads to a complete lack of lymphoid cells (Georgopoulos et al. 1994). On the contrary, inactivation of the gene encoding Aiolos (*Ikzf3*) in mice leads to an increase in Pre-B-cell numbers and development of B-lymphomas, owing to hyperproliferation of the B-cells (Wang et al. 1998). Moreover, Ikaros and Aiolos expression is often deregulated in human leukaemias (Winandy et al. 1995; Georgopoulos et al. 1997; Liippo et al. 2001; Takanashi et al. 2002). The two Ikaros and Aiolos-family members can activate or repress target genes by recruiting chromatin remodeling complexes such as NuRD or co-regulators such as CtBP.

Recently, Ikaros was shown to take part in a repressosome complex at the human β-globin locus, involving the histone deacetylase HDAC1, the chromatin remodeling protein Mi-2, GATA-1 and FOG-1 (Bottardi et al. 2009).

In this study, we investigated FOG-1 function in B-cells. We found that the transcriptional co-factor interacts with known key regulators of B-cell development, especially with Ikaros. Moreover, in Pre-B-cell lines, FOG-1 enforced expression induced a deregulation of gene expression and altered cellular proliferation. One of the identified target genes, *epithelial cell adhesion molecule* (*epcam*), also called *Tumor-associated calcium signal transducer 1* (*Tacstd1*) has been implicated in cellular proliferation. Its deregulation is consistent with the observed phenotype. We used this locus as a model to dissect the molecular mechanism of FOG-1 action in B-cells. Our biochemical and mechanistic analysis demonstrated that FOG-1 acts in concert with Ikaros and CtBP1 at the *epcam* locus, thus revealing a novel mechanism of action for FOG-1.

Furthermore, an analysis of FOG-1 post-translational modifications identified a novel critical residue for its function.

RESULTS

Previous results (Du Roure *et al*., in preparation) showed that FOG-1 was expressed from Pro-B-cell to immature B-cell stages at a relatively high level and was downregulated in mature B-cells and plasma cells.

This work identified regulated expression of FOG-1 in B-cells and suggests that this factor may play a role not previously appreciated in this lineage, especially in the early stages, on which we focused in this paper.

FOG-1 is part of two main epigenetic complexes in Pre-B-cells

To identify FOG-1 protein partners, we established stable PD31 Pro- and 230-238 Pre-B-cell lines with enforced expression of either GFP alone or GFP and FOG-1 fused or not with a FLAG-epitope tag (Fig. 1A). Using these cell lines, we performed whole and nuclear cell extractions and purified FLAG-FOG-1 complexes under two different salt stringency conditions (Fig. 1B). The control cell lines expressing only GFP or GFP together with FOG-1 lacking the FLAG tag were used as controls for nonspecific interactions.

Colloidal Coomassie staining of SDS-PAGE revealed one major band in the FLAG-FOG-1 soluble complex ; LC/MS/MS analysis of extracts fractionated by 1D-PAGE identified C-terminal binding protein 1 (CtBP1, Fig. 1C), a well characterized transcriptional co-activator or co-repressor of 48 kDa, known to interact with FOG-1 (Fox et al. 1999; Holmes et al. 1999; Snow and Orkin 2009). Using two-dimensional (2D)-gel electrophoresis, the strong specific interaction between FOG-1 and CtBP1 could be confirmed by immunoprecipitation (IP) of FLAG-FOG-1 in 230-238 Pre-B-cell line under high salt stringency (data not shown). IP experiments using PD31 cells demonstrated that FOG-1/CtBP1 interaction was independent of the used cell line (Supplementary Fig. 1). Moreover, since GATA factors are not expressed in B-cells and since FOG-1 is a co-factor lacking DNA binding activity, its recruitment to DNA in B-cells is likely to rely on other transcription factors. We therefore sought to identify new DNA-binding transcription factors associated to FOG-1 and performed whole lane experiment analysis by MS. Supplementary Table 1 summarizes the data obtained and lists complexes and potential partners of FOG-1. Of great interest, four transcription factors : Ikaros, Aiolos, GZP1 (GC-box-binding zinc finger protein 1) and

Runx-1 were identified. The interaction between Runx-1 and FOG-1 has already been described for its importance in megakaryocyte development (Huang et al. 2009). In contrast, Ikaros and Aiolos have not yet been shown to interact with FOG-1 and appeared as interesting candidates to mediate the function of FOG-1 in B-cells.

To further confirm the results obtained by the FLAG-tagging/proteomics approach, we performed IP experiments. By this approach we could demonstrate interaction between FOG-1 and GZP1, a member of the ZBP (Zinc finger DNA-binding protein) family of transcription factors, as well as with the critical haematopoietic factors Ikaros and Aiolos (Fig. 1D and Supplementary Fig. 2). We could also identify interactions with members of the two epigenetic complexes CtBP and NuRD (Fig. 1D, 1E and Supplementary Fig. 3). In particular, we could confirm by WB the interaction between FOG-1 and CtBP1, CtBP2, HDAC1, HDAC2 and CoREST (Fig. 1D and 1E). In addition, within B-lymphoid cells, we showed that MTA1, HDAC1, HDAC2 and RbAp46 (also known as RBBP7) were component of the NuRD complex specifically interacting with FOG-1 (Fig. 1D and 1E).

To determine more precisely the stoichiometry of FOG-1 within the two epigenetic complexes in 230-238 Pre-B-cells, we used the LC-Multiple Reaction Monitoring (MRM) relative protein quantification method. As already observed by WB, the MRM results confirmed CtBP1 as the main partner of FOG-1 in Pre-B-cells. Therefore, in pre-B-cells FOG-1 appears to be part of the CtBP complex rather than of the NuRD complex (Supplementary Fig. 3). Nevertheless, we can not exclude that the quantification in favour of the FOG-1/CtBP complex results from an experimental bias. With its nucleosome remodeling activity, the NuRD complex associated to FOG-1 might be more closely linked to the chromatin than the FOG-1/CtBP complex. In such a case, the FOG-1/NuRD complex would be more difficult to recover in the IP experiment.

In summary, we showed that, in early B-cell stages, the co-factor FOG-1 is part of two distinct epigenetic complexes: the CtBP and the NuRD complexes. In a system where FOG-1 should function independently of GATA factors, Ikaros and Aiolos were two identified transcription factors on which we further focused on.

Full-length or/and N-terminally truncated isoforms of endogenous FOG-1 interactions with CtBP1, MTA1, HDAC1, HDAC2, Ikaros and Aiolos

To assess the physiological relevance of partners recovered by LC/MS/MS, we performed several co-IP assays with endogenous FOG-1. As observed by Snow and Orkin (Snow and Orkin 2009) with MEL cell nuclear extracts, a polyclonal antibody to endogenous FOG-1 recognizes multiple distinct isoforms in 230-238 and 70Z/3 Pre-B-cell line nuclear extracts (Fig. 2A, lanes 1 and 2 and Fig. 2B, lanes 1, 2, 5, 6 and 7). Four bands could indeed be detected by WB: the two primary bands represented presumably post-translationally modified isoforms of FOG-1. According to Weiss *et al*., these could correspond to differentially phosphorylated species of the longest isoform of FOG-1 (Weiss and Orkin 1995). Two shorter isoforms of FOG-1, migrating faster, were also detected. Strikingly and consistently with what was demonstrated before (Hong et al. 2005; Rodriguez et al. 2005; Snow and Orkin 2009), the MTA1 antibody specifically immunoprecipitated the slower migrating band detected by the FOG-1 antibody (Supplementary Fig. 4, lane 6 and data not shown). On the contrary, as previously remarked (Snow and Orkin 2009), the CtBP1 antibody immunoprecipitated all isoforms of FOG-1 (Fig. 2A, lane 7 and Supplementary Fig. 4, lane 4). However, probably due to the displacement arising from the NuRD occupancy, the association between CtBP1 and the shortest isoform of FOG-1 was favoured compared to the association between CtBP1 and the longest isoform of FOG-1 (Supplementary Fig. 4, lane 5). In addition, IP by an Ikaros antibody only allowed the detection of the shortest isoforms of FOG-1 (data not shown).

In summary, we showed that in Pre-B-cells endogenous FOG-1 exists as multiple distinct isoforms and interacts with CtBP1, MTA1, HDAC1, HDAC2, Ikaros and Aiolos (Fig. 2 and data not shown).

FOG-1 regulates expression of the epcam gene in B-cells

Analysis of mice lacking FOG-1 in B-cells or with enforced expression of this factor demonstrated that modulation of FOG-1 expression has an impact at the pre-B-cell stage (our unpublished data). In agreement with these results, 230-238 pre-B-cells overexpressing FOG-1 showed an increase in proliferation compared to control cells only expressing GFP (Supplementary Fig. 5 and data not shown).

In order to determine the FOG-1 molecular mechanisms of action in Pre-B-cells, we performed a microarray analysis on established 230-238 stable Pre-B-cell line with enforced expression of FOG-1 (Supplementary Fig. 6). Among the genes with significant expression level changes, one was of particular interest: *epithelial cell adhesion molecule* (*epcam*), also called *Tumor-associated calcium signal transducer 1* (*Tacstd1*). This gene was upregulated four fold in Pre-B-cells overexpressing FOG-1 and we used it as a readout to understand the mode of action of FOG-1 in B-cells. As mentioned above, direct DNA-binding by FOG-1 has not been demonstrated before and the known main partners of FOG-1, GATA factors, are absent in B-cells. Since we have shown that FOG-1 can interact with Ikaros and Aiolos in pre-B-cells, we searched *in silico* possible transcription factor binding sites (fuzznuc software) in the *epcam* locus, focusing particularly on potential Ikaros/Aiolos binding motifs. We identified eight consensus DNA-binding sites for Ikaros and Aiolos (GGGAAGGGGAT) within the 5'-UTR and 4 kb upstream region of the *epcam* gene (Fig. 3A).

To test the contribution of Aiolos, Ikaros and FOG-1 to the regulation of the *epcam* gene, we established pre-B-cell lines stably expressing FLAG-tagged Ikaros1 or HA-tagged Aiolos together with FOG-1 (Fig. 3B) and measured by real-time RT-PCR assays the level of *epcam* expression. As shown in Figure 3C, overexpression of Aiolos by itself (be it alone or together with FOG-1) did not influence the transcription at the *epcam* locus compared with appropriate controls. Indeed, the mRNA expression level at the *epcam* locus in the four different 230-238 established cell lines was directly reflecting the FOG-1 mRNA expression level (Fig. 3C and Supplementary Fig. 7, left panels). Thus, activation of the endogenous *epcam* locus was FOG-1-dependent, but Aiolos-independent. In contrast, expression of FOG-1 together with Ikaros1 in 230-238 cells clearly showed a synergistic effect for activation of the *epcam* locus (Fig. 3C and Supplementary Fig. 7, right panels). In agreement with the microarray results, enforced expression of FOG-1 by itself led to activation of the *epcam* gene (14.8 fold). As predicted by *in silico* data, overexpression of Ikaros1 also activated this locus, albeit weakly (5.2 fold). Of great interest, upregulation of both FOG-1 and Ikaros1 resulted in a synergistic effect at the endogenous *epcam* locus (73.8 fold).

FOG-1 activation of the epcam locus is Ikaros1- and CtBP-dependent

To better understand how Ikaros1 and FOG-1 act in concert in B-cells, we performed pull-down experiments in 230-238 cells with enforced expression of Ikaros1 alone or together with FOG-1. The 230-238 cells only overexpressing both RFP and GFP were used as control for nonspecific partners. LC/MS/MS analysis of cell extracts fractionated by 1D-PAGE revealed that within 230-238 Pre-B-cells, Ikaros1 and FOG-1 were found in the same epigenetic complexes : the CtBP and the NuRD complexes (Supplementary Table 2). By WB, we further confirmed the results obtained by LC/MS/MS with specific antibodies against the FLAG tag, FOG-1, CtBP1 and Ikaros1. We described Ikaros1 interactions with FOG-1 and CtBP1 (Fig. 3D). This was consistent with our analysis of endogenous complexes, which showed that Ikaros1 might preferentially associate with the shortest isoforms of FOG-1, excluded from interaction with the NuRD complex (data not shown).

To assess the involvement of FOG-1 regarding the CtBP and NuRD complexes in the regulation at the *epcam* locus, we generated several mutants of the co-factor: a single mutant partially unable to bind the NuRD complex (HA-FOG-1K5A, (Hong et al. 2005)), a double mutant unable to bind the NuRD and CtBP complexes (HA-FOG-1dm) and FOG-1 truncated in one third of its length in the C-terminal part (FLAG-FOG-1trunc, see Figure 4A). We used a construction only allowing expression of the GFP marker as control for measuring the basal level of expression at the *epcam* locus. As expected, real-time RT-PCR assays at the *epcam* locus showed that the wild-type FLAG-tagged FOG-1 activated this locus (4.1 fold; Fig. 4A). The two FOG-1 mutants unable to bind CtBP1 (HA-FOG-1dm and FLAG-FOG-1trunc, Supplementary Fig. 8) failed to activate the *epcam* locus (1.6 fold and 0.6 fold respectively), unlike the wild-type FOG-1 protein (Fig. 4A). The FOG-1 mutant partially impaired in its interaction with the NuRD complex gave an arguable result, as activation of the *epcam* gene was higher (15.4 fold). This effect however, could be due to a higher expression level of the HA-FOG-1K5A mutant compared to the wild-type tagged protein (Fig. 4A, Supplementary Fig. 8).

To rigorously discard any role of GATA factors for FOG-1 function in B-cells, we generated FOG-1 mutants impaired in GATA binding. Based on the residues known to be the most important for GATA binding (Tsang et al. 1997; Fox et al. 1999), we

constructed different FOG-1 mutants with a single zinc finger mutated (FOG-1mut6), two zinc fingers mutated (FOG-1mut6, 9) or three zinc fingers mutated (FOG-1mut1, 6, 9). All three mutants were able to activate the *epcam* locus at least with the same efficiency as the wild-type tagged protein (Supplementary Fig. 9).

Together, these results demonstrate that activation of transcription by FOG-1 at the *epcam* locus is dependent on Ikaros1- and CtBP and independent of GATA factors.

FOG-1 phosphorylation is critical for activation of the epcam locus in B-cells

Our initial observations showed that in B-cells FOG-1 migrated as multiple bands on SDS-PAGE (Fig. 2A, lanes 1 and 2; Fig. 2B, lanes 1, 2, 5, 6 and 7) and recently Miccio *et al.* suggested that FOG-1 binding to NuRD might be regulated by phosphorylation (Miccio et al.). We therefore set out an experiment to identify post-translational modifications of FOG-1 in B-cells. For this, FOG-1 was immunoprecipitated from 230-238 pre-B-cell extracts and analyzed by LC/MS/MS. By this approach, twelve phosphorylation sites, which are on residues conserved between mouse and human FOG-1, were identified (Supplementary Fig. 10A, 10B and 10C). It is noteworthy that none of these sites overlaps with either of the nine zinc fingers or the NuRD and CtBP binding sites. Three of the identified sites (S397, S497 and S500 in FOG-1) were found within the region between zinc fingers 4 and 5, whose function has not yet been completely assigned despite a high conservation between mouse and human (Supplementary Fig. 10B and 10C). Snow *et al*. recently mapped this region as interacting with the CtBP complex (Snow et al.). Additional data from our LC/MS/MS analysis on FOG-1 mutants partially unable to bind the NuRD complex revealed a putative link between FOG-1/NuRD complex interaction and FOG-1 level of phosphorylation (data not shown).

To further dissect which pathways were involved in the regulation of the *epcam* gene, we generated mutants unable to be phosphorylated at two phosphoacceptor sites: serine S286 closest to the conserved Ikaros “double zinc finger”, which correspond to dimerization zinc finger sequences conserved between *Drosophila* Hunchbach and Ikaros proteins and serine S925 closest to the CtBP binding site, as well as the double mutant. We generated several 230-238 Pre-B-cell lines expressing these mutants together with Ikaros: first, we used a GFP marker to establish B-cell

lines expressing wild-type or mutant versions (FLAG-FOG-1S286A, FLAG-FOG-1S925A and FLAG-FOG-1S286A/S925A) of FOG-1 (Fig. 5A). We then used a H2Kk marker to introduce an Ikaros expression vector. Real-time RT-PCR assays were used to monitor expression of the *epcam* locus. Strikingly, the FOG-1 mutants unable to be phosphorylated at S286 -the phosphorylation site closest to the dimerization zinc finger domain mediating highly selective dimerization- failed to activate the *epcam* locus compared to the wild-type tagged protein (1.8 fold and 0.9 fold compared to 11.7 fold; Fig. 4B). Furthermore, these mutants could not mediate the synergistic effect resulting from wild-type FLAG-FOG-1 in conjunction with FLAG-Ikaros1 (1.1 fold and 1.2 fold compared to 30.9 fold; Fig. 5B and Supplementary Fig. 11).

Altogether these results reinforce the importance of the regulation mediated by FOG-1 at the *epcam* locus: apart from the recruitment of at least three proteins (Ikaros1, FOG-1 and CtBP1) at this locus, the requirement of FOG-1 phosphorylation at S286 represents an additional layer of control, which is summarized in the working model depicted in Figure 5C.

DISCUSSION

FOG-1 is a novel co-factor of Ikaros in B-cells

In this study, we have shown that FOG-1, a factor previously known to be crucial in the red blood cell lineage, is also expressed throughout B-cell development and particularly in early stages. We found by mass spectrometry and biochemical analysis that in pre-B–cells, FOG-1 interacts with the CtBP and NuRD complexes as well as with the zinc finger transcription factors Ikaros and Aiolos. Transcriptome analysis of pre-B-cells allowed us to identify genes regulated by FOG-1; among these genes, we used the *epcam* gene as an endogenous readout for FOG-1 function in B-cells. We could show that in this system, FOG-1 synergistically activates transcription in conjunction with Ikaros. The synergistic effect is not observed in presence of Aiolos. In addition, activation of the *epcam* gene depends on the recruitment of the CtBP complex by FOG-1. Remarkably, within the CtBP complex, we could identify LSD1, Lysine-Specific Demethylase 1, already known to play an important role in the biology of erythroid development (Saleque et al. 2007) and able to repress or activate transcription in a context-dependent manner (Metzger et al. 2005; Garcia-Bassets et al. 2007; Wang et al. 2007). Furthermore, we could identify several novel FOG-1 serine phosphoacceptor sites, with at least one being crucial for FOG-1 ability to activate transcription of *epcam*. Thus, our results identify FOG-1 as a novel transcriptional regulator in B–cells. Moreover, the Ikaros-FOG-1 interaction identified here might well also be relevant in other cellular settings outside the B-cell lineage.

Our biochemical studies demonstrated that FOG-1 was associated with the CtBP and NuRD complexes in Pre-B-cells, as was Ikaros1. While it was obvious from our results that interaction with CtBP is critical for activating the *epcam* locus, the biological significance of the FOG-1/NuRD association in Pre-B-cells remains to be further defined. Similarly, the importance of the interaction between FOG-1 and Aiolos could not be shown here. It has been previously suggested that at Pre-B stages of lymphocyte development, Ikaros and Aiolos might compete with or for regulatory factors. The balance between these two transcription factors is indeed critical for early B-cell development (Thompson et al. 2007). The outcome of this competition would depend on the amount of each of the players and may also be modulated by

the presence of FOG-1 and its partners. In addition, similarly to what FOG-1 does with regard to GATA-1 in red blood cells, it might fine-tune Ikaros function during early B-cell stages. Ikaros has been considered to be a nuclear architect in lymphocytes (Brown et al. 1997; Kim et al. 1999) ; it has been additionally shown to regulate the localization and propagation of heterochromatin and to potentiate the activities of complexes such as CtBP and NuRD by moving them away from or targeting them to the promoter of specific genes (Cobb et al. 2000; Koipally et al. 2002). Moreover, in the nucleus of proerythroblasts, FOG-1 mediates activation or silencing of genes by facilitating chromatin occupancy by GATA-1 (Pal et al. 2004), by recruiting epigenetic modifiers or by stabilizing tissue-specific chromatin loops (Vakoc et al. 2005). In addition, FOG-1 has been shown to be required for reconfiguring higher order chromatin organization *via* GATA factor exchange (Jing et al. 2008). Also, reminiscent of what was observed for Helios in T-cell nuclei (Hahm et al. 1998), FOG-1 might function as a limiting regulatory subunit for Ikaros within Pre-B-cell centromeric heterochromatin.

The findings reported here might not only be relevant in Pre-B-cells but also in other cellular settings. In absence of GATA factors, the regulatory outcome of Ikaros1-poised chromatin states of specific targets might depend on FOG-1 function. For instance, at the *epcam* locus, FOG-1 may represent the limiting factor for activation of the promoter in a chromatin environment defined by Ikaros1. We found that activation at this locus was indeed driven by FOG-1 in an Ikaros1- and CtBP-dependent manner (Fig. 3C, 4A and 5B). Furthermore, the LC/MS/MS results (Supplementary Table 2) identified the demethylase LSD1 as a component of the complex interacting with FOG-1. Thus, FOG-1 in conjunction with Ikaros1 might modify the epigenetic status of the *epcam* gene *via* LSD1, a component of the CtBP complex exhibiting enzymatic activity. Indeed, LSD1 has been shown to specifically demethylate mono- or di-methylated lysines 4 and 9 of histone 3, thus leading to gene repression or activation depending on the context (Shi et al. 2004; Metzger et al. 2005). The FOG-1 mutant partially impaired in its interaction with the NuRD complex showed a higher activation of the *epcam* gene compared to the wild-type protein. This observation might be due to the absence of HDAC1/2 activity, which results in repression.

Phosphorylation of FOG-1 is critical for transcriptional activation in Pre-B-cells

Our identification of several novel phosphorylation sites in FOG-1 gives novel insights into the function of this co-factor. Furthermore, a crosstalk between phosphorylation and SUMOylation of FOG-1 was identified by Snow *et al*. (Snow et al.). Of great interest, they demonstrated that SUMOylation of FOG-1 modulates its interaction with CtBP family members, specifically promoting CtBP1 binding. Otherwise, we showed that phosphorylation at S286 was necessary to activate the *epcam* locus (Fig. 4B and 5B). Further investigations will be required to determine the exact role of this serine: a plausible model would described this residue as being crucial within an activation domain of the protein; another one would suggest that it promotes FOG-1 and Ikaros1 interaction in a phosphorylation-dependent manner. The second hypothesis is rather attractive : Ikaros proteins are notably characterized by a C_2H_2 zinc-finger protein-protein interaction domain near the C terminus, that consists of two zinc finger motifs with an unusual structure that is most similar to those described in the *hunchback* gene in *Drosophila*. This domain mediates the formation of homo- or heterodimers and multimers between Ikaros family members, which is essential for the regulation of their functions (Georgopoulos et al. 1992; Hahm et al. 1994). The S286, which is close to the dimerization zinc finger domain could well be involved in FOG-1 and Ikaros1 interaction.

Reminiscent of recent studies (Jack and Crossley), FOGs and CtBPs as partners of Ikaros-family members might control lymphocyte proliferation and differentiation.

Our work leads to a reconsideration of the function of FOG-1 in haematopoiesis. Indeed, while the FOG-1/GATA-1/NuRD axis has now been well-described during erythropoiesis and megakaryopoiesis (Gao et al. ; Miccio et al.), the biological significance of FOG-1/CtBP interaction has never been rigorously addressed *in vivo*. In fact, knock-in mice bearing a deletion of the CtBP-binding sequence lacked a detectable phenotype in the tissues /cells examined (Katz et al. 2002) ; nevertheless, in this system, the CtBP interaction was not completely abolished by the mutation (Snow and Orkin 2009). Therefore, *in vivo* studies of the FOG-1/Ikaros1/CtBP axis during haematopoiesis and lymphopoiesis would be of great interest. This is

reminiscent of recent findings on the function of Ikaros, long considered as only playing a role in lymphopoiesis, and now known as being also crucial for red blood cell development (Dijon et al. 2008).

EXPERIMENTAL PROCEDURES

Plasmid construction

The cDNAs encoding murine FOG-1 or FLAG-FOG-1 were obtained from a pMT2-FOG-1 or a pcDNA3.1-FLAG-FOG-1 vector (kindly provided by M Crossley, Sydney). The pMSCV-IRES-RFP plasmid was constructed by substituting the cassette confering the resistance to puromycine by the cDNA sequence encoding the Red Fluorescent Protein (from the pβActinmRFP vector) into the pMSCV-IRES-Puro^r vector (kindly supplied by A Peters, Basel). For retroviral expression, the FOG-1 or FLAG-FOG-1 cDNAs were cloned into pMSCV-IRES-RFP or pMSCV-EGFP vectors, respectively. These constructions are referred as pMSCV-FOG-1-IRES-RFP and pMSCV-FLAG-FOG-1-EGFP in the legends of the figures.

The paragraph below describes the generation of several mutants of FOG-1: the partial mutant for the NuRD binding site, the double mutant (mutated in the CtBP binding site and partially mutated in the NuRD binding site) and a truncated protein in its C-terminal part. In addition to these mutants, partial mutants for the GATA-1 binding site and phosphoserine mutants were generated.

The pMSCV-HA-FOG-1(K5A)-EGFP plasmid was generated from a pcDNA3-HA-FOG-1(K5A) plasmid, kindly provided by Gerd A Blobel (Philadelphia). The pMSCV-HA-FOG-1(K5A)ΔCtBP-EGFP plasmid was obtained from a pcDNA3.1-FLAG-FOG-1ΔCtBP, previously generated by replacing the codons encoding the CtBP binding domain (Fox *et al.*, Embo J, 1999) by a mutated version of the pHA-FOG-1(K5A)ΔCtBPS706R vector (kind gift of W Hong, Philadelphia). The pMSCV-FLAG-FOG-1ΔC-term-EGFP plasmid was obtained by inverting a cDNA fragment located in the 3' region of the sequence encoding FOG-1, subsequently leading to a STOP codon in the open reading frame. The pMSCV-FLAG-FOG-1ΔZn6$_{GATA1}$-EGFP plasmid was constructed by subcloning of a cDNA fragment encoding a region around the zinc finger 6 of FOG-1 into the pEGFP-N1 vector (Clontech, pEGFP-N1Zn6(FOG-1)); a mutagenesis step led to the pEGFP-N1ΔZn6(FOG-1) $_{GATA1}$ plasmid. The pMSCV-FLAG-FOG-1ΔZn6Zn9$_{GATA1}$-EGFP and the pMSCV-FLAG-FOG-1S933A-EGFP plasmids were constructed by subcloning of a cDNA fragment encoding a region around the zinc finger 9 of FOG-1 into the pcDNA3.1 vector (Invitrogen, pcDNA3.1Zn9(FOG-1)); two independent

mutagenesis reactions generated either the pcDNA3.1ΔZn9(FOG-1)$_{GATA1}$ or the pcDNA3.1Zn9(FOG-1)$_{S933A}$ plasmids. The pMSCV-FLAG-FOG-1ΔZn1Zn6Zn9$_{GATA1}$-EGFP and the pMSCV-FLAG-FOG-1S294A-EGFP plasmids were constructed by subcloning of a cDNA fragment encoding a region around the zinc finger 1 of FOG-1 into the pEGFP-N1 vector (Clontech, pEGFP-N1Zn1(FOG-1)); two independent mutagenesis reactions led to either the pEGFP-N1ΔZn1(FOG-1)$_{GATA1}$ or the pEGFP-N1Zn1(FOG-1)$_{S294A}$ plasmids. The two cDNA fragments: ΔZn1(FOG-1)$_{GATA1}$ and Zn1(FOG-1)$_{S294A}$ were previously reinserted into pBSK^{+}-FLAG-FOG-1 vector to further generate the appropriate constructions for retroviral expression. The pMSCV-FLAG-FOG-1S294AS933A-EGFP plasmid was generated from pBSK^{+}-FLAG-FOG-1S294A *via* the reinsertion of the adequate insert into the pMSCV-FLAG-FOG-1S933A-EGFP vector. The pMSCV-FLAG-FOG-1S294A-IRES-RFP vector was obtained from the pMSCV-FLAG-FOG-1S294A-EGFP and the pMSCV-IRES-RFP. All details concerning the cloning strategies as well as the primer sequences used for mutagenesis are available upon request.

