PhD
Program in Molecular
and Translational Medicine
DIMET

(XXII cycle, year 2008-2009)

University of Milano-Bicocca
Department of Biotechnologies and Biosciences

The Sox6 transcription factor: its role in human and murine erythroid differentiation and mechanisms for its regulation

Claudio Cantu' **Matr. 036251**

PhD Coordinator: Prof. Andrea Biondi

Tutor: Prof. Sergio Ottolenghi

The research presented in this thesis was performed at the Department of Biotechnology and Biosciences, University of Milano-Bicocca, in the laboratory of genetics headed by Prof. Sergio Ottolenghi and Prof. Antonella Ronchi.

"I do remember one formative influence in my undergraduate life. There was an elderly professor in my department, who had been passionately keen on a particular theory for, oh, a number of years, and one day an American visiting researcher came and completely and utterly disproved our old's man hypothesis.
The old man strode to the front, shook his hand and said : - My dear fellow, I wish to thank you, I have been wrong these fifteen years-. And we all clapped our hands raw.
That was the scientific ideal, of somebody who had a lot invested, a lifetime almost invested in the theory, and he was rejoicing that he had been shown wrong, and the scientific truth had been advanced."

Richard Dawkins

TABLE OF CONTENTS

Chapter 1:

- GENERAL INTRODUCTION p.9
  - Hematopoiesis p.10
  - Erythropoiesis p.15
    - The globin switching p.18
    - Transcriptional control of the globin switching p.21
    - The regulation of erythroid differentiation p.24
    - Transcription factors p.26
    - The medical relevance: Hemoglobinopathies p.29
- SCOPE OF THE THESIS p.32
- REFERENCES p.35

Chapter 2:

- GENE EXPRESSION PROFILING DURING MOUSE ERYTHROID DIFFERENTIATION AND GLOBIN SWITCHING p.41
  - Introduction p.45
  - Results p.48
  - Data analysis p.50
  - How to select a candidate gene for functional studies? p.52
  - Future functional analysis of selected genes p.60
  - Discussion and future aims p.62
  - Materials and methods p.65
  - References p.66

Chapter3:

- SOX6 ENHANCES HUMAN ERYTHROID DIFFERENTIATION p.73
  - Introduction p.74
  - Experimental procedures p.77
  - Results p.83
    - Sox6 overexpression in human erythroleukemic K562 cells p.83
    - Sox6 overexpression in primary human erythroid progenitors p.89
    - Sox6 targets on DNA p.95
  - Discussion and future aims p.99
  - References p.104

Chapter 4:

- AUTOREGULATION OF SOX6 EXPRESSION DURING HUMAN ERYTHROID DIFFERENTIATION p.107
  - Abstract p.109
  - Introduction p.110
  - Experimental procedures p.112
  - Results p.119
    - Sox6 expression during erythoid maturation p.119
    - The Sox6 gene is repressed by high levels of exogenously overexpressed Sox6 p.121
    - The Sox6 gene contains Sox6 binding sites within its regulatory regions p.123
    - Sox6 binds in vitro and in vivo to the putative

Sox6 binding sequences identified within the Sox6 promoter p.125

- The -775 Sox6 binding sites within the Sox6 promoter mediate transcriptional repression in K562 cells. p.128

- Discussion p.132
- References p.137
- Footnotes p.140

Chapter 5:

- FINAL DISCUSSION: MOLECULAR AND CLINICAL CONSIDERATIONS
  - Summary p.142
  - Identification of new gene important for Erythropoiesis p.142
  - Sox6 in human erithroid system p.144
  - Sox6 and the globin switching p.145
  - Sox6 targets on DNA p.146
  - Sox6 blocks the progression through the cell cycle and activates SOCS3 p.148
  - Sox6 regulates its own transcription p.149
  - The clinical relevance p.151
- REFERENCES p.154

Chapter 1

GENERAL INTRODUCTION

HEMATOPOIESIS

The blood is one of the most fascinating tissues not only by its unique structure but also by its variety of cells. Each cell type has distinct appearance and performs specific and essential biological functions.
The blood is composed of different cell types: erythrocytes, platelets, lymphocytes, granulocytes, macrophages and mast cells; but among all of them, erythrocytes are the most abundant, occupying about 45% of the blood total volume. They are small and enucleated, with a biconcave discoid shape, completely filled with hemoglobin, the molecule responsible for their capability of transporting the oxygen and carbon dioxide throughout the body. The other cell types are less abundant, occupying about 1% of the blood volume, but performing important biological functions. Platelets are rod-shaped cell fragments derived from the differentiation of large cells denominated megakaryocytes, and have an essential role in blood coagulation. Lymphocytes, divided into B- and T-lymphocytes, play a crucial role in the specific immune response against microorganisms acting as direct killers either activating other cells that eliminate infected cells (T cells), or producing soluble antibodies that promote the elimination of the infected microorganism (B cells). Granulocytes are divided in neutrophils, eosinophils and basophils, and they represent the innate immune system having a role in the inflammatory response and phagocytosis. Macrophages and mast cells are not found in the circulation but within the tissues. Macrophages arise from monocytes wich are circulating in the blood before they migrate to the tissues were they terminally differentiate. Mast cells originate from an

unidentified progenitor in the bone marrow that migrates through the blood into mucosal tissues where they differentiate into mature mast cells.

All these specialized cells originate from a single cell, the hematopoietic stem cell (HSC), a relatively quiescent cell, that rarely divides to generate another stem cell and a progenitor, a very highly proliferative cell, that is able to initiate the whole hematopoiesis process (Fig.1).

HSCs are extremely rare cells: about 1 to 10 HSCs per 100.000 cells in the mouse bone marrow, where they reside (Abkowitz *et al.*, 2000). They have the ability to renew themselves and to differentiate into all different blood cell populations. All progenitors originated from the division of the HSC lose the self renewal ability and commit to differentiation, becoming multilineage precursors (MPL). Then the progenitors differentiation process occurs through a series of commitment steps, each leading to further restriction to a certain hematopoietic lineage. Which are these steps, and the existence of which lineages restricted progenitors, is a controversial issue, and different models are proposed (Katsura *et al.*, 2002; Prohaska *et al.*, 2002): the most widly accepted model claims that the first commitment step separates limphoyd from myeloid potential and is supported by the identification of a common lymphoid progenitor (CLP) and a common myeloid precursor (CMP). CLPs are restricted to the lymphoid lineage and can give rise exclusively to B-, T-cells and natural killer cells, while CMPs can give rise to granulocytes, erythrocytes, megakaryocytes and macrophages (Akashi *et al.*, 2000; Johnson *et al.*, 1977). These CMPs will undergo further lineage

restriction when the granulocytes/monocyte potential is separated from erythroid/magakaryocyte potential. The existence of CLP is now controversial because of the identification of myeloid/T cell (MTP) and myeloid/B-cell (MBP) bipotent precursors. The existence of such cells suggests an alternative model in which myeloid potential is maintained in the early commitment stages of all hematopoietic lineages. According to this model MPLs can give rise to common myeloid lymphoid precursors (CMPLs). CMPLs further commit into either MBPs or MTPs. Later on, the myeloid potential is separated from the megakaryocytic/erythroid and the B- and T-cell potential (Kawamoto H. and Katsura Y., 2009).

During the last few years, however, the idea of this rigid hierarchy between the HSC and the different hematopoietic cells has been questioned. In fact, an increasing numbers of reports suggest the occurrence, in particular in-vitro conditions, of a transdifferentiation process between different hematopoietic precursors (Graf *et al.*, 2002).

But why a particular progenitor cell choose its commitment is still now another controversial issue, and is a matter of speculation. Two contrasting models have been put forward: the instructive model and the permissive model. The instructive model states that specific exogenous signals induce commitment step. These signals are ligands and cytokines that interact with specific receptors on cell surface and thus activate intracellular pathways leading to the expression of genes specific of a given lineage. This model reinforces the importance of the stroma and the niche in the cell differentiation.

On the other hand, the permissive or selective model proposes that a cell-autonomous process, such as a stochastic change in the expression of some critical gene, drives commitment of blood cells to a distinct cell lineage. The role of the stroma, in this case, would be only to provide a selective environment for the growth and the survival of the committed progenitor cells (Enver T., Jacobsen S., 2009).

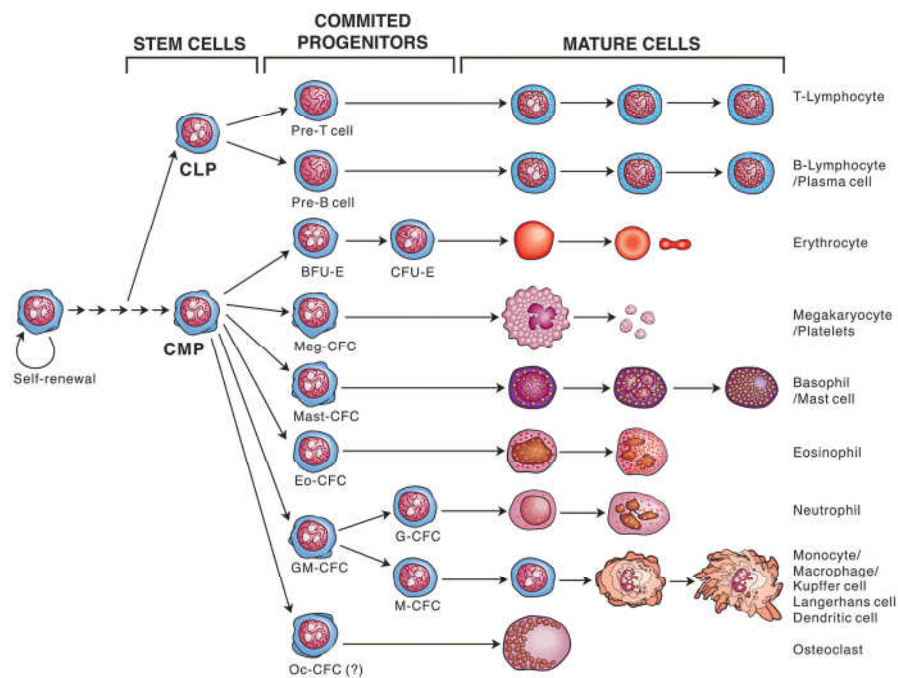


Figure 1. Schematic representation of hematopoiesis. The the stem cell, origin of all other cells, is on the left, and the differentiation occurs proceeding right way. All known lineages committed progenitors are indicated (From www.bloodlines.stemcells.com/chapters.html by Donald Metcalf).

In man, as in mouse, the site of hematopoiesis changes during embryonic development. In the early stage of mammalian development it takes place within the yolk sac. The first wave of hematopoiesis consists mainly on the production of large, nucleated, primitive erythrocytes that synthesize embryonic globins. Primitive hematopoiesis starts between embryonic day 7.0 (E7.0) and E7.5 in mice, or day 15-18 in humans, within the blood islands of the yolk sac, which cellular cluster composed by a central macrophage surrounded by masses of cells that gradually differentiate into primitive erythroblasts. These primitive erythroblasts enter the vascular system of the embryo proper where they continue to divide for several days and finally enucleate into the blood stream (Palis J., 2008).

Definitive hematopoiesis is responsible of the production of all hematopoietic lineages present in the adult organism including the definitive erythrocytes, characterized by the absence of nucleus and by the expression of adult globins. The embryonic origin of definitive hematopoiesis and of the HSC is a controversial issue and different models have been proposed so far.

It is now widely accepted that the yolk sac synthesizes a second transient wave of erythroid progenitors (that will give rise to "definitive erythrocytes"), that enter the bloodstream and seed the fetal liver. At the same time, hematopoietic stem cells emerging from the AGM (Aorta-Gonad-Mesonephros) region within the embryo, seed the liver and are the presumed source of long-term erythroid potential (Mikkola HKA. And Orkin SH., 2006). In humans the fetal liver becomes the site of definitive hematopoiesis around day 42, and in

mouse around E10.5. Fetal definitive erythroid precursors mature in macrophage islands within the liver, enucleate, and enter the bloodstream as erythrocytes (McGrath K., Palis J., 2008).

Later in gestation thymus and spleen are formed and colonized by hematopoietic progenitors involved in the production of differentiated lymphoid cells in fetal and adult stages. It has been also reported how the placenta is another major hematopoietic tissue during embryonic development (Gekas C. et al., 2008; Alvarez-Silva *et al.*, 2003). In mice, hematopoiesis also occurs in the spleen (Fig.2B).

Finally, around birth (around the 22th week of human gestation), the HSCs migrate from the fetal liver to the bone marrow, which becomes the principal hematopoietic tissue throughout the whole adult life (Palis and Segel 1998).

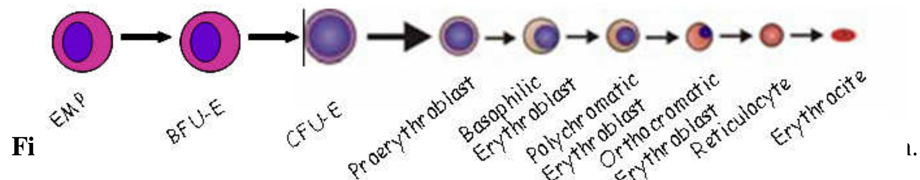
ERYTHROPOIESIS

The average lifespan of erythrocytes is 60 days in mice and 120 in humans. To guarantee a constant supply of oxygen to the body tissues, a continued generation of mature erythrocytes is necessary. The red cells volume is about the 45% of blood volume, and every day in our body 2×10^{11} red blood cells are produced, that means 2 million erythrocytes per second.

Like any other cell in the hematopoietic system, erythrocytes derives from the HSC. Erythropoiesis involves several steps of programmed differentiation between the HSC and the mature erythrocyte (Fig. 2).

First of all, the HSC differentiates into a MLP, also called CFU-S, which is a pluripotent cell that does not retain the capability to self-

renew. CMPs are derived from this CFU-S and can give rise to MEPs, which are fully restricted to the megakaryocytic and erythroid lineages.



Fi l.
Erythroid/megakaryocyte progenitors (EMP) can give rise to both erythrocytes and megakaryocytes. Burst forming units-erythroid (BFU-e) and colony forming units-erythroid (CFU-e) are erythroid restricted progenitors identifiable with functional assays. Terminal erythroid differentiation begins with the proerythroblast, the first morphologically identifiable progenitor and terminates upon the hemoglobin protein accumulation and nucleus extrusion .

The Burst Forming Unit-Erythroid (BFU-E) is the most primitive erythroid-restricted progenitor, and we are able to identify it by functional assays (Wong *et al.* 1986): in the presence of erythropoietin (Epo), interleukin 3 (IL 3), granulocyte-macrophage colony stimulating factor (GM-CSF), thrombopoietin and stem cell factor (SCF), this cell can give rise to large colonies of some thousands of hemoglobinized erythroblasts after 5-7 (mouse cells) or 14-16 (human cells) days of culture. The BFU-E further differentiates into the Colony Forming Unit-Erythroid (CFU-E), a more mature erythroid progenitor closely related to the proerythroblast, that, if plated in a semi-solid medium gives rise to small colonies, containing few hundreds hemoglobinized erythroblasts after 2-4 (mouse cells) or 5-8 (human cells) days of culture (Whyatt *et al.*, 2000; Wong *et al.*, 1986).

The proerythroblast is a large cell (14-19 µm and 11-13 µ in diameter in human and mice respectively) with a large nucleus, visible nucleoli, surrounded by a basophilic cytoplasm. Basophilic erythroblasts are slightly smaller cells (12-17 µm and 9-10 µm in diameter, in human and mouse respectively) as are their nuclei. A strong hemoglobinization occurs at the orthochromatic erythroblast stage. These cells are the smallest nucleated erythrocyte precursors (8-12 µm and 7-8 µm in diameter in human and mouse respectively) and their nuclei undergo pycnotic degeneration: the chromatin becomes very condensed and the nucleus becomes smaller. The nucleus is then extruded from the cell that becomes a reticulocyte. Reticulocytes are a little bit larger than fully mature erythrocytes (7-8 µm in man and 3.5-4.5 µm in mouse diameter) have irregular shapes and still contain few cytoplasmic organelles, necessary to complete the globin protein synthesis (Allen *et al.*, 1982; Bondurantand Koury, 1998).
It takes 48 to 72 hours from the proerythroblast to reticulocyte stage, and this process occurs in the erythroblastic island: a particular anatomic compartment able to provide a unique environment. It is a structure, consisting of a centrale macrophage surrounded by erythroid precursors at different stages of maturation. The more immature precursors are located to the centre, and as they mature they move away from the body of the macrophage. Erythroid precursors keep in contact with the cytoplasmic extensions of the macrophage during all stages of maturation until occurrence of enucleation (Bessis *et al.*, 1983), and the macrophage has the final role to phagocytise the extruded nucleus after the terminal maturation (Allen and Dexter, 1982; Allen and Testa, 1991).