Besides, the pMSCV-FLAG-Ikaros1-IRES-GFP and the pMSCV-FLAG-Ikaros1-IRES-H-2Kk vectors and their corresponding control constructions (respectively pMSCV-IRES-GFP and pMSCV-IRES-H-2Kk) were kindly provided by S Winandy (Chicago). The pMSCV-HA-Aiolos-IRES-GFP and pMSCV-IRES-GFP vectors were a kind gift of B S Cobb (London).

Cell culture, transfection and retroviral infection

Cells were cultured in a humidified tissue culture incubator set up at 5% CO_2 and 37°C. Phoenix ecotropic cells were maintained in classical Dubelcco's Modified Eagle's Medium supplemented with 10% HI FCS, 1% penicillin-streptomycin and 2mM L-glutamine. PD31, 70Z/3 and 230-238 cell lines were cultured in RPMI 1640 completed with 10% HI FCS, 1% penicillin-streptomycin and 4mM L-glutamine. The generation of PD31 and 230-238 cells stably expressing the Green or Red Fluorescent Proteins (GFP or RFP) or/and the (tagged) wild-type or mutant FOG-1 followed the procedures described below. Phoenix ecotropic cells were transfected with the appropriate constructions *via* Lipofectamine™ 2000 reagent (Invitrogen™) according to the manufacturer's instructions. The subsequent retroviral supernatants

produced were used for infection of either PD31 or 230-238 cells by centrifugation at 1260 g for 90 min at 25°C in the presence of 5 mg/mL of polybrene (Sigma). The cells were sorted by FACS for being either RFP^+ or GFP^+ 72 h later, so as to get rid of the not infected cells. The establishment of 230-238 expressing both the RFP and the GFP and FOG-1 or/and Ikaros1 or FOG-1 or/and Aiolos was achieved by two successive retroviral infections: a first infection of 230-238 cells with retroviral supernatants allowing stable expression of the RFP alone or the RFP together with FOG-1 was followed by a FACS of RFP^+ cells. Then, a second infection of *in vitro* expanded 230-238 RFP^+ cells with retroviral supernatants allowing stable expression of the GFP alone or the GFP together with Ikaros1 or Aiolos was followed by a sort of double positive cells, both RFP^+ and GFP^+; the GFP was used as a marker for the eventual expression of Ikaros1 or Aiolos. The establishment of 230-238 expressing both the GFP and H-2Kk marker and wild-type or mutant FOG-1 or/and Ikaros1 was achieved by 2 successive retroviral infections : first, the cells were selected for being GFP^+ by FACS, attesting of the eventual expression of FOG-1 as well ; they were then put back in culture before the second infection. The cells were further sorted for being double positive, both GFP^+ and H-$2Kk^+$; the H-2Kk marker was used for monitoring the eventual expression of Ikaros1.

Real-time RT-PCR for gene expression

Total RNA from 230-238 Pre-B-cells sorted for being RFP^+, GFP^+ or both, or GFP^+, H-$2Kk^+$ or both was purified using RNeasy Micro or Mini Kits (QIAGEN) and was primed with oligo(dT) for first-strand cDNA synthesis (ThermoScript™ RT-PCR System, Invitrogen) according to the manufacturer's instructions. A MESA GREEN qPCR MasterMix Plus for SYBR® Assay (Eurogentec) was used for quantitative real-time PCR (qRT-PCR) and results were quantified with an ABI PRISM 7000® instrument (Applied Biosystems). Sequences of primer pairs are in Supplementary Table3. Thermal cycling conditions were 50°C for 2 min, 95°C for 10 min, then, 40 cycles at 95°C for 15 seconds and 60°C for 1 min. Signals were captured during the polymerization step (60°C). A threshold was set in the linear part of the amplification curve and the number of cycles needed to reach it was calculated for each gene. Melting curve analysis was done to assess the purity of the amplified cDNA. Results

were normalized using Glyceraldehyde 3-phosphate dehydrogenase (Gapdh) mRNA as an internal control for each sample.

FOG-1 and Ikaros1 multiprotein complex purification and proteomic analysis

The methods for purification of tagged transcriptional (co-)factor complexes and LC/MS/MS of associated proteins were performed as described below. Briefly, PD31 or 230-238 cells stably expressing the GFP alone or the GFP and the FLAG- or HA-tagged wild-type or mutant FOG-1 were expanded up to 5×10^8 cells. Whole cell extracts were prepared into either a high or a low salt stringency lysis buffer (50mM Hepes pH 7.4, 1mM EDTA, either 300mM or 150mM NaCl, 70mM KCl, 1% Triton X-100, 0.1% CHAPS, 1:25 complete EDTA-free protease inhibitor cocktail (Roche), 1μg/mL of aprotinin, leupeptin, pepstatin, 1mM PMSF, 25mM NaF, 25mM glycerol phosphate, 5% sucrose) and centrifuged. Alternatively, a nuclear fractionation protocol was followed: so as to get rid of the cytosolic fraction, cells were first lysed in a lysis buffer containing 10mM Hepes pH 7.4, 1.5mM $MgCl_2$, 10mM KCl, 0.5mM DTT, 0.05% NP40, 1:25 complete EDTA-free protease inhibitor cocktail (Roche), 1μg/mL of aprotinin, leupeptin, pepstatin, 1mM PMSF, 25mM NaF, 25mM glycerol phosphate. Nuclear extracts were then prepared into the low salt stringency lysis buffer described above. First, the FLAG-tagged FOG-1 complexes were immunoprecipitated using anti-FLAG® M2 affinity gel (Sigma). The beads and bound material were washed with the first buffer specified above and eluted by incubation with excess 3XFLAG peptide (0.3 mg/mL). 230-238 cells expressing only the RFP and the GFP or the RFP, the GFP, FOG-1 or/and FLAG-tagged Ikaros1 were expanded up to 5×10^8 cells. Nuclear extracts were generated following the steps described previously in the text. Immunoprecipitation, washes and elution were performed like for FLAG-tagged FOG-1 complexes. The samples were reduced and alkylated by DTT and iodoacetamide respectively, and boiled before loading onto a NuPAGE® Novex® Bis-Tris Mini Gel (4-12%, Invitrogen). The Coomassie stained were cut, washed and in-gel digested overnight at 37°C with sequencing grade modified trypsin (Promega). The peptides were separated by nano-HPLC (Agilent 1100 nanoLC system, Agilent Technologies) coupled to a 4000Q-Trap mass spectrometer (Applied

Biosystems) or an LTQ/Orbi-trap VELOS hybrid mass spectrometer (Thermo Fisher Scientific) and analyzed in the discovery mode of operation. Peptides were identified searching UniProt 15.14 using Mascot Distiller 2.3 and Mascot 2.3 (Matrix Science). Results were compiled in Scaffold 2.06 (Proteome Software). To assess the specific enrichment of proteins relative quantification LC-MRM was partially performed considering the peptides found in the sample and its corresponding control bands.

Co-immunoprecipitation assays

Several different co-immunoprecipitation (co-IP) assays were performed in this study. Nuclear cell extracts were prepared from 10^9 cells (70Z/3 or 230-238) using the same mode of fractionation described in the paragraph focusing on multiprotein complex purification. After pre-clearing the cell lysate, the proteins of interest were immunoprecipitated using specific antibodies coupled to Protein G sepharose 4 Fast Flow Slurry beads (GE Healthcare). Endogenous FOG-1, CtBP1, MTA1, HDAC1, HDAC2, Aiolos or Ikaros pull-down assays were performed with anti-FOG-1 (Santa Cruz, A-20), anti-CtBP1 (Santa Cruz, K-15), anti-MTA1 (Santa Cruz, H-166 or A-18), anti-HDAC1, anti-HDAC2 (both gift from C Seiser, Vienna), anti-Aiolos (gift from K Georgopoulos, Charlestown) or anti-Ikaros (Santa Cruz, H-100) antibodies; anti-HA (Santa Cruz, Y-11) or anti-FLAG (Sigma-aldrich, F7425) antibodies were used as negative controls. The beads and bound material were washed with the low salt stringency lysis buffer (see above) and eluted by boiling for 5 min at 95°C in 2X loading buffer.

Western blot analysis

Proteins were separated on either a NuPAGE® Novex® Bis-Tris Mini Gel (4-12%, Invitrogen) or 5-10% SDS-PAGE, transferred to a PROTRAN® Nitrocellulose Transfer Membrane (Whatman®) and immunoblotted with the appropriate primary and secondary antibodies. The following primary antibodies were used : anti-FLAG (Sigma-aldrich, F7425), anti-FOG-1 (Santa Cruz, A-20), anti-GZP1 (Abnova, H00023528-M02), anti-Aiolos (gift from K Georgopoulos, Charlestown – different batch of antibody than the one used for co-IP assays ; Santa Cruz, S-21), anti-Ikaros (Santa Cruz, H-100), anti-CtBP1 (Santa Cruz, K15), anti-CtBP2 (Santa Cruz, C-16), anti-

CoREST (Upstate, 07-455), anti-HDAC1, anti-HDAC2 (both gift from C Seiser, Vienna), anti-MTA1 (Santa Cruz, H-166 or A-18) and anti-RbAp46 (Abcam, ab3535). As control, anti-Brg1 (Santa Cruz, G-7) and anti-p15INK4B (Cell signaling, 4822) antibodies were used. The following secondary antibodies were used : Donkey polyclonal to Goat IgG HRP (Abcam, ab7125), ECL anti-mouse IgG HRP linked whole Ab (GE Healthcare, NA931V) and ECL anti-Rabbit IgG HRP linked whole Ab (GE Healthcare, NA934V). Signals were detected with Amersham Hyperfilm™ ECL (GE Healthcare).

Post-translational modification analysis

For PTM analysis, as described in FOG-1 multiprotein complex purification paragraph, overexpressed FLAG- or HA-tagged wild-type or mutant FOG-1 were immunopurified using anti-FLAG® M2 affinity gel (Sigma) or anti-HA affinity matrix (Roche). In this latter case, the elution was remarkably achieved by adding an excess of HA peptide (0.3 mg/mL, Bachem). After reduction (DTT) and alkylation (iodoacetamide) treatments, eluted material was separated by NuPAGE® Novex® Bis-Tris Mini Gel (4-12%, Invitrogen) and stained with Colloidal Blue Staining Kit (Invitrogen). Selected bands migrating around 180 kDa were excised, in gel digested with trypsin (Promega) and analyzed for PTM by LC-MS/MS on a 4000Q-Trap (Applied Biosystems) in the IDA mode: 1 EMC, 4 ER, 4 EPI, 60 seconds exclusion after 2^{nd} EPI. FOG-1 containing bands were analyzed and evaluated for phosphorylation using Mascot.

ACKNOWLEDGEMENTS

We thank M Crossley for providing the pMT2-FOG-1 and the pcDNA3.1-FLAG-FOG-1 vectors, A Peters for providing the pMSCV-IRES-Puro^r vector, G A Blobel for providing the pcDNA3-HA-FOG-1(K5A) plasmid, W Hong for the pHA-FOG-1(K5A)ΔCtBPS706R construction, S Winandy for the pMSCV-IRES-GFP, the pMSCV-FLAG-Ikaros1-IRES-GFP, the pMSCV-IRES-H-2Kk and the pMSCV-FLAG-Ikaros1-IRES-H-2Kk plasmids and B S Cobb for the pMSCV-IRES-GFP and the pMSCV-HA-Aiolos-IRES-GFP vectors. We also thank M Busslinger for useful discussions. In addition, we thank R Portmann, former member of the protein analysis platform for the 2D-gel electrophoresis, H Angliker and E J Oakeley, both alumni from the functional genomics platform for scanning the arrays and helping in analysis, respectively. Moreover, we thank H-R Hotz from the bioinformatics platform for *in silico* search of possible transcription factor binding sites. We finally thank D Zingg for excellent assistance and M A Choukrallah for critical reading of the manuscript.

FUNDING

This work was supported by the Novartis Research Foundation.

REFERENCES

Amigo, J.D., Ackermann, G.E., Cope, J.J., Yu, M., Cooney, J.D., Ma, D., Langer, N.B., Shafizadeh, E., Shaw, G.C., Horsely, W. et al. 2009. The role and regulation of friend of GATA-1 (FOG-1) during blood development in the zebrafish. *Blood* 114(21): 4654-4663.

Bartholdy, B., Du Roure, C., Bordon, A., Emslie, D., Corcoran, L.M., and Matthias, P. 2006. The Ets factor Spi-B is a direct critical target of the coactivator OBF-1. *Proc Natl Acad Sci U S A* 103(31): 11665-11670.

Bottardi, S., Ross, J., Bourgoin, V., Fotouhi-Ardakani, N., Affar el, B., Trudel, M., and Milot, E. 2009. Ikaros and GATA-1 combinatorial effect is required for silencing of human gamma-globin genes. *Mol Cell Biol* 29(6): 1526-1537.

Brown, K.E., Guest, S.S., Smale, S.T., Hahm, K., Merkenschlager, M., and Fisher, A.G. 1997. Association of transcriptionally silent genes with Ikaros complexes at centromeric heterochromatin. *Cell* 91(6): 845-854.

Cantor, A.B. and Orkin, S.H. 2005. Coregulation of GATA factors by the Friend of GATA (FOG) family of multitype zinc finger proteins. *Semin Cell Dev Biol* 16(1): 117-128.

Cobb, B.S., Morales-Alcelay, S., Kleiger, G., Brown, K.E., Fisher, A.G., and Smale, S.T. 2000. Targeting of Ikaros to pericentromeric heterochromatin by direct DNA binding. *Genes Dev* 14(17): 2146-2160.

Dijon, M., Chabannon, C., and Tonnelle, C. 2008. [Ikaros and erythropoiesis]. *Med Sci (Paris)* 24(12): 1065-1070.

Fox, A.H., Liew, C., Holmes, M., Kowalski, K., Mackay, J., and Crossley, M. 1999. Transcriptional cofactors of the FOG family interact with GATA proteins by means of multiple zinc fingers. *EMBO J* 18(10): 2812-2822.

Gao, Z., Huang, Z., Olivey, H.E., Gurbuxani, S., Crispino, J.D., and Svensson, E.C. FOG-1-mediated recruitment of NuRD is required for cell lineage re-enforcement during haematopoiesis. *EMBO J* 29(2): 457-468.

Garcia-Bassets, I., Kwon, Y.S., Telese, F., Prefontaine, G.G., Hutt, K.R., Cheng, C.S., Ju, B.G., Ohgi, K.A., Wang, J., Escoubet-Lozach, L. et al. 2007. Histone methylation-dependent mechanisms impose ligand dependency for gene activation by nuclear receptors. *Cell* 128(3): 505-518.

Georgopoulos, K., Bigby, M., Wang, J.H., Molnar, A., Wu, P., Winandy, S., and Sharpe, A. 1994. The Ikaros gene is required for the development of all lymphoid lineages. *Cell* 79(1): 143-156.

Georgopoulos, K., Moore, D.D., and Derfler, B. 1992. Ikaros, an early lymphoid-specific transcription factor and a putative mediator for T cell commitment. *Science* 258(5083): 808-812.

Georgopoulos, K., Winandy, S., and Avitahl, N. 1997. The role of the Ikaros gene in lymphocyte development and homeostasis. *Annu Rev Immunol* 15: 155-176.

Hahm, K., Cobb, B.S., McCarty, A.S., Brown, K.E., Klug, C.A., Lee, R., Akashi, K., Weissman, I.L., Fisher, A.G., and Smale, S.T. 1998. Helios, a T cell-restricted Ikaros family member that quantitatively associates with Ikaros at centromeric heterochromatin. *Genes Dev* 12(6): 782-796.

Hahm, K., Ernst, P., Lo, K., Kim, G.S., Turck, C., and Smale, S.T. 1994. The lymphoid transcription factor LyF-1 is encoded by specific, alternatively spliced mRNAs derived from the Ikaros gene. *Mol Cell Biol* 14(11): 7111-7123.

Holmes, M., Turner, J., Fox, A., Chisholm, O., Crossley, M., and Chong, B. 1999. hFOG-2, a novel zinc finger protein, binds the co-repressor mCtBP2 and modulates GATA-mediated activation. *J Biol Chem* 274(33): 23491-23498.

Hong, W., Nakazawa, M., Chen, Y.Y., Kori, R., Vakoc, C.R., Rakowski, C., and Blobel, G.A. 2005. FOG-1 recruits the NuRD repressor complex to mediate transcriptional repression by GATA-1. *EMBO J* 24(13): 2367-2378.

Huang, H., Yu, M., Akie, T.E., Moran, T.B., Woo, A.J., Tu, N., Waldon, Z., Lin, Y.Y., Steen, H., and Cantor, A.B. 2009. Differentiation-dependent interactions between RUNX-1 and FLI-1 during megakaryocyte development. *Mol Cell Biol* 29(15): 4103-4115.

Jack, B.H. and Crossley, M. GATA proteins work together with friend of GATA (FOG) and C-terminal binding protein (CTBP) co-regulators to control adipogenesis. *J Biol Chem* 285(42): 32405-32414.

Jing, H., Vakoc, C.R., Ying, L., Mandat, S., Wang, H., Zheng, X., and Blobel, G.A. 2008. Exchange of GATA factors mediates transitions in looped chromatin organization at a developmentally regulated gene locus. *Mol Cell* 29(2): 232-242.

Katz, S.G., Cantor, A.B., and Orkin, S.H. 2002. Interaction between FOG-1 and the corepressor C-terminal binding protein is dispensable for normal erythropoiesis in vivo. *Mol Cell Biol* 22(9): 3121-3128.

Kim, J., Sif, S., Jones, B., Jackson, A., Koipally, J., Heller, E., Winandy, S., Viel, A., Sawyer, A., Ikeda, T. et al. 1999. Ikaros DNA-binding proteins direct formation of chromatin remodeling complexes in lymphocytes. *Immunity* 10(3): 345-355.

Koipally, J., Heller, E.J., Seavitt, J.R., and Georgopoulos, K. 2002. Unconventional potentiation of gene expression by Ikaros. *J Biol Chem* 277(15): 13007-13015.

Kurata, H., Lee, H.J., McClanahan, T., Coffman, R.L., O'Garra, A., and Arai, N. 2002. Friend of GATA is expressed in naive Th cells and functions as a repressor of GATA-3-mediated Th2 cell development. *J Immunol* 168(9): 4538-4545.

Liippo, J., Nera, K.P., Veistinen, E., Lahdesmaki, A., Postila, V., Kimby, E., Riikonen, P., Hammarstrom, L., Pelkonen, J., and Lassila, O. 2001. Both normal and leukemic B lymphocytes express multiple isoforms of the human Aiolos gene. *Eur J Immunol* 31(12): 3469-3474.

Matthias, P. and Rolink, A.G. 2005. Transcriptional networks in developing and mature B cells. *Nat Rev Immunol* 5(6): 497-508.

Metzger, E., Wissmann, M., Yin, N., Muller, J.M., Schneider, R., Peters, A.H., Gunther, T., Buettner, R., and Schule, R. 2005. LSD1 demethylates repressive histone marks to promote androgen-receptor-dependent transcription. *Nature* 437(7057): 436-439.

Miccio, A., Wang, Y., Hong, W., Gregory, G.D., Wang, H., Yu, X., Choi, J.K., Shelat, S., Tong, W., Poncz, M. et al. NuRD mediates activating and repressive functions of GATA-1 and FOG-1 during blood development. *EMBO J* 29(2): 442-456.

Orkin, S.H. 1995. Hematopoiesis: how does it happen? *Curr Opin Cell Biol* 7(6): 870-877.

Pal, S., Cantor, A.B., Johnson, K.D., Moran, T.B., Boyer, M.E., Orkin, S.H., and Bresnick, E.H. 2004. Coregulator-dependent facilitation of chromatin occupancy by GATA-1. *Proc Natl Acad Sci U S A* 101(4): 980-985.

Querfurth, E., Schuster, M., Kulessa, H., Crispino, J.D., Doderlein, G., Orkin, S.H., Graf, T., and Nerlov, C. 2000. Antagonism between C/EBPbeta and FOG in eosinophil lineage commitment of multipotent hematopoietic progenitors. *Genes Dev* 14(19): 2515-2525.

Rodriguez, P., Bonte, E., Krijgsveld, J., Kolodziej, K.E., Guyot, B., Heck, A.J., Vyas, P., de Boer, E., Grosveld, F., and Strouboulis, J. 2005. GATA-1 forms distinct activating and repressive complexes in erythroid cells. *EMBO J* 24(13): 2354-2366.

Saleque, S., Kim, J., Rooke, H.M., and Orkin, S.H. 2007. Epigenetic regulation of hematopoietic differentiation by Gfi-1 and Gfi-1b is mediated by the cofactors CoREST and LSD1. *Mol Cell* 27(4): 562-572.

Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstine, J.R., Cole, P.A., and Casero, R.A. 2004. Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* 119(7): 941-953.

Snow, J.W., Kim, J., Currie, C.R., Xu, J., and Orkin, S.H. Sumoylation regulates interaction of FOG1 with C-terminal-binding protein (CTBP). *J Biol Chem* 285(36): 28064-28075.

Snow, J.W. and Orkin, S.H. 2009. Translational isoforms of FOG1 regulate GATA1-interacting complexes. *J Biol Chem* 284(43): 29310-29319.

Takanashi, M., Yagi, T., Imamura, T., Tabata, Y., Morimoto, A., Hibi, S., Ishii, E., and Imashuku, S. 2002. Expression of the Ikaros gene family in childhood acute lymphoblastic leukaemia. *Br J Haematol* 117(3): 525-530.

Tanaka, M., Zheng, J., Kitajima, K., Kita, K., Yoshikawa, H., and Nakano, T. 2004. Differentiation status dependent function of FOG-1. *Genes Cells* 9(12): 1213-1226.

Thompson, E.C., Cobb, B.S., Sabbattini, P., Meixlsperger, S., Parelho, V., Liberg, D., Taylor, B., Dillon, N., Georgopoulos, K., Jumaa, H. et al. 2007. Ikaros DNA-binding proteins as integral components of B cell developmental-stage-specific regulatory circuits. *Immunity* 26(3): 335-344.

Tsang, A.P., Fujiwara, Y., Hom, D.B., and Orkin, S.H. 1998. Failure of megakaryopoiesis and arrested erythropoiesis in mice lacking the GATA-1 transcriptional cofactor FOG. *Genes Dev* 12(8): 1176-1188.

Tsang, A.P., Visvader, J.E., Turner, C.A., Fujiwara, Y., Yu, C., Weiss, M.J., Crossley, M., and Orkin, S.H. 1997. FOG, a multitype zinc finger protein, acts as a cofactor for transcription factor GATA-1 in erythroid and megakaryocytic differentiation. *Cell* 90(1): 109-119.

Vakoc, C.R., Letting, D.L., Gheldof, N., Sawado, T., Bender, M.A., Groudine, M., Weiss, M.J., Dekker, J., and Blobel, G.A. 2005. Proximity among distant regulatory elements at the beta-globin locus requires GATA-1 and FOG-1. *Mol Cell* 17(3): 453-462.

Wang, J., Scully, K., Zhu, X., Cai, L., Zhang, J., Prefontaine, G.G., Krones, A., Ohgi, K.A., Zhu, P., Garcia-Bassets, I. et al. 2007. Opposing LSD1 complexes

function in developmental gene activation and repression programmes. *Nature* 446(7138): 882-887.

Wang, J.H., Avitahl, N., Cariappa, A., Friedrich, C., Ikeda, T., Renold, A., Andrikopoulos, K., Liang, L., Pillai, S., Morgan, B.A. et al. 1998. Aiolos regulates B cell activation and maturation to effector state. *Immunity* 9(4): 543-553.

Weiss, M.J. and Orkin, S.H. 1995. GATA transcription factors: key regulators of hematopoiesis. *Exp Hematol* 23(2): 99-107.

Winandy, S., Wu, P., and Georgopoulos, K. 1995. A dominant mutation in the Ikaros gene leads to rapid development of leukemia and lymphoma. *Cell* 83(2): 289-299.

Xie, H., Ye, M., Feng, R., and Graf, T. 2004. Stepwise reprogramming of B cells into macrophages. *Cell* 117(5): 663-676.

Zhou, M., Ouyang, W., Gong, Q., Katz, S.G., White, J.M., Orkin, S.H., and Murphy, K.M. 2001. Friend of GATA-1 represses GATA-3-dependent activity in CD4+ T cells. *J Exp Med* 194(10): 1461-1471.

FIGURE LEGENDS

Figure 1. Purification of FOG-1 multiprotein complexes from 230-238 Pre-B-cells

(A) From top to bottom, schematic diagram of the pMSCV vectors used for stable expression of GFP, alone or in combination with FOG-1 or FLAG-FOG-1 in 230-238 cells. 5'LTR=5'Long Terminal Repeat, Psi=extended viral Packaging signal, $_{PGK}$Pr=murine phosphoglycerate kinase Promoter, eGFP=enhanced Green Fluorescent Protein, 3'LTR=3'Long Terminal Repeat. (B) Schematic diagram showing anti-FLAG affinity purification of FLAG-FOG-1 containing complexes. (C) Colloïdal Coomassie blue stained NuPAGE® Novex® Bis-Tris Mini Gel of purified material from a representative affinity purification. MW=Molecular Weight in kilo Dalton. (D) Western blot analysis of immunoprecipitated material. Cells expressing FLAG-FOG-1 and GFP (right), or GFP only (left, control cells) were immunoprecipitated under low stringency condition with an anti-FLAG antibody. Subsequently, co-immunoprecipitating proteins were detected by western blotting with antibodies against the indicated factors. (E) Western blot analysis of immunoprecipitated material. Cells expressing FLAG-FOG-1 (right), or FOG-1 (left, control cells) were immunoprecipitated under high stringency condition with an anti-FLAG antibody. Subsequently, co-immunoprecipitated proteins were detected by western blotting with antibodies against the indicated factors. Anti-Brg1 (D) and anti-p15 (E) were used as negative controls and Ponceau S staining of the membranes is shown as a loading control.

Figure 2. Interaction between endogenous FOG-1 and CtBP1, Aiolos

(A) 230-238 cell extracts were used for immunoprecipitation with antibodies against HA (control), FOG-1 or CtBP1, as indicated at the bottom. The precipitated material was then probed by western blotting for the presence of FOG-1 or CtBP1. (B) Immunoprecipitation of 70Z/3 cell extracts was done with antibodies against HA (control), FOG-1 or CtBP and the precipitates were analyzed for the presence of FOG-1, Aiolos or CtBP1. In all cases (A-B) the IgG light chain is shown as a loading control.

Figure 3. FOG-1 and Ikaros1 activate the endogenous *epcam* locus.