THE GLOBIN SWITCHING

The globin switching is a complex molecular mechanism that leads to the production of different hemoglobin molecules at different stages of life during the development. The physiological meaning of this process, is the production of hemoglobin molecules with different affinity for the oxygen. For example, the fetus needs a hemoglobin with higher oxygen affinity than the mother's hemoglobin, to be able to extract oxygen from the mother's bloodstream.

The hemoglobin is a tetrameric protein, composed by two α and two β chains. In man, there are different α-like (ζ, α_1 e α_2) and β-like (ε, ${}^{G}\gamma$, ${}^{A}\gamma$, δ, β) genes, arranged on the chromosomes (11 and 16 respectively) in the same order they are temporally expressed during the development (Fig 3A and B).

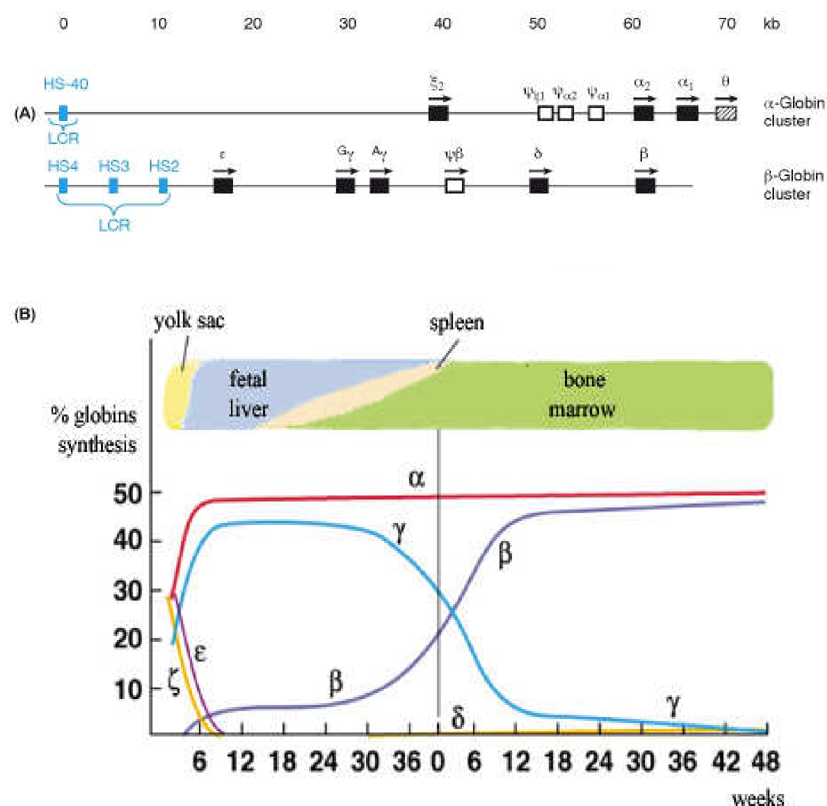


Figure 3. A) Schematic representation of the human globin genes clusters and their relative regulatory elements. **B)** The human genetic globin switching, and the change of hematopoietic organ during the ontogenesis.

Two globin gene switches occur during human development: the embryonic to fetal hemoglobin switch, which coincides with the transition from embryonic (yolk sac) to definitive (fetal liver) hematopoiesis; and the fetal to adult switch, which occurs at the perinatal period. The switches from ε to γ, from γ to β and from ζ to α gene expression are controlled predominantly at the transcriptional level. Unlike humans, most species (for example the mouse) have only one switch, from embryonic to definitive globin expression, occurring early in development. The expression of an exclusive globin gene during the fetal period is a rather recent event, which took place 35 to 55 million years ago during primate evolution (Reviewed in Stamatoyannopoulos G., 2005).

In mouse, the β-like globin genes cluster (ε^Y, βh1, β major, β minor) is located on chromosome 7, and the α-like globin genes (x, α_1, α_2) are on chromosome 11. The embryonic ε^Y and βh1 genes are expressed during embryonic life until they are substituted by adult β major and β minor genes between E11.5 – E13.5, when mouse globin switching occurs (Fantoni *et al.*, 1969) (Fig.3).

Although the mouse does not possess fetal globin genes, mice transgenic for the human globin locus linked to a LCR element express it "appropriately" during development: they start to transcribe fetal γ globin gene during the fetal life around E10.5 and silence it around E16 (Strouboulis *et al.*, 1992). This suggests that human and

mouse primitive erythroblasts undergo the similar differentiation mechanisms and for this reason mouse is a good model for studying the human globin switching.

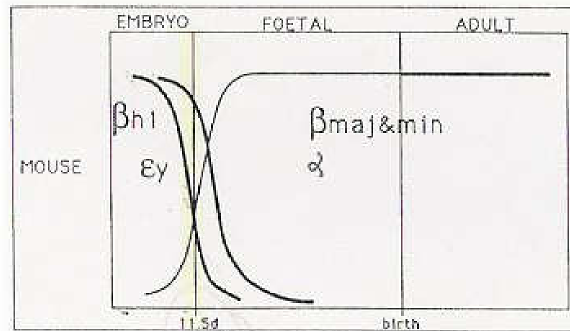


Figure 3: mouse globin switching.

The medical relevance of hemoglobin switching became apparent before any systematic research on its mechanism, and clinical experimentation started.

In a group of genetic disorders called hereditary persistence of fetal hemoglobin (HPFH), the expression of the γ-globin gene may persist at high levels in adult erythroid cells. Molecular studies of the HPFH syndromes have identified several important regulatory elements for the normal pattern of γ-globin gene expression. Increased levels of γ-globin, and so the production of a fetal hemoglobin (HbF) can considerably ameliorate the clinical course of inherited disorders of β-globin gene expression, such as β thalassemia and sickle cell anemia. HPFH is usually due to deletions of different sizes involving the β-globin gene cluster or to point mutations in the γ-globin gene promoters. There are some Mendelian forms of HPFH that are rare and do not explain the common form of heterocellular HPFH,

characterized by relatively low levels of fetal hemoglobin (2-4%), which represents the maiority of normal HbF variation, and is clearly inherited as a quantitative genetic trait (Forget BG., 1998; Thein SL. et al., 2009).

TRANSCRIPTIONAL CONTROL OF THE GLOBIN SWITCHING

The complex globin genes expression pattern is controlled by regulatory cis-acting regions on the DNA, that reside thousands of nucleotides far from the globin genes. The most studied is the β-globin genomic locus, whose expression is regulated by the Locus Control Region (LCR), a fundamental region for the globin genes expression. In fact, its absence leads to a complete loss of the expression of these genes in some β-thalassemic patients and to an inactive chromatin configuration (Forrester *et al.*, 1990). The LCR is located in man 6-22 kb upstream to the ε-globin gene, and is characterized by five DNAse I hypersensitive sites (HS), regions containing each a core of 200-300 nucleotides accessible to transcription factors and chromatin remodeling factors (Grosveld *et al.*, 1987). There are three other important distal HSs: one is located 20 kb downstream the human β-globin gene, and two are located 110 kb upstream the β-globin cluster (Bulger *et al.*, 2000). Many models have been proposed to explain how the LCR acts in regulating globin genes expression. One model is that LCR forms loops interacting with distal HSs and creates an “active chromatin hub” (ACH) which, together with promoters of the globin genes, is able to assemble a

“transcription factory” active during development (Fig. 4) (de Laat and Grosveld, 2003; Palstra *et al.*, 2003). Using the “chromosome conformation capture” (3C) technique it has been demonstrated that the transition from a “chromatin hub”, to an ACH requires stage specific transcription factors: for example EKLF is necessary for the formation on an active chromatin hub on the β-globin gene promoter (Reviewed in Noordermeer D. Wouter de Laat W., 2008).

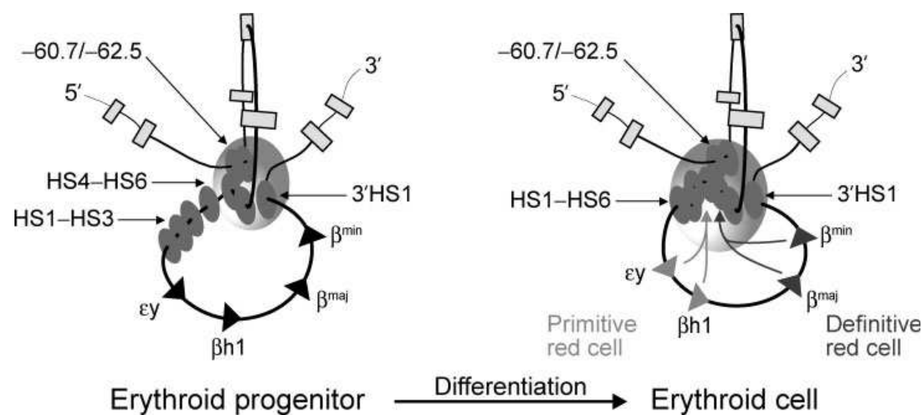


Figure 4. The structure of the mouse CH, and its activation to ACH through erythroid differentiation

In the murine primitive erythroid cells the ACH is mainly formed with embryonic globin genes promoters, whether in definitive erythroid cells it’s formed with adult globin genes promoters. The globin switching is then due to a chromatin remodeling that leads to the transcriptional activation in mutually exclusive manner of globin genes that need to be expressed at different specific moments during the development (Palstra *et al.*, 2003).

In the last fifteen years, al lot of transcription factors were demonstrated to bind different cis-regulatory region within the LCR locus, and to every single globin gene promoter, and the most

important ones are summarized in the figure 5, taken from Noordermeer D. and de Laat W., 2008.

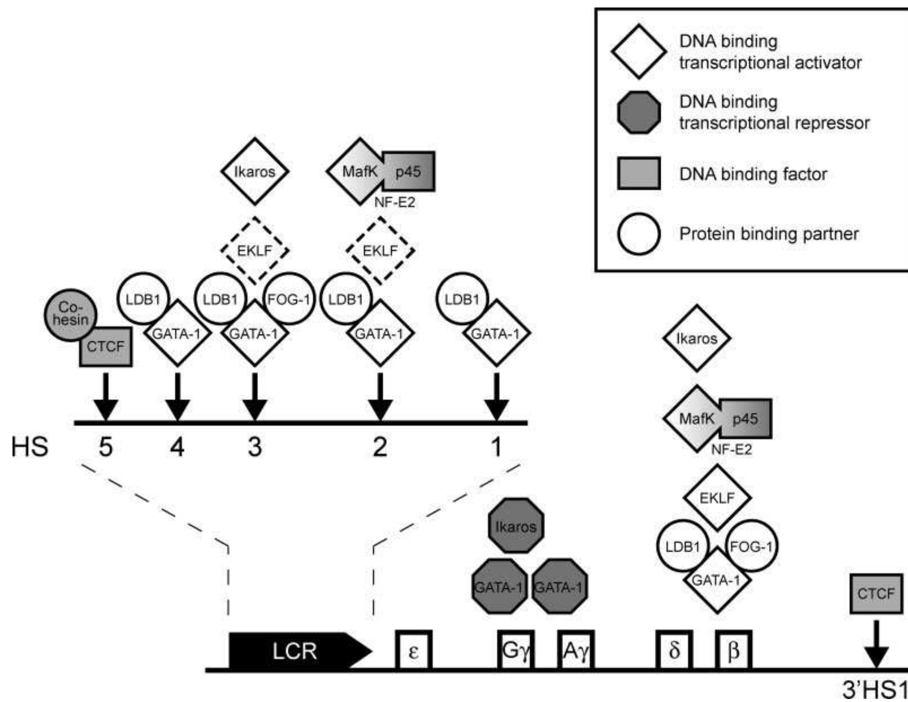


Figure 5. This figure combine several factors able to bind the b-globin locus (studied in human or mouse models). Globin genes are depicted by boxes, and HSs on the LCR are indicated by arrows. The dashed line that shows EKLF, indicates that its binding on the HS3 and HS2 is contradictory (From Noordermeer D. and de Laat W., 2008).

THE REGULATION OF ERYTHROID DIFFERENTIATION

The erythropoiesis needs to be regulated by external factors, produced by other cells, and by internal regulators that guarantee the correct genes expression at the different stages of the differentiation pathway.
Erythropoietin (Epo) is the principal cytokine able to regulate the definitive red cell maturation. It is produced by the liver during fetal development, and by the kidney through the adult life. The importance of the Epo pathway, is exemplified in Epo-/- and Epo receptor -/- mice, which die in utero of anemia, for a complete failure of definitive erythropoiesis beyond CFU-E stage. The Epo receptor is expressed on the surface of erythroid progenitors cells from the BFU-E stage, through to orthochtomatic erythroblasts (nearly 1000 receptors/cell). The cytoplasmic domain of Epo receptor binds the tyrosine Janus kinase (JAK2), which is activated by transphosphorilation following receptor dimerization. JAK2 is able to phosphorilate different substrates, but among them STAT5a and STAT5b (signal transducer and activator of transcription) are activated in the Epo receptor signaling cascade. Upon phosphorilation, STAT5a and b are able to dimerize, translocate into the nucleus, and function as transactivators of several genes important for the erythroid differentiation. The dephosphorilation of the Epo receptor is the principal mechanism for its downregulation, and is achieved by the phosphatase SHP-1, which is indeed able to dephosphorilate also JAK2. Another mechanism for downregulating the Epo receptor involves the SOCS family of proteins (Suppressor of Cytokine

Signaling). This proteins are able to inhibit JAK2 by directly binding to its kinase domain, and targeting it for ubiquitin-mediated proteosomal degradation. In the erythroid system, in particular SOCS1, CIS and SOCS3 are then activated by STATs factors following Epo binding to its receptor, leading to a negative feedback mechanism of STATs regulation. This complex regulatory network is summarized in the figure 6 (The complex Epo pathway is reviewed in Ingley et al., 2004).

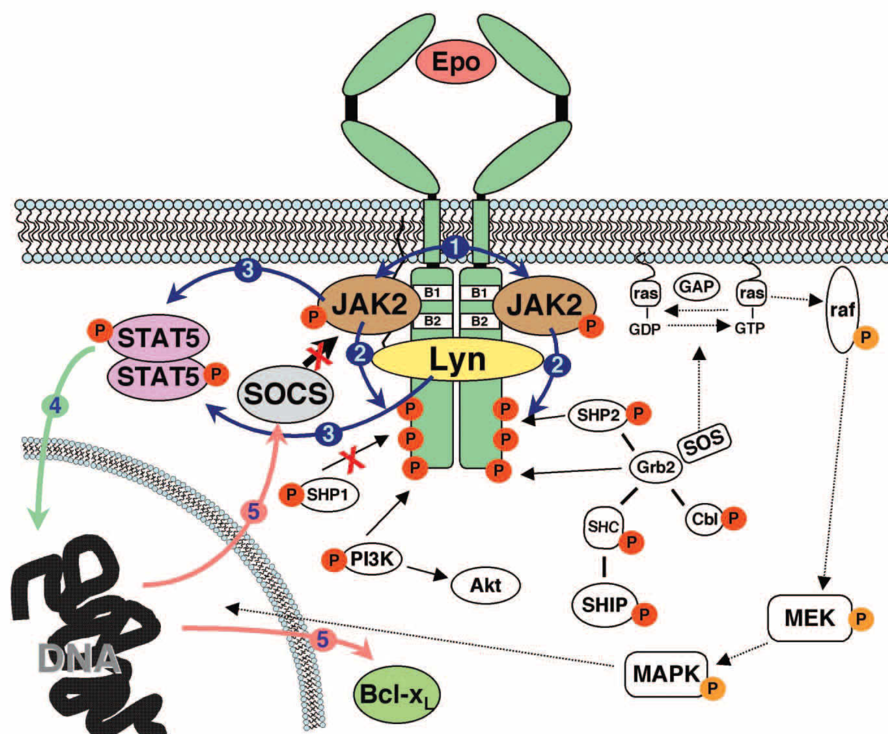


Figure 6. The complex pathway starting with the binging of Epo to its receptor. After Epo receptor phosphorilation, the pre-associated JAK2 becomes active and phosphorilates STAT5, which translocates to the nucleus and starts the transcription of several genes. One target is the antiapoptotic Bcl-X_L, and the SOCS family, inhibitors of this signaling pathway. Other pathways activated by this receptor are ras, raf, MAPK and PI3 kinase/Akt (From Ingley E. et al, 2004).

TRANSCRIPTION FACTORS

A host of transcriptional regulators that control erythropoiesis have already been identified, and now, major efforts are focused on elucidating the underlying molecular mechanisms. Canonical transcriptional mechanisms involve the binding of sequence-specific *trans*-acting factors (transcription factors) to particular DNA sequences, the *cis*-elements, followed by recruitment of coregulators via direct protein–protein interactions. Coregulators typically form large multiprotein complexes and mediate either activation or repression, in a context-dependent manner. Among the variety of transcriptional factors and coregulators of erythroid differentiation, there are some whose particular importance is worth studying:

-**GATA1** is a member of a family of six transcription factors, GATA-1 to GATA6 able to bind to the DNA consensus sequence (A/T)GATA(A/G) by two zinc-finger motifs. A common sequence motif [(A/T)GATA(A/G)] (GATA motif) exists within transcriptional regulatory regions at most, if not all, erythroid cell-specific genes. This founding member of the GATA transcription factor family is expressed in erythroid, megakaryocytic, eosinophil, and mast cell lineages. Targeted disruption of *Gata1* in mice provided evidence for its essential function in stimulating erythropoiesis (Pevny *et al.*, 1991, 1995; Weiss *et al.*, 1994). GATA-1 null erythroid cells fail to mature beyond the proerythroblast stage; GATA-1 null mouse embryos die due to severe anemia between E10.5 and 11.5; GATA-1 knock-down embryos (GATA1.05) which express only about 5% of GATA-1

levels relatively to normal, show an arrest of the primitive erythropoiesis and die between E11.5 and E12.5 (Kim SI. Bresnick EH., 2007).

High levels of GATA-1 arrests cellular proliferation (Rylski *et al.*, 2003; Munugalavadla *et al.*, 2005); GATA-1-mediated survival of erythroid precursors is obtained by $Bcl\text{-}_{xL}$ antiapoptotic protein expression (Weiss and Orkin, 1995; Gregory *et al.*, 1999) and by the activation of the transcription of Epo receptor (Chiba *et al.*, 1993), whose signaling is known to be an important progenitors survival pathway (Lacombe *et al.*, 1999).

- **EKLF**, Erythroid Kruppel-like transcription factor binds CACCC motifs, analogous to other KLFs that have diverse roles during cellular differentiation and development (Bieker, 2001). The targeted disruption of EKLF in mice demonstrates that EKLF is crucial for erythropoiesis and for adult β-like globin gene transcription (Nuez *et al.*, 1995; Perkins *et al.*, 1995). It was initially considered to regulate exclusively the definitive erythropoiesis (Nuez *et al.*, 1995), but a recent work has demonstrated that EKLF also is able to promote primitive erythropoiesis (Hodge *et al.*, 2006). Transcriptional profiling studies revealed that EKLF regulates diverse genes, including those encoding heme biosynthesis enzymes and cytoskeletal proteins (Drissen *et al.*, 2005; Hodge *et al.*, 2006; Nilson *et al.*, 2006). EKLF is an important determinant of hemoglobin switching (Donze *et al.*, 1995; Perkins *et al.*, 1996; Wijgerde *et al.*, 1995; Gillemans *et al.*, 1998). EKLF occupies HS1–HS3 of the β-globin LCR and the β*major* promoter (Im *et al.*, 2005). The results of EKLF ablation in mouse,

and EKLF occupancy at the endogenous β-globin locus provide strong evidences that EKLF is an essential regulator of adult β-like globin gene regulation.

-p45/NF-E2, which consists of a hematopoietic-specific subunit, p45/NF-E2 (Ney *et al.*, 1993; Andrews *et al.*, 1993a), and a member of the small Maf protein family (Andrews *et al.*, 1993b), is expressed in erythroid and megakaryocytic cells. Despite the enhancer activity mediated by NF-E2-binding sites, targeted disruption of *p45/NF-E2* in mice did not appear to significantly perturb erythropoiesis or β-globin transcription, but showed that p45/NF-E2 is crucial for megakaryopoiesis (Shivdasani and Orkin, 1995; Shivdasani *et al.*, 1995b). CB3 murine erythroleukemia cells, which lack p45/NF-E2 expression due to retroviral insertion within its locus, do not express β *major*, and p45/NF-E2 expression reactivates β*major* expression (Kotkow and Orkin, 1995; Kiekhaefer *et al.*, 2004). Endogenous p45/NF-E2 occupies the LCR, and the β*major* promoter in erythroid cells (Sawado *et al.*, 2001). NF-E2 expression in fact rescues Pol II occupancy at the β-globin promoter in CB3 cells, resulting in a considerable transcriptional activation (Johnson *et al.*, 2001). Taken together these data indicate that NF-E2 is an important regulator of hemoglobin synthesis. The lack of a large defect in β-like globin expression in p45/NF-E2-null mice might be due to the existence of some NF-E2-related factors, able to compensate for the defect.

THE MEDICAL RELEVANCE: HEMOGLOBINOPATHIES

Hemoglobinopathies are the most common inherited monogenic diseases, and affect million people in the world; among them, in particular, β-thalassemia and Sickle cell anemia. These diseases are caused by mutations involving the adult β-globin gene, responsible for the synthesis of over 97% of the adult-type haemoglobin HBA ($\alpha_2\beta_2$). Patients carrying heterozygous mutations on one β gene develop a mild anemia; mutations on both β genes cause a very severe anemia and require intensive clinical treatments (Stamatoyannopoulos and Nienhuis, 1994; Weatherall and Clegg, 1981; Forget, 1998; Thein, 1998).

The only possible clinical treatment is blood transfusion, that causes another series of symptoms, for example a toxic accumulation of Fe^{2+} within liver and heart: patients can overcome the surplus of Fe^{2+} only with daily perfusions with specific iron-chelating drugs.

Nowadays the researchers are trying to find different approaches for new therapeutical trials heading to heal thalassemia. Two different strategies are now considered.

The first one is gene therapy, whose aim is to introduce a normal β-globin gene within hemopoietic stem cells of the patient using viral vectors that can drive the expression of a correct variant of the gene in a tissue specific manner. At the moment, virus derived from HIV seems to be the only way to transduce the β-globin gene into hemopoietic progenitors, but they still cannot be used for human gene

therapy for safety reasons, although some progress toward safe and effective gene therapy have been recently reported (Lebensburger J., Persons DA., 2008).

The second one is the maintenance of fetal γ-globin gene expression during the whole adult life.

The fetuses carrying genetic mutations that lead to β-thalassemia or sickle cell anemia, are not affected by the disease, since during the embryonic-fetal stages of development, other genes (ε and γ globin) rapresent with α-globin the largest majority of haemoglobin molecules $\alpha 2\varepsilon 2$ (Gower2) and $\alpha 2\gamma 2$ (HbF). Numerous clinical observations show how coinheritance of mutations causing high expression of γ globin after birth (Hereditary Persistance of Fetal Hemoglobin, HPFH), significantly ameliorates the clinical condition of patients affected by thalassemia or Sickle cell anemia (Stamatoyannopoulos and Nienhuis, 1994; Weatherall and Clegg, 1981; Forget, 1998; Thein, 1998).

Many efforts have been made to increase γ-globin expression in β thalassemic and HbS patients using drugs as butyrates, 5-azaC or their derivates. Butyrates are non-specific inhibitors of histone deacetylase that lead to the maintenance of the histone acetylation on promoters, with the consequent transcriptional activation of the γ-globin gene.

5-azaC and its derivates inhibits the DNA methylation; since the γ-globin promoter is ipermethylated and indeed silenced during adult life, 5-azaC leads to an ipomethylation of this promoter and a consequent upregulation of γ-globin gene expression (Atweh *et al.*, 2003).

The modest results obtained with these drugs encouraged the study of new therapeutical approaches with a particular attention to transcription factors and molecules that bind γ-globin promoter and that can modulate its transcriptional activity.

In the past our group identified some HPFH mutations within the γ-globin promoter that alter the binding of different nuclear factors (NF-E3, GATA-1, Sp1, NF-Y) involved in the regulation of γ-globin gene expression (Nicolis *et al.*, 1989; Ottolenghi *et al.*, 1989; Ronchi *et al.*, 1995; Ronchi *et al.*, 1996; Ronchi *et al.*, 1989).

Very recently genome-wide association studies in man identified three loci containing a series of five single-nucleotide polymorphisms (SNP), able to explain the 20% of variation in HbF expression (Menzel S., 2007). The SNP with the largest effect was localized in the locus of the transcription factor BCL11A, whose protein product was then demonstrated to be able to repress the γ-globin transcription by direct binding on its regulatory regions, in a stage-specific manner (Sankaran VG., 2008, Sankaran VG, 2009).

In conclusion, the identification of proteins involved in the regulation of globin genes through proteomics and functional genomics approaches, represents the rational therapeutical basis for possible future treatments of thalassemia.

SCOPE OF THE THESIS

The aim of my thesis, is to discover new genes involved in the regulation of erythroid differentiation, and hemoglobin switching process. After their identification, our goal is to study their function, and, possibly, their molecular mechanism of action. Basic research about erythropoiesis, and globin chains transcriptional regulation may have a strong impact on the medical treatment of a great variety of blood disease, such as β-thalassemia or sickle cell anemia: genetic disorders cause by impaired production of the hemoglobin. Many efforts have been done to study which genes are involved in the physiologic downregulation of human embryonic/fetal globin genes, in order to provide molecular approaches directed to the activation of these genes in thalassemic patients, in which the adult globin production is impaired. This was first suggested by the clinical observation that a group of genetic aberrations called hereditary persistence of fetal haemoglobin (HPFH), in which the level of fetal γ-globin is maintained at high levels in the blood of adult individuals, the clinical conditions of thalassemic patients were significantly ameliorated.

- In chapter 2 I will describe the genomic approach we used, based on the DNA microarrays technology, that allowed us to study the differential gene expression profiles of different cell populations in mouse fetal liver, in the period in which the hemoglobin switching process occurs (E11.5 to E13.5). Genes emerging from this study are genes differentially expressed during erythropoiesis and the globin

switching, and represent potential targets of pharmacological treatment, for therapeutical approaches of hemoglobinopathies, based on the reactivation of globin fetal genes in thalassemic patients.

- In chapter 3 are represented all the experiments we used to study the molecular function of one gene identified and selected with methods described in chapter 2: the transcription factor Sox6. The main idea is to overexpress the selected gene in a human model context (a cell line and a primary culture extracted by normal or thalassemic patients), then measure and describe the phenotypic changes addressable to its enforced expression. Further we addressed the molecular mechanism of action of this protein, looking for its early target genes on the DNA, by using either bioinformatic tools, that reveal its possible target binding sites within the human genome, and molecular tools, necessary to validate in-vitro what was found in-silico.

- Finally, in chapter 4, we demonstrated that, among Sox6 targets, there is Sox6 itself. I will describe the precise genomic sequences, within the Sox6 human promoter, to which Sox6 is able to bind in-vitro and in-vivo, and how it is indeed able to repress its own transcription, suggesting a feedback mechanism of auto-regulation during erythroid maturation, either using a series of reporter gene studies of this region linked to a reporter gene, and a quantitative analysis of the endogenous Sox6 transcript and protein, strongly downregulated upon Sox6 forced expression in a cell line in culture.

- In chapter 5, all the results presented in this thesis are briefly discussed as well as the future perspectives of this research in molecular and translational medicine.

REFERENCES

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

Allen, T. D. and T. M. Dexter. 1982. Ultrastructural aspects of erythropoietic differentiation in long-term bone marrow culture. Differentiation **21**:86-94.

Allen, T. D. and N. G. Testa. 1991. Cellular interactions in erythroblastic islands in long-term bone marrow cultures, as studied by time-lapse video. Blood Cells **17** :29-38.

Alvarez-Silva, M., P. Belo-Diabangouaya, J. Salaun, and F. Dieterlen-Lievre. 2003. Mouse placenta is a major hematopoietic organ. Development **130**:5437-5444.

Andrews, N. C., K. J. Kotkow, P. A. Ney, H. Erdjument-Bromage, P. Tempst, and S. H. Orkin. 1993. The ubiquitous subunit of erythroid transcription factor NF-E2 is a small basic-leucine zipper protein related to the v-maf oncogene. Proc.Natl.Acad.Sci.U.S.A **90**:11488-11492.

Andrews, N. C., K. J. Kotkow, P. A. Ney, H. Erdjument-Bromage, P. Tempst, and S. H. Orkin. 1993. The ubiquitous subunit of erythroid transcription factor NF-E2 is a small basic-leucine zipper protein related to the v-maf oncogene. Proc.Natl.Acad.Sci.U.S.A **90**:11488-11492.

Atweh, G. F., J. DeSimone, Y. Saunthararajah, H. Fathallah, R. S. Weinberg, R. L. Nagel, M. E. Fabry, and R. J. Adams. 2003. Hemoglobinopathies. Hematology Am.Soc.Hematol.Educ.Program.14-39.

Bieker, J. J. 2001. Kruppel-like factors: three fingers in many pies. J.Biol.Chem. **276**:34355-34358.

Bulger, M., M. A. Bender, J. H. van Doorninck, B. Wertman, C. M. Farrell, G. Felsenfeld, M. Groudine, and R. Hardison. 2000. Comparative structural and functional analysis of the olfactory

receptor genes flanking the human and mouse beta-globin gene clusters. Proc.Natl.Acad.Sci.U.S.A **97**:14560-14565.

Bulger, M., M. A. Bender, J. H. van Doorninck, B. Wertman, C. M. Farrell, G. Felsenfeld, M. Groudine, and R. Hardison. 2000. Comparative structural and functional analysis of the olfactory receptor genes flanking the human and mouse beta-globin gene clusters
129. Proc.Natl.Acad.Sci.U.S.A **97**:14560-14565.

Chiba, T., Y. Nagata, A. Kishi, K. Sakamaki, A. Miyajima, M. Yamamoto, J. D. Engel, and K. Todokoro. 1993. Induction of erythroid-specific gene expression in lymphoid cells. Proc.Natl.Acad.Sci.U.S.A **90**:11593-11597.

de Laat, W. and F. Grosveld. 2003. Spatial organization of gene expression: the active chromatin hub
130. Chromosome.Res. **11**:447-459.

de Laat, W. and F. Grosveld. 2003. Spatial organization of gene expression: the active chromatin hub. Chromosome.Res. **11**:447-459.

Donze, D., T. M. Townes, and J. J. Bieker. 1995. Role of erythroid Kruppel-like factor in human gamma- to beta-globin gene switching. J.Biol.Chem. **270**:1955-1959.

Enver, T. and S. E. Jacobsen. 2009. Developmental biology: Instructions writ in blood. Nature **461**:183-184.

Fantoni, A., C. A. De la, D. Chui, R. A. Rifkind, and P. A. Marks. 1969. Control mechanisms of the conversion from synthesis of embryonic to adult hemoglobin. Ann.N.Y Acad.Sci. **165**:194-204.

Forget, B. G. 1998. Molecular basis of hereditary persistence of fetal hemoglobin. Ann.N.Y Acad.Sci. **850**:38-44.

Forrester, W. C., E. Epner, M. C. Driscoll, T. Enver, M. Brice, T. Papayannopoulou, and M. Groudine. 1990. A deletion of the human beta-globin locus activation region causes a major alteration in chromatin structure and replication across the entire beta-globin locus. Genes Dev. **4**:1637-1649.

Gekas, C., E.Rhodes K., K. A. Mikkola H. 2008. Isolation and analysis of hematopoietic stem cells from the placenta. J Vis Exp. 10:3791/742.

Gillemans, N., R. Tewari, F. Lindeboom, R. Rottier, T. de Wit, M. Wijgerde, F. Grosveld, and S. Philipsen. 1998. Altered DNA-binding specificity mutants of EKLF and Sp1 show that EKLF is an activator of the beta-globin locus control region in vivo. Genes Dev. **12**:2863-2873.

Graf, T. 2002. Differentiation plasticity of hematopoietic cells. Blood **99**:3089-3101.