(A) Scheme depicting the *epcam* promoter region with the Transcription Start Site (TSS) and consensus binding sites for Ikaros1. (B) Scheme of the experimental design. 230-238 cells were first infected with viruses expressing FOG-1 or control, together with RFP and positive cells were isolated. Subsequently, cells were infected with viruses expressing Ikaros, Aiolos or control, together with GFP and positive cells were isolated. (C) SYBR green real-time RT-PCR analysis of endogenous *epcam* gene expression in the different FACS sorted 230-238 $RFP^{+}GFP^{+}$ cells, as indicated. Note that the scales of the y axis used for the experiments involving Aiolos (left panel) or Ikaros1 (right panel) are different. Gapdh mRNA transcript levels were used for normalization. (D) Extracts from 230-238 cell expressing RFP and GFP only (control), Ikaros1 (Ik1), or FOG-1 and Ikaros1 (FOG-1/Ik1) were used for immunoprecipitation experiments with an anti-FLAG antibody in order to precipitate the FLAG-tagged Ikaros protein. The precipitates were probed for the indicated proteins. Anti-Brg1 was used as negative control. 230-238 Pre-B-cell line overexpressing RFP and GFP was used as control for non specific interactions and an unspecific band is shown as loading control.

Figure 4. FOG-1 molecular mechanism of action at the endogenous *epcam* locus in Pre-B-cells

(A) Analysis of FOG-1 proteins impaired in the interaction with NuRD or CtBP. SYBR green real-time RT-PCR analysis of FOG-1 and *epcam* mRNA transcript levels in 230-238 Pre-B-cells retrovirally infected with the empty vector or the vectors encoding the wild-type or mutant proteins. Left panel: from top to bottom, schematic representation of the wild-type (WT) FOG-1 protein, the NuRD partial mutant, the double mutant (partially mutated in the NuRD binding site and mutated in the CtBP binding site) and the C-terminally truncated protein. Middle panel: FOG-1 mRNA transcript levels, to assess expression of the different constructs. Right panel: endogenous *epcam* locus mRNA transcript levels. Gapdh mRNA transcript levels were used for normalization. (B) Analysis of FOG-1 proteins mutated in some of the identified phosphorylation sites. SYBR real-time RT-PCR analysis of FOG-1 and *epcam* mRNA transcript levels in 230-238 Pre-B-cells retrovirally infected with the empty

vector or the vectors encoding the wild-type or phosphoacceptor site mutant proteins. Left panel : from top to bottom, schematic representation of the wild-type protein, the FLAG-FOG-1pS286A mutant, the FLAG-FOG-1pS925A mutant and the FLAG-FOG-1pS286AS925A double mutant. Middle panel : FOG-1 mRNA transcript levels, to assess expression of the different constructs. Right panel : endogenous *epcam* mRNA transcript levels. Gapdh mRNA transcript levels were used for normalization.

Figure 5. Requirement of FOG-1 serine 286 for the synergistic effect with Ikaros1 to activate the *epcam* locus

(A) Scheme of the experimental design. 230-238 cells were first infected with viruses expressing FOG-1 or mutant versions, together with GFP and positive cells were isolated. Subsequently, cells were infected with viruses expressing H2Kk as a marker and Ikaros1 or control. Double positive cells (GFP/H2Kk) were used for the analysis. (B) SYBR green real-time RT-PCR analysis of endogenous *epcam* gene expression in FACS sorted 230-238 $GFP^{+}H2Kk^{+}$ cells. Gapdh mRNA transcript levels were used for normalization. (C) Model of activation at the *epcam* locus.

Figure 1
Versavel *et al*.

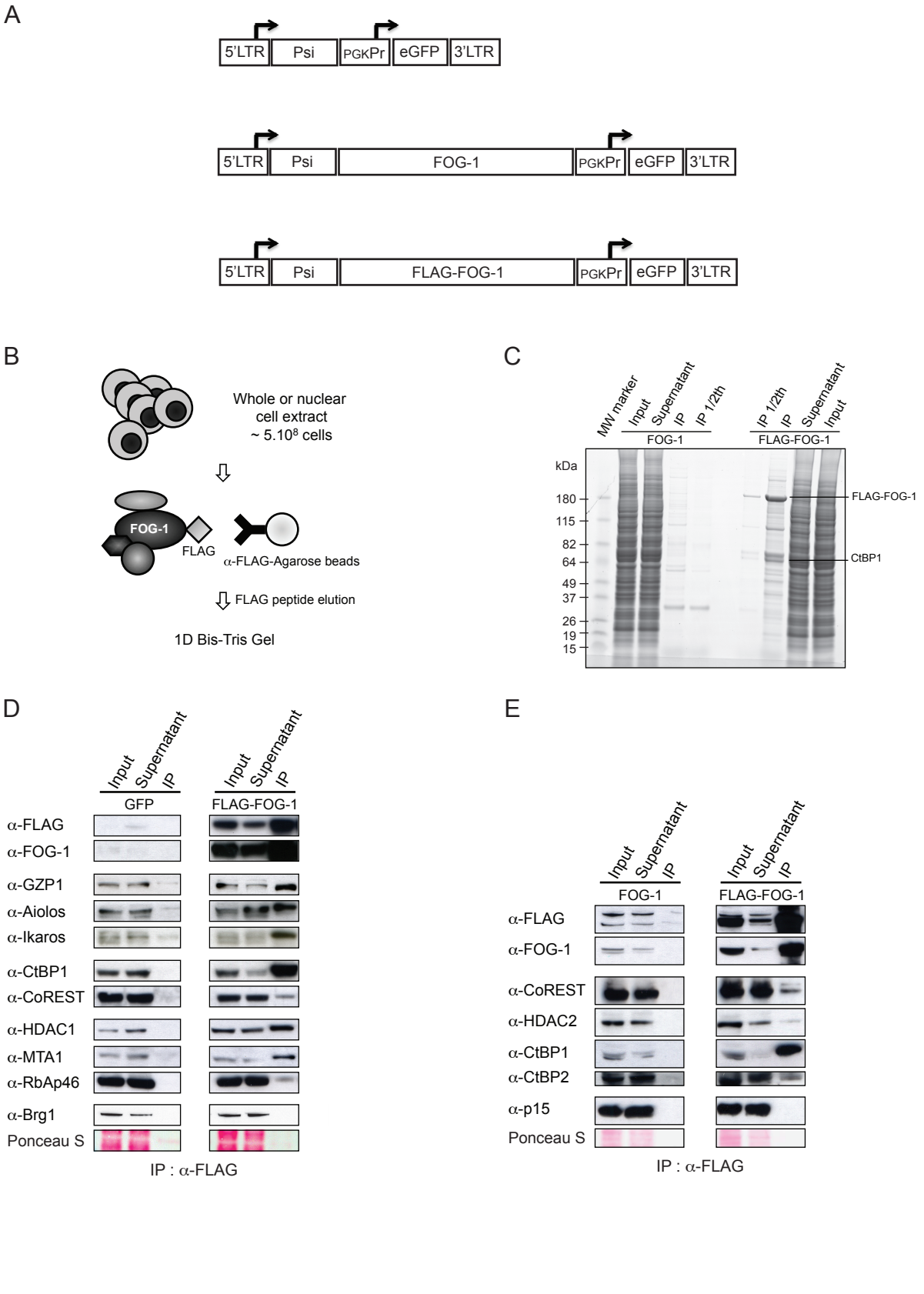
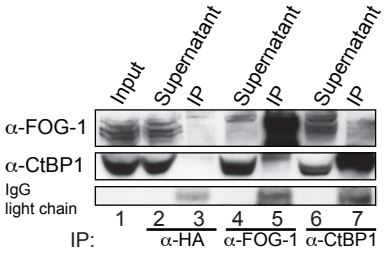


Figure 2
Versavel *et al*.

A



B

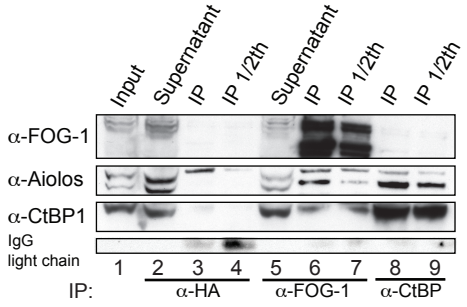
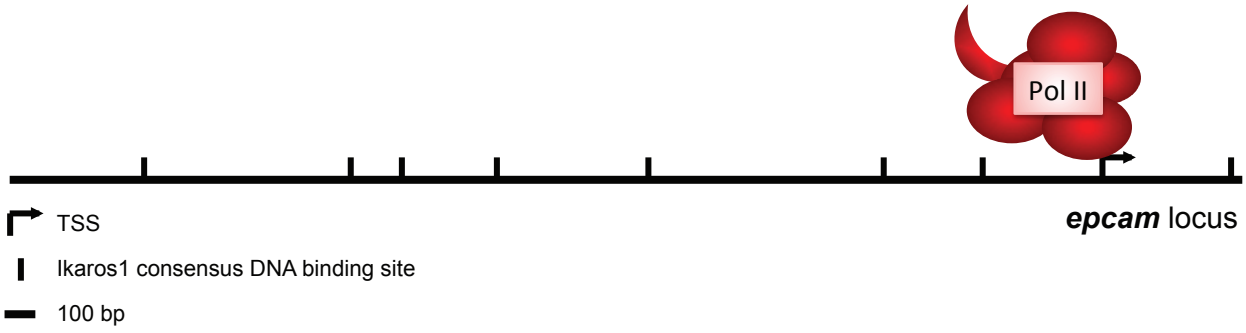
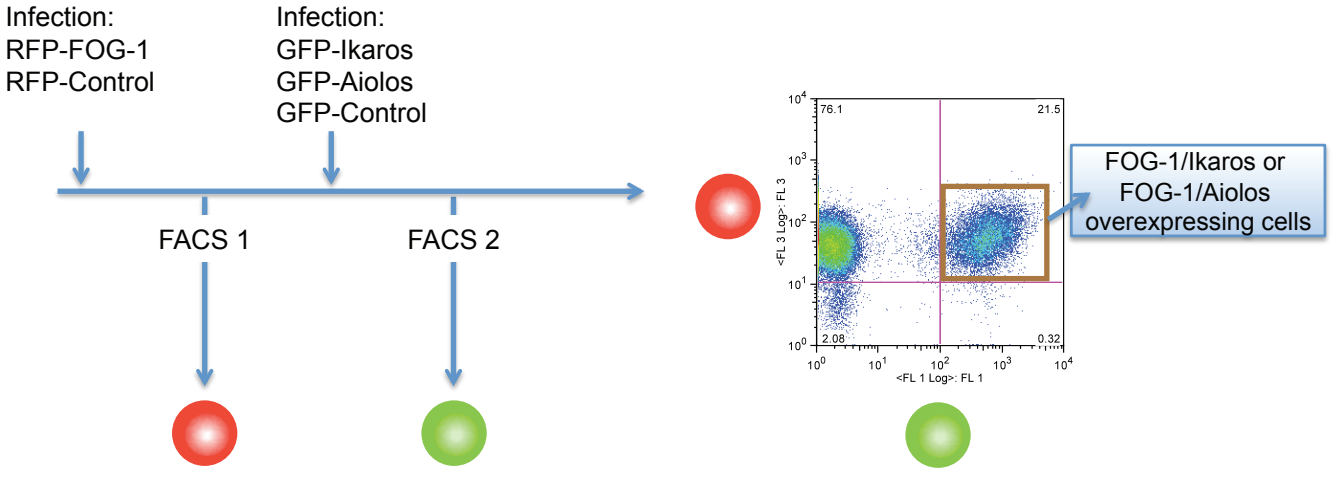


Figure 3
Versavel *et al*.

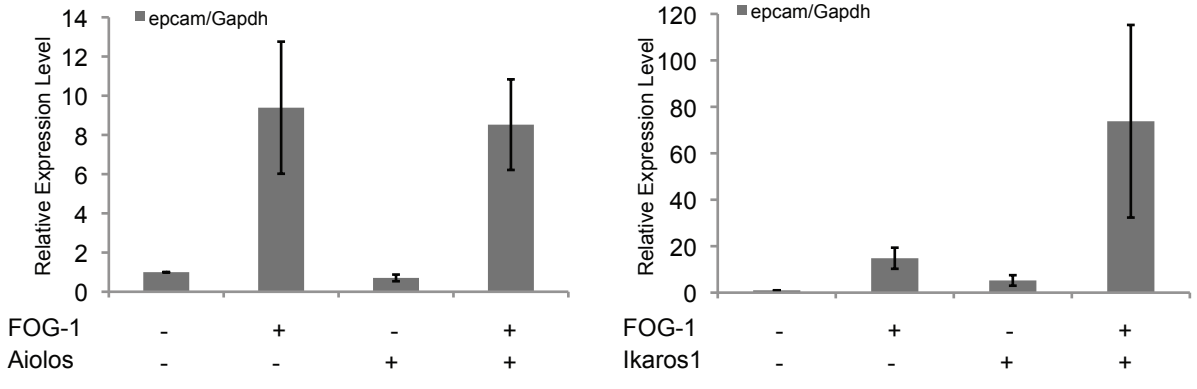
A



B



C



D

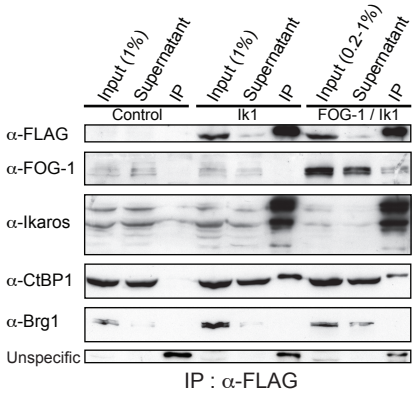


Figure 4
Versavel *et al*.

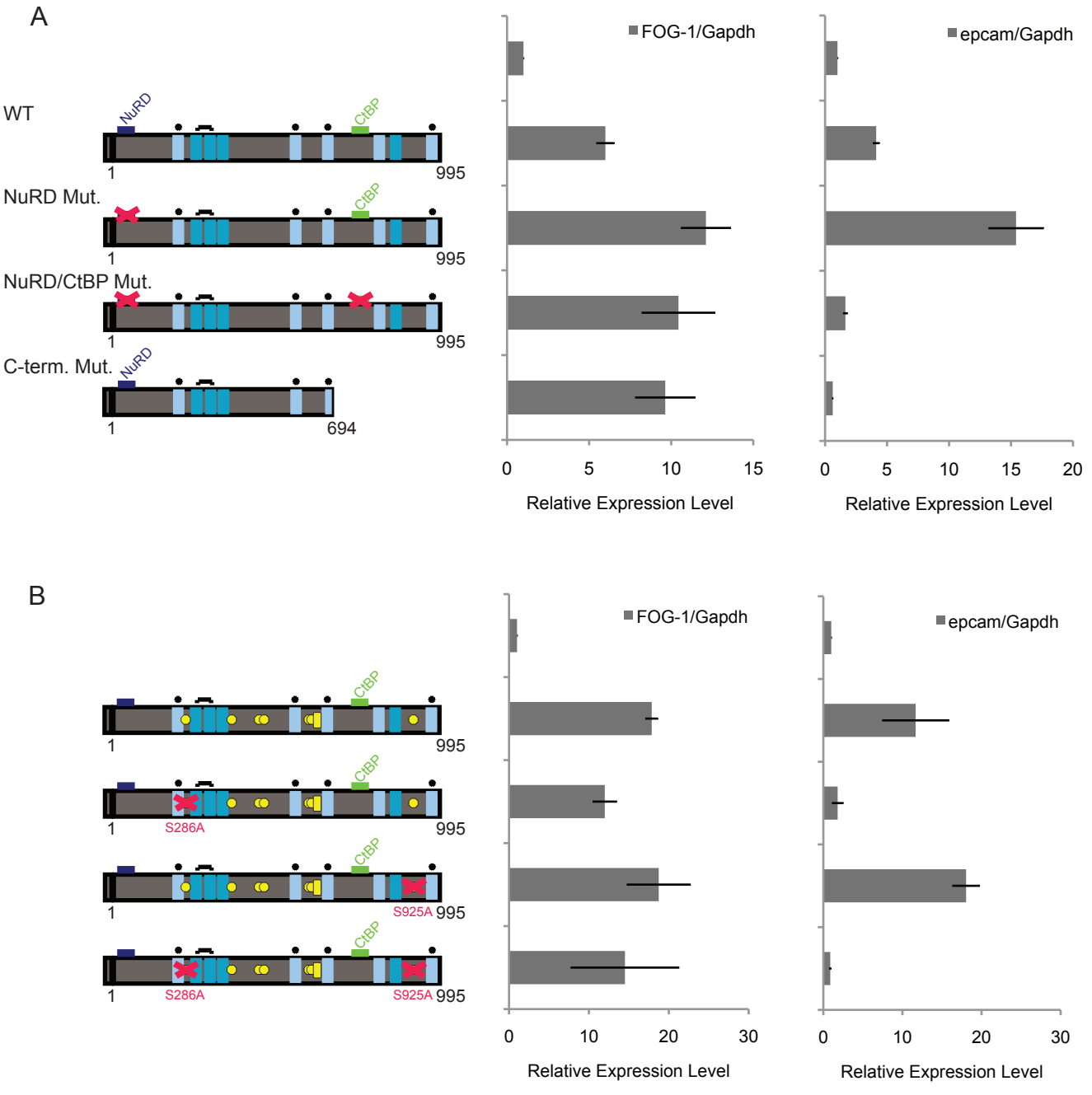
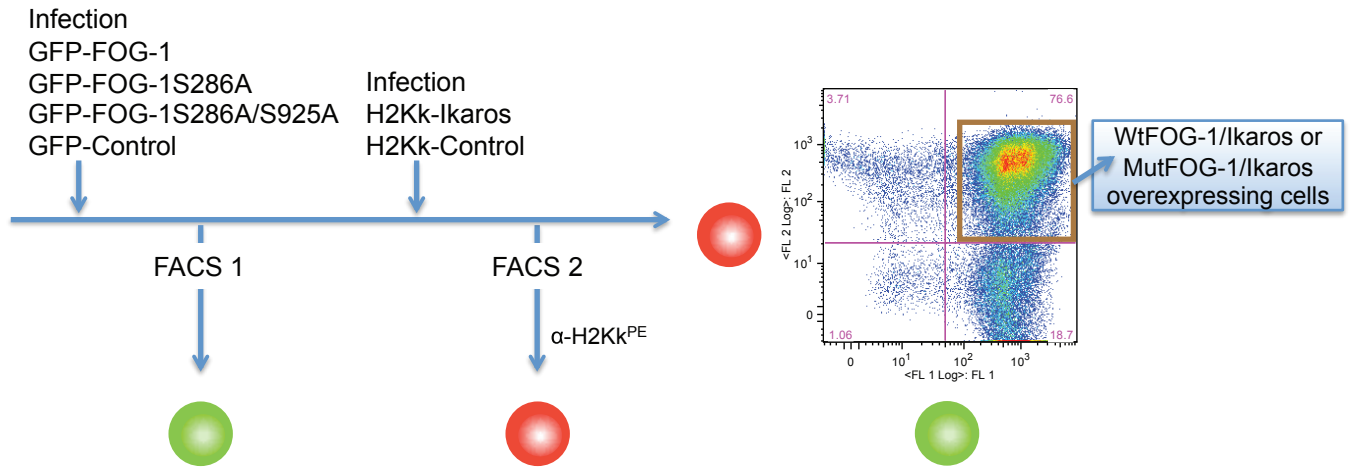
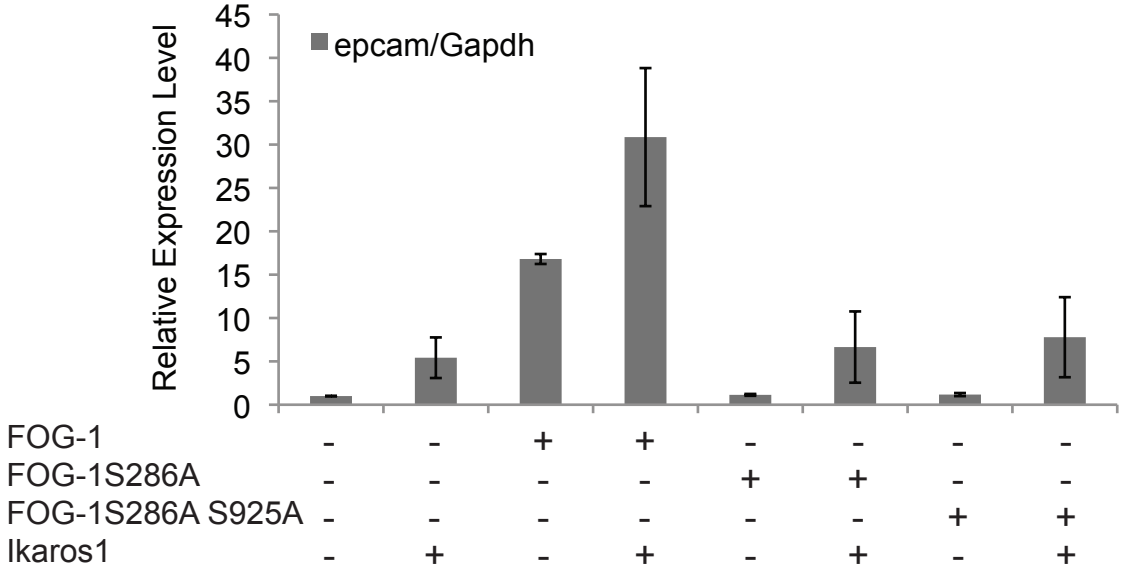


Figure 5
Versavel *et al*.

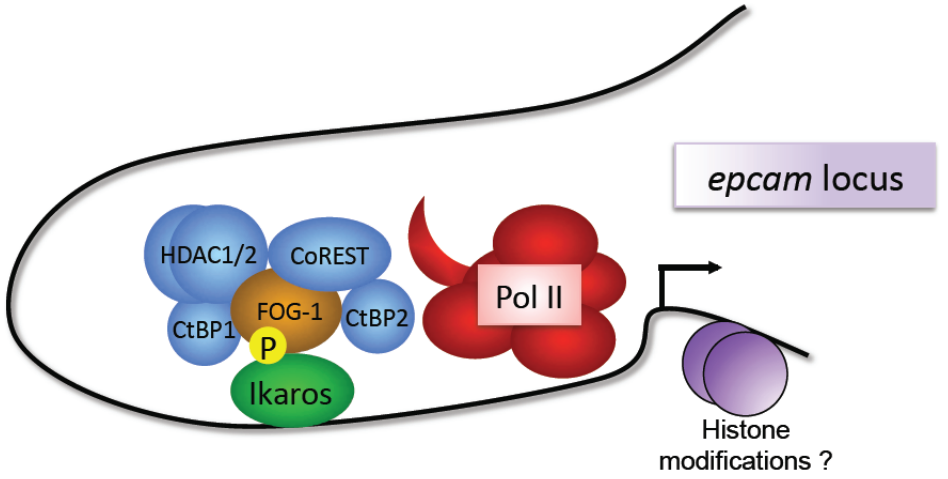
A



B



C



SUPPLEMENTARY FIGURE LEGENDS

Supplementary Figure 1. Specific interaction between FOG-1 and CtBP1 at the Pro-/Pre-B-cell stage of lymphocyte differentiation

IP under low salt stringency of endogenous CtBP1 by FLAG-FOG-1 in the Pro-/Pre-B-cell line PD31. PD31 Pro-/Pre-B-cell line overexpressing the GFP with FOG-1 was used as control for non specific interactions.

Supplementary Figure 2. Interaction between FOG-1 and Aiolos or Ikaros isoforms

IP under low salt stringency of endogenous Aiolos and Ikaros by FLAG-FOG-1 in 230-238 Pre-B-cell line ; full picture showing interactions between FOG-1 and different isoforms of the Ikaros-family members. 230-238 Pre-B-cell line overexpressing only the GFP was used as control for non specific interactions.

Supplementary Figure 3. Stoichiometry in favor of the CtBP/FOG-1 complex compared to the NuRD/FOG-1 complex in 230-238 Pre-B-cells

(A) Diagram showing the procedure for MRM analysis. LC/MS/MS=Liquid Chromatography with tandem Mass Spectrometry ; IDA=Information Dependent Acquisition ; MRM=Multiple Reaction Monitoring (B) MRM chromatograms of control and study samples, label-free experiment, corrected for difference in amount loaded. Dotted line : control ; straight line : FLAG-FOG-1. A - FOG-1_MOUSE peptide LVTEPHGAPR (transition 538.9/634.1) ; B - KPYM_MOUSE peptide GSGTAEVELK (transition 495.9/688.1), reference peptide ; C - CtBP1_MOUSE peptide GETLGIIGLGR (transition 543.3/685.2) ; D - HNRPK_MOUSE peptide IILDLISESPIK (transition 670.9/1114.6), reference peptide. (C) Table showing the results arising from anti-FLAG affinity purification of FLAG-FOG-1 containing complexes in tagged FOG-1 overexpressing cells. Analyses were performed from whole cell extracts. After gel separation, the proteins were identified by LC/MS/MS followed by MASCOT database search against Uniprot (2010_09). For LC-MRM analyses, peptide lists were generated out of the Mascot.dat file using MRM Buddy, a home made software. Peak areas were integrated with Analyst and peak ratios (sample versus control) as well as standard errors were calculated in Excel. Shown is the result list for Exp.1/Cellular extract (see Supplementary Table 1).

Supplementary Figure 4. Interaction between endogenous FOG-1 and CtBP1 or MTA1

Immunoprecipitation of 230-238 cell extracts was done with antibodies against HA (control), CtBP1 or MTA1 and the precipitates were probed for the presence of FOG-1 or CtBP1. The IgG light chain is shown as a loading control.

Supplementary Figure 5. Increase in the total number of cells overexpressing FOG-1 or FLAG-FOG-1 compared to cells only overexpressing GFP

(A) Growth curves of 230-238 GFP^+ and 230-238 GFP^+ FLAG-FOG-1 cells. The latter were favored in culture than 230-238 GFP^+ cells. A graph representative of several experiments is depicted. (B) Cell cycle analysis by flow cytometry. Following a 30-min pulse of BrdU, cells were collected and fixed according to the manufacturer's instructions (APC BrdU Flow Kit, BD Pharmingen^{TM}). After staining with antibodies against BrdU and 7-AAD, cells were analyzed by FACS. Representative dot plots are shown for the different cell lines. A diagram representing the different states of the cells : the resting state called GO and the cell cycle phases (G1, S, M and G2) is shown on the left of the dotplots.

Supplementary Figure 6. Microarray analysis on established 230-238 stable Pre-B-cell line with enforced expression of FOG-1

Volcano plot showing *epcam* upregulation upon FOG-1 upregulation (upper panel).

Heatmap showing *epcam* upregulation upon *Zfpm1* upregulation (lower panel).

Supplementary Figure 7. FOG-1, Aiolos and Ikaros1 expression level

SYBR green real-time RT-PCR analysis of FOG-1, Aiolos and Ikaros1 gene expression in the different FACS sorted 230-238 RFP^+GFP^+ cells. Gapdh mRNA transcript levels were used for normalization.

Supplementary Figure 8. Two FOG-1 mutants unable to bind CtBP1

Western blot analysis of immunoprecipitated material. Cells expressing HA-FOG-1K5A or HA-FOG-1dm and GFP (left panel: right) or FOG-1 only (left panel: left, control cells) were immunoprecipitated with an anti-HA antibody. Cells expressing

FLAG-FOG-1 or FLAG-FOG-1trunc and GFP (right panel: right) or GFP only (right panel: left, control cells) were immunoprecipitated with an anti-FLAG antibody. Subsequently, co-immunoprecipitating proteins were detected by western blotting with antibodies against FOG-1 or CtBP1.