Gregory, T., C. Yu, A. Ma, S. H. Orkin, G. A. Blobel, and M. J. Weiss. 1999. GATA-1 and erythropoietin cooperate to promote erythroid cell survival by regulating bcl-xL expression. Blood **94**:87-96.

Grosveld, F., G. B. van Assendelft, D. R. Greaves, and G. Kollias. 1987. Position-independent, high-level expression of the human beta-globin gene in transgenic mice. Cell **51**:975-985.

Hock, H., M. J. Hamblen, H. M. Rooke, J. W. Schindler, S. Saleque, Y. Fujiwara, and S. H. Orkin. 2004. Gfi-1 restricts proliferation and preserves functional integrity of haematopoietic stem cells. Nature **431**:1002-1007.

Hodge, D., E. Coghill, J. Keys, T. Maguire, B. Hartmann, A. McDowall, M. Weiss, S. Grimmond, and A. Perkins. 2006. A global role for EKLF in definitive and primitive erythropoiesis. Blood **107**:3359-3370.

Ingley, E., P. A. Tilbrook, and S. P. Klinken. 2004. New insights into the regulation of erythroid cells. IUBMB.Life **56**:177-184.

Johnson, G. R. and D. Metcalf. 1977. Pure and mixed erythroid colony formation in vitro stimulated by spleen conditioned medium with no detectable erythropoietin. Proc.Natl.Acad.Sci.U.S.A **74**:3879-3882.

Johnson, K. D., H. M. Christensen, B. Zhao, and E. H. Bresnick. 2001. Distinct mechanisms control RNA polymerase II recruitment to

a tissue-specific locus control region and a downstream promoter. Mol.Cell **8**:465-471.

Katsura, Y. 2002. Redefinition of lymphoid progenitors. Nat.Rev.Immunol. **2**:127-132.

Kawamoto, H. and Katsura Y.. 2009. A new paradigm for hematopoietic cell lineages: revision of the classical concept of the myeloid–lymphoid dichotomy.Trends in Immunology. **30**:193-200

Kiekhaefer, C. M., M. E. Boyer, K. D. Johnson, and E. H. Bresnick. 2004. A WW domain-binding motif within the activation domain of the hematopoietic transcription factor NF-E2 is essential for establishment of a tissue-specific histone modification pattern. J.Biol.Chem. **279**:7456-7461.

Kim, I., T. L. Saunders, and S. J. Morrison. 2007. Sox17 dependence distinguishes the transcriptional regulation of fetal from adult hematopoietic stem cells. Cell **130**:470-483.

Kim, S. I. and E. H. Bresnick. 2007. Transcriptional control of erythropoiesis: emerging mechanisms and principles. Oncogene **26**:6777-6794.

Kotkow, K. J. and S. H. Orkin. 1995. Dependence of globin gene expression in mouse erythroleukemia cells on the NF-E2 heterodimer. Mol.Cell Biol. **15**:4640-4647.

Lacombe, C. and P. Mayeux. 1999. The molecular biology of erythropoietin. Nephrol.Dial.Transplant. **14 Suppl 2**:22-28.

Lebensburger, J. and D. A. Persons. 2008. Progress toward safe and effective gene therapy for beta-thalassemia and sickle cell disease. Curr.Opin.Drug Discov.Devel. **11**:225-232.

McGrath, K. E., T. P. Bushnell, and J. Palis. 2008. Multispectral imaging of hematopoietic cells: where flow meets morphology. J.Immunol.Methods **336**:91-97.

K. A. Mikkola H. and S. H. Orkin. The journey of developing hemopoietic stem cells. 2006. Development. **133**:3733-3744

Munugalavadla, V., L. C. Dore, B. L. Tan, L. Hong, M. Vishnu, M. J. Weiss, and R. Kapur. 2005. Repression of c-kit and its downstream substrates by GATA-1 inhibits cell proliferation during erythroid maturation. Mol.Cell Biol. **25**:6747-6759.

Nicolis, S., C. Bertini, A. Ronchi, S. Crotta, L. Lanfranco, E. Moroni, B. Giglioni, and S. Ottolenghi. 1991. An erythroid specific enhancer upstream to the gene encoding the cell-type specific transcription factor GATA-1. Nucleic Acids Res. **19** :5285-5291.

Nilson, D. G., D. E. Sabatino, D. M. Bodine, and P. G. Gallagher. 2006. Major erythrocyte membrane protein genes in EKLF-deficient mice. Exp.Hematol. **34**:705-712.

Noordermeer, D. and W. de Laat . 2008. Joining the loops: beta-globin gene regulation. IUBMB.Life **60**:824-833.

Nuez, B., D. Michalovich, A. Bygrave, R. Ploemacher, and F. Grosveld. 1995. Defective haematopoiesis in fetal liver resulting from inactivation of the EKLF gene. Nature **375**:316-318.

Ottolenghi, S., R. Mantovani, S. Nicolis, A. Ronchi, N. Malgaretti, B. Giglioni, and J. Gilman. 1989. Altered binding to the gamma-globin promoter of two erythroid specific nuclear proteins in different HPFH syndromes. Prog.Clin.Biol.Res. **316A**:229-236.

Palis, J. and G. B. Segel. 1998. Developmental biology of erythropoiesis. Blood Rev. **12**:106-114.

Palis, J. 2008. Ontogeny of erythropoiesis. Curr.Opin.Hematol. **15**:155-161.

Palstra, R. J., B. Tolhuis, E. Splinter, R. Nijmeijer, F. Grosveld, and W. de Laat. 2003. The beta-globin nuclear compartment in development and erythroid differentiation. Nat.Genet. **35**:190-194.

Park, I. K., D. Qian, M. Kiel, M. W. Becker, M. Pihalja, I. L. Weissman, S. J. Morrison, and M. F. Clarke. 2003. Bmi-1 is required for maintenance of adult self-renewing haematopoietic stem cells. Nature **423**:302-305.

Perkins, A. C., A. H. Sharpe, and S. H. Orkin. 1995. Lethal beta-thalassaemia in mice lacking the erythroid CACCC-transcription factor EKLF. Nature **375**:318-322.

Pevny, L., M. C. Simon, E. Robertson, W. H. Klein, S. F. Tsai, V. D'Agati, S. H. Orkin, and F. Costantini. 1991. Erythroid differentiation in chimaeric mice blocked by a targeted mutation in the gene for transcription factor GATA-1. Nature **349**:257-260.

Prohaska, S. S., D. C. Scherer, I. L. Weissman, and M. Kondo. 2002. Developmental plasticity of lymphoid progenitors. Semin.Immunol. **14**:377-384.

Ronchi, A., S. Nicolis, C. Santoro, and S. Ottolenghi. 1989. Increased Sp1 binding mediates erythroid-specific overexpression of a mutated (HPFH) gamma-globulin promoter. Nucleic Acids Res. **17**:10231-10241.

Ronchi, A., M. Berry, S. Raguz, A. Imam, N. Yannoutsos, S. Ottolenghi, F. Grosveld, and N. Dillon. 1996. Role of the duplicated CCAAT box region in gamma-globin gene regulation and hereditary persistence of fetal haemoglobin. EMBO J. **15**:143-149.

Rylski, M., J. J. Welch, Y. Y. Chen, D. L. Letting, J. A. Diehl, L. A. Chodosh, G. A. Blobel, and M. J. Weiss. 2003. GATA-1-mediated proliferation arrest during erythroid maturation. Mol.Cell Biol. **23**:5031-5042.

Sankaran, V. G., T. F. Menne, J. Xu, T. E. Akie, G. Lettre, B. Van Handel, H. K. Mikkola, J. N. Hirschhorn, A. B. Cantor, and S. H. Orkin. 2008. Human fetal hemoglobin expression is regulated by the developmental stage-specific repressor BCL11A. Science **322**:1839-1842.

Sankaran, V. G., J. Xu, T. Ragoczy, G. C. Ippolito, C. R. Walkley, S. D. Maika, Y. Fujiwara, M. Ito, M. Groudine, M. A. Bender, P. W. Tucker, and S. H. Orkin. 2009. Developmental and species-divergent globin switching are driven by BCL11A. Nature **460**:1093-1097.

Sawado, T., K. Igarashi, and M. Groudine. 2001. Activation of beta-major globin gene transcription is associated with recruitment of

NF-E2 to the beta-globin LCR and gene promoter. Proc.Natl.Acad.Sci.U.S.A **98**:10226-10231.

Shivdasani, R. A., E. L. Mayer, and S. H. Orkin. 1995. Absence of blood formation in mice lacking the T-cell leukaemia oncoprotein tal-1/SCL. Nature **373**:432-434.

Shivdasani, R. A., M. F. Rosenblatt, D. Zucker-Franklin, C. W. Jackson, P. Hunt, C. J. Saris, and S. H. Orkin. 1995. Transcription factor NF-E2 is required for platelet formation independent of the actions of thrombopoietin/MGDF in megakaryocyte development. Cell **81**:695-704.

Stamatoyannopoulos, G., C. A. Blau, B. Nakamoto, B. Josephson, Q. Li, E. Liakopoulou, B. Pace, T. Papayannopoulou, S. W. Brusilow, and G. Dover. 1994. Fetal hemoglobin induction by acetate, a product of butyrate catabolism. Blood **84**:3198-3204.

Stamatoyannopoulos, G. 2005. Control of globin gene expression during development and erythroid differentiation. Exp.Hematol. **33**:259-271.

Strouboulis, J., N. Dillon, and F. Grosveld. 1992. Developmental regulation of a complete 70-kb human beta-globin locus in transgenic mice. Genes Dev. **6**:1857-1864.

Thein, S. L. and J. E. Craig. 1998. Genetics of Hb F/F cell variance in adults and heterocellular hereditary persistence of fetal hemoglobin. Hemoglobin **22**:401-414.

Thein, S.L., Menzel S., Lathrop M., Garner C. 2009. Control of fetal hemoglobin: new insights emerging from genomics and clinical implications.Hum Mol Genet. 15;18(R2):R216-23.

Weatherall, D. J., L. Pressley, W. G. Wood, D. R. Higgs, and J. B. Clegg. 1981. Molecular basis for mild forms of homozygous beta-thalassaemia. Lancet **1**:527-529.

Weiss, M. J., G. Keller, and S. H. Orkin. 1994. Novel insights into erythroid development revealed through in vitro differentiation of GATA-1 embryonic stem cells. Genes Dev. **8**:1184-1197.

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

Whyatt, D., F. Lindeboom, A. Karis, R. Ferreira, E. Milot, R. Hendriks, M. de Bruijn, A. Langeveld, J. Gribnau, F. Grosveld, and S. Philipsen. 2000. An intrinsic but cell-nonautonomous defect in GATA-1-overexpressing mouse erythroid cells. Nature **406**:519-524.

Wijgerde, M., F. Grosveld, and P. Fraser. 1995. Transcription complex stability and chromatin dynamics in vivo. Nature **377**:209-213.

Wong, P. M., S. W. Chung, D. H. Chui, and C. J. Eaves. 1986. Properties of the earliest clonogenic hemopoietic precursors to appear in the developing murine yolk sac. Proc.Natl.Acad.Sci.U.S.A **83**:3851-3854.

Chapter 2

GENE EXPRESSION PROFILING DURING MOUSE ERYTHROID DIFFERENTIATION AND GLOBIN SWITCHING

GENE EXPRESSION PROFILING DURING MOUSE GLOBIN SWITCH

Claudio Cantu'[1], FrancescaBosè[1], Cristina Fugazza[2], Maura Casalgrandi[1], Sergio Ottolenghi[1] and Antonella Ronchi[1].

Dipartimento di Biotecnologie e Bioscienze, Università di Milano-Bicocca, Milano, Italy[1]; John Radcliffe Hospital, Weatheral Institute of Molecular Medicine, Oxford, UK[2].

In mouse erythropoiesis, the crucial events responsible for the correct fetal-adult globin “switching” take place between E10.5 and E13.5 in cells undergoing erythroid maturation. This work proposes a genomic approach to identify new candidate genes able regulate the globin switching process and the erythroid maturation. To this end we analyzed by microarray analysis, the global expression profile of cell populations purified from mouse fetal liver at different day of embryonic development (E11.5, E12.5, E13.5) and at different stages of maturation (on the basis of their expression of the two specific surface antigens, c-Kit, the Stem Cell Factor tyrosine-kinase receptor, that marks hemopoietic high progenitors, and Ter 119, a protein associated to the glycophorin A, the most important erythroid marker. Here we present the method that allowed us to choose the transcription factor Sox6 for our further studies, among a list of the more promising genes, that were expressed at low level in pluripotent hematopoietic progenitors (cKit+/TER119-), and their expression

level rises in erythroid committed early progenitors (cKit+/TER119+), and in the mature erythroblasts (cKit-/TER119+).

INTRODUCTION

The hematopoietic system sustains the production of all different blood cell types, through a complex and regulated process in which self renewing pluripotent Hematopoietic Stem Cells (HSC) progressively differentiate to progenitor cells of gradually more restricted potential. A fraction of this process, the erythropoiesis, leads to the generation of mature red blood cells that guarantee the supply of oxygen to the body tissues by the synthesis of the different globin chains required at different stages of development.

During mouse development, the first wave of erythropoiesis arises in the blood islands of the extra-embryonic yolk sac (E7-E7.5) and is sustained by progenitors with a restricted differentiation potential. Primitive erythroblasts enter the embryo’s bloodstream where they undergo terminal differentiation and at last enucleate. The yolk sac synthesizes a second transient wave of "definitive" erythroid progenitors with a megakaryocytic-erythroid restricted potential, that seed the liver of the fetus. By E11.5, hematopoietic stem cells within the embryo colonizes the liver and are the presumed source of long-term erythroid potential. Finally, around birth, HSCs migrate to the bone marrow which becomes the principal hematopoietic tissue throughout adult life (McGrath K., Palis J., 2008).

The transition from primitive to definitive erythropoiesis coincides with the switching in globin gene expression. In the mouse there are

four functional β globin genes: εY and βH1, expressed during primitive erythropoiesis, the β major and β minor adult genes whose expression becomes predominant starting from the establishment of the adult erythropoiesis (which takes place in fetal liver between days E11.5 and E13.5). Several results suggest that common mechanisms underlie the hemoglobin switching process in mouse and man. In fact, although a "fetal" stage, like that in man, is not apparent in the mouse, a transgenic human β locus is correctly regulated in the mouse (The ε globin is followed by γ and then β globin expression). Whereas at E10.5 ε and γ globin gene transcription predominates, at E13.5 the switch to β globin transcription is essentially complete (Palis J., 2008).

Despite the enormous efforts done in the last decades of molecular biology studies in this field, many of the molecules controlling globin genes switching remains unknown.

This is almost in part due to the complexity of the switching process, which is composted by two aspects:

- a cellular one, that involves the replacement of an early population of embryonic progenitors, deriving from the yolk sac, with definitive intraembryonic-derived hematopoietic cells (developmental switching).
- an intracellular component is the modulation within each cell of globin expression with progressive extinction of fetal hemoglobin and replacement by adult hemoglobin (Papayannopoulou *et al.*, 1979, Kingsley PD. Et al., 2006).

Also in mouse embryo ontogeny, globin switching from βH1- to εY is dynamically regulated at the transcriptional level, and it is not only due to the sequential appearance of primitive and definitive lineages. In adult human, for example, it is well known that during erythroblast maturation the proportion of γ-globin is much higher in immature than in more differentiated cells where the γ-globin gene is essentially suppressed.

This complex and mostly unknown mechanism suggests that, presumably, several categories of genes, controlling both intracellular events and extracellular signals, are involved in the switching process.

The availability of gene expression profiles techniques has been proved to be a very powerful tool for the high-throughput analysis of a very high number of genes expressed in different populations. The characterization of terminal erythropoiesis suffers from a series of limitations, mainly related to the low numbers of cells, their heterogeneity, and their transcriptional inactivity due to the absence of the nucleus in the last final stage of differentiation. For this reason, the use of immortalized cell lines (expecially obtained from human or murine leukemic clones) allows the availability of large number of homogeneous cells, but with aberrations regarding factor/hormone dependence, signal transduction, and altered cell cycle progression that might limit the physiological significance of the data obtained in these cellular systems.

In this study we propose a genomic approach for analysing the global expression profile of cell populations purified from mouse fetal liver at different stages of development (E11.5, E12.5, E13.5) and at

different stages of erythroid maturation (on the basis of their expression of the specific antigens c-kit and TER119).
This approach provides a database of mRNA's expression in "ex vivo" cells during mouse hemoglobin switch, very useful to discover new genes involved in the molecular mechanisms underlying normal and pathological erythropoiesis.

RESULTS

With the purpose of analysing the transcriptional profile of mouse globin switching, we purified hemopoietic cells from ex vivo fetal liver cell populations at days E11.5, E12.5 and E13.5 of embryonic development, and sorted them in three cell populations with a FACS (Fluorescence Activated Cell Sorter) by the expression of cKit (Stem Cell Factor tyrosine-kinase receptor) and TER119 (that recognize a protein associated to the glycophorin A, expressed only on the surface of erythrocytes) antigens (Fig.1):
- $cKit^{+}/TER119^{-}$: pluripotent hematopoietic progenitors (together with a small number of progenitors of other lineages).
- $cKit^{+}/TER119^{+}$: erythroid committed early progenitors.
- $cKit^{-}/TER119^{++}$: more differentiated erythroblasts and mature erythrocytes.

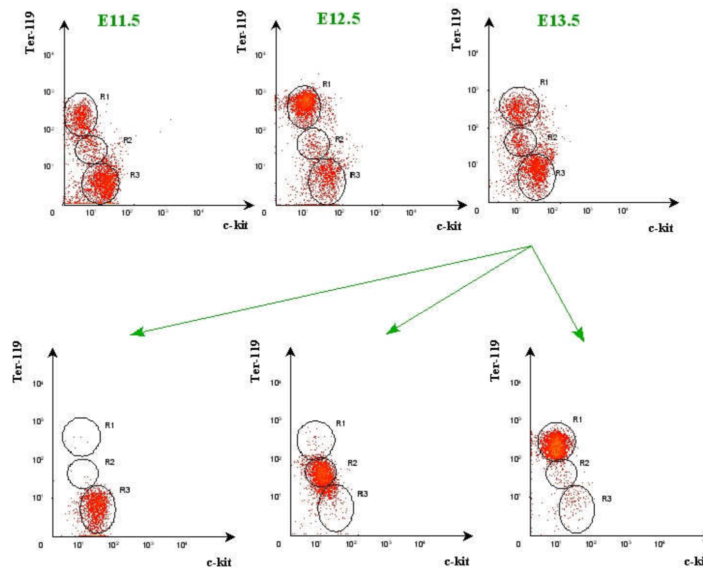


Figure 1. The FACS analysis of the three mouse fetal liver cell populations at the three different days of mouse embryonic development (E11.5, E12.5, E13.5). Lower panel: cell populations sorted at day E13.5; R1 represents cKit-/TER119+ cells, R2 cKit+/TER119+ and R3 cKit+/TER119- cells respectively. The profile of the purified populations is exactly the same for every one of the three days analyzed.

Several pregnant mice were used and the livers from numerous embryos were pooled. This helped to minimize the unavoidable slight differences of individual embryos. TER119^{+} cell population is mainly constituted of mature enucleated erythrocytes that contain almost exclusively hemoglobin protein and have already completed their transcriptional program. To eliminate these more mature erythroid cells we used a differential osmotic lysis buffer, able to lyse only mature red cells.

We obtained the global mRNA from each of these population, that was then analysed by Affimetrix technology (GeneChip Mouse Genome 430A). Each experiment was done in duplicate.
At day E11.5, the limited amount of cells in the small fetal livers, permitted us to purify only two populations: the $cKit^{+}/TER119^{-}$ cell population and a TER positive population, corresponding to a merged version of $cKit^{+}/TER119^{+}$ and the $cKit^{-}/TER119^{++}$ populations of the other two days: E12.5 and E13.5.

DATA ANALYSIS

The differentially expressed genes –DEG- (with an extent of variation of 3-4 fold changes -log2-) are classified in "functional families" using KEGG (Kyoto Enciclopedia of Genes and Genome) and Gene Ontology databases.
Analysing the data, we had a lot of internal controls: several genes, already known to be important for the erythroid development, were indeed differentially expressed. Among them for example, as expected, a lot were upregulated in the more differentiated cells: genes involved in HEME synthesis (ALAS2, UroPorphobilinogen synthase), erythroid-specific surface antigens and cytosckeletal components (Erythropoietin receptor, Glycophorin A, Spectrin alpha 1 and beta 1, Duffy Blood group, erythrocyte protein band 4.9 and 4.2, solute carrier family 4, member 1, ankyrin 1) (data not shown).
Another important example was constituted by the expression changes in globin genes. The developmental analysis shows a sharp decline in the transcription of the embryonic globin chains (βH1 and εY) from day 11.5 to day 13.5. The adult β major globin gene is already

expressed in the most immature progenitors indicating that, at day E11.5 the switching of the adult transcripts has already consistently begun. Their fold increase expression during maturation as compared to the embryonic transcripts was lower than what was expected, probably because a technical reason, for example the saturation of the probes, due to the too high level of adult globin mRNA in the cKit$^-$ Ter119$^+$ cells (Fig.2).

	E 11,5		E 12,5			E 13,5		
	ckit $^+$ TER119 $^-$	ckit $^-$ TER119$^+$	ckit $^+$ TER119 $^-$	ckit $^+$ TER119 $^+$	ckit $^-$ TER119 $^{++}$	ckit $^+$ TER119 $^-$	ckit $^+$ TER119 $^+$	ckit $^-$ TER119 $^{++}$
βH1	684	9148	269	455	1711	20	325	1275
εY	1284	6807	362	475	1859	17	229	432
β major	6362	6807	8067	8560	12194	7016	10001	8797

Figure 2. Embryonic and adult globin changes as analyzed by the DNA microarrays analysis carried out in sorted cell populations. Numbers represent arbitrary units.

Our experimental design allows three different possible and significant analyses focused on the different components of the switching process: development and maturation.

- Horizontal analysis: to study developmental-related changes. Similar cell populations are compared at the three different developmental times.

- Vertical analysis: to reveal changes during erythroid maturation.

  Cell populations at different stages of erythroid maturation are compared within the same developmental day. This comparison will allow to pinpoint changes related to cell differentiation. What is of a certain interest, is that the change observed for a given factor in the vertical analysis, for example at E11.5, could be far different from the change observed at E13.5.
- Diagonal analysis: it tries to consider both the developmental and the maturation related components of the switching. In fact, the progenitors $cKit^{+}/Ter119^{-}$ cells analysed at day E 11.5 will give rise at the more differentiated cell population at later days. This comparison has the aim to follow immature progenitors while maturing in vivo during the switching period.

HOW TO SELECT A CANDIDATE GENE FOR FUNCTIONAL STUDIES?

Among genes which expression varies the most in our transcriptional analyses, as expected, many are related to heme synthesis, protein degradation, red cell membrane biogenesis and their changes are similar at all developmental times. Theoretically, every gene category could be involved in globin switching, or important for differentiation. We could than select interesting target for further studies, among Differentially Expressed Genes (DEGs) in our experiment, from different categories.

Cytoskeleton proteins

During erythroid differentiation, a lot of genes encoding for cytoskeletal components (Erythropoietin receptor, Glycophorin A, Spectrin alpha 1 and beta 1, Duffy Blood group, erythrocyte protein band 4.9 and 4.2, solute carrier family 4, member 1, ankyrin 1) are upregulated. This is of course expected, since cytoskeletal remodelling is a fundamental process to produce a mature red blood cell.

Among DEGs there is gelsolin, an 82-84 kDa calcium-binding protein belonging to the gelsolin protein superfamily, which is involved in the remodelling of cytoskeleton, by the ability to sever and cap actin filaments (Janmey *et al.*, 1986). Gelsolin severs the actin polymer, and then remains attached to the fast growing end forming a cap. It is able to generate this way a large number of very short actin filaments which cannot reanneal or elongate at their barbed ends. The actin network is disassembled and can then be rapidly re-assembled by selective uncapping of gelsolin and liberation of polymerization competent filament ends after particular cellular stimuli (Yin and Stull, 1999). Gelsolin plays a role in different actin-mediated processes: cell motility, phagocytosis and apoptosis (Silacci *et al.*, 2004). It has also been implicated in a number of disease pathways: oncogenic transformation (Kwiatkowski *et al.*, 1999), inflammation and amyloidosis (Maury *et al.*, 2000; Maury *et al.*, 2001).

Gelsolin exists in two isoforms: one cytoplasmic, expressed in a wide varity of cell types (platelets, macrophages and neutrophils) (Witke *et al.*, 1995); the other is a plasma isoform, generated mainly in muscle cells by an alternative splicing that includes 25 amino acids (leader peptide) in the N-terminus of the protein (Yin *et al.*, 1984).

In our DNA microarray experiment, gelsolin expression decreases from cKit+/TER119- cells to cKit+/TER119+ cells, but increases again in mature cKit-/TER119+ cells (Fig. 3).

Cell to cell interaction proteins

Cell-cell interactions are fundamental since from the beginning of erythroid differentiation, expecially for a proper formation of the erythroblastic island where erythroid cells develop and mature.

One example is the ICAM4 integrin adhesion molecule, required for erythroblasic island formation and regulation of apoptosis in the bone marrow microenvironment (Lee *et al.*, 2006).

In this category of genes, we focused our attention on Map17 whose expression increased during erythroid maturation, especially at E11.5 (Fig. 3).

Map17 is a 17kDa membrane-associated protein, upregulated in various human colon, breast, lung and kidney carcinomas, it is normally expressed in the renal proximal tubules, but its physiological role is not already known (Kocher *et al.*, 1995; Kocher *et al.*, 1996). Map17 is adjacent to the gene encoding for SCL/Tal1 (Stem Cell Leukemia), a fundamental transcription factor of hemopoietic stem cells, and they are co-regulated by a common erythroid restricted enhancer located 40kb downstream of SCL exon 1a (Delabesse *et al.*, 2005).

Interestingly, the expression of Map17 in a colon carcinoma cell line, markedly decreases cell proliferation in vitro and tumor growth in vivo (Kocher *et al.*, 1996).

Cell cycle proteins

Many changes in our microarray data are related to the cell cycle (cyclines, cdk, cell cycle transcription factors). This is expected since it is know how erythroid precursors cells undergo several proliferation cycles, and the terminal erythroid differentiation is stricktly linked with the cell cycle withdrawal.

Of this category, in our experiment, Cdc25B and Btg2 increase their expression level during erythroid maturation (Fig. 3).

Cdc25B gene encodes for a G2 phase of the cell cycle phosphatase. Cdc25B3 recruits histone acetyl-transferase complexes such as CBP/p300, PCAF and SRCs (Ma *et al*., 2001).

Btg2 controls cell growth, differentiation and survival. It is an antiproliferative gene, important for the differentiation of neuronal precursors and osteoblasts in the developing mouse (Iacopetti *et al*., 1999; Raouf and Seth, 2002). Btg2 has also been described as a potential GATA-1 target gene (Rylsky *et al*., 2003).

	E 11,5		E 12,5			E 13,5		
	ckit $^{+}$ TER119 $^{-}$	ckit $^{-}$ TER119 $^{++}$	ckit $^{+}$ TER119 $^{-}$	ckit $^{+}$ TER119 $^{+}$	ckit $^{-}$ TER119 $^{++}$	ckit $^{+}$ TER119 $^{-}$	ckit $^{+}$ TER119 $^{+}$	ckit $^{-}$ TER119 $^{++}$
Gelsolin	121	38	175	80	201	150	56	83
MAP17	19	1050	45	45	150	50	45	250
Cdc25B3	78	982	300	411	1297	109	570	2251
Btg2	67	359	142	221	571	72	315	793

Figure 3. The changes in the expression of the selected candidate genes from the DNA microarrays. They all increase their expression (red) during erythroid

differentiation. Numbers represent arbitrary units. Changes in the expression of all the selected genes are validated by RT-PCR (Data not shown).

Since our aim is to find genes involved in the regulation of erythroid specific mechanisms, we focused our attention to a specific classes of transcripts: one encoding for transcriptional factors.

Transcription factors

Transcription factors are fundamental proteins for all lineage specific differentiation pathways. They are able influence the choice of the appropriate lineage by driving the expression of genes specific of the different cell types, also inhibiting inappropriate gene expression programs.

In our experiment many transcription factors show remarkable changes from the more immature (cKit^{+}/TER119^{-}) cells to the more mature cells (cKit^{+}/TER119^{+} and cKit^{-}/TER119^{+}): among them, the majority decreases, while only few increase their expression. The first trend is common to transcription factors known to be relevant for immature cells maintenance or for the specification of non erythroid lineages (GATA2, members of the CEBP family, myb, Hoxa9, runx1 and runx2, Pu1, Meis) (Fig. 4B).

On the other side some well known transcription factors, important for the fulfilment of the erythroid program such as GATA1 and its partners tal1, lmo2, ldb1, or Eklf, increase their expression level. We could observe only a slight increase, since they are already transcribed in the cKit positive/ TER119 negative cell fraction (and are not therefore picked up as DEG genes in the vertical analysis) (Fig. 4A).

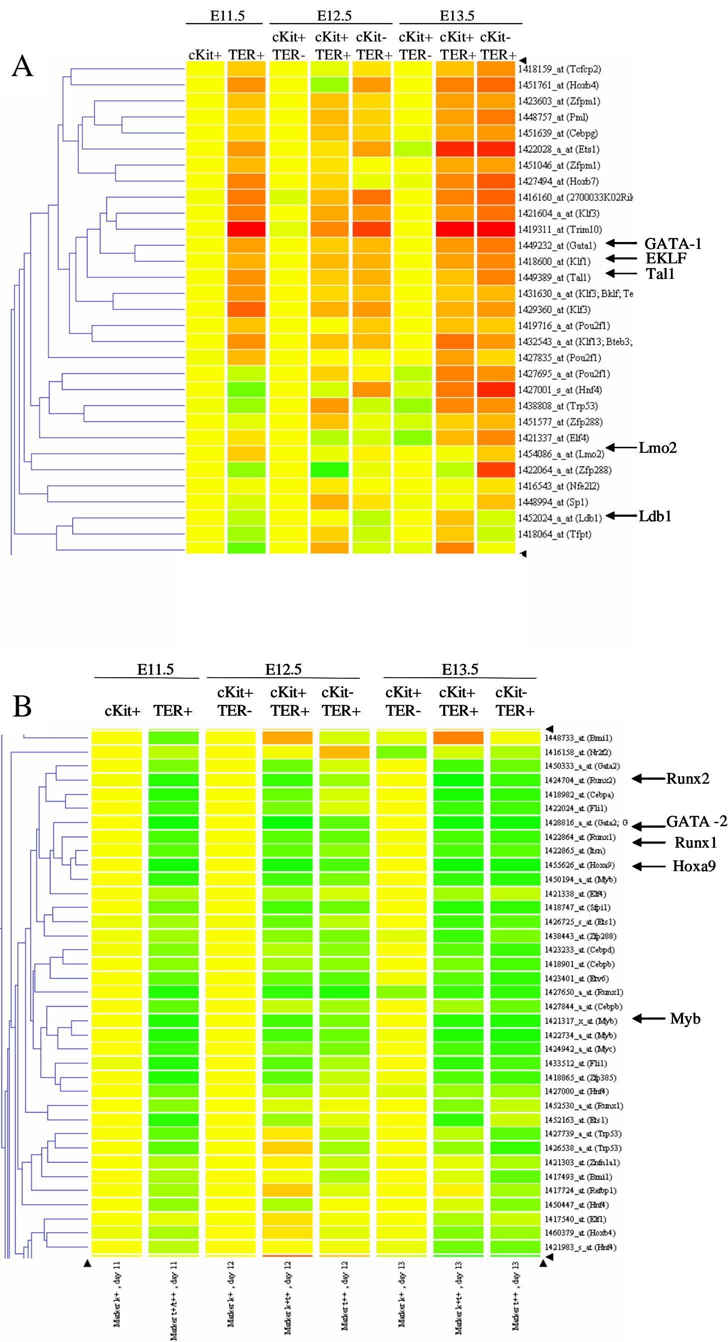


Figure 4. Clustering of DEG transcription factors. A: Genes whose expression mostly increases (in red). B: those whose expression decreases (in green).

Among the transcription factors sharply changing their level of expression in parallel with the erythroid commitment, we focused our attention to two members of the Sox (Sry-type HMG box) family: Sox6 and Sox4.

No role for these two transcription factors, at the moment of our choice, were know in the erythroid system.

Sox6 increases its expression while Sox4 decreases (Fig. 5).