Supplementary Figure 9. GATA-1 independent transcriptional activation at the *epcam* locus

SYBR green real-time RT-PCR analysis for FOG-1 and *epcam* mRNA transcript levels in 230-238 Pre-B-cells retrovirally infected with the empty vector (pMSCV-EGFP) or the vectors encoding the wild-type (pMSCV-FLAG-FOG-1-EGFP) or mutant (pMSCV-FLAG-FOG-1ΔZn6$_{GATA1}$-EGFP, pMSCV-FLAG-FOG-1ΔZn6Zn9$_{GATA1}$-EGFP, pMSCV-FLAG-FOG-1ΔZn1Zn6Zn9$_{GATA1}$-EGFP) proteins. Left panel : from top to bottom, schematic representation of the wild-type (WT) protein and the partial mutants for GATA-1 binding sites (Zn6$_{T26A}$, Zn6/9$_{T26A}$, Zn1/6/9$_{T26A}$). Middle panel : FOG-1 mRNA transcript levels. Right panel : endogenous *epcam* mRNA transcript levels. Gapdh mRNA transcript levels were used for normalization.

Supplementary Figure 10. Mass Spectrometry analysis of (FLAG-)FOG-1 phosphorylation in 230-238 Pre-B-cell line

(A) Tryptic peptides in the sequence of full-length FOG-1 resulting from the excision of overexpressed FLAG-FOG-1 from the colloïdal blue-stained gel and analysis by LC/MS/MS. The 12 phosphoacceptor sites are depicted by asterisks and the consensus sequence for 7 among them is underlined in green. The table precises the positions of the phosphoserines in full-length FOG-1. « or » means « p-sites are not unambiguously identified » ; only the peptide containing one phosphorylation could be identified, the alignment of the p-sites was not possible. (B) Positions of identified phosphorylation sites (yellow circles or rectangle) within the FLAG-FOG-1 protein. (C) Sequence alignment between human FOG-1 (hFOG-1, top) and mouse FOG-1 (mFOG-1, bottom). Blue boxes are underlying not conserved residues between hFOG-1 and mFOG-1. Green bar plots are displaying the percentage of conservation between hFOG-1 and mFOG-1. Orange boxes are localizing serine residues identified as phosphoacceptor sites by MS.

Supplementary Figure 11. mRNA encoding wild-type and mutant FOG-1 proteins and Ikaros1 expression level

SYBR green real-time RT-PCR analysis of FOG-1 and Ikaros1 gene expression in the different FACS sorted 230-238 $GFP^{+}H2Kk^{+}$ cells. Gapdh mRNA transcript levels were used for normalization.

Supplementary Table 1. FOG-1 complexes and associated proteins identified by MS

Table showing the results arising from anti-FLAG affinity purification of FLAG-FOG-1 containing complexes in tagged FOG-1 overexpressing cells. Analyses were performed from nuclear or cellular extracts. After gel separation, the proteins were identified by LC/MS/MS followed by MASCOT database search against Uniprot (2010_09). Number of assigned spectra and sequence coverage are compiled with Scaffold (min. protein ID probability : 50% ; min. number of peptides : 1 ; min. peptide ID probability : 90%).

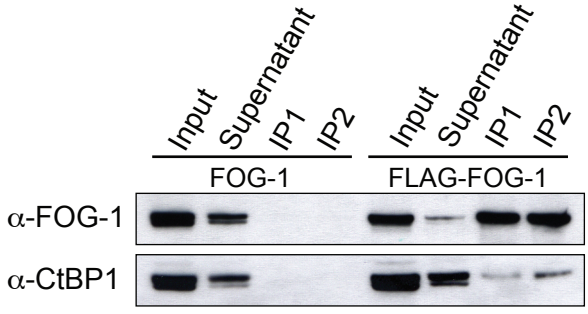
Supplementary Table 2. Ikaros1 complexes and associated proteins identified by MS

Table showing the results arising from anti-FLAG affinity purification of FLAG-Ikaros1 containing complexes in FOG-1 or/and tagged Ikaros1 overexpressing cells. After gel separation, the proteins were identified by LC/MS/MS followed by MASCOT database search against Uniprot (2010_09). Numbers of assigned spectra and sequence coverage are compiled with Scaffold (min. protein ID probability : 50% ; min. number of peptides : 1 ; min. peptide ID probability : 90%).

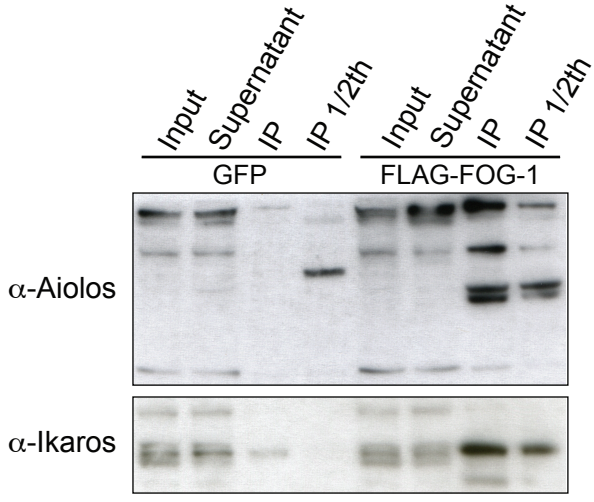
Supplementary Table 3. Primers forward and reverse for SYBR green real-time RT-PCR

Aiolos, epcam, FOG-1, GAPDH and Ikaros primer pair sequences are given.

Supplementary Figure 1
Versavel *et al*.

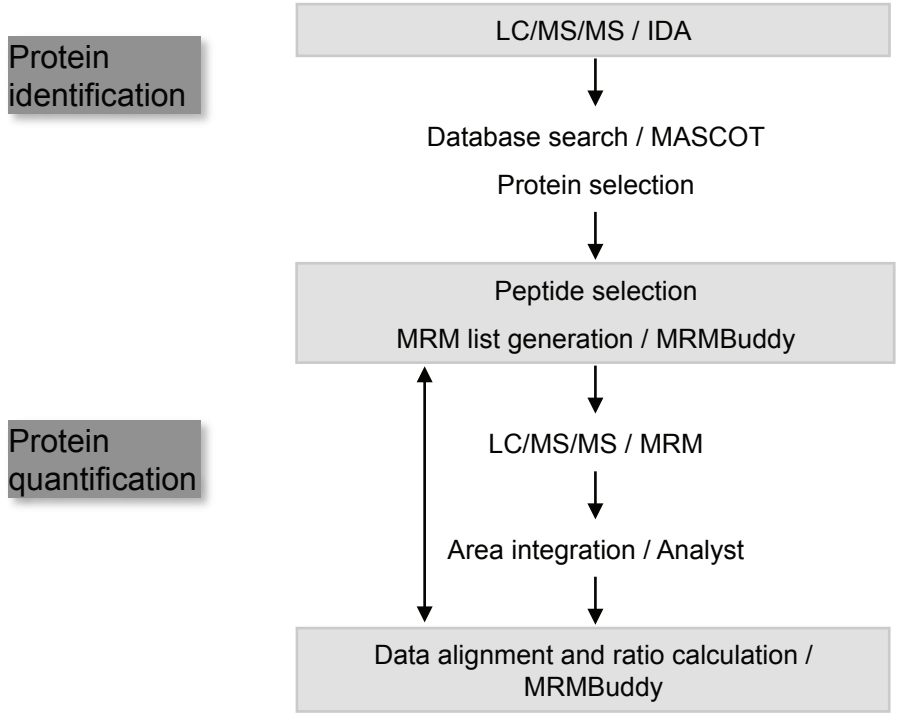


Supplementary Figure 2
Versavel *et al*.

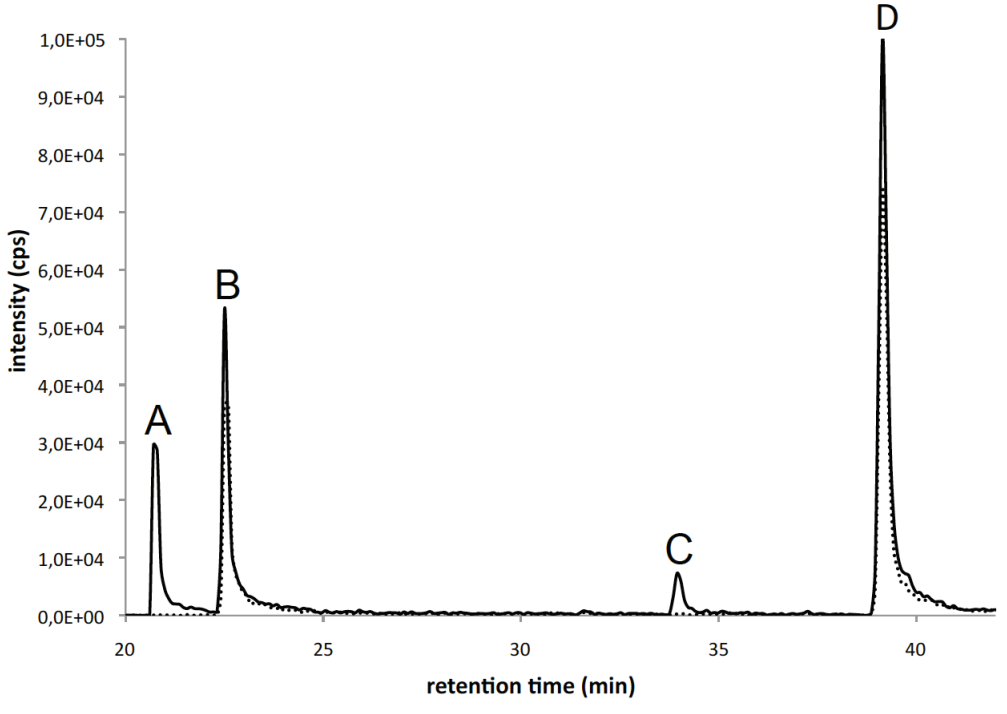


Supplementary Figure 3
Versavel *et al*.

A



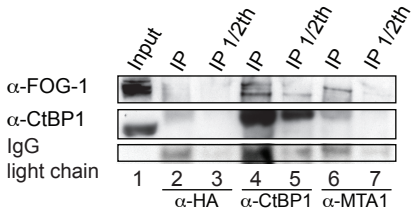
B



C

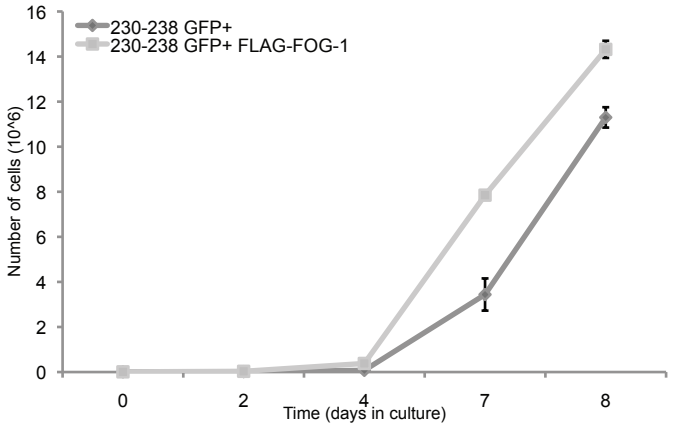
Protein name	Accession number	Molecular Weight	Fold change sample / control	Standard error
CtBP1	O88712	48 kDa	3448	716
FOG-1	O35615	106 kDa	2348	191
Chd4	Q6PDQ2	218 kDa	3.4	0.1
HDAC2	P70288	55 kDa	3.2	0.5
HDAC1	O09106	55 kDa	3.1	0.3
TIF1B	Q62318	89 kDa	1.9	0.3

Supplementary Figure 4
Versavel *et al*.

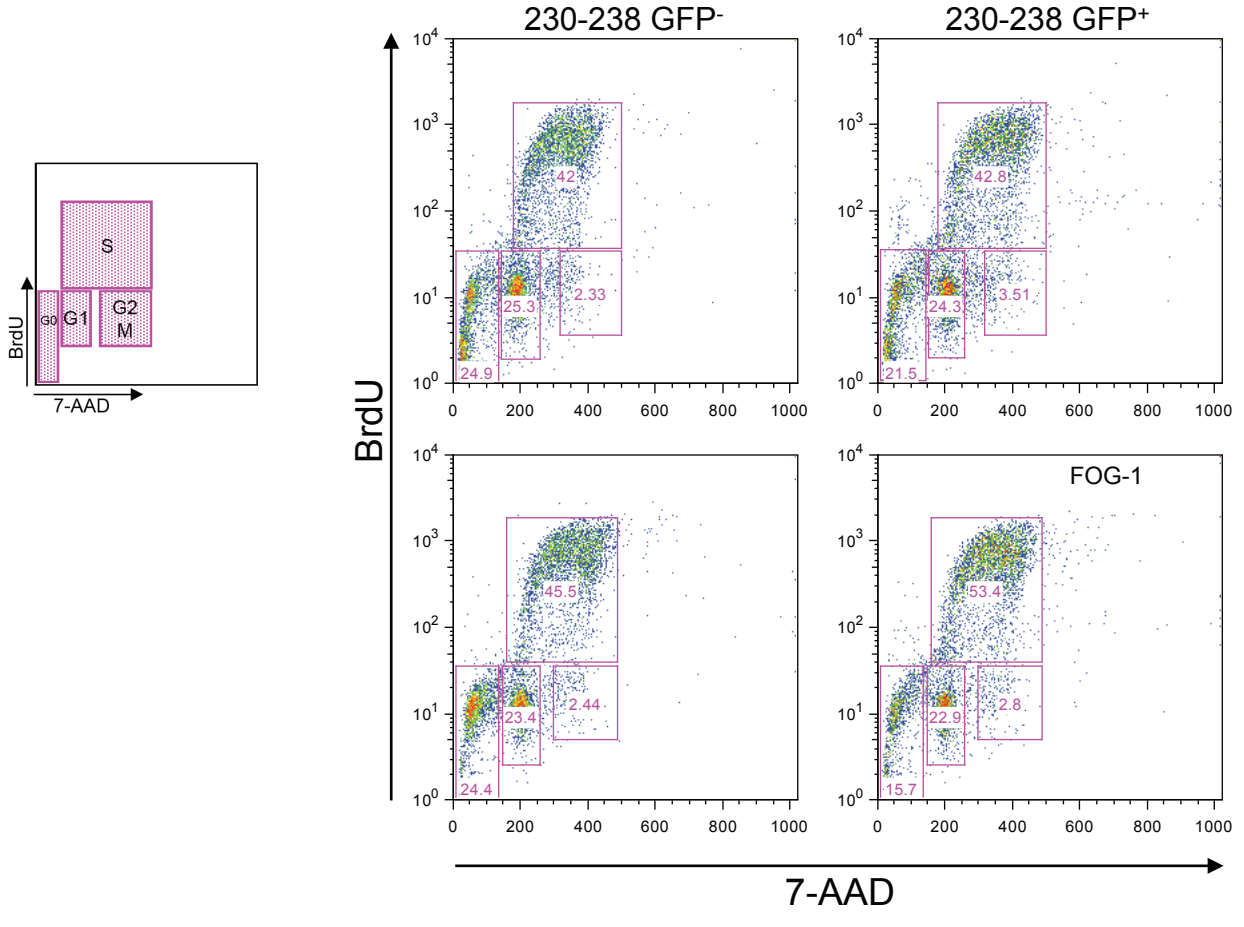


Supplementary Figure 5
Versavel *et al*.

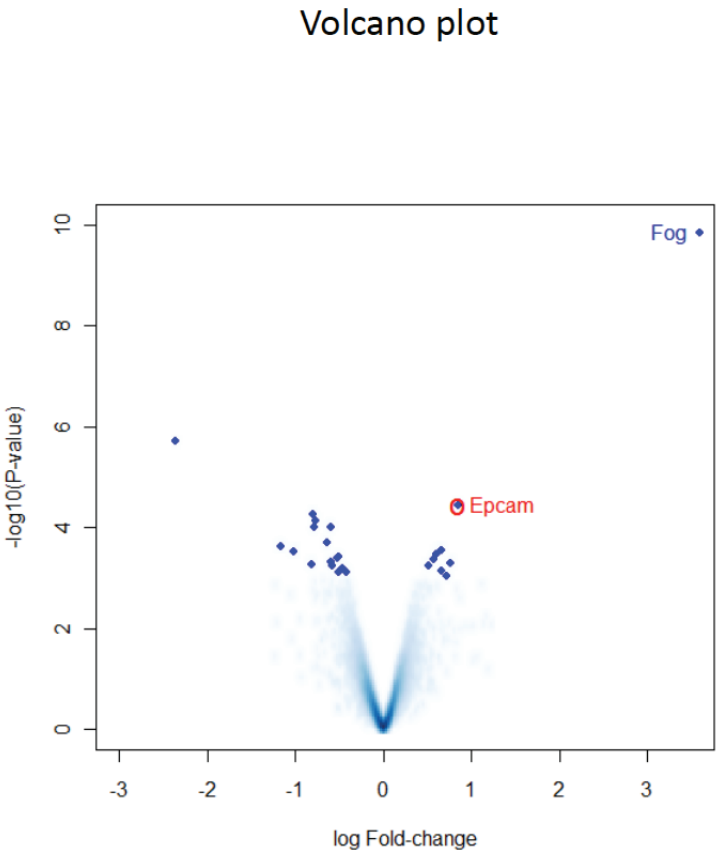
A

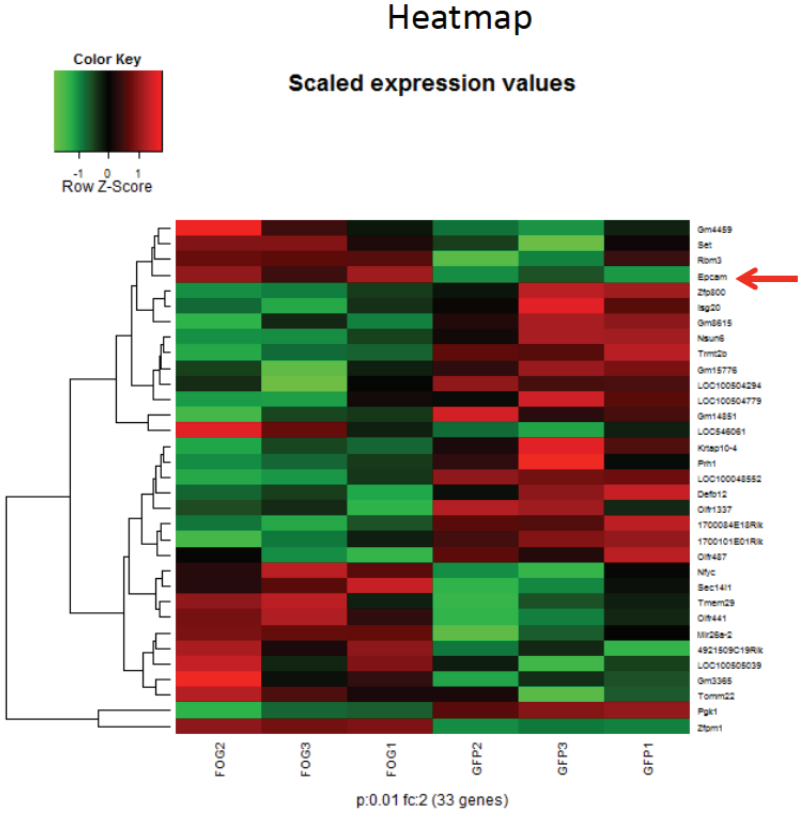


B

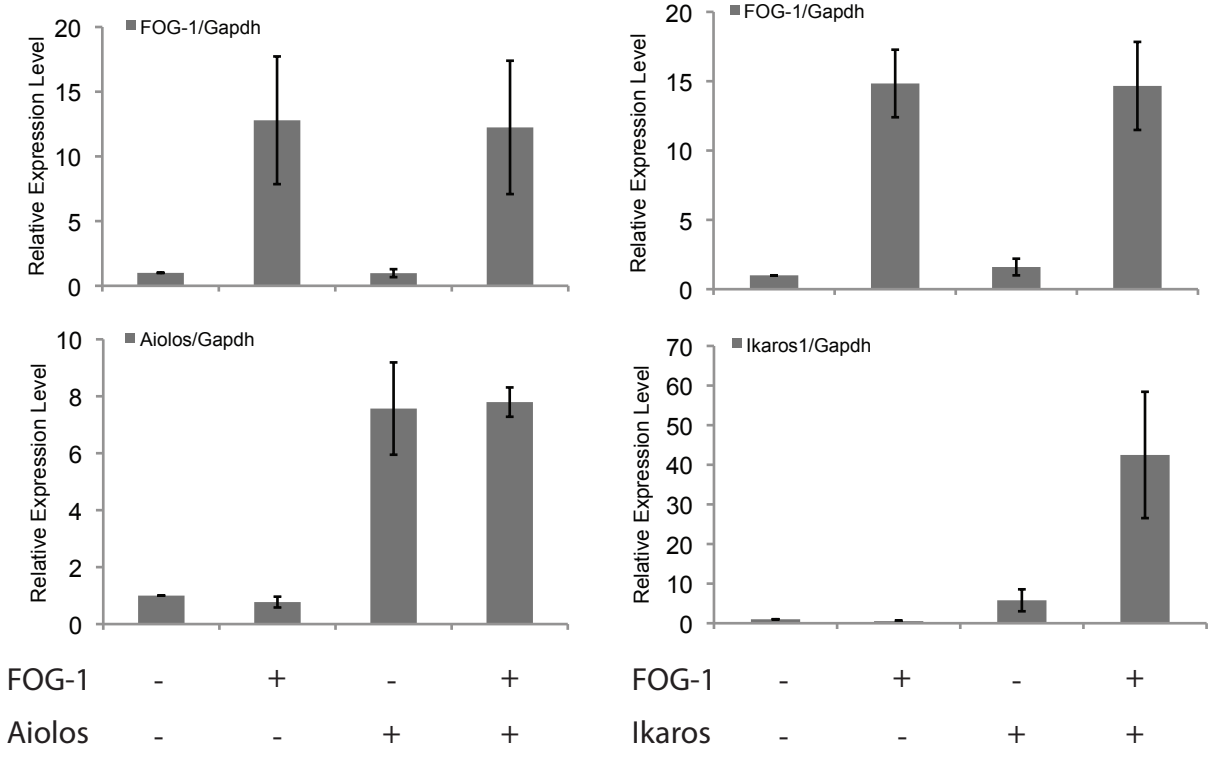


Supplementary Figure 6
Versavel *et al*.

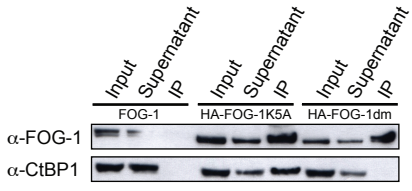


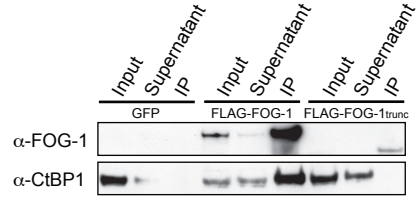


Supplementary Figure 7
Versavel *et al*.

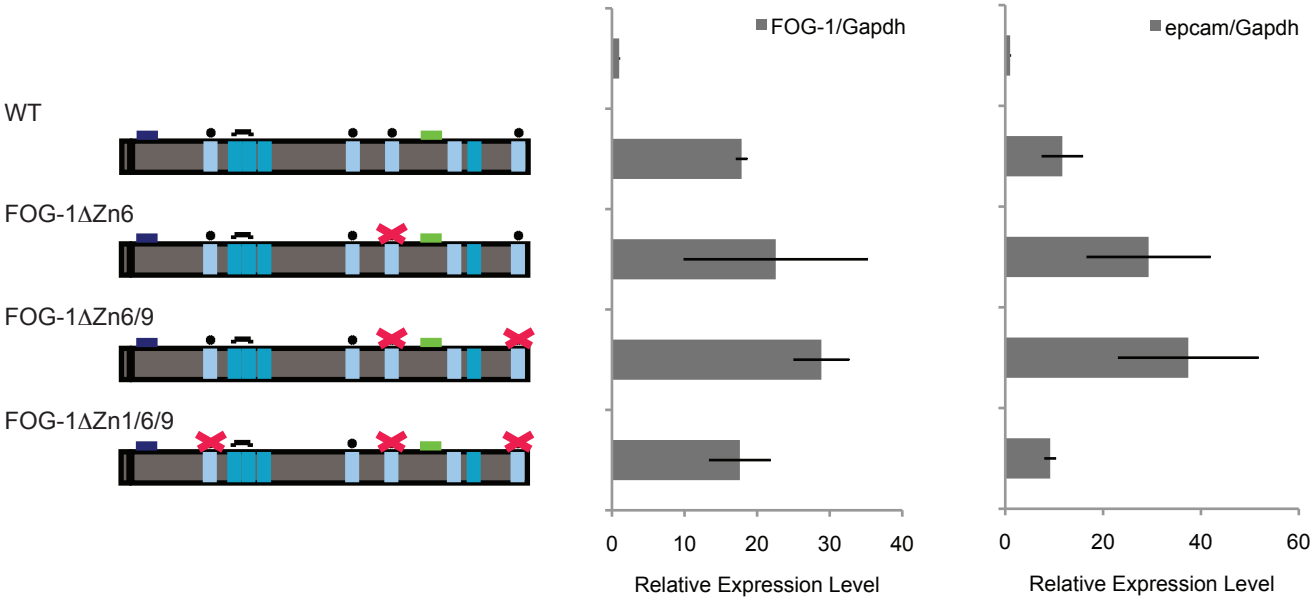


Supplementary Figure 8
Versavel *et al*.





Supplementary Figure 9
Versavel *et al*.



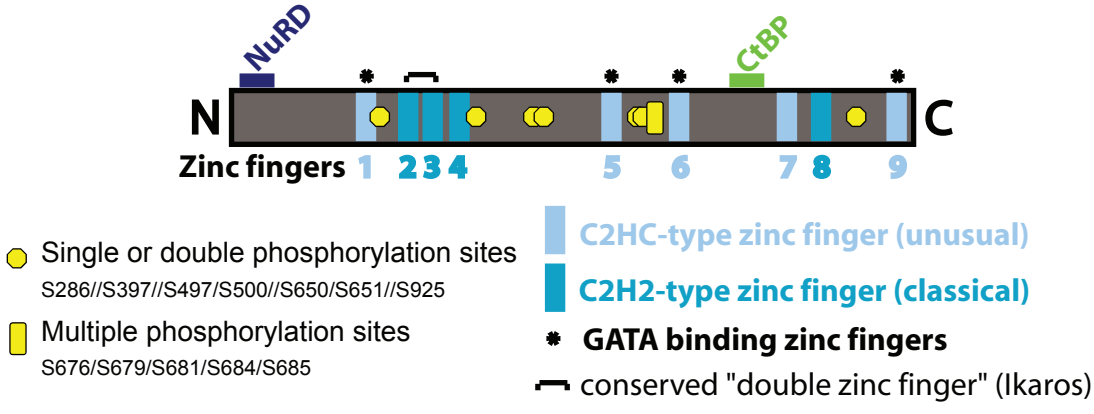
Supplementary Figure 10
Versavel *et al*.

A

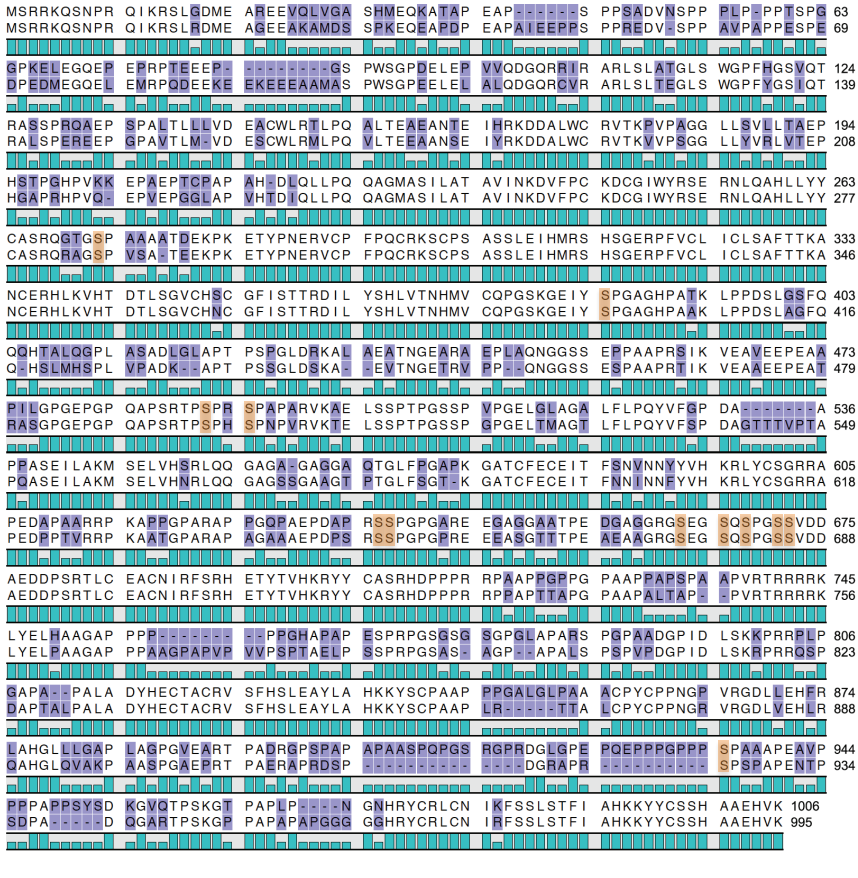
Tryptic peptides in the sequence of full-length FOG-1	Positions of the phosphoacceptor sites (pS) in full-length FOG-1
284-AGS*PVSATEEKPK-296	pS286
393-GEIYS*PGAGHPAAK-406	pS397
495-TPS*PHS*PNPVR-505	pS497 or pS500
637-APAGAAAEPDPSRS*S*PGPGPR-657	pS650 or pS651
675-GS*EGS*QS*PGS*S*VDDAEDDPSR-695	pS676 or pS679 or pS681 or pS684 or pS685
925-S*PSPAPENTPSDPADQGAR-943	pS925

The numbers preceding and following each peptide denote the position of the first and last amino acids of the tryptic peptides.
S*P= Consensus sequence of phosphoacceptor sites (asterisks)
pS= phosphoSerine

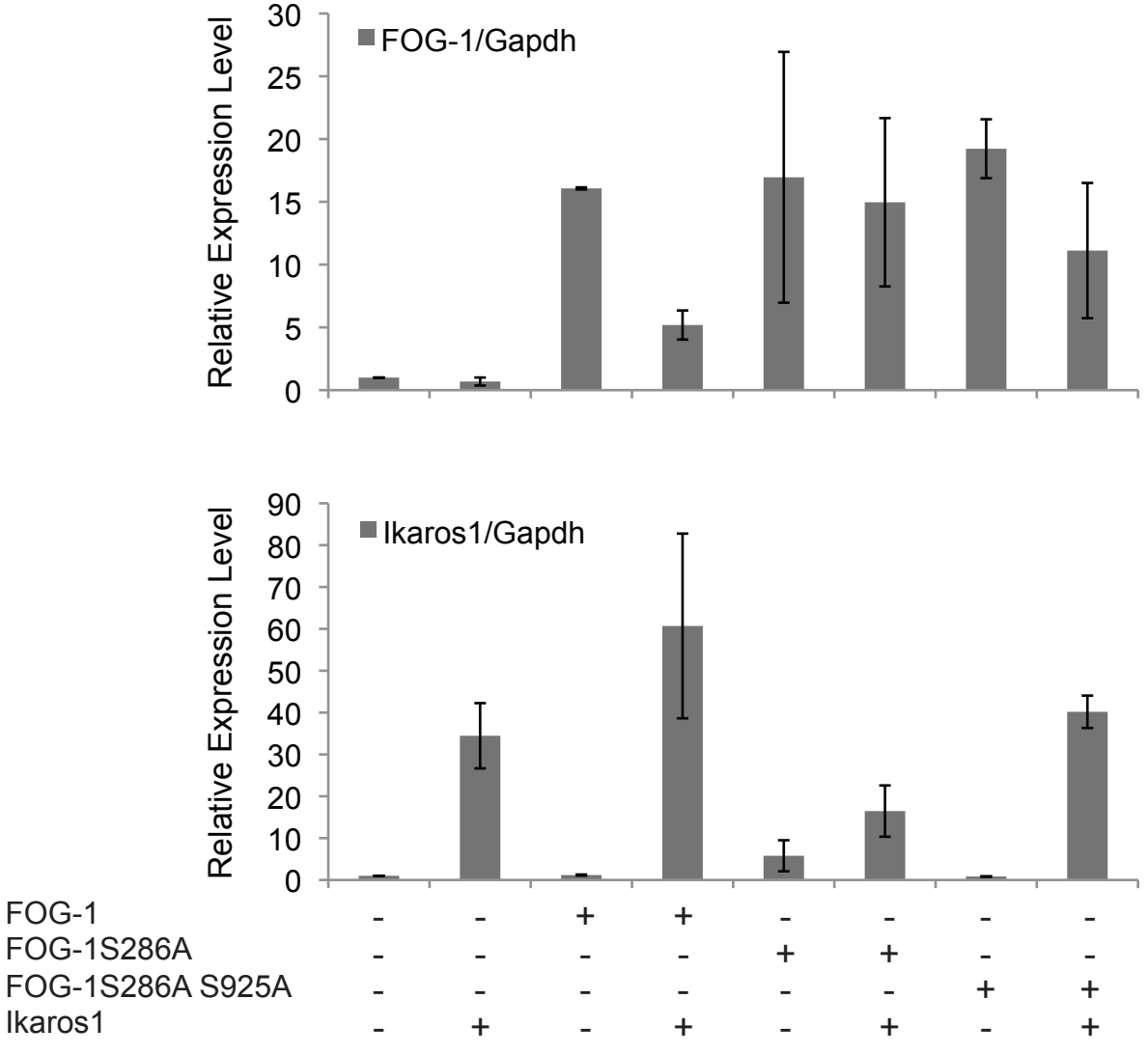
B



C



Supplementary Figure 11
Versavel *et al*.



Supplementary Table 1
Versavel *et al*.

			Nuclear extracts						Cellular extracts			
			Exp.1		Exp.2		Exp.3		Exp.1		Exp.2	
Protein name	**Accession number**	**Molecular Weight**	**#Assigned spectra**	**Coverage**	**#Assigned spectra**	**Coverage**	**#Assigned spectra**	**Coverage**	**#Assigned spectra**	**Coverage**	**#Assigned spectra**	**Coverage**
FOG-1	O35615	106 kDa	31	13,0%	78	30,0%	131	39,0%	130	35,0%	220	40,0%
CtBP complex												
CtBP1	O88712	48 kDa	3	9,3%	8	10,0%	17	27,0%	25	13,0%	76	20,0%
HDAC1	O09106	55 kDa					2*)	5,0%	4	10,0%	1	1,9%
HDAC2	P70288	55 kDa							1	6,1%	1	3,7%
RCOR1	Q8CFE3	52 kDa									1	2,5%
ZEB2	Q9R0G7	136 kDa	1	0,7%			5	2,7%				
NuRD complex												
Chd4	Q6PDQ2	218 kDa			2	0,9%	3	1,1%	5	3,4%		
HDAC1	O09106	55 kDa					2*)	5,0%	4	10,0%	1	1,9%
HDAC2	P70288	55 kDa							1	6,1%	1	3,7%
MBD3	Q9Z2D8	32 kDa			1	5,6%			3	5,3%		
MTA1	Q8K4B0	81 kDa			3	4,9%	1	1,4%				
MTA2	Q9R190	75 kDa					2	4,2%				
RbAp48 (RBBP4)	Q60972	48 kDa	1	2,8%	2	3,5%	2	8,5%				
RbAp46 (RBBP7)	Q60973	48 kDa							2	3,8%		
Transcription machinery												
Aiolos	B1AQE6	58 kDa	1	2,2%								
Ikaros1	Q03267	57 kDa	1	3,3%	1	3,3%			1	3,3%	2	3,3%
RUNX1	Q03347	49 kDa			1	3,3%			2	3,3%	5	12,0%
GZP1	Q3U063	97 kDa	1	1,8%								
TIF1B	Q62318	89 kDa			1	1,9%	2	6,5%			4	9,1%

*) cannot be distinguished between HDAC1 and HDAC2

			Ikaros-1		FOG-1/Ikaros-1	
Protein name	**Accession number**	**Molecular Weight**	**#Assigned spectra**	**Coverage**	**#Assigned spectra**	**Coverage**
Ikaros1	Q03267	57 kDa	136	26,0%	72	24,0%
CtBP complex						
CtBP1	O88712	48 kDa	18	21,0%	4	8,4%
HDAC1	O09106	55 kDa	5	6,4%	3	5,0%
HDAC2	P70288	55 kDa	1	4,3%	1	5,0%
KDM1 (LSD1)	Q6ZQ88	93 kDa	1	1,8%		
NuRD complex						
Chd4	Q6PDQ2	218 kDa	7	4,8%	2	1,9%
HDAC1	O09106	55 kDa	5	6,4%	3	5,0%
HDAC2	P70288	55 kDa	1	4,3%	1	5,0%
MTA2	Q9R190	75 kDa	5	9,7%	4	4,2%
RBBP4	Q60972	48 kDa	7	11,0%	5	6,4%
RBBP7	Q60973	48 kDa	6	8,9%	4	4,9%
Transcription machinery						
Aiolos	B1AQE6	58 kDa	66	42,0%	22	19,0%
FOG-1	O35615	106 kDa			3	1,7%
TIF1B	Q62318	89 kDa	24	21,0%	29	20,0%

Primer	Forward	Reverse
Aiolos	5'-ACA GCA GAC CAA CCG GTG GGA A-3'	5'-ACT GGA ACG GGC GTT CGC-3'
epcam	5'-CAC AGA ATA CTG TCA TTT GC-3'	5'-AGC ACG TGG CGG TGC CGT TG-3'
FOG-1	5'-CCA ACT GTG AAC GCC ATC TC-3'	5'-GAT CTC ACC CTT GGA GCC TG-3'
GAPDH	5'-GCC TCG TCC CGT AGA CAA AAT-3'	5'-TTC CCA TTC TCG GCC TTG A-3'
Ikaros	5'-TCG GGA GAG AAA ATG AAT GG-3'	5'-AGG CCG TTC ACC AGT ATG AC-3'

Chapter 3 : Discussion

The two papers previously presented already discuss the function of FOG-1 : in the haematopoietic system as far as the first article is concerned ; with *in vitro* work in the B cell lineage for the second one. Nevertheless, some points and additional experiments remain to be described and explained. First of all, a summary of the findings already available on the *epcam* gene, the gene found as a target of FOG-1 in B cell lines is discussed. Second, general data arising from the experiments done in mice overexpressing FOG-1 in the entire haematopoietic tree and from the proteomics performed in B cell lines give rise to further investigations.

1. *epcam*, a gene involved in cancer

A recent meeting report (Baeuerle and Gires, 2007) pointed out the importance of understanding the regulation of *epcam* at the transcriptional level. Key progress on *epcam* biological significance are recapitulated below. EpCAM (CD326; Gires, Atlas Genet Cytogenet Oncol Haematol, 2008) is a transmembrane glycoprotein showing frequent and high-level expression in numerous human malignancies including colon and breast cancers (Spizzo et al., 2004; Went et al., 2006). It is indeed known to induce cell proliferation via the induction of the proto-oncogene *c-myc* and the *cyclin A* and *E* genes (Munz et al., 2004). Moreover, it has been previously shown to be expressed at low level in a fraction of murine B and pre B lymphomas (Bergsagel et al., 1992) and frequently as well as widely expressed on Cancer Stem Cells (Gires et al., 2009). Recently, tumour-associated antigen EpCAM was shown to mediate nuclear signaling explaining how this protein could function in cell proliferation (Maetzel et al., 2009).

2. Investigation of FOG-1 function in the haematopoietic system

2.1. *In vivo*

To investigate FOG-1 function *in vivo*, we used $R26^{FOG\text{-}1}$:VaviCre mice. Western blot analysis were performed to estimate the level of FOG-1 overexpression in splenocytes and thymocytes derived from $R26^{FOG\text{-}1}$:VaviCre animals. As shown in figure 18, an up-regulation in the overexpressing cells as compared to control cells was observed. We tried this assay on bone marrow cells but we unfortunately never could get proper results.

Investigations in more detail of $R26^{FOG\text{-}1}$:VaviCre mice would be required. Nevertheless, some preliminary results of blood analysis were quite promising. This was performed on 4 males (2 control and 2 $R26^{FOG\text{-}1}$:VaviCre mice) and revealed a decrease of one third of the white blood cells, one third of the lymphocytes and one third of the monocytes in $R26^{FOG\text{-}1}$:VaviCre mice (Fig. 19). Females were not analyzed yet. A preliminary explanation could be reminiscent of the phenotype observed in GATA-1.05/X mice. Indeed, these latter contain two types of haematopoietic cells, owing to the fact that GATA-1 gene is located on the X chromosome and to the process of X-inactivation. In

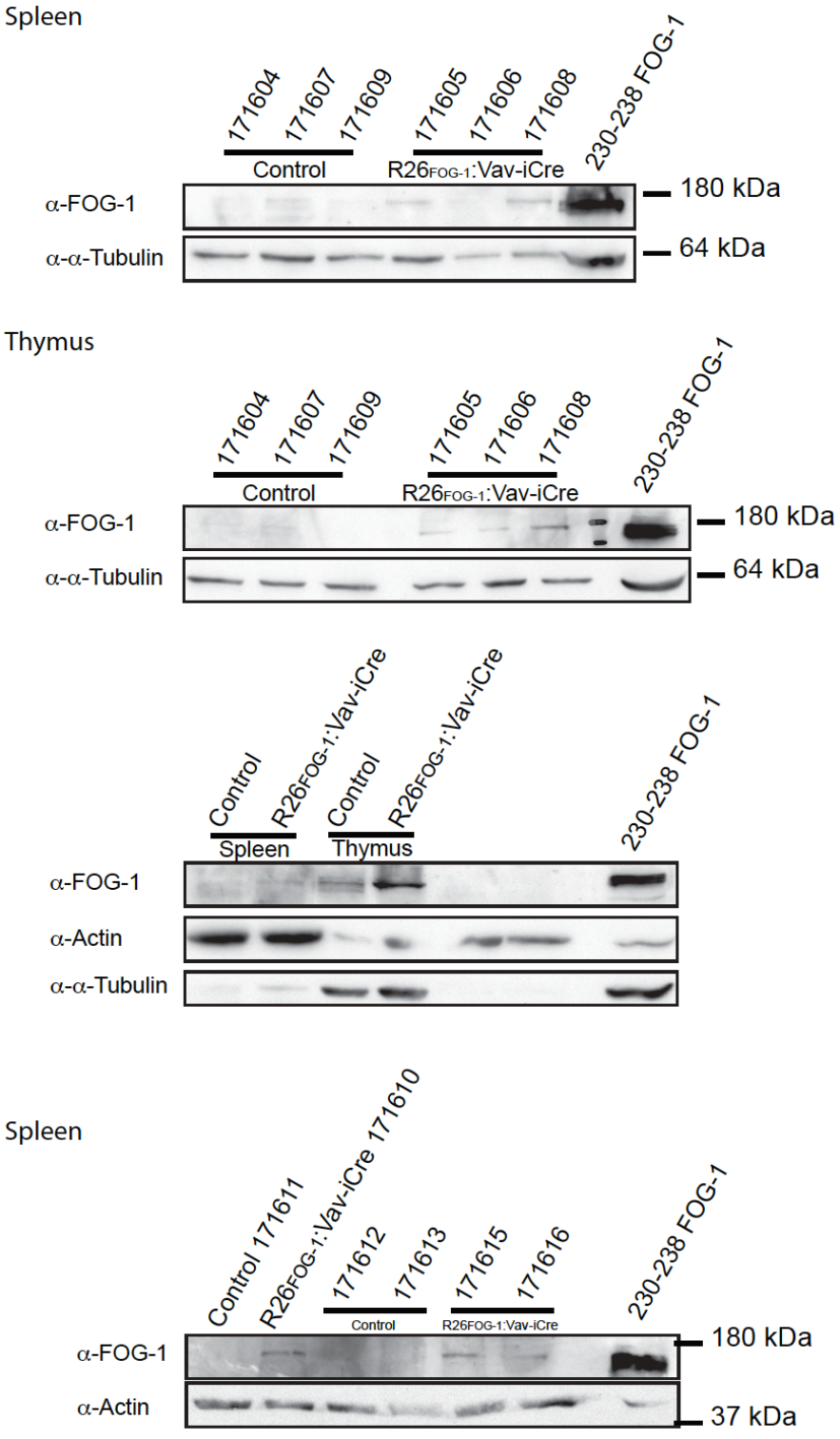


Figure 18 : Up-regulation of FOG-1 protein in splenocytes and thymocytes derived from R26^{FOG-1}:VaviCre mice. FOG-1, α-Tubulin and Actin were detected by western blotting in splenocytes derived from 7 control and 7 R26^{FOG-1}:VaviCre mice and thymocytes derived from 4 control and 4 R26^{FOG-1}:VaviCre mice. Representative data are shown.

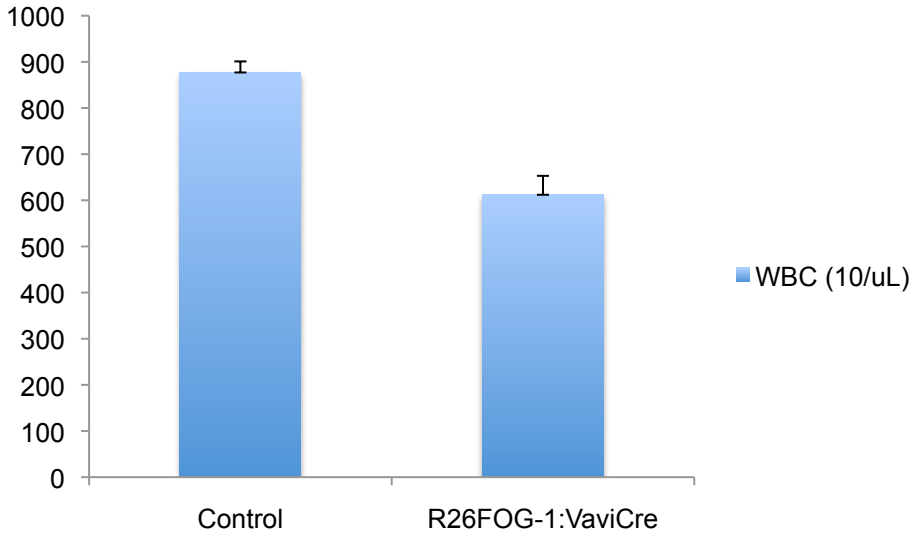
one cell type, the X chromosome bearing the GATA-1.05 allele is inactivated, but the one bearing the wild-type GATA-1 allele is active. These haematopoietic progenitors express normal amounts of GATA-1 and are able to differentiate into their appropriate cell types, including enucleated erythrocytes and platelets. In the other cell type, the X chromosome bearing the GATA-1.05 allele is active, but the one bearing the wild-type GATA-1 allele is inactive. In the latter case, the expression level of GATA-1 in erythroid and megakaryocytic progenitors is very low. This reduction appears to cause both arrest of differentiation and stimulation of proliferation (Shimizu et al., 2004). Following X-inactivation in female $R26^{FOG\text{-}1}$:VaviCre mice, GATA-1 is expressed randomly from the maternal X chromosome or from the paternal one ; which might lead to cells expressing GATA-1 at slightly different levels. On the contrary, in male $R26^{FOG\text{-}1}$:VaviCre mice, GATA-1 is always expressed from the maternal X chromosome. Therefore, in all cells when FOG-1 is overexpressed, the pool of GATA-1 protein available is exactly the same leading to reproducible not attenuated effects arising from changes in the balance of FOG-1 partners available.

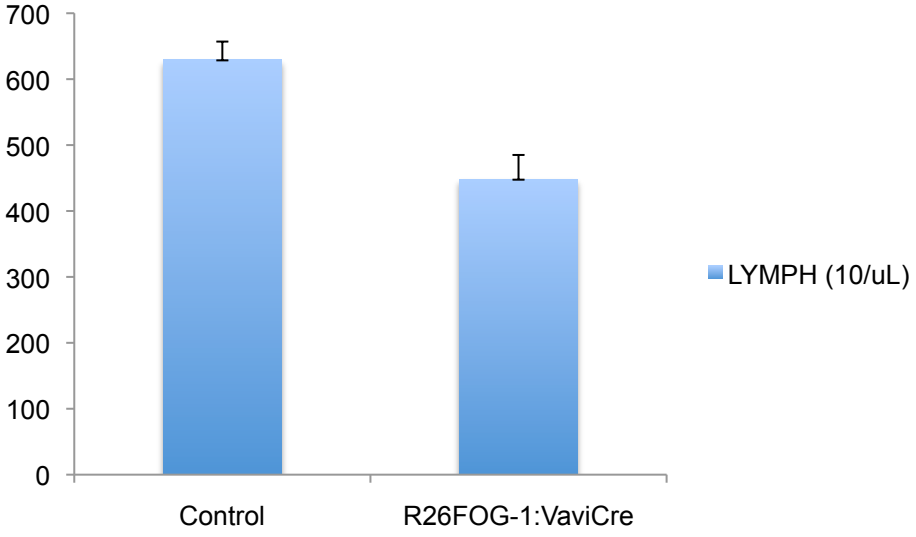
Not surprisingly since FOG-1 is required for the differentiation of erythrocytes (Tsang et al., 1998; Tsang et al., 1997), we observed a moderately altered erythropoiesis in the spleen of some $R26^{FOG\text{-}1}$:VaviCre animals.

As discussed in the article “Inappropriate expression of FOG-1 does not affect B-cells but reduces the number of circulating eosinophils”, the highly significant diminution of the total number of circulating eosinophils should be investigated further. Experiments similar to those described in Mori *et al*.’s work (Mori et al., 2009) could be performed, especially colony forming assays.

2.2. *In vitro*

Biochemical experiments on FOG-1 in B cells revealed a number of interesting interactors that could actually further explain FOG-1 mechanism of action in red blood cells, where it is essential for differentiation as well as in other cellular settings. Among these factors, Aiolos, GZP1, Ikaros and TIF1B are of particular interest because their role in association to FOG-1 have never been addressed. In B lymphocytes, we showed that FOG-1 acts in a GATA-1 independent manner and more importantly in an Ikaros dependent manner at the transcriptional level. This is of great interest since so far in red blood cells, FOG-1 association with Ikaros has been suggested but always in conjunction with GATA-1 (Bottardi et al., 2009; Ross et al.). In addition, the proteomic analysis performed on the interaction of FOG-1 and Ikaros revealed that in all the pull-down experiments, the same peptide ($_{257}$SNHSAQDAVDNLLLLSK$_{273}$) inside Ikaros sequence was always recovered. It maps between Ikaros zinc finger 4 and 5 within the transactivation domain of the protein. It is also noteworthy that this peptide contains a compositional bias from position 268 till position 271 with 4 poly-leucine, which might be involved in a hydrophobic pocket. To strengthen these results, *in silico* search with Pymol to see whether this peptide might be necessary and sufficient to interact with FOG-1 would be interesting. Depending on the results, crystalization assays with Ikaros peptide in complex with FOG-1 could be performed. In addition, inside FOG-1 sequence, there





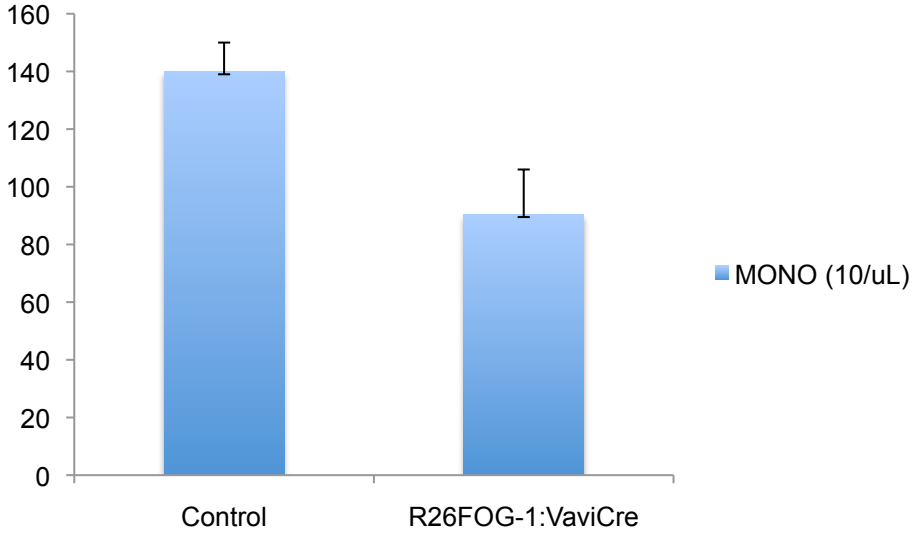


Figure 19 : Reduction of circulating white blood cells, lymphocytes and monocytes. The numbers of white blood cells, lymphocytes and monocytes obtained in full blood count analysis of 2 control and 2 $R26^{FOG-1}$:VaviCre mice were averaged. Standard error of the mean is shown.

are two zinc finger motifs with an unusual structure that is most similar to those encoded by the *hunchbach* gene in *Drosophila*. This domain mediates the formation of homo- or heterodimers and multimers between Ikaros family members, an essential feature for the regulation of their functions. Therefore, to further investigate the interaction between FOG-1 and Ikaros, generating a FOG-1 mutant in this dimerization zinc finger domain (zinc fingers 2 and 3) would be engaging. In parallel to FOG-1 partner search, a phosphorylation site analysis within FOG-1 protein was done by proteomics. Studies in the B cells showed that the FOG-1 serine 286 might be important for its function together with Ikaros. The pS286A FOG-1 mutant and the phosphomimetic pS286E FOG-1 mutant could be used to adress this question in other cellular settings where FOG-1 is known to be a critical factor. Considering the high number of phosphorylation sites identified by MS, the generation of different FOG-1 mutants for phosphoacceptor sites would be interesting. Notably, 3 of them (S397, S497 and S500) were found within the region between zinc fingers 4 and 5, whose function has not yet been assigned despite a high conservation between mouse and human. In addition, Snow *et al*. identified three phosphorylation sites near the canonical CtBP binding site (residues 811-815) located at S803, S805 and S822 (Snow et al.). Given FOG-1/CtBP1 interaction in B cells, this appears interesting.

Chapter 4 : References

1. Chapter 1 : Introduction

Abul K. Abbas, A.H.L., Shiv Pillai (2010). Cellular and Molecular Immunology, Sixth edition.

Arulampalam, V., Eckhardt, L., and Pettersson, S. (1997). The enhancer shift: a model to explain the developmental control of IgH gene expression in B-lineage cells. Immunol Today *18*, 549-554.

Avitahl, N., Winandy, S., Friedrich, C., Jones, B., Ge, Y., and Georgopoulos, K. (1999). Ikaros sets thresholds for T cell activation and regulates chromosome propagation. Immunity *10*, 333-343.

Azuara, V., Brown, K.E., Williams, R.R., Webb, N., Dillon, N., Festenstein, R., Buckle, V., Merkenschlager, M., and Fisher, A.G. (2003). Heritable gene silencing in lymphocytes delays chromatid resolution without affecting the timing of DNA replication. Nat Cell Biol *5*, 668-674.