	E 11,5		E 12,5			E 13,5		
	ckit $^{+}$ TER119 $^{-}$	ckit $^{-}$ TER119 $^{++}$	ckit $^{+}$ TER119 $^{-}$	ckit $^{+}$ TER119 $^{+}$	ckit $^{-}$ TER119 $^{++}$	ckit $^{+}$ TER119 $^{-}$	ckit $^{+}$ TER119 $^{+}$	ckit $^{-}$ TER119 $^{++}$
Sox4	51	15	113	28	47	83	16	2
Sox6	16	72	30	81	89	19	63	168

Figure 5. The expression level of the selected transcription factors: Sox6 transcription is always rising (red) during maturation, while Sox4 mRNA follows an opposite behaviour (green) These data were validated by RT-PCR on the same RNA preparations, and normalized on the expression of an housekeeping hene (HPRT)(not shown).

The Sox (Sry-type HMG box) family of transcription factors is characterized by a conserved HMG (High Mobility Group) DNA binding domain. SOX proteins bind to the minor groove of the DNA, and bend it of 70-80°, acting as chromatin architectural proteins, also by assembling multiprotein chromatin structures (Wegner, 1999; Wilson and Koopman, 2002). Sox proteins are important transcriptional regulators of various tissues specific differentiation processes during development (reviewed in Lefebvre V. Et al., 2007).

Sox6 null mutant mice (both Knock out mice, or homozygous for an inversion disrupting the Sox6 gene, p^{100H} mice, show delayed growth, myopathy and atrioventricular heart block, and the few that arrive to birth, die within 2 weeks after (Hagiwara *et al*., 2000). The expression of Sox6 is also required for the correct formation of nervous system (Connor *et al*., 1995; Hamada-Kanazawa *et al*., 2004; Hamada-Kanazawa *et al*., 2004 b), cartilage (Lefebvre *et al*., 1998; Ikeda *et al*., 2005) and muscle (Hagiwara *et al*., 2000; Cohen-Barak *et al*., 2003; Hagiwara *et al*., 2005)

Sox6 contains a leucine-zipper motif required for its dimerization with other proteins, including other members of the Sox family, in particular with Sox5 in the carthilage system (Lefebvre *et al*., 1998). Sox6 acts either as an activator or a repressor of transcription, obviously depending on its interactors on target sequences (Murakami *et al*., 2001; de Crombrugghe *et al*., 2000). Sox6 has also been shown to act as a general splicing factor (Ohe *et al*., 2002).

Recently, Sox6 has been shown to increase between E10.5 and E11.5, and to repress εY-globin transcription in the mouse, by directly binding on its promoter (Yi Z. *et al*., 2006). It has been shown that in the p^{100H} mouse, the transition from the definitive Erythropoiesis is impaired: an increased number of nucleated and misshapen definitive red blood cells is present in the fetal circulation, at the end of the gestation (Dumitriu B. *et al*., 2006).

On the other side, Sox4 is expressed in various tissues (Farr *et al*., 1993; Wilson *et al*., 2005). Sox4 knockout mice die from cardiac failure around day E14 (Ya *et al*., 1998). In the hematopoietic system,

Sox4 is only known to be required for B lymphocyte development (Schilham and Clevers, 1998).

Sox proteins bind similar sequences on DNA. The opposite pattern of expression of Sox6 and Sox4 during erythroid differentiation observed in our experiment lead us to hypothesize that these two genes might regulate the same targets, in an opposite fashion, and in different stages of differentiation. Hypothesizing an antagonistic function of these two proteins, we aim to study their role in murine and human erythropoiesis.

FUTURE FUNCTIONAL ANALYSIS OF SELECTED GENES

For the functional analysis of the selected genes, we plan to exploit cellular systems that may allow us to rapidly analyse the effects of the overexpression and/or the downregulation of a relatively large number of genes selected.

We will adopt the following cellular systems:

- K562, a human erythroleukemic cell line, extracted from a patient with Chronic myelogenous leukemia in blast crisis (Lozzio CB. and Lozzio BB., 1977). This cell line is able to differentiate toward an erythroid lineage only by induction with some chemical inducers, such as the Hemin, an Heme precursor molecule, the TPA or butyrate derivatives (Villeval JL. Et al., 1983). Nonetheless, this cell line is quite a useful model when a great number of cells is necessary for genetic of proteomic assays.

- A better and more physiologic model of primary culture of erythroid progenitors, directly extracted from human peripheral blood (Ronzoni et al., 2008), or from donor cord blood (Migliaccio G., et al. 2002). Hemopoietic $CD34^+$ progenitor cells are extracted and immunopurified. With the appropriate cytokine "cocktail" I will describe in the Matherials and Methods section, they expand in culture, and following the expression of erythroid specific markers (Glycophorin A, globins etc.), you can check their capability to retrace the physiologic erythroid differentiation.

To evaluate the role of candidate genes by overexpression, we are constructed lentiviral vectors based on pHR SIN BX IR/EMW (derived from pHR SIN CSGW, Demaison C. Et al., 2002), kindly provided by Dr. Cristina Fugazza, working in Prof. Tariq Enver's lab in Oxford. In this construct the candidate cDNA is cloned upstream to an internal ribosome entry site (IRES) sequence linked to the Emerald GFP coding sequence. Cells expressing the exogenous cDNA also express GFP and can therefore be FACS sorted for further in vitro culture and expansion or for colony assays.
K562 infected cells will be used for biochemical studies and to analyse their differentiation potential in the presence or absence of the differentiating drugs, TPA or hemin.
Although I will not present any results about this, we are setting up an experimental system able to silence the expression of the candidate gene, using the RNAi technology (Elbashir *et al.*, 2001; Sharp *et al.*, 2001). We will use lentiviral vectors in which the specific short-

hairpin cassette is cloned in the pSUPER retroviral vector (Oligoengine, Seattle).

Target cells (K562 and CD34^{+} progenitors cells), at various stages of maturation, will be infected, selected for their expression of GFP by FACS sorting and plated for colony assays or for long term cultures. Globin gene analysis and phenotypic changes will be then evaluated at the single colony level or in mass cultures.

DISCUSSION AND FUTURE AIMS

Gene expression profiling of mouse fetal liver cell populations

To identify new genes potentially able to regulate the synthesis of globin genes and that can represent potential target for therapeutical approaches to pharmacological treatment of hemoglobinopathies, we studied the gene expression in mouse fetal liver cell populations during the switch process (between E11.5 and 13.5), with a DNA microarrays.

Although an embryonic/fetal switching, like that in man, is not apparent in the mouse, transgenic ε, γ and β human globin genes are appropriately regulated in the mouse: ε expression is followed by γ and then β globin transcription between days E10.5 and E13.5 (Trimborn *et al.*, 1999; Wijgerde *et al.*, 1995).

We purified by FACS analysis, three selected cell populations from mouse fetal liver at days E11.5, E12.5 and E13.5 of embryonic development on the basis of their expression of cKit and TER119 antigens: the pluripotent hematopoietic progenitors (cKit^{+}/TER119^{-}),

the erythroid committed early progenitors ($cKit^+/TER119^+$) and the differentiated erythroblasts plus mature erythrocytes ($cKit^-/TER119^+$). The differences among the mRNAs of each population, were finally analysed by Affimetrix technology (GeneChip mouse Genome 430A).

Data analysis

Many categories of genes were found to vary; among transcription factors for example the expression of GATA-2, normally highly transcribed in more immature cells, drops in TER119 positive cells. Other transcription factors important in Erythropoiesis, such as GATA-1, tal1, lmo2, ldb1 and Eklf, are already transcribed at high levels both in immature cKit+ and mature TER119+ cells and for this reason are not scored among DEG genes.

The category of genes encoding for enzymes involved in the heme synthesis (ALAS2, UroPBGS), cell surface antigens, cytoskeleton components (Erythropoietin receptor, Glycophorin A, Spectrin alpha 1 and beta 1, Duffy Blood group, erythrocyte protein band 4.9 and 4.2, solute carrier family 4, member 1, ankyrin 1) and Globin genes, are upregulated in the most mature cell populations, and since they were already known to vary in this precise way, we can consider them as internal controls of our experiment.

It is interesting to observe how the β major globin gene is already expressed within the more immature cell populations at the beginning of the globin switch (E11.5) in the same moment βH1 and εY embryonic globin genes start to be downregulated (horizontal analysis), thus defining the fetal-adult globin switching.

Selection of candidate genes

Any category of gene, in principle, may be relevant for globin switching. Nonetheless, among DEGs we selected for further analysis genes belonging to categories of transcription factors, signalling molecules and cell cycle regulators, that showed a strong and robust variation (>3-4 fold change-log2-) during development and/or during maturation (vertical and horizontal analysis respectively). Among them we focused our efforts in two proteins belonging to the same group of transcription factors: Sox6 and Sox4. These two Sox proteins bind the same sequences on DNA, but have opposite pattern of expression in our experiment: Sox6 is upregulated while Sox4 is downregulated during maturation. This permitted us to hypothesize that they might regulate the same target genes in an opposite fashion, and in different moment of the development. Our purpose, is to study their role in murine and principally in human erythroid systems.

MATERIALS AND METHODS

DNA microarray experiment

CD1 pregnant mice were sacrificed and fetal liver of embryos at the same stage of development (E11.5, E12.5 or E13.5) were taken and pooled. After differential osmotic lysis (RBC buffer pH 7.2-7.4: Nh_4Cl 150mM, $KHCO_3$ 1mM, Na_2EDTA 0,1mM) that eliminate the more mature erythroid cells, they were stained with biotinilated anti cKit antibody (streptavidin-FITC coniugated as secondary antibody) and anti TER119-PE antibody and sorted by FACS Calibur (all antibodies are from BD biosciences).

Total RNA was extract from the three sorted cell populations (cKit+/TER119-, cKit+/TER119+, cKit-/TER119++); hybridizable biotinilated RNA targets were producted using ENZO BioArray HighYield RNA transcript labelling kit (distributed by Affimetrix, Inc.) and analyzed by GeneChip Mouse Genome 430A (Affimetrix).

REFERENCES

Cohen-Barak, O., Z. Yi, N. Hagiwara, K. Monzen, I. Komuro, and M. H. Brilliant. 2003. Sox6 regulation of cardiac myocyte development. Nucleic Acids Res. **31**:5941-5948.

Connor, F., E. Wright, P. Denny, P. Koopman, and A. Ashworth. 1995. The Sry-related HMG box-containing gene Sox6 is expressed in the adult testis and developing nervous system of the mouse. Nucleic Acids Res. **23**:3365-3372.

de Crombrugghe, B., V. Lefebvre, R. R. Behringer, W. Bi, S. Murakami, and W. Huang. 2000. Transcriptional mechanisms of chondrocyte differentiation. Matrix Biol. **19**:389-394.

Delabesse, E., S. Ogilvy, M. A. Chapman, S. G. Piltz, B. Gottgens, and A. R. Green. 2005. Transcriptional regulation of the SCL locus: identification of an enhancer that targets the primitive erythroid lineage in vivo. Mol.Cell Biol. **25**:5215-5225.

Demaison, C., K. Parsley, G. Brouns, M. Scherr, K. Battmer, C. Kinnon, M. Grez, and A. J. Thrasher. 2002. High-level transduction and gene expression in hematopoietic repopulating cells using a human immunodeficiency [correction of imunodeficiency] virus type 1-based lentiviral vector containing an internal spleen focus forming virus promoter. Hum.Gene Ther. **13**:803-813.

Dumitriu, B., M. R. Patrick, J. P. Petschek, S. Cherukuri, U. Klingmuller, P. L. Fox, and V. Lefebvre. 2006. Sox6 cell-autonomously stimulates erythroid cell survival, proliferation, and terminal maturation and is thereby an important enhancer of definitive erythropoiesis during mouse development. Blood **108**:1198-1207.

Elbashir, S. M., J. Harborth, W. Lendeckel, A. Yalcin, K. Weber, and T. Tuschl. 2001. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. Nature **411**:494-498.

Farr, C. J., D. J. Easty, J. Ragoussis, J. Collignon, R. Lovell-Badge, and P. N. Goodfellow. 1993. Characterization and mapping of the human SOX4 gene. Mamm.Genome **4** :577-584.

Hagiwara, N., S. E. Klewer, R. A. Samson, D. T. Erickson, M. F. Lyon, and M. H. Brilliant. 2000. Sox6 is a candidate gene for p100H myopathy, heart block, and sudden neonatal death. Proc.Natl.Acad.Sci.U.S.A **97**:4180-4185.

Hagiwara, N., B. Ma, and A. Ly . 2005. Slow and fast fiber isoform gene expression is systematically altered in skeletal muscle of the Sox6 mutant, p100H. Dev.Dyn. **234**:301-311.

Hamada-Kanazawa, M., K. Ishikawa, K. Nomoto, T. Uozumi, Y. Kawai, M. Narahara, and M. Miyake. 2004. Sox6 overexpression causes cellular aggregation and the neuronal differentiation of P19 embryonic carcinoma cells in the absence of retinoic acid HAMADAKANAZAWA2004A. FEBS Lett. **560**:192-198.

Hamada-Kanazawa, M., K. Ishikawa, D. Ogawa, M. Kanai, Y. Kawai, M. Narahara, and M. Miyake. 2004. Suppression of Sox6 in P19 cells leads to failure of neuronal differentiation by retinoic acid and induces retinoic acid-dependent apoptosis HAMADAKANAZAWA2004. FEBS Lett. **577**:60-66.

Hamada-Kanazawa, M., K. Ishikawa, D. Ogawa, M. Kanai, Y. Kawai, M. Narahara, and M. Miyake. 2004. Suppression of Sox6 in P19 cells leads to failure of neuronal differentiation by retinoic acid and induces retinoic acid-dependent apoptosis. FEBS Lett. **577**:60-66.

Hamada-Kanazawa, M., K. Ishikawa, K. Nomoto, T. Uozumi, Y. Kawai, M. Narahara, and M. Miyake. 2004. Sox6 overexpression causes cellular aggregation and the neuronal differentiation of P19 embryonic carcinoma cells in the absence of retinoic acid. FEBS Lett. **560**:192-198.

Iacopetti, P., M. Michelini, I. Stuckmann, B. Oback, E. Aaku-Saraste, and W. B. Huttner. 1999. Expression of the antiproliferative gene TIS21 at the onset of neurogenesis identifies single neuroepithelial cells that switch from proliferative to neuron-generating division. Proc.Natl.Acad.Sci.U.S.A **96**:4639-4644.

Ikeda, T., H. Kawaguchi, S. Kamekura, N. Ogata, Y. Mori, K. Nakamura, S. Ikegawa, and U. I. Chung. 2005. Distinct roles of

Sox5, Sox6, and Sox9 in different stages of chondrogenic differentiation. J.Bone Miner.Metab **23**:337-340.

Janmey, P. A. and T. P. Stossel. 1986. Kinetics of actin monomer exchange at the slow growing ends of actin filaments and their relation to the elongation of filaments shortened by gelsolin. J.Muscle Res.Cell Motil. **7**:446-454.

Kingsley, P. D., J. Malik, R. L. Emerson, T. P. Bushnell, K. E. McGrath, L. A. Bloedorn, M. Bulger, and J. Palis. 2006. "Maturational" globin switching in primary primitive erythroid cells. Blood **107**:1665-1672.

Kocher, O., P. Cheresh, L. F. Brown, and S. W. Lee. 1995. Identification of a novel gene, selectively up-regulated in human carcinomas, using the differential display technique. Clin.Cancer Res. **1**:1209-1215.

Kocher, O., P. Cheresh, and S. W. Lee. 1996. Identification and partial characterization of a novel membrane-associated protein (MAP17) up-regulated in human carcinomas and modulating cell replication and tumor growth. Am.J.Pathol. **149**:493-500.

Kwiatkowski, D. J. 1999. Functions of gelsolin: motility, signaling, apoptosis, cancer. Curr.Opin.Cell Biol. **11**:103-108.

Lee, G., A. Lo, S. A. Short, T. J. Mankelow, F. Spring, S. F. Parsons, K. Yazdanbakhsh, N. Mohandas, D. J. Anstee, and J. A. Chasis. 2006. Targeted gene deletion demonstrates that the cell adhesion molecule ICAM-4 is critical for erythroblastic island formation. Blood **108**:2064-2071.