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

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

Bartholdy, B., and Matthias, P. (2004). Transcriptional control of B cell development and function. Gene *327*, 1-23.

Batten, M., Groom, J., Cachero, T.G., Qian, F., Schneider, P., Tschopp, J., Browning, J.L., and Mackay, F. (2000). BAFF mediates survival of peripheral immature B lymphocytes. J Exp Med *192*, 1453-1466.

Baumann, M., Pontiller, J., and Ernst, W. Structure and basal transcription complex of RNA polymerase II core promoters in the mammalian genome: an overview. Mol Biotechnol *45*, 241-247.

Beckmann, H., Su, L.K., and Kadesch, T. (1990). TFE3: a helix-loop-helix protein that activates transcription through the immunoglobulin enhancer muE3 motif. Genes Dev *4*, 167-179.

Berger, S.L., Kouzarides, T., Shiekhattar, R., and Shilatifard, A. (2009). An operational definition of epigenetics. Genes Dev *23*, 781-783.

Bottardi, S., Ross, J., Bourgoin, V., Fotouhi-Ardakani, N., Affar el, B., Trudel, M., and Milot, E. (2009). Ikaros and GATA-1 combinatorial effect is required for silencing of human gamma-globin genes. Mol Cell Biol *29*, 1526-1537.

Bross, L., Fukita, Y., McBlane, F., Demolliere, C., Rajewsky, K., and Jacobs, H. (2000). DNA double-strand breaks in immunoglobulin genes undergoing somatic hypermutation. Immunity *13*, 589-597.

Brown, K.E., Guest, S.S., Smale, S.T., Hahm, K., Merkenschlager, M., and Fisher, A.G. (1997). Association of transcriptionally silent genes with Ikaros complexes at centromeric heterochromatin. Cell *91*, 845-854.

Burley, S.K., and Roeder, R.G. (1996). Biochemistry and structural biology of transcription factor IID (TFIID). Annu Rev Biochem *65*, 769-799.

Calame, K.L. (2001). Plasma cells: finding new light at the end of B cell development. Nat Immunol *2*, 1103-1108.

Cantor, A.B., Iwasaki, H., Arinobu, Y., Moran, T.B., Shigematsu, H., Sullivan, M.R., Akashi, K., and Orkin, S.H. (2008). Antagonism of FOG-1 and GATA factors in fate choice for the mast cell lineage. J Exp Med *205*, 611-624.
Cantor, A.B., and Orkin, S.H. (2005). Coregulation of GATA factors by the Friend of GATA (FOG) family of multitype zinc finger proteins. Semin Cell Dev Biol *16*, 117-128.
Cariappa, A., Tang, M., Parng, C., Nebelitskiy, E., Carroll, M., Georgopoulos, K., and Pillai, S. (2001). The follicular versus marginal zone B lymphocyte cell fate decision is regulated by Aiolos, Btk, and CD21. Immunity *14*, 603-615.
Carpenter, A.C., and Bosselut, R. Decision checkpoints in the thymus. Nat Immunol *11*, 666-673.
Casellas, R., Jankovic, M., Meyer, G., Gazumyan, A., Luo, Y., Roeder, R., and Nussenzweig, M. (2002). OcaB is required for normal transcription and V(D)J recombination of a subset of immunoglobulin kappa genes. Cell *110*, 575-585.
Chinnadurai, G. (2007). Transcriptional regulation by C-terminal binding proteins. Int J Biochem Cell Biol *39*, 1593-1607.
Cinamon, G., Matloubian, M., Lesneski, M.J., Xu, Y., Low, C., Lu, T., Proia, R.L., and Cyster, J.G. (2004). Sphingosine 1-phosphate receptor 1 promotes B cell localization in the splenic marginal zone. Nat Immunol *5*, 713-720.
Ciofani, M., and Zuniga-Pflucker, J.C. (2007). The thymus as an inductive site for T lymphopoiesis. Annu Rev Cell Dev Biol *23*, 463-493.
Cobb, B.S., Morales-Alcelay, S., Kleiger, G., Brown, K.E., Fisher, A.G., and Smale, S.T. (2000). Targeting of Ikaros to pericentromeric heterochromatin by direct DNA binding. Genes Dev *14*, 2146-2160.
Cockerill, P.N., and Garrard, W.T. (1986). Chromosomal loop anchorage of the kappa immunoglobulin gene occurs next to the enhancer in a region containing topoisomerase II sites. Cell *44*, 273-282.
Cockerill, P.N., Yuen, M.H., and Garrard, W.T. (1987). The enhancer of the immunoglobulin heavy chain locus is flanked by presumptive chromosomal loop anchorage elements. J Biol Chem *262*, 5394-5397.
Corcoran, L.M., and Karvelas, M. (1994). Oct-2 is required early in T cell-independent B cell activation for G1 progression and for proliferation. Immunity *1*, 635-645.
Corcoran, L.M., Karvelas, M., Nossal, G.J., Ye, Z.S., Jacks, T., and Baltimore, D. (1993). Oct-2, although not required for early B-cell development, is critical for later B-cell maturation and for postnatal survival. Genes Dev *7*, 570-582.
Crepieux, P., Coll, J., and Stehelin, D. (1994). The Ets family of proteins: weak modulators of gene expression in quest for transcriptional partners. Crit Rev Oncog *5*, 615-638.
Crispino, J.D., Lodish, M.B., MacKay, J.P., and Orkin, S.H. (1999). Use of altered specificity mutants to probe a specific protein-protein interaction in differentiation: the GATA-1:FOG complex. Mol Cell *3*, 219-228.
D'Eustachio, P., Pravtcheva, D., Marcu, K., and Ruddle, F.H. (1980). Chromosomal location of the structural gene cluster encoding murine immunoglobulin heavy chains. J Exp Med *151*, 1545-1550.
Dahl, R., Ramirez-Bergeron, D.L., Rao, S., and Simon, M.C. (2002). Spi-B can functionally replace PU.1 in myeloid but not lymphoid development. EMBO J *21*, 2220-2230.

Dooner, G.J., Colvin, G.A., Dooner, M.S., Johnson, K.W., and Quesenberry, P.J. (2008). Gene expression fluctuations in murine hematopoietic stem cells with cell cycle progression. J Cell Physiol *214*, 786-795.
Duan, B., and Morel, L. (2006). Role of B-1a cells in autoimmunity. Autoimmun Rev *5*, 403-408.
Eastman, Q.M., Leu, T.M., and Schatz, D.G. (1996). Initiation of V(D)J recombination in vitro obeying the 12/23 rule. Nature *380*, 85-88.
Engel, I., and Murre, C. (2001). The function of E- and Id proteins in lymphocyte development. Nat Rev Immunol *1*, 193-199.
Ernst, P., and Smale, S.T. (1995). Combinatorial regulation of transcription II: The immunoglobulin mu heavy chain gene. Immunity *2*, 427-438.
Esteller, M. (2008). Epigenetics in cancer. N Engl J Med *358*, 1148-1159.
Fang, M., Li, J., Blauwkamp, T., Bhambhani, C., Campbell, N., and Cadigan, K.M. (2006). C-terminal-binding protein directly activates and represses Wnt transcriptional targets in Drosophila. EMBO J *25*, 2735-2745.
Ford, E., Strubin, M., and Hernandez, N. (1998). The Oct-1 POU domain activates snRNA gene transcription by contacting a region in the SNAPc largest subunit that bears sequence similarities to the Oct-1 coactivator OBF-1. Genes Dev *12*, 3528-3540.
Forrester, W.C., van Genderen, C., Jenuwein, T., and Grosschedl, R. (1994). Dependence of enhancer-mediated transcription of the immunoglobulin mu gene on nuclear matrix attachment regions. Science *265*, 1221-1225.
Fox, A.H., Liew, C., Holmes, M., Kowalski, K., Mackay, J., and Crossley, M. (1999). Transcriptional cofactors of the FOG family interact with GATA proteins by means of multiple zinc fingers. EMBO J *18*, 2812-2822.
Fujita, N., Jaye, D.L., Geigerman, C., Akyildiz, A., Mooney, M.R., Boss, J.M., and Wade, P.A. (2004). MTA3 and the Mi-2/NuRD complex regulate cell fate during B lymphocyte differentiation. Cell *119*, 75-86.
Fujiwara, Y., Browne, C.P., Cunniff, K., Goff, S.C., and Orkin, S.H. (1996). Arrested development of embryonic red cell precursors in mouse embryos lacking transcription factor GATA-1. Proc Natl Acad Sci U S A *93*, 12355-12358.
Gao, Z., Huang, Z., Olivey, H.E., Gurbuxani, S., Crispino, J.D., and Svensson, E.C. FOG-1-mediated recruitment of NuRD is required for cell lineage re-enforcement during haematopoiesis. EMBO J *29*, 457-468.
Garriga-Canut, M., and Orkin, S.H. (2004). Transforming acidic coiled-coil protein 3 (TACC3) controls friend of GATA-1 (FOG-1) subcellular localization and regulates the association between GATA-1 and FOG-1 during hematopoiesis. J Biol Chem *279*, 23597-23605.
Gearhart, P.J., and Wood, R.D. (2001). Emerging links between hypermutation of antibody genes and DNA polymerases. Nat Rev Immunol *1*, 187-192.
Georgopoulos, K., Bigby, M., Wang, J.H., Molnar, A., Wu, P., Winandy, S., and Sharpe, A. (1994). The Ikaros gene is required for the development of all lymphoid lineages. Cell *79*, 143-156.
Georgopoulos, K., Moore, D.D., and Derfler, B. (1992). Ikaros, an early lymphoid-specific transcription factor and a putative mediator for T cell commitment. Science *258*, 808-812.

Georgopoulos, K., Winandy, S., and Avitahl, N. (1997). The role of the Ikaros gene in lymphocyte development and homeostasis. Annu Rev Immunol *15*, 155-176.
Ghosh, S., May, M.J., and Kopp, E.B. (1998). NF-kappa B and Rel proteins: evolutionarily conserved mediators of immune responses. Annu Rev Immunol *16*, 225-260.
Gilbert, S.F. (2010). Developmental Biology, Ninth Edition.
Godfrey, D.I., Stankovic, S., and Baxter, A.G. Raising the NKT cell family. Nat Immunol *11*, 197-206.
Gregory, G.D., Miccio, A., Bersenev, A., Wang, Y., Hong, W., Zhang, Z., Poncz, M., Tong, W., and Blobel, G.A. FOG1 requires NuRD to promote hematopoiesis and maintain lineage fidelity within the megakaryocytic-erythroid compartment. Blood *115*, 2156-2166.
Gstaiger, M., Georgiev, O., van Leeuwen, H., van der Vliet, P., and Schaffner, W. (1996). The B cell coactivator Bob1 shows DNA sequence-dependent complex formation with Oct-1/Oct-2 factors, leading to differential promoter activation. EMBO J *15*, 2781-2790.
Guinamard, R., Okigaki, M., Schlessinger, J., and Ravetch, J.V. (2000). Absence of marginal zone B cells in Pyk-2-deficient mice defines their role in the humoral response. Nat Immunol *1*, 31-36.
Hahm, K., Cobb, B.S., McCarty, A.S., Brown, K.E., Klug, C.A., Lee, R., Akashi, K., Weissman, I.L., Fisher, A.G., and Smale, S.T. (1998). Helios, a T cell-restricted Ikaros family member that quantitatively associates with Ikaros at centromeric heterochromatin. Genes Dev *12*, 782-796.
Hahm, K., Ernst, P., Lo, K., Kim, G.S., Turck, C., and Smale, S.T. (1994). The lymphoid transcription factor LyF-1 is encoded by specific, alternatively spliced mRNAs derived from the Ikaros gene. Mol Cell Biol *14*, 7111-7123.
Havran, W.L., and Allison, J.P. (1988). Developmentally ordered appearance of thymocytes expressing different T-cell antigen receptors. Nature *335*, 443-445.
Hayday, A.C., Barber, D.F., Douglas, N., and Hoffman, E.S. (1999). Signals involved in gamma/delta T cell versus alpha/beta T cell lineage commitment. Semin Immunol *11*, 239-249.
Herr, W., and Cleary, M.A. (1995). The POU domain: versatility in transcriptional regulation by a flexible two-in-one DNA-binding domain. Genes Dev *9*, 1679-1693.
Hess, J., Nielsen, P.J., Fischer, K.D., Bujard, H., and Wirth, T. (2001). The B lymphocyte-specific coactivator BOB.1/OBF.1 is required at multiple stages of B-cell development. Mol Cell Biol *21*, 1531-1539.
Heyworth, C., Pearson, S., May, G., and Enver, T. (2002). Transcription factor-mediated lineage switching reveals plasticity in primary committed progenitor cells. EMBO J *21*, 3770-3781.
Ho, L., and Crabtree, G.R. Chromatin remodelling during development. Nature *463*, 474-484.
Hochheimer, A., and Tjian, R. (2003). Diversified transcription initiation complexes expand promoter selectivity and tissue-specific gene expression. Genes Dev *17*, 1309-1320.
Holtmeier, W., and Kabelitz, D. (2005). gammadelta T cells link innate and adaptive immune responses. Chem Immunol Allergy *86*, 151-183.

Holtschke, T., Lohler, J., Kanno, Y., Fehr, T., Giese, N., Rosenbauer, F., Lou, J., Knobeloch, K.P., Gabriele, L., Waring, J.F., *et al.* (1996). Immunodeficiency and chronic myelogenous leukemia-like syndrome in mice with a targeted mutation of the ICSBP gene. Cell *87*, 307-317.
Hong, W., Nakazawa, M., Chen, Y.Y., Kori, R., Vakoc, C.R., Rakowski, C., and Blobel, G.A. (2005). FOG-1 recruits the NuRD repressor complex to mediate transcriptional repression by GATA-1. EMBO J *24*, 2367-2378.
Honjo, T., Nagaoka, H., Shinkura, R., and Muramatsu, M. (2005). AID to overcome the limitations of genomic information. Nat Immunol *6*, 655-661.
Hozumi, K., Negishi, N., Suzuki, D., Abe, N., Sotomaru, Y., Tamaoki, N., Mailhos, C., Ish-Horowicz, D., Habu, S., and Owen, M.J. (2004). Delta-like 1 is necessary for the generation of marginal zone B cells but not T cells in vivo. Nat Immunol *5*, 638-644.
Hu, J.S., Olson, E.N., and Kingston, R.E. (1992). HEB, a helix-loop-helix protein related to E2A and ITF2 that can modulate the DNA-binding ability of myogenic regulatory factors. Mol Cell Biol *12*, 1031-1042.
Jacobs, H., and Bross, L. (2001). Towards an understanding of somatic hypermutation. Curr Opin Immunol *13*, 208-218.
Jenuwein, T., Forrester, W.C., Fernandez-Herrero, L.A., Laible, G., Dull, M., and Grosschedl, R. (1997). Extension of chromatin accessibility by nuclear matrix attachment regions. Nature *385*, 269-272.
Jiang, H., and Chess, L. (2009). How the immune system achieves self-nonself discrimination during adaptive immunity. Adv Immunol *102*, 95-133.
Jing, H., Vakoc, C.R., Ying, L., Mandat, S., Wang, H., Zheng, X., and Blobel, G.A. (2008). Exchange of GATA factors mediates transitions in looped chromatin organization at a developmentally regulated gene locus. Mol Cell *29*, 232-242.
Kang, J., and Raulet, D.H. (1997). Events that regulate differentiation of alpha beta TCR+ and gamma delta TCR+ T cells from a common precursor. Semin Immunol *9*, 171-179.
Kantor, A.B., Stall, A.M., Adams, S., and Herzenberg, L.A. (1992). Differential development of progenitor activity for three B-cell lineages. Proc Natl Acad Sci U S A *89*, 3320-3324.
Kazatchkine, M.D., and Kaveri, S.V. (2001). Immunomodulation of autoimmune and inflammatory diseases with intravenous immune globulin. N Engl J Med *345*, 747-755.
Kee, B.L., Quong, M.W., and Murre, C. (2000). E2A proteins: essential regulators at multiple stages of B-cell development. Immunol Rev *175*, 138-149.
Kenneth M. Murphy, P.T., Mark Walport (2008a). Janeway' s Immunobiology, Seventh Edition.
Kenneth M. Murphy, P.T., Mark Walport (2008b). Janeway's Immunobiology, Seventh Edition.
Kim, J., Sif, S., Jones, B., Jackson, A., Koipally, J., Heller, E., Winandy, S., Viel, A., Sawyer, A., Ikeda, T., *et al.* (1999). Ikaros DNA-binding proteins direct formation of chromatin remodeling complexes in lymphocytes. Immunity *10*, 345-355.
Kim, U., Qin, X.F., Gong, S., Stevens, S., Luo, Y., Nussenzweig, M., and Roeder, R.G. (1996). The B-cell-specific transcription coactivator OCA-

B/OBF-1/Bob-1 is essential for normal production of immunoglobulin isotypes. Nature *383*, 542-547.
Klein, J., and Sato, A. (2000). The HLA system. First of two parts. N Engl J Med *343*, 702-709.
Klug, C.A., Morrison, S.J., Masek, M., Hahm, K., Smale, S.T., and Weissman, I.L. (1998). Hematopoietic stem cells and lymphoid progenitors express different Ikaros isoforms, and Ikaros is localized to heterochromatin in immature lymphocytes. Proc Natl Acad Sci U S A *95*, 657-662.
Koipally, J., Renold, A., Kim, J., and Georgopoulos, K. (1999). Repression by Ikaros and Aiolos is mediated through histone deacetylase complexes. EMBO J *18*, 3090-3100.
Kong, Q., and Maizels, N. (2001). DNA breaks in hypermutating immunoglobulin genes: evidence for a break-and-repair pathway of somatic hypermutation. Genetics *158*, 369-378.
Krantz, S.B., and Goldwasser, E. (1965a). On the mechanism of erythropoietin-induced differentiation. II. T he effect on RNA synthesis. Biochim Biophys Acta *103*, 325-332.
Krantz, S.B., and Goldwasser, E. (1965b). On the mechanism of erythropoietin-induced differentiation. IV. Some characteristics of erythropoietin action on hemoglobin synthesis in marrow cell culture. Biochim Biophys Acta *108*, 455-462.
Kulessa, H., Frampton, J., and Graf, T. (1995). GATA-1 reprograms avian myelomonocytic cell lines into eosinophils, thromboblasts, and erythroblasts. Genes Dev *9*, 1250-1262.
Kurata, H., Lee, H.J., McClanahan, T., Coffman, R.L., O'Garra, A., and Arai, N. (2002). Friend of GATA is expressed in naive Th cells and functions as a repressor of GATA-3-mediated Th2 cell development. J Immunol *168*, 4538-4545.
Lai, A.Y., and Wade, P.A. Cancer biology and NuRD: a multifaceted chromatin remodelling complex. Nat Rev Cancer *11*, 588-596.
Latchman, D.S. (1999). POU family transcription factors in the nervous system. J Cell Physiol *179*, 126-133.
Letting, D.L., Chen, Y.Y., Rakowski, C., Reedy, S., and Blobel, G.A. (2004). Context-dependent regulation of GATA-1 by friend of GATA-1. Proc Natl Acad Sci U S A *101*, 476-481.
Lin, H., and Grosschedl, R. (1995). Failure of B-cell differentiation in mice lacking the transcription factor EBF. Nature *376*, 263-267.
Liu, H., Schmidt-Supprian, M., Shi, Y., Hobeika, E., Barteneva, N., Jumaa, H., Pelanda, R., Reth, M., Skok, J., and Rajewsky, K. (2007). Yin Yang 1 is a critical regulator of B-cell development. Genes Dev *21*, 1179-1189.
Luo, Y., and Roeder, R.G. (1995). Cloning, functional characterization, and mechanism of action of the B-cell-specific transcriptional coactivator OCA-B. Mol Cell Biol *15*, 4115-4124.
Maeda, K., Nishiyama, C., Tokura, T., Nakano, H., Kanada, S., Nishiyama, M., Okumura, K., and Ogawa, H. (2006). FOG-1 represses GATA-1-dependent FcepsilonRI beta-chain transcription: transcriptional mechanism of mast-cell-specific gene expression in mice. Blood *108*, 262-269.
Malone, C.S., and Wall, R. (2002). Bob1 (OCA-B/OBF-1) differential transactivation of the B cell-specific B29 (Ig beta) and mb-1 (Ig alpha) promoters. J Immunol *168*, 3369-3375.

Manis, J.P., Tian, M., and Alt, F.W. (2002). Mechanism and control of class-switch recombination. Trends Immunol *23*, 31-39.
Martin, F., and Kearney, J.F. (2000). B-cell subsets and the mature preimmune repertoire. Marginal zone and B1 B cells as part of a "natural immune memory". Immunol Rev *175*, 70-79.
Martin, F., Oliver, A.M., and Kearney, J.F. (2001). Marginal zone and B1 B cells unite in the early response against T-independent blood-borne particulate antigens. Immunity *14*, 617-629.
Mayer, R., and Zaghouani, H. (1991). The contribution of the Ly-1 B-cell subset to self-reactivity. Immunol Ser *55*, 61-79.
McKercher, S.R., Torbett, B.E., Anderson, K.L., Henkel, G.W., Vestal, D.J., Baribault, H., Klemsz, M., Feeney, A.J., Wu, G.E., Paige, C.J.*, et al.* (1996). Targeted disruption of the PU.1 gene results in multiple hematopoietic abnormalities. EMBO J *15*, 5647-5658.
Mebius, R.E., and Kraal, G. (2005). Structure and function of the spleen. Nat Rev Immunol *5*, 606-616.
Mendez, L.M., Polo, J.M., Yu, J.J., Krupski, M., Ding, B.B., Melnick, A., and Ye, B.H. (2008). CtBP is an essential corepressor for BCL6 autoregulation. Mol Cell Biol *28*, 2175-2186.
Miccio, A., Wang, Y., Hong, W., Gregory, G.D., Wang, H., Yu, X., Choi, J.K., Shelat, S., Tong, W., Poncz, M.*, et al.* NuRD mediates activating and repressive functions of GATA-1 and FOG-1 during blood development. EMBO J *29*, 442-456.
Michael, N., Martin, T.E., Nicolae, D., Kim, N., Padjen, K., Zhan, P., Nguyen, H., Pinkert, C., and Storb, U. (2002). Effects of sequence and structure on the hypermutability of immunoglobulin genes. Immunity *16*, 123-134.
Molnar, A., and Georgopoulos, K. (1994). The Ikaros gene encodes a family of functionally diverse zinc finger DNA-binding proteins. Mol Cell Biol *14*, 8292-8303.
Morgan, B., Sun, L., Avitahl, N., Andrikopoulos, K., Ikeda, T., Gonzales, E., Wu, P., Neben, S., and Georgopoulos, K. (1997). Aiolos, a lymphoid restricted transcription factor that interacts with Ikaros to regulate lymphocyte differentiation. EMBO J *16*, 2004-2013.
Murre, C., McCaw, P.S., and Baltimore, D. (1989). A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, MyoD, and myc proteins. Cell *56*, 777-783.
Nelsen, B., Tian, G., Erman, B., Gregoire, J., Maki, R., Graves, B., and Sen, R. (1993). Regulation of lymphoid-specific immunoglobulin mu heavy chain gene enhancer by ETS-domain proteins. Science *261*, 82-86.
Nichogiannopoulou, A., Trevisan, M., Friedrich, C., and Georgopoulos, K. (1998). Ikaros in hemopoietic lineage determination and homeostasis. Semin Immunol *10*, 119-125.
Nutt, S.L., Heavey, B., Rolink, A.G., and Busslinger, M. (1999). Commitment to the B-lymphoid lineage depends on the transcription factor Pax5. Nature *401*, 556-562.
O'Neill, D.W., Schoetz, S.S., Lopez, R.A., Castle, M., Rabinowitz, L., Shor, E., Krawchuk, D., Goll, M.G., Renz, M., Seelig, H.P.*, et al.* (2000). An ikaros-containing chromatin-remodeling complex in adult-type erythroid cells. Mol Cell Biol *20*, 7572-7582.

Oracki, S.A., Walker, J.A., Hibbs, M.L., Corcoran, L.M., and Tarlinton, D.M. Plasma cell development and survival. Immunol Rev *237*, 140-159.
Orkin, S.H., and Zon, L.I. (2008). SnapShot: hematopoiesis. Cell *132*, 712.
Pal, S., Cantor, A.B., Johnson, K.D., Moran, T.B., Boyer, M.E., Orkin, S.H., and Bresnick, E.H. (2004). Coregulator-dependent facilitation of chromatin occupancy by GATA-1. Proc Natl Acad Sci U S A *101*, 980-985.
Papavasiliou, F.N., and Schatz, D.G. (2000). Cell-cycle-regulated DNA double-stranded breaks in somatic hypermutation of immunoglobulin genes. Nature *408*, 216-221.
Papavasiliou, F.N., and Schatz, D.G. (2002). Somatic hypermutation of immunoglobulin genes: merging mechanisms for genetic diversity. Cell *109 Suppl*, S35-44.
Paule, M.R., and White, R.J. (2000). Survey and summary: transcription by RNA polymerases I and III. Nucleic Acids Res *28*, 1283-1298.
Pennington, D.J., Silva-Santos, B., Shires, J., Theodoridis, E., Pollitt, C., Wise, E.L., Tigelaar, R.E., Owen, M.J., and Hayday, A.C. (2003). The inter-relatedness and interdependence of mouse T cell receptor gammadelta+ and alphabeta+ cells. Nat Immunol *4*, 991-998.
Perdomo, J., Holmes, M., Chong, B., and Crossley, M. (2000). Eos and pegasus, two members of the Ikaros family of proteins with distinct DNA binding activities. J Biol Chem *275*, 38347-38354.
Phillips, K., and Luisi, B. (2000). The virtuoso of versatility: POU proteins that flex to fit. J Mol Biol *302*, 1023-1039.
Pillai, S., and Cariappa, A. (2009). The follicular versus marginal zone B lymphocyte cell fate decision. Nat Rev Immunol *9*, 767-777.
Ptashne, M., and Gann, A. (1997). Transcriptional activation by recruitment. Nature *386*, 569-577.
Querfurth, E., Schuster, M., Kulessa, H., Crispino, J.D., Doderlein, G., Orkin, S.H., Graf, T., and Nerlov, C. (2000). Antagonism between C/EBPbeta and FOG in eosinophil lineage commitment of multipotent hematopoietic progenitors. Genes Dev *14*, 2515-2525.
Rebollo, A., and Schmitt, C. (2003). Ikaros, Aiolos and Helios: transcription regulators and lymphoid malignancies. Immunol Cell Biol *81*, 171-175.
Remenyi, A., Tomilin, A., Pohl, E., Lins, K., Philippsen, A., Reinbold, R., Scholer, H.R., and Wilmanns, M. (2001). Differential dimer activities of the transcription factor Oct-1 by DNA-induced interface swapping. Mol Cell *8*, 569-580.
Rettig, M.B., Ma, H.J., Vescio, R.A., Pold, M., Schiller, G., Belson, D., Savage, A., Nishikubo, C., Wu, C., Fraser, J.*, et al.* (1997). Kaposi's sarcoma-associated herpesvirus infection of bone marrow dendritic cells from multiple myeloma patients. Science *276*, 1851-1854.
Reya, T., O'Riordan, M., Okamura, R., Devaney, E., Willert, K., Nusse, R., and Grosschedl, R. (2000). Wnt signaling regulates B lymphocyte proliferation through a LEF-1 dependent mechanism. Immunity *13*, 15-24.
Reynaud, C.A., Aoufouchi, S., Faili, A., and Weill, J.C. (2003). What role for AID: mutator, or assembler of the immunoglobulin mutasome? Nat Immunol *4*, 631-638.
Rivera, R.R., Stuiver, M.H., Steenbergen, R., and Murre, C. (1993). Ets proteins: new factors that regulate immunoglobulin heavy-chain gene expression. Mol Cell Biol *13*, 7163-7169.

Rodriguez, P., Bonte, E., Krijgsveld, J., Kolodziej, K.E., Guyot, B., Heck, A.J., Vyas, P., de Boer, E., Grosveld, F., and Strouboulis, J. (2005). GATA-1 forms distinct activating and repressive complexes in erythroid cells. EMBO J *24*, 2354-2366.
Rolink, A., and Melchers, F. (1993). Generation and regeneration of cells of the B-lymphocyte lineage. Curr Opin Immunol *5*, 207-217.
Rolink, A.G., Andersson, J., and Melchers, F. (1998). Characterization of immature B cells by a novel monoclonal antibody, by turnover and by mitogen reactivity. Eur J Immunol *28*, 3738-3748.
Rolink, A.G., Tschopp, J., Schneider, P., and Melchers, F. (2002). BAFF is a survival and maturation factor for mouse B cells. Eur J Immunol *32*, 2004-2010.
Sabbattini, P., and Dillon, N. (2005). The lambda5-VpreB1 locus--a model system for studying gene regulation during early B cell development. Semin Immunol *17*, 121-127.
Sabbattini, P., Lundgren, M., Georgiou, A., Chow, C., Warnes, G., and Dillon, N. (2001). Binding of Ikaros to the lambda5 promoter silences transcription through a mechanism that does not require heterochromatin formation. EMBO J *20*, 2812-2822.
Saito, T., Chiba, S., Ichikawa, M., Kunisato, A., Asai, T., Shimizu, K., Yamaguchi, T., Yamamoto, G., Seo, S., Kumano, K.*, et al.* (2003). Notch2 is preferentially expressed in mature B cells and indispensable for marginal zone B lineage development. Immunity *18*, 675-685.
Samardzic, T., Marinkovic, D., Nielsen, P.J., Nitschke, L., and Wirth, T. (2002). BOB.1/OBF.1 deficiency affects marginal-zone B-cell compartment. Mol Cell Biol *22*, 8320-8331.
Saunders, T.W.M.a.M.E. (2006). The Immune Response, Basic and Clinical Principles.
Schatz, D.G., and Ji, Y. Recombination centres and the orchestration of V(D)J recombination. Nat Rev Immunol *11*, 251-263.
Schiemann, B., Gommerman, J.L., Vora, K., Cachero, T.G., Shulga-Morskaya, S., Dobles, M., Frew, E., and Scott, M.L. (2001). An essential role for BAFF in the normal development of B cells through a BCMA-independent pathway. Science *293*, 2111-2114.
Schilham, M.W., Oosterwegel, M.A., Moerer, P., Ya, J., de Boer, P.A., van de Wetering, M., Verbeek, S., Lamers, W.H., Kruisbeek, A.M., Cumano, A.*, et al.* (1996). Defects in cardiac outflow tract formation and pro-B-lymphocyte expansion in mice lacking Sox-4. Nature *380*, 711-714.
Schubart, D.B., Rolink, A., Schubart, K., and Matthias, P. (2000). Cutting edge: lack of peripheral B cells and severe agammaglobulinemia in mice simultaneously lacking Bruton's tyrosine kinase and the B cell-specific transcriptional coactivator OBF-1. J Immunol *164*, 18-22.
Schubart, D.B., Sauter, P., Massa, S., Friedl, E.M., Schwarzenbach, H., and Matthias, P. (1996). Gene structure and characterization of the murine homologue of the B cell-specific transcriptional coactivator OBF-1. Nucleic Acids Res *24*, 1913-1920.
Schubart, K., Massa, S., Schubart, D., Corcoran, L.M., Rolink, A.G., and Matthias, P. (2001). B cell development and immunoglobulin gene transcription in the absence of Oct-2 and OBF-1. Nat Immunol *2*, 69-74.

Scott, E.W., Simon, M.C., Anastasi, J., and Singh, H. (1994). Requirement of transcription factor PU.1 in the development of multiple hematopoietic lineages. Science *265*, 1573-1577.
Seder, R.A., and Ahmed, R. (2003). Similarities and differences in CD4+ and CD8+ effector and memory T cell generation. Nat Immunol *4*, 835-842.
Sha, W.C., Liou, H.C., Tuomanen, E.I., and Baltimore, D. (1995). Targeted disruption of the p50 subunit of NF-kappa B leads to multifocal defects in immune responses. Cell *80*, 321-330.
Shapiro-Shelef, M., and Calame, K. (2005). Regulation of plasma-cell development. Nat Rev Immunol *5*, 230-242.
Sharrocks, A.D. (2001). The ETS-domain transcription factor family. Nat Rev Mol Cell Biol *2*, 827-837.
Shirai, T., Hirose, S., Okada, T., and Nishimura, H. (1991). CD5+ B cells in autoimmune disease and lymphoid malignancy. Clin Immunol Immunopathol *59*, 173-186.
Sigvardsson, M., Clark, D.R., Fitzsimmons, D., Doyle, M., Akerblad, P., Breslin, T., Bilke, S., Li, R., Yeamans, C., Zhang, G.*, et al.* (2002). Early B-cell factor, E2A, and Pax-5 cooperate to activate the early B cell-specific mb-1 promoter. Mol Cell Biol *22*, 8539-8551.
Sigvardsson, M., O'Riordan, M., and Grosschedl, R. (1997). EBF and E47 collaborate to induce expression of the endogenous immunoglobulin surrogate light chain genes. Immunity *7*, 25-36.
Snapper, C.M., Marcu, K.B., and Zelazowski, P. (1997). The immunoglobulin class switch: beyond "accessibility". Immunity *6*, 217-223.
Stall, A.M., Adams, S., Herzenberg, L.A., and Kantor, A.B. (1992). Characteristics and development of the murine B-1b (Ly-1 B sister) cell population. Ann N Y Acad Sci *651*, 33-43.
Staudt, L.M., Singh, H., Sen, R., Wirth, T., Sharp, P.A., and Baltimore, D. (1986). A lymphoid-specific protein binding to the octamer motif of immunoglobulin genes. Nature *323*, 640-643.
Stavnezer, J. (1996). Antibody class switching. Adv Immunol *61*, 79-146.
Strubin, M., Newell, J.W., and Matthias, P. (1995). OBF-1, a novel B cell-specific coactivator that stimulates immunoglobulin promoter activity through association with octamer-binding proteins. Cell *80*, 497-506.
Sugai, M., Gonda, H., Kusunoki, T., Katakai, T., Yokota, Y., and Shimizu, A. (2003). Essential role of Id2 in negative regulation of IgE class switching. Nat Immunol *4*, 25-30.
Sugiyama, D., Tanaka, M., Kitajima, K., Zheng, J., Yen, H., Murotani, T., Yamatodani, A., and Nakano, T. (2008). Differential context-dependent effects of friend of GATA-1 (FOG-1) on mast-cell development and differentiation. Blood *111*, 1924-1932.
Sun, J., Matthias, G., Mihatsch, M.J., Georgopoulos, K., and Matthias, P. (2003). Lack of the transcriptional coactivator OBF-1 prevents the development of systemic lupus erythematosus-like phenotypes in Aiolos mutant mice. J Immunol *170*, 1699-1706.
Sun, L., Liu, A., and Georgopoulos, K. (1996). Zinc finger-mediated protein interactions modulate Ikaros activity, a molecular control of lymphocyte development. EMBO J *15*, 5358-5369.
Sun, X.H. (1994). Constitutive expression of the Id1 gene impairs mouse B cell development. Cell *79*, 893-900.

Takanashi, M., Yagi, T., Imamura, T., Tabata, Y., Morimoto, A., Hibi, S., Ishii, E., and Imashuku, S. (2002). Expression of the Ikaros gene family in childhood acute lymphoblastic leukaemia. Br J Haematol *117*, 525-530.
Tanaka, M., Lai, J.S., and Herr, W. (1992). Promoter-selective activation domains in Oct-1 and Oct-2 direct differential activation of an snRNA and mRNA promoter. Cell *68*, 755-767.
Tanaka, M., Zheng, J., Kitajima, K., Kita, K., Yoshikawa, H., and Nakano, T. (2004). Differentiation status dependent function of FOG-1. Genes Cells *9*, 1213-1226.
Tang, H., and Sharp, P.A. (1999). Transcriptional regulation of the murine 3' IgH enhancer by OCT-2. Immunity *11*, 517-526.
Tanigaki, K., Han, H., Yamamoto, N., Tashiro, K., Ikegawa, M., Kuroda, K., Suzuki, A., Nakano, T., and Honjo, T. (2002). Notch-RBP-J signaling is involved in cell fate determination of marginal zone B cells. Nat Immunol *3*, 443-450.
Tenen, D.G., Hromas, R., Licht, J.D., and Zhang, D.E. (1997). Transcription factors, normal myeloid development, and leukemia. Blood *90*, 489-519.
Thompson, E.C., Cobb, B.S., Sabbattini, P., Meixlsperger, S., Parelho, V., Liberg, D., Taylor, B., Dillon, N., Georgopoulos, K., Jumaa, H.*, et al.* (2007). Ikaros DNA-binding proteins as integral components of B cell developmental-stage-specific regulatory circuits. Immunity *26*, 335-344.
Tomilin, A., Remenyi, A., Lins, K., Bak, H., Leidel, S., Vriend, G., Wilmanns, M., and Scholer, H.R. (2000). Synergism with the coactivator OBF-1 (OCA-B, BOB-1) is mediated by a specific POU dimer configuration. Cell *103*, 853-864.
Tonegawa, S. (1983). Somatic generation of antibody diversity. Nature *302*, 575-581.
Tsang, A.P., Fujiwara, Y., Hom, D.B., and Orkin, S.H. (1998). Failure of megakaryopoiesis and arrested erythropoiesis in mice lacking the GATA-1 transcriptional cofactor FOG. Genes Dev *12*, 1176-1188.
Tsang, A.P., Visvader, J.E., Turner, C.A., Fujiwara, Y., Yu, C., Weiss, M.J., Crossley, M., and Orkin, S.H. (1997). FOG, a multitype zinc finger protein, acts as a cofactor for transcription factor GATA-1 in erythroid and megakaryocytic differentiation. Cell *90*, 109-119.
Urbanek, P., Wang, Z.Q., Fetka, I., Wagner, E.F., and 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.
Vakoc, C.R., Letting, D.L., Gheldof, N., Sawado, T., Bender, M.A., Groudine, M., Weiss, M.J., Dekker, J., and Blobel, G.A. (2005). Proximity among distant regulatory elements at the beta-globin locus requires GATA-1 and FOG-1. Mol Cell *17*, 453-462.
Wang, J.H., Avitahl, N., Cariappa, A., Friedrich, C., Ikeda, T., Renold, A., Andrikopoulos, K., Liang, L., Pillai, S., Morgan, B.A.*, et al.* (1998). Aiolos regulates B cell activation and maturation to effector state. Immunity *9*, 543-553.
Wang, J.H., Nichogiannopoulou, A., Wu, L., Sun, L., Sharpe, A.H., Bigby, M., and Georgopoulos, K. (1996). Selective defects in the development of the fetal and adult lymphoid system in mice with an Ikaros null mutation. Immunity *5*, 537-549.
Welner, R.S., Pelayo, R., and Kincade, P.W. (2008). Evolving views on the genealogy of B cells. Nat Rev Immunol *8*, 95-106.

Winandy, S., Wu, P., and Georgopoulos, K. (1995). A dominant mutation in the Ikaros gene leads to rapid development of leukemia and lymphoma. Cell *83*, 289-299.
Wu, L., Maillard, I., Nakamura, M., Pear, W.S., and Griffin, J.D. (2007). The transcriptional coactivator Maml1 is required for Notch2-mediated marginal zone B-cell development. Blood *110*, 3618-3623.
Xie, H., Ye, M., Feng, R., and Graf, T. (2004). Stepwise reprogramming of B cells into macrophages. Cell *117*, 663-676.
Yuan, D., Dang, T., Hawley, J., Jenuwein, T., and Grosschedl, R. (1995). Role of the OCTA site in regulation of IgH chain gene transcription during B cell activation. Int Immunol *7*, 1163-1172.
Zhou, M., Ouyang, W., Gong, Q., Katz, S.G., White, J.M., Orkin, S.H., and Murphy, K.M. (2001). Friend of GATA-1 represses GATA-3-dependent activity in CD4+ T cells. J Exp Med *194*, 1461-1471.
Zhuang, Y., Soriano, P., and Weintraub, H. (1994). The helix-loop-helix gene E2A is required for B cell formation. Cell *79*, 875-884.

2. Chapter 3 : Discussion

Baeuerle, P.A., and Gires, O. (2007). EpCAM (CD326) finding its role in cancer. Br J Cancer *96*, 417-423.
Bergsagel, P.L., Victor-Kobrin, C., Timblin, C.R., Trepel, J., and Kuehl, W.M. (1992). A murine cDNA encodes a pan-epithelial glycoprotein that is also expressed on plasma cells. J Immunol *148*, 590-596.
Bottardi, S., Ross, J., Bourgoin, V., Fotouhi-Ardakani, N., Affar el, B., Trudel, M., and Milot, E. (2009). Ikaros and GATA-1 combinatorial effect is required for silencing of human gamma-globin genes. Mol Cell Biol *29*, 1526-1537.
Gires, O., Klein, C.A., and Baeuerle, P.A. (2009). On the abundance of EpCAM on cancer stem cells. Nat Rev Cancer *9*, 143; author reply 143.
Maetzel, D., Denzel, S., Mack, B., Canis, M., Went, P., Benk, M., Kieu, C., Papior, P., Baeuerle, P.A., Munz, M.*, et al.* (2009). Nuclear signalling by tumour-associated antigen EpCAM. Nat Cell Biol *11*, 162-171.
Mori, Y., Iwasaki, H., Kohno, K., Yoshimoto, G., Kikushige, Y., Okeda, A., Uike, N., Niiro, H., Takenaka, K., Nagafuji, K.*, et al.* (2009). Identification of the human eosinophil lineage-committed progenitor: revision of phenotypic definition of the human common myeloid progenitor. J Exp Med *206*, 183-193.
Munz, M., Kieu, C., Mack, B., Schmitt, B., Zeidler, R., and Gires, O. (2004). The carcinoma-associated antigen EpCAM upregulates c-myc and induces cell proliferation. Oncogene *23*, 5748-5758.
Ross, J., Mavoungou, L., Bresnick, E.H., and Milot, E. GATA-1 utilizes Ikaros and the Polycomb Repressive Complex 2 to suppress Hes1 and to promote erythropoiesis. Mol Cell Biol.
Shimizu, R., Kuroha, T., Ohneda, O., Pan, X., Ohneda, K., Takahashi, S., Philipsen, S., and Yamamoto, M. (2004). Leukemogenesis caused by incapacitated GATA-1 function. Mol Cell Biol *24*, 10814-10825.
Snow, J.W., Kim, J., Currie, C.R., Xu, J., and Orkin, S.H. Sumoylation regulates interaction of FOG1 with C-terminal-binding protein (CTBP). J Biol Chem *285*, 28064-28075.
Spizzo, G., Went, P., Dirnhofer, S., Obrist, P., Simon, R., Spichtin, H., Maurer, R., Metzger, U., von Castelberg, B., Bart, R.*, et al.* (2004). High Ep-CAM

expression is associated with poor prognosis in node-positive breast cancer. Breast Cancer Res Treat *86*, 207-213.
Tsang, A.P., Fujiwara, Y., Hom, D.B., and Orkin, S.H. (1998). Failure of megakaryopoiesis and arrested erythropoiesis in mice lacking the GATA-1 transcriptional cofactor FOG. Genes Dev *12*, 1176-1188.
Tsang, A.P., Visvader, J.E., Turner, C.A., Fujiwara, Y., Yu, C., Weiss, M.J., Crossley, M., and Orkin, S.H. (1997). FOG, a multitype zinc finger protein, acts as a cofactor for transcription factor GATA-1 in erythroid and megakaryocytic differentiation. Cell *90*, 109-119.
Went, P., Vasei, M., Bubendorf, L., Terracciano, L., Tornillo, L., Riede, U., Kononen, J., Simon, R., Sauter, G., and Baeuerle, P.A. (2006). Frequent high-level expression of the immunotherapeutic target Ep-CAM in colon, stomach, prostate and lung cancers. Br J Cancer *94*, 128-135.

Chapter 5 : Appendix

1. Impairment of B cell development in mice knock-out for FOG-1 or overexpressing FOG-1 in the B cell compartment

To investigate FOG-1 function *in vivo*, we used cFOG-1KO:*mb-1*Cre and $R26^{FOG-1}$:*mb-1*Cre mice. Mice homozygous for FOG-1 conditional deletion were born with the expected Mendelian frequency. The same observation was done with mice overexpressing FOG-1 from the Pro B cell stage, indicating respectively that FOG-1 expression or reduced level of expression in B cells is not essential for survival *in utero*. cFOG-1KO:*mb-1*Cre animals and wild-type littermates were indistinguishable. Similarly, $R26^{FOG-1}$:*mb-1*Cre mice showed no apparent differences with their wild-type littermates. However, after whole bone marrow cell population extraction, analysis of the expression of B220, a B cell specific surface marker revealed some differences between cFOG-1KO:*mb-1*Cre mice and wild-type littermates as well as between $R26^{FOG-1}$:*mb-1*Cre mice and wild-type littermates. The B220 quantification was performed by FACS within the lymphocyte population of the 2 different types of mice. In male cFOG-1KO:*mb-1*Cre mice, the $B220^{high}$ cell population was increased by 25% (Fig. 20A). These results were confirmed in independent experiments (2 males and 2 females cFOG-1KO:*mb-1*Cre mice exhibited an increase in $B220^{high}$ cells or mature B cells). Nevertheless, in some of the experiments, cFOG-1KO:*mb-1*Cre mice surprisingly exhibited a decrease in $B220^{high}$ cells (1 male and 2 females). On the contrary, in male $R26^{FOG-1}$:*mb-1*Cre mice, the $B220^{high}$ cell population was decreased by around 70% (Fig. 20B). In independent experiments, female $R26^{FOG-1}$:*mb-1*Cre mice exhibited a decrease in $B220^{high}$ cell population (n=5). Besides, out of independent experiments, we observed that male cFOG-1KO: *mb-1*Cre mice showed less Pre B cells than the control mice (n=3). On the contrary, female cFOG-1KO:*mb-1*Cre mice exhibited Pre B lymphoma (n=3). Further investigations would be required to statistically demonstrate the relevance of these phenotypes and to fully define them. A bone marrow transplantation assay should help to figure out how primary B cells deleted for FOG-1 react and develop in an entire organism. This could also be considered with primary B cells overexpressing FOG-1 under the *Rosa26* locus. As discussed in the article “Inappropriate expression of FOG-1 does not affect B-cells but reduces the number of circulating eosinophils”, since targeted *Rosa26* homozygous mice are viable and appear normal, the expression level of the transgene could be doubled by generating $R26^{FOG-1}$/$R26^{FOG-1}$ homozygous mice (Soriano, 1999).

Considering that Pre BI cells are firmly committed to become $B220^{high}$ cells (Welner et al., 2008), *in vitro* cultures of primary Pre BI cells showed results consistent with the previous data. Indeed, male Pre BI cells knock-out for FOG-1 showed an increase in proliferation compared to the control cells whereas male overexpressing FOG-1 Pre BI cells exhibited a growth defect compared to the control cells (Fig. 20C, D). These observations were independent of the gender of mice since almost similar results were obtained

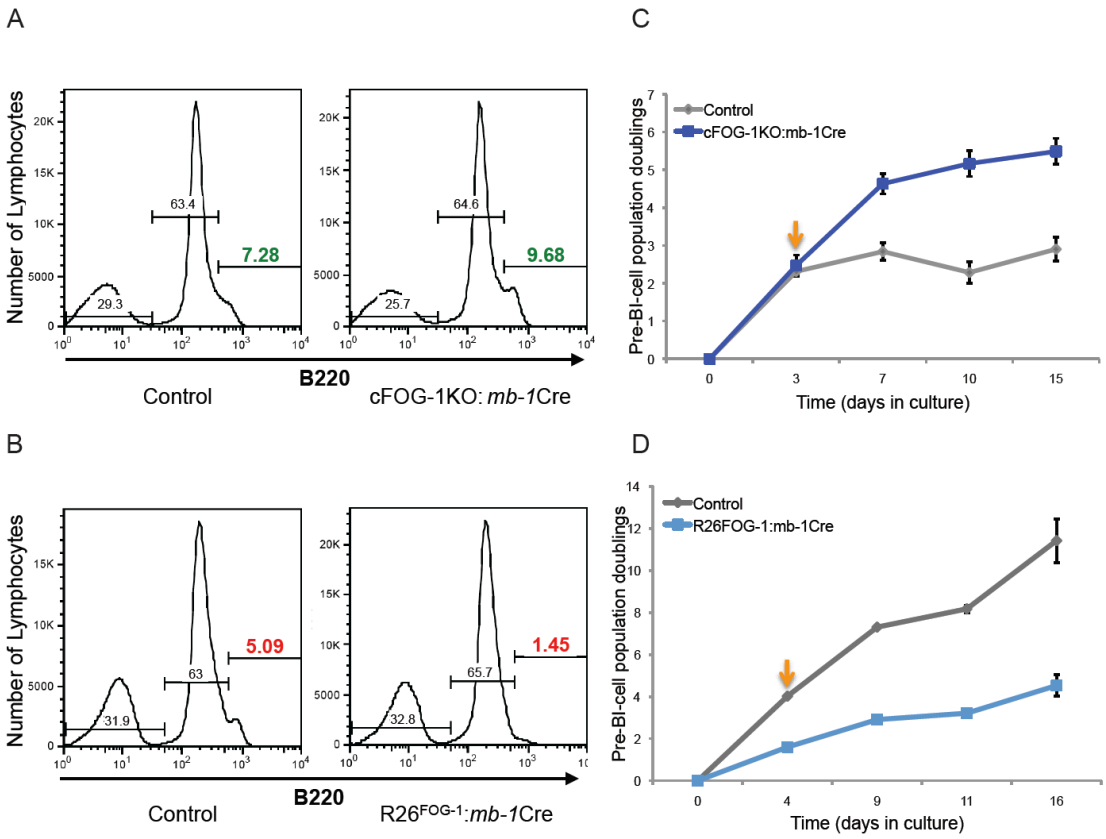


Figure 20 : Phenotypic analyses of lymphocyte population. (A) Cell surface expression of B220 was analyzed by flow cytometry in lymphocyte population of control and cFOG-1KO:*mb-1*Cre mice. (B) Cell surface expression of B220 was analyzed by flow cytometry in lymphocyte population of control and R26^{FOG-1}:*mb-1*Cre mice. Graphs representative of several experiments are shown. (C) Growth curve of Pre BI cells isolated from control and cFOG-1KO:*mb-1*Cre mice. (D) Growth curve of Pre BI cells isolated from control and R26^{FOG-1}:*mb-1*Cre mice. Orange arrows are precising the days at which the microarrays were performed.

in experiments involving female mice or a mixture of male and female mice (data not shown).

It is noteworthy that 4 males and 1 female cFOG-1KO:*mb-1*Cre mice showed a decrease in spleen weight. Paler spleens were observed in cFOG-1KO:*mb-1*Cre mice.

This is reminiscent of the phenotypes observed for GATA-1.05/X mice which carry an allele that express GATA-1 at approximately 5% of the normal level. Indeed, these mice suffer from two distinct types of acute leukemia, an early-onset c-Kit positive nonlymphoid leukemia and a late-onset B lymphocytic leukemia (Shimizu et al., 2004). It is noteworthy that the ages of onset of leukemia are markedly different between the mice with nonlymphoid leukemia and those with B cell leukemia. The average age of onset of mice with c-Kit^{+}/CD19^{-} leukemia is 143 days after birth, whereas that of mice with c-Kit^{-}/CD19^{+} leukemia is 387 days. In this paper, they claimed that residual GATA-1 activity in GATA-1.05/X mice is necessary for the propensity to develop leukemias. Even if GATA factors are below detectable level in B cells, knockdown of FOG-1 from early stages of B cell development changes the balance normally observed between FOG-1 partners. For instance, FOG-1/GATA-1, FOG-1/Ikaros and GATA-1/Ikaros stoechiometries could be displaced. It is therefore plausible that the targeted knockdown or overexpression of FOG-1 creates a situation where the differentiation potential of the precursor is shifted and an environment permissive for the transformation of normal B cells.

2. Microarray experiments

Cellular proliferation and differentiation are 2 different aspects of the development, corresponding to a dynamic equilibrium. To assess whether primary Pre BI cell differentiation was impaired in cFOG-1KO:*mb-1*Cre mice or in R26^{FOG-1}:*mb-1*Cre mice, we performed different microarray analysis.