Lefebvre, V., P. Li, and B. de Crombrugghe. 1998. A new long form of Sox5 (L-Sox5), Sox6 and Sox9 are coexpressed in chondrogenesis and cooperatively activate the type II collagen gene. EMBO J. **17**:5718-5733.

Lefebvre, V., B. Dumitriu, A. Penzo-Mendez, Y. Han, and B. Pallavi. 2007. Control of cell fate and differentiation by Sry-related high-mobility-group box (Sox) transcription factors. Int.J.Biochem.Cell Biol. **39**:2195-2214.

Lozzio, B. B. and C. B. Lozzio . 1977. Properties of the K562 cell line derived from a patient with chronic myeloid leukemia. Int.J.Cancer **19**:136.

Ma, Z. Q., Z. Liu, E. S. Ngan, and S. Y. Tsai. 2001. Cdc25B functions as a novel coactivator for the steroid receptors. Mol.Cell Biol. **21**:8056-8067.

Maury, C. P., M. Liljestrom, G. Boysen, T. Tornroth, C. A. de la, and E. L. Nurmiaho-Lassila. 2000. Danish type gelsolin related amyloidosis: 654G-T mutation is associated with a disease pathogenetically and clinically similar to that caused by the 654G-A mutation (familial amyloidosis of the Finnish type). J.Clin.Pathol. **53**:95-99.

Maury, C. P., M. Liljestrom, S. Tiitinen, K. Laiho, K. Kaarela, and C. Ehnholm. 2001. Apolipoprotein E phenotypes in rheumatoid arthritis with or without amyloidosis. Amyloid. **8**:270-273.

Migliaccio, G., R. Di Pietro, G. di, V, A. Di Baldassarre, A. R. Migliaccio, L. Maccioni, R. Galanello, and T. Papayannopoulou. 2002. In vitro mass production of human erythroid cells from the blood of normal donors and of thalassemic patients. Blood Cells Mol.Dis. **28**:169-180.

Murakami, A., S. Ishida, J. Thurlow, J. M. Revest, and C. Dickson. 2001. SOX6 binds CtBP2 to repress transcription from the Fgf-3 promoter. Nucleic Acids Res. **29**:3347-3355.

Ohe, K., E. Lalli, and P. Sassone-Corsi. 2002. A direct role of SRY and SOX proteins in pre-mRNA splicing. Proc.Natl.Acad.Sci.U.S.A **99**:1146-1151.

Palis, J. 2008. Ontogeny of erythropoiesis. Curr.Opin.Hematol. **15**:155-161.

Papayannopoulou, T., T. Kalmantis, and G. Stamatoyannopoulos. 1979. Cellular regulation of hemoglobin switching: evidence for inverse relationship between fetal hemoglobin synthesis and degree of maturity of human erythroid cells. Proc.Natl.Acad.Sci.U.S.A **76**:6420-6424.

Raouf, A. and A. Seth. 2002. Discovery of osteoblast-associated genes using cDNA microarrays. Bone **30**:463-471.

Ronzoni, L., P. Bonara, D. Rusconi, C. Frugoni, I. Libani, and M. D. Cappellini. 2008. Erythroid differentiation and maturation from peripheral CD34+ cells in liquid culture: cellular and molecular characterization. Blood Cells Mol.Dis. **40**:148-155.

Rylski, M., J. J. Welch, Y. Y. Chen, D. L. Letting, J. A. Diehl, L. A. Chodosh, G. A. Blobel, and M. J. Weiss. 2003. GATA-1-mediated proliferation arrest during erythroid maturation. Mol.Cell Biol. **23**:5031-5042.

Schilham, M. W. and H. Clevers . 1998. HMG box containing transcription factors in lymphocyte differentiation. Semin.Immunol. **10**:127-132.

Sharp, P. A. 2001. RNA interference--2001. Genes Dev. **15**:485-490.

Silacci, P., L. Mazzolai, C. Gauci, N. Stergiopulos, H. L. Yin, and D. Hayoz. 2004. Gelsolin superfamily proteins: key regulators of cellular functions. Cell Mol.Life Sci. **61**:2614-2623.

Trimborn, T., J. Gribnau, F. Grosveld, and P. Fraser. 1999. Mechanisms of developmental control of transcription in the murine alpha- and beta-globin loci. Genes Dev. **13**:112-124.

Villeval, J. L., P. G. Pelicci, A. Tabilio, M. Titeux, A. Henri, F. Houesche, P. Thomopoulos, W. Vainchenker, M. Garbaz, H. Rochant, J. Breton-Gorius, P. A. Edwards, and U. Testa. 1983. Erythroid properties of K562 cells. Effect of hemin, butyrate and TPA induction. Exp.Cell Res. **146**:428-435.

Wegner, M. 1999. From head to toes: the multiple facets of Sox proteins. Nucleic Acids Res. **27**:1409-1420.

Wijgerde, M., F. Grosveld, and P. Fraser. 1995. Transcription complex stability and chromatin dynamics in vivo. Nature **377**:209-213.

Wilson, M. and P. Koopman. 2002. Matching SOX: partner proteins and co-factors of the SOX family of transcriptional regulators. Curr.Opin.Genet.Dev. **12**:441-446.

Wilson, M. E., K. Y. Yang, A. Kalousova, J. Lau, Y. Kosaka, F. C. Lynn, J. Wang, C. Mrejen, V. Episkopou, H. C. Clevers, and M. S. German. 2005. The HMG box transcription factor Sox4 contributes to the development of the endocrine pancreas. Diabetes **54**:3402-3409.

Witke, W., A. H. Sharpe, J. H. Hartwig, T. Azuma, T. P. Stossel, and D. J. Kwiatkowski. 1995. Hemostatic, inflammatory, and fibroblast responses are blunted in mice lacking gelsolin. Cell **81**:41-51.

Ya, J., M. W. Schilham, P. A. de Boer, A. F. Moorman, H. Clevers, and W. H. Lamers. 1998. Sox4-deficiency syndrome in mice is an animal model for common trunk. Circ.Res. **83**:986-994.

Yi, Z., O. Cohen-Barak, N. Hagiwara, P. D. Kingsley, D. A. Fuchs, D. T. Erickson, E. M. Epner, J. Palis, and M. H. Brilliant. 2006. Sox6 directly silences epsilon globin expression in definitive erythropoiesis. PLoS.Genet. **2**:e14.

Yin, H. L., D. J. Kwiatkowski, J. E. Mole, and F. S. Cole. 1984. Structure and biosynthesis of cytoplasmic and secreted variants of gelsolin. J.Biol.Chem. **259**:5271-5276.

Yin, H. L. and J. T. Stull. 1999. Proteins that regulate dynamic actin remodeling in response to membrane signaling minireview series. J.Biol.Chem. **274**:32529-32530.

Chapter3

SOX6 ENHANCES HUMAN ERYTHROID DIFFERENTIATION

(Manuscript in preparation)

SOX6 ENHANCES HUMAN ERYTHROID DIFFERENTIATION

Cantu' Claudio *, Ierardi Rosella ^, Alborelli Ilaria *, Fugazza Cristina [§], Grande Vito *, Bosè Francesca[#], Ottolenghi Sergio *, Ferrari Giuliana ^ °, Ronchi Antonella *.

* Dipartimento di Biotecnologie e Bioscienze, Università di Milano-Bicocca, Italy. ^ San Raffaele Telethon Institute for Gene Therapy (HSR-TIGET), San Raffaele Scientific Institute, 20132 Milan, Italy.
° Vita-Salute San Raffaele University, Milan, Italy. [§] John Radcliffe Hospital, MHU-Molecular Hematology Unit, Oxford, UK . [#] INGM-Fondazione Istituto Nazionale dio Genetica molecolare, via Francesco Sforza 28, 20122 Milano, Italy.

INTRODUCTION

Sox proteins are important transcriptional regulators of many developmental processes where they control the specification and differentiation of many cell types, by organizing local chromatin structures and assembling multiprotein complexes on DNA (Wolffe A.P. et at., 1994, Werner M.H. et al., 1997).

In particular, Sox6, originally isolated form adult mouse testis (Connor F., 1995) and required for the development of the central nervous system (Hamada-Kanazawa M., et al., 2004a, Hamada-Kanazawa M., et al., 2004b), for chondrogenesis (Lefebvre V., 2009, Han Y. And Lefebvre V., 2008, Ikeda T. et al., 2004, Smits P. et al.,

2001), and for cardiac and skeletal muscle formation (Hagiwara N., et al., 2000, Cohen-Barak O., 2001), also contributes to control definitive erythropoiesis, the process leading to the formation of mature RBCs (Dumitriu B. et al., 2006).

Primitive erythropoiesis originates in the yolk sac and produce large, nucleated red blood cells, while definitive erythropoiesis, producing small enucleated RBCs starts in fetal liver and moves perinatally to bone marrow, the major erythropoietic tissue in the adult life.

In definitive hematopoiesis, progenitors committed to erythropoiesis progressively differentiate into BF-UE () and CF-UE, that in turn, give rise to proerythroblasts, erythroblasts and finally to mature RBCs. These differentiation stages are accompanied by the progressive accumulation of erythroid-specific markers, including the globins chains required at the different stages of development.

In humans, globins are encoded by two gene clusters: the α locus, which contains the embryonic ζ gene and the two adult α genes and the β locus consisting of the ε embryonic, Gγ, Aγ fetal and the adult δ and β genes. Two globin switching occur in man: embryonic to fetal (ε to γ) and fetal to adult (γ to β), whereas other species (including mouse) undergo only one switch from embryonic to adult globins expression.

Several transcription factors are essential for erythroid commitment to occur and for differential globins gene expression during development: their absence is associated with a wide spectrum of phenotypes ranging from mild anemia to death due to a complete failure of erythropoiesis.

Among them, Sox6 has been recently shown to stimulate erythroid cell survival, proliferation and terminal maturation of definitive murine Erythropoiesis (Dumitriu B. et al. 2006). Accordingly, Sox6 null fetuses and pups mice are anemic and present defective RBCs. Moreover, Sox6 directly acts on globin genes, silencing the embryonic ey gene in definitive erythroid cells by binding to its promoter (Yi Z. Et al., 2006). Accordingly with the proposed role of εy silencing in adulthood, Sox6 null embryonic liver stem cells engrafted into lethally irradiated wild type adult mice show high levels of εy expression in bone marrow, spleen and blood (Cohen.Barak O. et al., 2007). All the above data derive from the analysis two mouse models in which Sox6 is absent due to a spontaneous mutation (Hagiwara N., 2000) or to a KO (Dumitriu B., 2006) and nothing is known so far on the positive effect of Sox6 and on its role in human erythropoiesis.

In this paper we explore the role of Sox6 in human erythropoiesis by its overexpression in both the erytholeukemic K562 cell line (Lozzio BB. and Lozzio CB., 1977) and in primary erythroid cultures from CD34+ cells purified from Cord Blood (Migliaccio G.et al., 2002). We demonstrate that Sox6 induces significantly erythroid differentiation in both models as shown by morphological, RT-PCR and FACS analysis.

Sox6 overexpression in K562 forces their terminal maturation: despite their erithroleukemic origin, K562 overexpressing Sox6 grow at a very low rate when compared with control cells, and die within about ten days after transduction. Upon Sox6 overexpression, several erythroid specific transcripts are greatly increased , i.e. mRNAs for enzymes controlling the heme-biosynthetic pathway, for transcription

factors and for all globins (although the ε/γ-globin ratio is decreased), suggesting that Sox6 acts as a general positive regulator of erythroid maturation and erythroid genes expression.

Accordingly to the phenotypic changes observed in K562 cells, Sox6 overexpression in primary cells is accompanied by an accelerated kinetic of maturation and in an increased number of cells that reach the final enucleation step.

EXPERIMENTAL PROCEDURES

SOX6 overexpressing vector

The Sox6 murine cDNA was kindly provided by Prof. Michiko Hamada-Kanazawa, Kobe-Gakuin University, Japan. The Sox6 cDNA was cloned in frame with a 3' FLAG epitope to generate a Sox6-FLAG cassette (S6F) in which Sox6 cDNA lacks the 49 C-ter aminoacids still retaining its biological proprieties (Hamada-Kanazawa et al., 2004). The Sox6-FLAG cassette (EcoRI – KpnI blunted sites) was then cloned immediately upstream to the IRES-Emerald GFP cassette (blunted BamHI site) of the pHR SIN BX IR/EMW (derived from pHR SIN CSGW, (Demaison C., et al., 2002)) lentiviral vector (Vec). The two packaging plasmids psPAX2 and pMD-VSVG were used to produce Lentiviral pseudo-particles in 293T cells (www.lentiweb.com). The expression level of the exogenous Sox6-FLAG was estimated in both K562 cells and in the primary Human Cord Blood cells (at day 10) by comparing the relative expression of either the exogenous (in S6F-transduced cells) and the endogenous (in cells transduced with the corresponding empty vector)

Sox6 transcripts with GAPDH expression, considered as internal standard.

Lentiviral harvesting protocol

Exponentially growing HEK 293T cells were transfected using Lipofectamine 2000 (Invitrogen) with the three vectors lentiviral system. About 48 h after transfection, the supernatant containing recombinant viruses was harvested, filtered (0.45 µm), centrifuged at 20.000g for 8 hours, and then frozen at −80 °C. Lentiviruses were titrated on HEK 293T cells by measuring the percentage of GFP positive cells by FACS analysis.

Cell cultures

K562 were cultured in RPMI medium supplemented by 10% Fetal Bovine Serum, PenStrep and L-glutamine. Transduction was performed overnight, by adding the vector stock at multiplicity of infection (MOI) 50, using a virus at a concentration of $5x10^8$ TU/ml.

$CD34^+$ cells were purified from cord blood by positive selection from MNC, using anti-CD34 microbeads (Miltenyi Biotech) according to the manufacturer's instructions. $CD34^+$ cells were plated at a concentration of $0.5\text{-}1x10^6$ cells/ml and prestimulated for 30 hours in CellGro medium (Cell Genix) supplemented with 300 ng/ml human stem cell factor (hSCF), 300 ng/ml human Flt-3-ligand (hFlt3-l), 100 ng/ml human thrombopoietin (TPO) and 60 ng/ml human IL-3 (hIL-3) (all PeproTech) on plates coated with retronectin (Takara Shuzo). Transduction was performed overnight, by adding the vector stock at multiplicity of infection (MOI) 100 and at a viral concentration of

$5x10^8$ TU/ml. The following day cells were washed and either grown in suspension as erythroid cultures or plated in methylcellulose medium for colony forming unit (CFU) assay. Erythroblasts (day 6) derived from cord blood $CD34^+$ cells were transduced overnight by addition of lentiviral vector at MOI 50 and at a viral concentration of $5x10^8$ TU/ml. The following day cells were washed and grown in suspension as erythroid culture.

$CD34^+$ cells were cultured for 2 weeks in StemSpan (Stem Cell Technologies) containing 20% of fetal bovine serum (FBS, Hyclone) and supplemented with hSCF (10 ng/ml), human erythropoietin (EPO, 1 U/ml), hIL-3 (1 ng/ml), 10^{-6} M dexamethasone (Sigma), and 10^{-6} M β-estradiol (Sigma). $CD34^+$ cells were seeded at a concentration of 10^5 cells/ml and diluted over time to maintain the concentration at 1–2 x 10^6 cells/ml. The progression toward erythroid differentiation was evaluated by staining with PE-conjugated anti-GpA antibody (DakoCytomation (DakoCytomation) and FACS analysis. Morphological analysis and differential counting was performed on cytospins by MayGrünwald-Giemsa staining and microscope inspection.

CFU assay for human progenitors

$CD34^+$ cells were washed with PBS and plated at a density of 1000 cells/ml in methylcellulose medium containing hSCF, hGM-CSF, hIL3 and hEPO (GF H4434; Stem Cell Technologies). After 2 weeks, BFU-E, CFU-GM, and CFU-GEMM colonies were counted, and single colonies (20–30 for each experiment) were isolated. DNA and

RNA were extracted by TRIzol REAGENT (Invitrogen) for molecular analysis.

Chromatin Immunoprecipitation (ChIP) assay

Briefly, 10^6 K562 cells for each Immunoprecipitation reaction were fixed with 0.4% formaldehyde for 10 minutes at room temperature, and chromatin was sonicated to a size of about 500 bp. Immunoprecipitation was performed after overnight incubation with anti-FLAG antibody (Sigma F-7425), and subsequent incubation with protein A agarose (Upstate biotechnology). Immunoprecipitated DNA was then analysed by amplifying an equivalent of 10^4 cells DNA with the following oligonucleotides:

Human ε promoter Fw: 5' GTTGCAGATAGATGAGGAGCC 3'

Human ε promoter Rw 5' GTCAAGGCTGACCTGTGTCC 3'.