2.1. On primary Pre BI cells isolated from R26^{FOG-1}:*mb-1*Cre mice

The microarray analysis realized using stringent conditions (P-value: 0.01, Absolute linear fold change: 2.0) on primary Pre BI cells isolated from R26^{FOG-1}:*mb-1*Cre mice revealed a change in expression of 99 genes compared to control Pre BI cells (Fig. 21 and Table 1). Among these misregulated genes, 78 genes were downregulated and 21 genes were upregulated. Of great interest, some contained several Ikaros binding sites in their promoter: for instance, *Lrrn4cl* and *Ctse* were upregulated in Pre BI cells with enforced expression of FOG-1; on the contrary, *Gbp6* and *Sh2d5* were downregulated in Pre BI cells overexpressing FOG-1. Moreover, at least 19 genes, which were misregulated in the Pre BI cells with enforced expression of FOG-1, exhibited GATA factor binding sites in their promoter. *Ctsa*, *Capg*, *Cdc42ep4*, *Sgpl1*, *Plk2*, *Zfand2a* and *Klf3* are some examples of such upregulated genes in Pre BI cells overexpressing FOG-1. *Klf3*, whose product binds to the CACCC box of erythroid cell-expressed genes and may play a role in haematopoiesis is a particularly interesting target gene. *Gfra2*, *Usp11*,

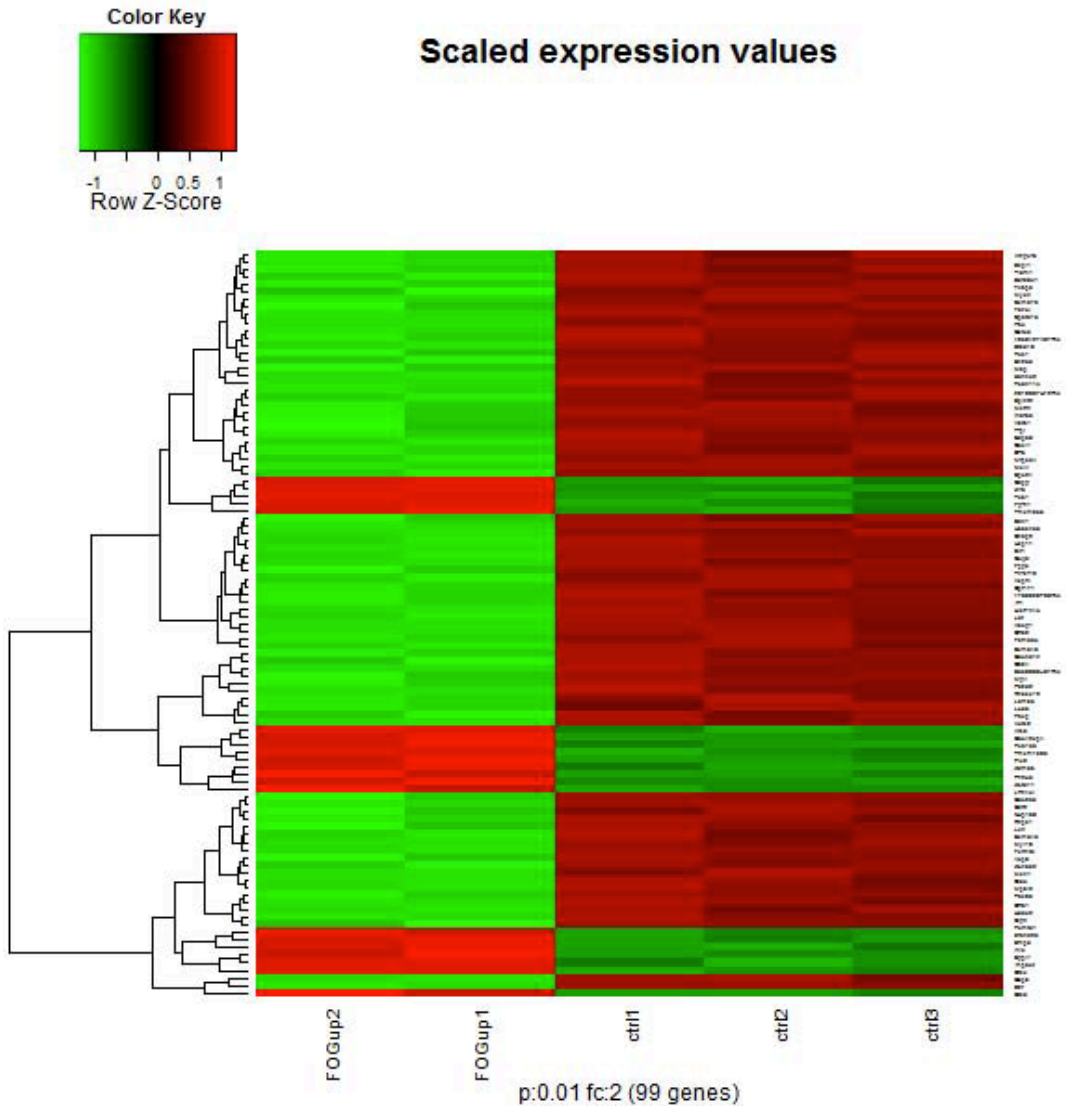


Figure 21 : Gene expression profiles of Pre BI cells with enforced expression of FOG-1 compared to control cells. After extraction from bone marrows of 2 males $R26^{FOG\text{-}1}$:*mb-1*Cre mice and from 2 male control animals, $B220^{+}$, $CD19^{+}$, $NK1.1^{+}$, c-kit^{+} and 493^{+} cells considered as Pre BI cells were sorted. The cells were put back in culture for 4 days after seeding. Ctrl1, 2 and 3 stand for Pre BI cells that do not express the Cre under the *mb-1* promoter and therefore do not overexpress FOG-1. FOGup1 and 2 stand for Pre BI cells that overexpress FOG-1 under the *Rosa26* promoter.

Ubash3a, *Aspm*, *Pdxp*, *Btbd10* and *Zdhhc2* were downregulated in Pre BI cells with enforced expression of FOG-1. More interestingly, *Tiam1* (*T cell lymphoma invasion and metastasis 1*), *Irf1* (*Interferon regulatory factor 1*) and *Pde8a* (*Phosphodiesterase 8A*), a gene involved in immune cell activation were also downregulated in Pre BI cells isolated from $R26^{FOG\text{-}1}$:*mb-1*Cre mice. Additionally, interesting epigenetic regulator genes, *Sap30* and *Hist1h4c* were identified as downregulated genes in the array. Besides, *Emp3*, *Fgfrl1* and *Tob1*, three genes involved in cell proliferation were upregulated in Pre BI cells overexpressing FOG-1. In addition, *Fgfrl1* and *Tob1* have been shown to have a negative effect on cell proliferation. Otherwise, *Sema4d*, *Sema7a*, *Tfpi*, *Spint2*, *Tnfsf10*, *Pbk* and *Fermt3*, which were downregulated in Pre BI cells with enforced expression of FOG-1, are genes important for blood function and immune responses. More specifically, *Lat* and *Sit1* also found downregulated in Pre BI cells overexpressing FOG-1, have been shown to be required for proper T cell signalling. *Abtb1*, another downregulated gene has been shown to play a role in developmental processes. Finally, three genes encoding epigenetic factors exhibiting enzymatic activities were identified as misregulated genes in Pre BI cells overexpressing FOG-1: *Kdm5b* (*Lysine-specific demethylase 5b*) was upregulated, whereas *Men1* (*Multiple endocrine neoplasia 1*) and *Hdac2* (*Histone deacetylase 2*) were downregulated in Pre BI cells isolated from $R26^{FOG\text{-}1}$:*mb-1*Cre mice.

2.2. On primary Pre BI cells isolated from cFOG-1KO:*mb-1*Cre mice

The microarray analysis realized using stringent conditions (P-value: 0.0001, Absolute linear fold change: 2) on primary Pre BI cells isolated from cFOG-1KO:*mb-1*Cre mice revealed a change in expression of 22 genes compared to control Pre BI cells (Fig. 22 and Table 2). Among these misregulated genes, 9 genes were downregulated and 13 genes were upregulated. Of great interest, *Fmn2* a gene already known to be expressed during B cell development (in GC B cells) was downregulated in Pre BI cells KO for FOG-1. *Rgs18* was another interesting gene downregulated in the Pre BI cells KO for FOG-1. It is noteworthy that these two latter genes contain regulatory GATA-1 binding sites in their promoter. Finally, *Prelid2*, a gene involved in mouse embryonic development was also downregulated in Pre BI cells KO for FOG-1. *Ptpn3*, *Ms4a6b*, *Tlr7*, *Itgam* and *Zhx2* were some of the genes found upregulated in the array. *Ptpn3*, *Ms4a6b* and *Zhx2* were genes notably involved in T cell development regulation. *Itgam* encodes a receptor known to play a role in B cell development. It has been shown to be expressed in precursor B cell acute lymphoblastic leukaemia. This experiment should have probably been done later after appearance of the phenotype in culture.

Gene abbreviation	Gene full name	Misregulation
Abtb1	Ankyrin repeat and BTB (POZ) domain containing 1	↘
Aspm	Asp (abnormal spindle) homolog, microcephaly associated	↘
Btbd10	BTB (POZ) domain containing 10	↘
Capg	Capping protein, gelsolin-like	↗
Cdc42ep4	CDC42 effector protein (Rho GTPase binding) 4	↗
Ctsa	Cathepsin A	↗
Ctse	Cathepsin E	↗
Emp3	Epithelial membrane protein 3	↗
Fermt3	Fermitin family member 3	↘
Fgfrl1	Fibroblast growth factor receptor-like 1	↗
Gbp6	Guanylate binding protein family, member 6	↘
Gfra2	GDNF family receptor alpha 2	↘
Hdac2	Histone deacetylase 2	↘
Hist1h4c	Histone cluster 1, H4c	↘
Irf1	Interferon regulatory factor 1	↘
Kdm5b	Lysine-specific demethylase 5B	↗
Klf3	Kruppel-like factor 3 (basic)	↗
Lat	Linker for activation of T cells	↘
Lrrn4cl	LRRN4 C-terminal like	↗
Men1	Multiple endocrine neoplasia 1	↘
Pbk	PDZ binding kinase	↘
Pde8a	Phosphodiesterase 8A	↘
Pdxp	Pyridoxal (pyridoxine, vitamin B6) phosphatase	↘
Plk2	Polo-like kinase 2	↗
Sap30	Sin3A-associated protein, 30 kDa	↘
Sema4d	Sema domain, immunoglobulin domain (Ig), transmembrane domain (TM) and short cytoplasmic domain, (semaphorin) 4D	↘
Sema7a	Semaphorin 7A, GPI membrane anchor	↘
Sgpl1	Spingosine-1-phosphate lyase 1	↗
Sh2d5	SH2 domain-containing protein 5	↘
Sit1	Signaling threshold regulating transmembrane adaptor 1	↘
Spint2	Serine peptidase inhibitor, Kunitz type, 2	↘
Tfpi	Tissue factor pathway inhibitor (lipoprotein-associated coagulation inhibitor)	↘
Tiam1	T-cell lymphoma invasion and metastasis 1	↘
Tnfsf10	Tumor necrosis factor (ligand) superfamily, member 10	↘
Tob1	Transducer of ERBB2, 1	↗
Ubash3a	Ubiquitin associated and SH3 domain containing A	↘
Usp11	Ubiquitin specific peptidase 11	↘
Zdhhc2	Zinc finger, DHHC-type containing 2	↘
Zfand2a	Zinc finger, AN1-type domain 2A	↗

Table 1: List of a number of genes found misregulated in the microarray performed in Pre BI cells overexpressing FOG-1. The arrows are indicating whether the gene is down or upregulated.

2.3. FOG-1 target genes in primary Pre BI cells

A comparison of the gain (R26^{FOG-1} mice) and loss-of-function (cFOG-1KO mice) analysis reveals an overlap of genes that are reciprocally regulated in both experimental approaches. They are downregulated in primary Pre BI cells isolated from R26^{FOG-1}:*mb-1*Cre mice and they are upregulated in primary Pre BI cells isolated from cFOG-1KO:*mb-1*Cre mice. These genes, namely *Hdac2*, *Igll1*, *Ikzf3* and *Rag1* can be regarded as direct FOG-1 targets with high confidence.

HDAC2 has already been shown to play a role in B cells in different biological systems. In the chicken DT40 cell line, HDAC2 controls the amount of the IgM H-chain at the steps of both transcription of its gene and alternative processing of its pre-mRNA (Takami et al., 1999). In this system, HDAC2 downregulates IgM light chain gene promoter activity (Takechi et al., 2002) and modulates immunoglobulin gene conversion frequency and distribution (Lin et al., 2008). Experiments in the B cells of conditionally targeted mice for HDAC1 and HDAC2 demonstrated that B cell development strictly requires the presence of at least one of these enzymes. When both enzymes are ablated, B cell development is indeed blocked at an early stage, and the rare remaining Pre B cells show a block in G1 accompanied by the induction of apoptosis (Yamaguchi et al.).

Igll1 encodes one of the surrogate light chain subunits, lambda5 and is a member of the immunoglobulin gene superfamily. The Pre B cell receptor is found on the surface of Pro B and Pre B cells, where it is involved in transduction of signals for cellular proliferation, differentiation from the Pro B cell to the Pre B cell stage, allelic exclusion at the Ig heavy chain gene locus, and promotion of Ig light chain gene rearrangements. Ikaros proteins regulate the Pre BCR component lambda5 in a stage-specific manner. In Pre BI cells, Ikaros modulated lambda5 expression in competition with the transcriptional activator EBF. At the transition from the Pre BI to Pre BII stage, the expression of the Ikaros family member Aiolos was upregulated and required for the efficient silencing of *Igll1* (Thompson et al., 2007).

Ikzf3 encodes Aiolos a transcription factor that plays an important role in the regulation of lymphocyte differentiation. It plays an essential role in regulation of B cell differentiation, proliferation and maturation to an effector state (Wang et al., 1998). Notably, Aiolos is upregulated in chronic lymphocytic leukaemia (Duhamel et al., 2008). Aiolos together with Ikaros inhibit Pre B cell proliferation by directly suppressing c-Myc expression (Ma et al.).

Rag1 is a catalytic component of the Rag complex, a multiprotein complex that mediates the DNA cleavage phase during V(D)J recombination (Grundy et al., 2009). In the Rag complex, Rag1 mediates the DNA-binding to the conserved RSS and catalyzes the DNA cleavage activities by introducing a double-strand break between the RSS and the adjacent coding segment. The Rag complex also plays a role in the Pre B cell allelic exclusion, a process leading to expression of a single immunoglobulin heavy chain allele to enforce clonality and monospecific recognition by the BCR expressed on individual B lymphocytes (Franklin, 2006).

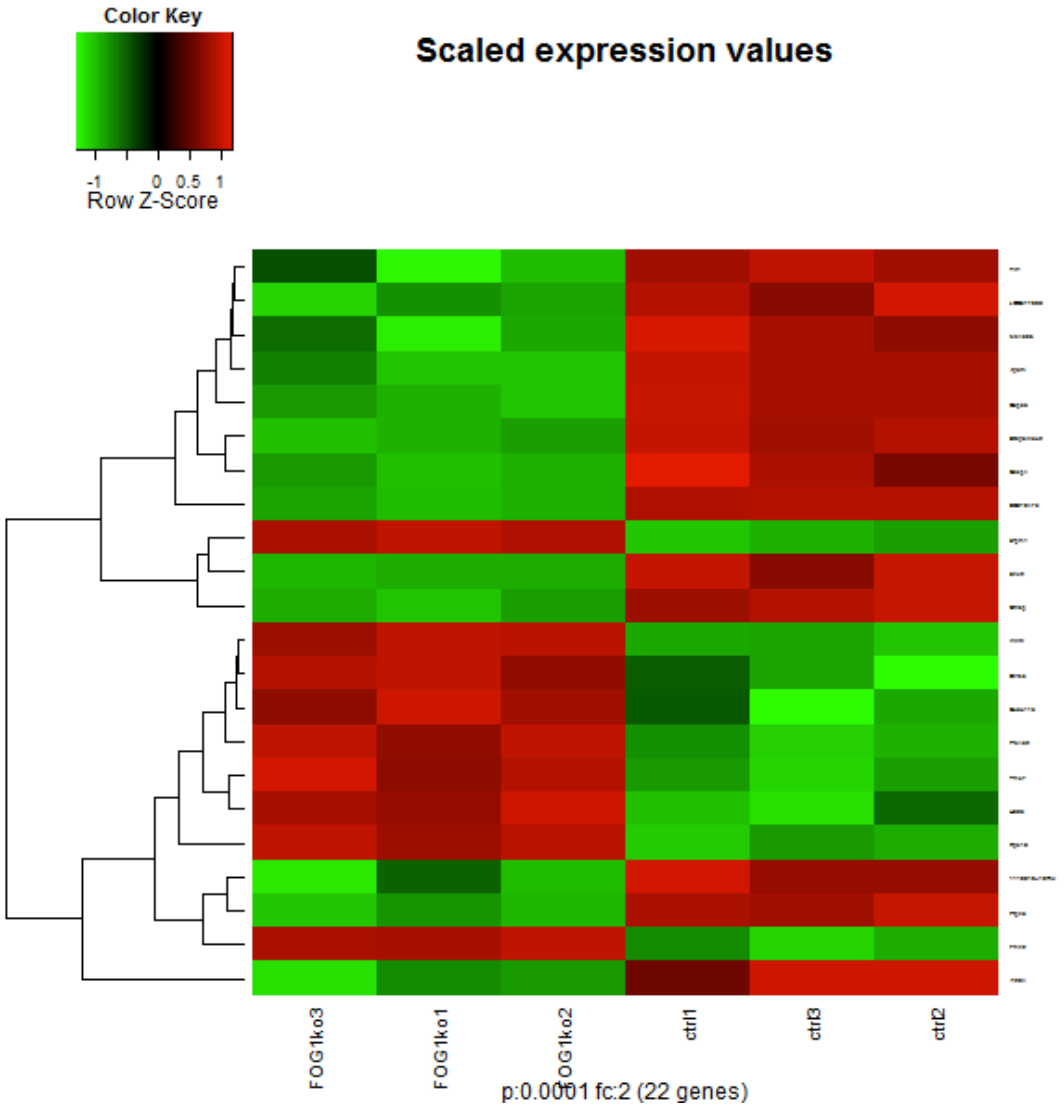


Figure 22 : Gene expression profiles of Pre BI cells KO for FOG-1 compared to control cells. After extraction from bone marrows of 2 males cFOG-1KO:*mb-1*Cre mice and from 1 male control animal. Ctrl1, 2 and 3 stand for Pre BI cells that do not express the Cre under the *mb-1* promoter and therefore are not deleted for FOG-1 expression. FOG-1ko1, 2 and 3 stand for Pre BI cells that do not anymore express FOG-1.

Gene abbreviation	Gene full name	Misregulation
Fmn2	Formin 2	↘
Itgam	Integrin alpha M	↗
Ms4a6b	Membrane-spanning 4-domains, subfamily A, member 6B	↗
Prelid2	Preli domain containing 2	↘
Ptpn3	Protein tyrosine phosphatase, non-receptor type 3	↗
Rgs18	Regulator of G-protein signaling 18	↘
Tlr7	Toll-like receptor 7	↗
Zhx2	Zinc fingers and homeoboxes 2	↗

Table 2 : List of a number of genes found misregulated in the microarray performed in Pre BI cells deleted for FOG-1. The arrows are indicating whether the gene is down or upregulated.

References

Duhamel, M., Arrouss, I., Merle-Beral, H., and Rebollo, A. (2008). The Aiolos transcription factor is up-regulated in chronic lymphocytic leukemia. Blood *111*, 3225-3228.

Franklin, A. (2006). Hypothesis: a biological role for germline transcription in the mechanism of V(D)J recombination--implications for initiation of allelic exclusion. Immunol Cell Biol *84*, 396-403.

Grundy, G.J., Ramon-Maiques, S., Dimitriadis, E.K., Kotova, S., Biertumpfel, C., Heymann, J.B., Steven, A.C., Gellert, M., and Yang, W. (2009). Initial stages of V(D)J recombination: the organization of RAG1/2 and RSS DNA in the postcleavage complex. Mol Cell *35*, 217-227.

Lin, W., Hashimoto, S., Seo, H., Shibata, T., and Ohta, K. (2008). Modulation of immunoglobulin gene conversion frequency and distribution by the histone deacetylase HDAC2 in chicken DT40. Genes Cells *13*, 255-268.

Ma, S., Pathak, S., Mandal, M., Trinh, L., Clark, M.R., and Lu, R. Ikaros and Aiolos inhibit pre-B-cell proliferation by directly suppressing c-Myc expression. Mol Cell Biol *30*, 4149-4158.

Shimizu, R., Kuroha, T., Ohneda, O., Pan, X., Ohneda, K., Takahashi, S., Philipsen, S., and Yamamoto, M. (2004). Leukemogenesis caused by incapacitated GATA-1 function. Mol Cell Biol *24*, 10814-10825.

Soriano, P. (1999). Generalized lacZ expression with the ROSA26 Cre reporter strain. Nat Genet *21*, 70-71.

Takami, Y., Kikuchi, H., and Nakayama, T. (1999). Chicken histone deacetylase-2 controls the amount of the IgM H-chain at the steps of both transcription of its gene and alternative processing of its pre-mRNA in the DT40 cell line. J Biol Chem *274*, 23977-23990.

Takechi, S., Adachi, M., and Nakayama, T. (2002). Chicken HDAC2 down-regulates IgM light chain gene promoter activity. Biochem Biophys Res Commun *299*, 263-267.

Thompson, E.C., Cobb, B.S., Sabbattini, P., Meixlsperger, S., Parelho, V., Liberg, D., Taylor, B., Dillon, N., Georgopoulos, K., Jumaa, H.*, et al.* (2007). Ikaros DNA-binding proteins as integral components of B cell developmental-stage-specific regulatory circuits. Immunity *26*, 335-344.

Wang, J.H., Avitahl, N., Cariappa, A., Friedrich, C., Ikeda, T., Renold, A., Andrikopoulos, K., Liang, L., Pillai, S., Morgan, B.A.*, et al.* (1998). Aiolos regulates B cell activation and maturation to effector state. Immunity *9*, 543-553.

Welner, R.S., Pelayo, R., and Kincade, P.W. (2008). Evolving views on the genealogy of B cells. Nat Rev Immunol *8*, 95-106.

Yamaguchi, T., Cubizolles, F., Zhang, Y., Reichert, N., Kohler, H., Seiser, C., and Matthias, P. Histone deacetylases 1 and 2 act in concert to promote the G1-to-S progression. Genes Dev *24*, 455-469.

3. Abbreviations

Ab Antibody
ACF Chromatin-Assembly Factor
AFC Ab-Forming Cell
Ag Antigen
AID Activation-Induced Deaminase
Aire Autoimmune regulator
APC Ag Presenting Cell
BAFFR B cell Activating Factor Receptor
BCL-6 B Cell Lymphoma 6
BCR B Cell Ag Receptor
bHLH basic Helix-Loop-Helix
Blimp-1 B lymphocyte induced maturation protein 1
bp base pair
btk Bruton's tyrosine kinase
CD Cluster of Differentiation
CDR Complementarity Determining Region
C/EBP CAAT Enhancer Binding Protein
CFU-MK Colony-Forming Unit-MegaKaryocytes
CHD Chromodomain Helicase DNA binding protein
CHRAC Chromatin Accessibility Complex
CLP Common Lymphoid Progenitor
CMJ Cortico-Medullary Junction
CMP Common Myeloid Progenitor
CSR Class Switch Recombination
CtBP C-terminal Binding Protein
CTD C-Terminal Domain
cTEC cortical Thymic Epithelial Cell
CXCL13 CXC-chemokine Ligand 13
DC Dendritic Cell
DN Double Negative
DNA DeoxyriboNucleic Acid
DNA-PK DNA-dependent Protein Kinase
DP Double Positive
DSB Double Strand Break
EBF Early B cell Factor
Elf-1 E74-like factor 1
ES Embryonic Stem cell
ETP Early T cell Progenitor
FACS Fluorescence-Activated Cell Sorting
FOG-1 Friend Of GATA-1
FR Framework Region
GC Germinal Center
GMP Granulocyte/Monocyte Progenitor
granB granzymeB
GTF General Transcription Factor
HDAC Histone Deacetylase
HLH Helix-Loop-Helix
HMG High Mobility Group

HS Hypersensitive site
HSC Haematopoietic Stem Cell
IFN-γ Interferon-gamma
Ig Immunoglobulin
Igh Ig heavy chain gene
Igll1 Immunoglobulin lambda-like polypeptide 1
IgM Immunoglobulin M
IKKβ Inhibitor of nuclear factor Kappa-B Kinase subunit beta
IL Interleukin
iNKT invariant Natural Killer T
Inr Initiator
IRF Interferon-Regulatory Factor
ISP Immature Single Positive
kb kilobase
LCR Locus Control Region
Lef-1 Lymphoid enhancer binding factor 1
LPS Lipopolysaccharide
LSD1 Lysine (K)-Specific Demethylase 1
M Myeloid
Mac-1 Macrophage receptor 1
MAR Matrix Attachment Region
MBD Methyl-CpG-Binding Domain
MEP Megakaryocyte/Erythroid Progenitor
MHC Major Histocompatibility Complex
MITF Microphtalmia-associated Transcription Factor
MK Megakaryocyte
mRNA messenger RNA
MTA Metastasis-Associated
mTEC medullary Thymic Epithelial Cell
MZ Marginal Zone
MZ B Marginal Zone B cell
NF-κB Nuclear Factor-kappaB
NHEJ Non-Homologous End Joining
NK Natural Killer
NKT Natural Killer T cell
NuRD Nucleosome Remodeling and Deacetylase complex
NuRF Nucleosome Remodelling Factor
p mast cell protease
OBF-1 Oct-Binding Factor 1
PALS Periarteriolar Lymphoid Sheath
PAX-5 Paired box protein 5
PCR Polymerase Chain Reaction
PIC Pre-Initiation Complex
Pol RNA Polymerase
PRC2 Polycomb Repressive Complex 2
Pre B cell B cell Precursor
Pro B cell B cell Progenitor
Rag Recombination activating gene
RbBP Retinoblastoma-associated Binding Protein
RBC Red Blood Cell

RBP-Jκ Recombination signal Binding Protein for Ig kappa J region
rRNA ribosomal RNA
RNA RiboNucleic Acid
RSS Recombination Signal Sequence
S1PR1 Sphingosine 1-Phosphate Receptor 1
SCZ Subcapsular Zone
SHM Somatic Hypermutation
SOX4 Sex-determining region Y box 4
SP Single Positive
Stat6 Signal transducer and activator of transcription 6
TACC3 Transforming Acidic Coiled-Coil 3
TAF_n TBP-Associated Factor
TBP TATA-Binding Protein
TCR T Cell Ag Receptor
TdT Terminal deoxynucleotidyl Transferase
Th T helper cell
tph tryptophan hydroxylase
tRNA transfer RNA
XBP-1 X-Box-binding Protein 1
XRCC X-ray Repair Cross-Complementing protein

4. Acknowledgements

First of all, I would like to acknowledge my family, and notably my parents and my brother, who always gave me support during my studies and especially during my thesis.

Second, I wish to acknowledge my friends, who were always present when I needed during my PhD. Among them, I would particularly like to acknowledge Christine, Lynda, Marie, Charlène, Sylvia and Caoimhe.

Then, I also wish to acknowledge Dr. Patrick Matthias, my supervisor, who gave me the opportunity to do my PhD in a great scientific environment at the FMI.

I would like to thank as well the members of my thesis committee Susan Gasser and Antonius Rolink. Ragna Sack and Daniel Hess from the proteomic facility in addition to Hubertus Kohler from the FACS facility were of a great help for discussions. I should also cite Jean-François Spetz and Patrick Kopp from the transgenic mouse facility as well as Thierry Doll and Matthias Muller from Novartis.

I want to thank a former colleague in the laboratory, Camille Du Roure, who was very supportive and helped me to push things forward.

I also want to thank Fabien Cubizolles for making a funny atmosphere in the laboratory.

Finally, I want to thank Oliver Truee, Nina Reichert, Karen Cornille, three former lab members and the current lab members, including Gabriele Matthias, Chun Cao, Daniel Weis, Vincent Pillonel, Mohamed Amin Choukrallah, Anand Manoharan and Yasuyuki Miyake for stimulating discussions and a good atmosphere in the laboratory.

5. *Curriculum Vitae*

Aude VERSAVEL

403, Rue Charles Péguy
76520 Franqueville Saint Pierre
France
33 6 73 80 72 72
aude_versavel@yahoo.fr
Born November 6th 1983, in France.

DEGREES

Current **PhD in Cellular and Molecular Biology**
University of Basel
2007 **Master 2 in Cellular and Molecular Biology**
University Pierre et Marie Curie (with honours)
2006 **Master 1 in Cellular and Molecular Biology**
University Pierre et Marie Curie (with honours)
2005 **Bachelor** of science in Biology
University Pierre et Marie Curie (with honours)

LANGUAGES

Native in French
Fluent in English
Studied Spanish for 8 years

PROFESSIONAL EXPERIENCES

Sept. 2007-Current **PhD position** – FMI, Basel, Switzerland
FOG-1, a transcriptional regulator within the haematopoietic system
Pr Dr Patrick Matthias
Sept. 2006-July 2007 **Research Internship** – Curie Institute, Paris, France
Function of the Asf1 histone chaperones during *Xenopus laevis* development
Dr Geneviève Almouzni
July-Sept. 2006 **Research Internship** – Experimental Biology Group, Development Biology Laboratory, CNRS UMR 7622, Paris, France
Function of Tbx6 gene during *Xenopus laevis* development
Dr Clémence Carron
July-Sept. 2005 **Research Internship** – Centre de Recherche des Cordeliers, INSERM U872, Team 16, Paris, France
Mechanisms by which human IntraVenous Immunoglobulins (IVIg) regulate the cell mediated auto-immune response
Dr Srini Kaveri
Aug. 2002 **Summer administrative positions** – AXA Assurances, Belbeuf, France

SOME OF MY PERSONAL INTERESTS

Love travelling and discovering new horizons. Enjoy skiing, swimming and classical dancing. Keen on cooking and tasting creative cuisine.