GAPDH locus Fw: 5' CGGAGTCAACGGATTTGGTCGTAT 3'

GAPDH locus Rw: 5' AGCCTTCTCCATGGTGGTGAAGAC 3'

Immunoprecipitation was repeated 3 times on independent Chromatin preparations.

RNA isolation and RT-PCR

Total RNA from 10^5 cells (from K562 or differentiating human CD34+ cells at the different time points) was purified with TRI Reagent (Applied Biosystem AM9738), treated with RQ1 DNase (Promega) for 30 min at 37°C and retrotranscribed (High Capacity cDNA Reverse Transcription Kit, Applied Biosystem cat n°4368814). Negative control reactions (without Reverse Transcriptase) gave no signal. Real time analysis was performed using ABI Prism 7500, (PE

Applied Biosystems). Primers were designed to amplify 100 to 150bp amplicons on the basis of sequences from the Ensembl database (http://www.ensembl.org/Homo_sapiens/Info/Index). Samples from three or more independent experiments were analyzed in triplicate. Specific PCR product accumulation was monitored by SYBR Green dye fluorescence in 25-µl reaction volume. Dissociation curves confirmed the homogeneity of PCR products.

All primers used are listed in the table below:

Epo Fw	CTGTATCATGGACCACCTCGG
Epo Rw	TGAAGCACAGAAGCTCTTCGG
Jak2 Fw	ATCTGACCTTTCCATCTGGGG
Jak2 Rw	TGGTTGGGTGGATACCAGATC
Stat5A Fw	TTACTGAAGATCAAGCTGGGG
Stat5A Rw	TCATTGTACAGAATGTGCCGG
Stat5B Fw	CATTTTCCCATTGAGGTGCG
Stat5B Rw	GGGTGGCCTTAATGTTCTCC
GAPDH Fw	ACGGATTTGGTCGTATTGGG
GAPDH Rw	TGATTTTGGAGGGATCTCGC
ALAS-E Fw	CAACATCTCAGGCACCAGTA
ALAS-E Rw	CTCCACTGTTACGGATACCT
Fech Fw	ATCCAGCAGCTGGAGGGTCT
Fech Rw	TGAATCTTGGGGGTTCGGCG
GATA1 Fw	AGTTTGTGGATCCTGCTCTG
GATA1 Rw	GCAATGGGTACACCTGAAAG
GATA2 Fw	AGCGTCTCCAGCCTCATCTTCCGCG
GATA2 Rw	CGAGTCTTGCTGCGCCTGCTT
EKLF Fw	CCTGTTTGGTGGTCTCTTCACA
EKLF Rw	AGGGTCCATTCGTGGGAAA
NF-E2p45 Fw	AGTGTCAGCTCAGGCTCAGC
NF-E2p45 Rw	GCAGCTCGGTGATGGACATG

p53 Fw	CTGTCATCTTCTGTCCCTTC
p53 Rw	TGGAATCAACCCACAGCTGCA
p21 Fw	GTCACTGTCTTGTACCCTTGTG
p21 Rw	CGGCGTTTGGAGTGGTAGAAA
p27 Rw	AAGCGACCTGCAACCGACGATTCTT
p27 Fw	GCTCCACAGAACCGGCATTT
alpha globin Fw	GAGGCCCTGGAGAGGATGTTCC
alpha globin Rw	ACAGCGCGTTGGGCATGTCGTC
beta globin Fw	TACATTTGCTTCTGACACAAC
beta globin Rw	ACAGATCCCCAAAGGAC
gamma globin Fw	CTTCAAGCTCCTGGGAAATGT
gamma globin Rw	GCAGAATAAAGCCTACCTTGAAAG
epsilon globin Fw	GCCTGTGGAGCAAGATGAAT
epsilon globin Rw	GCGGGCTTGAGGTTGT

RESULTS

Sox6 overexpression in human erythroleukemic K562 cells

To get insight into the role of Sox6 in human erythropoiesis, we first overexpressed Sox6 by lentiviral transduction in K562 erythroleukemic cells.

K562 cells were transduced either with a vector containing a Sox6FLAG overexpressing cassette upstream to an IRES EGFP - element (S6F-GFP) or with the corresponding Empty Vector (Vec-GFP) (Fig.1, panel A).

The efficiency of transduction was similar (>95% of GFP-positive cells, not shown) for both vectors and the expression of the exogenous Sox6 was tested by Western blot (Fig.1 panel B). In K562, 48h after S6F-GFP transduction, we observed a marked reduction of proliferation when compared to that of Vec-GFP transduced cells. The K562 S6-GFP culture declined to exhaustion within day 9 after transduction (Fig.1, panel C).

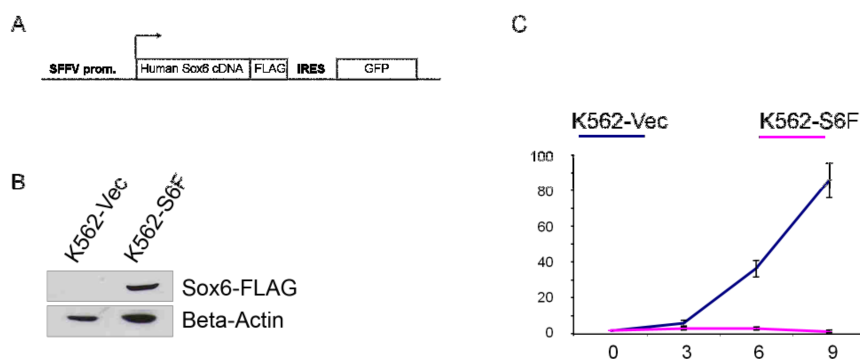


Figure 1: Growth Kinetic of K562 cells overexpressing Sox6.

The Sox6 overexpressing vector (A), managed to produce abundant exogenous protein in transduced K562 cells, as was assessed by Western Blot analysis using an anti-Flag antibody to detect the exogenous protein (B).
(C) 1x 10^6 exponentially growing K562 cells were transduced at day 0 either with the S6F or the corresponding empty vector (Vec) and cells were washed and replated in fresh medium 24h after transduction. S6F-transduced cells stop growing 3 days after transduction and the culture die within day 9. Error bars refer to 3 independent experiments.

Since Epo is the major cytokine required for survival, proliferation and differentiation of committed erythroid progenitor cells, we looked for Epo pathway activation in Sox6 overexpressing cells. Among genes whose expression we tested by RT-PCR, only one had a robust increase in its expression, SOCS3 (Fig 2). SOCS3 belongs to the family of Suppressor Of Cytokine Signaling proteins , and it is know to repress several pathways activated by the cytokines binding to their receptors. SOCS3 directly binds to the kinase domain of JAK2, thereby inhibiting its tyrosine-kinase activity (Nicola NA. and Greenhalgh CJ., 2000). The activation of SOCS3 by Sox6 its very early, since it happens just three hours after K562 transduction with Sox6 containing vector. Since it is already know the anti-proliferative potential of SOCS3 protein, and its protective role in tumors of different tissues (Haan S. et al., 2009; Rigby RJ. et al., 2007; Ogata H. et al., 2006a, Ogata H. et al., 2006b), its activation alone could be able to explain, almost in part, the stop in the K562 cells growth.

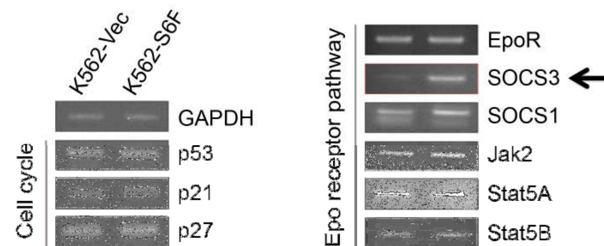


Figure 2: A semi-quantitative RT-PCR analysis, to evaluate the expression level of some genes involved in the erythroid progenitors cell cycle regulation. The black arrow indicates the only one we found whose expression level was significantly changed, SOCS3.

In these same cells, Sox6 overexpression is accompanied by a profound phenotypic change (Fig.2): K562-S6F transduced cells show a red pellet upon centrifugation, indicating an accumulation of haemoglobin (Fig.2, panel A), confirmed by the increased number of O-Dianisidine positive cells on cytospin preparations (Fig. 2, panel B). Moreover, FACS analysis on these cells reveals the emergence of a double positive $CD71^{+}$(Transferrine Receptor) $CD235^{++}$(Glycophorin A) cell population, representing a more mature erythroid population if compared to the Vec-GFP-transduced cells (Fig. 2, panel C).

To better characterize K562-S6F cells, we performed a semi quantitative RT-PCR analysis 72 h after transduction, comparing the expression of a series of genes known to vary during erythoid maturation between K562-Vec and K562-S6F cells. The result of this analysis is summarized in Fig 2D. As expected on the basis of the observed increased hemoglobinization, among genes whose

expression is strongly increased upon Sox6 overexpression there are globin genes (see below, detailed analysis), together with genes encoding key enzymes of the heme biosynthetic pathway (ALAS-E, FECH, PBGD). In these same cells, the appearance of a more mature erythroid phenotype, is mirrowed by a drop in the expression of the "megakaryocytic genes" GPIIB and GPIIIA, which are already expressed in control K562 cells.

Genes encoding the major transcription factors responsible for erythroid cell type specification, i.e. GATA1, GATA2, EKLF, and p45-NFE2, do not show significant variations, although only this latter shows a very small but reproducibly increase.

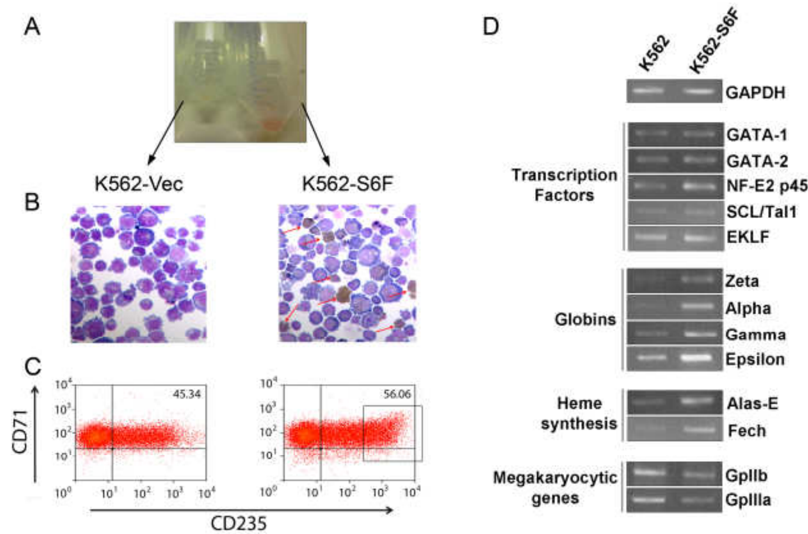


Figure 3: Phenotipic changes of K562 cells upon Sox6 overexpression.

S6F-transduced cells (left panels), when compared with cells transduced with the empty vector(Vec ,right panels), show: (A) a reddish pellet .(B) an increased number of O-Dianisidine positive cells (brown staining), indicating hemoglobin accumulation. (C) an emerging population of strongly CD71/GpA (CD235) double

positive cells (FACS analysis) (D) a change in their gene expression expression profile (RT-PCR analysis).

Since Sox6 has been proposed to specifically repress εy globin expression in mouse, we carefully analysed by Real time PCR the relative changes of globin genes transcription upon Sox6 overexpression, considering GAPDH as internal standard. As show Fig 3 both β-like genes (ε and γ) and α-like gene (ζ and α) normally expressed by K562 cells, are induced by Sox6 overexpression (Fig 4, panel A). Of interest, the γ/ε transcripts ratio varies from 2,5 in K562-Vec to 5,5 in K562-S6F (Fig. 4, panel B), although, in the observed context of general globins genes activation induced by Sox6, the level of ε globin gene expression results, in terms of absolute amount of transcript, increased. On this basis and because Sox6 has been shown to repress the εy globin gene in mouse by direct binding on its promoter, we performed a Chromatin Immunopecipitation experiment to ascertain whether Sox6 is able to bind to the human epsilon promoter although the Sox binding site seems to be not completely conserved between mouse and man, and in man, is quite different from the classical consensus for SOX proteins (Fig. 4, panel C). By the use of an anti-FLAG antibody that recognizes the transduced S6FLAG, we demonstrated that Sox6 is indeed able to bind to the human ε gene promoter (Fig. 4, panel D.).

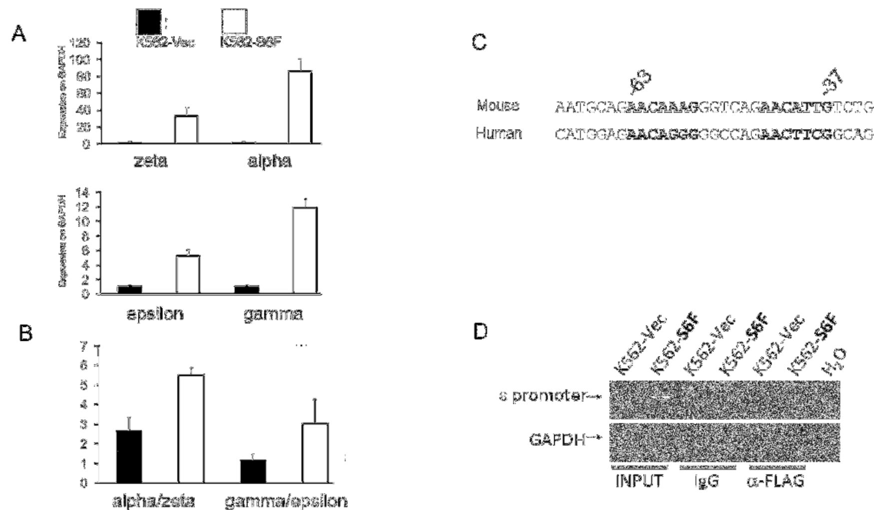


Figure 4: All Globins chains normally expressed in K562 are upregulated in K562-S6F-transduced cells

(A)Left panels: Real time PCR analysis on cDNAs from K562-S6F and K562-Vec cells. Histograms show the relative levels of expression (in one representative experiment) of β-like (ε and γ) and α-like (ζ and α) globin genes compared to GAPDH, considered as 1. Right panels: γ/ε and α/ζ ratio changes of expression in K562-Vec vs K562 S6F cells

(B) ChIP analysis confirms the ability of Sox6 to bind the human ε globin promoter Upper panel: sequence conservation of the double Sox6 binding sites between mouse and human promoters. Lower Panel: the anti-FLAG or rabbit IgG antibodies were used to immunoprecipitate chromatin from K562-Vec or K562-S6F cells and the region containing the double Sox6 binding site was amplified. Upper panel: primers designed on the ε-globin promoter region. Lower panel: primers on the GAPDH gene as a negative control

Together these data suggest that Sox6 overexpression strongly induces erythroid differentiation, as revealed by the increaded hemoglobinization (O-Dianosideine positivity), the accumulation of specific surface markers (FACS analysis), and by a general increase

in the expression of erythroid specific genes (RT-PCR). Moreover, all globins genes normally expressed in K562 are overexpressed upon Sox6 transduction, although the increase of ε-globin is marginal if compared to the others. Finally, the observed enhanced erythroid differentiation is accompanied by a dramatic reduction of cell proliferation.

Sox6 overexpression in primary human erythroid progenitors

Because of the profound effect induced by Sox6 overexpression in K562 cells, we moved to a better model of human erythroid differentiation, i.e. primary CD34+ cells purified for cord blood, expanded in vitro and induced to differentiate in liquid culture.

Real time PCR analysis on these cultures (Fig. 4) shows that Sox6 expression is absent in the most immature CD34+ progenitors (day 0 to day 6), starts to be detectable at the beginning of the erythroblasts differentiation phase (day 8, 72% GlicophorinA positive cells), reaches a peak around day 12 of culture (85% GlicophorinA positive cells), and finally decline at the end of the culture, when a maximum of Gpa positivity (nearly 90%) is reached (Fig. 5, panel A). On the basis of this expression pattern, we transduced the culture at day 6, corresponding to the end of the erythroblasts expansion phase and immediately prior to the onset of Sox6 expression, as is outlined in the table in figure 5B.

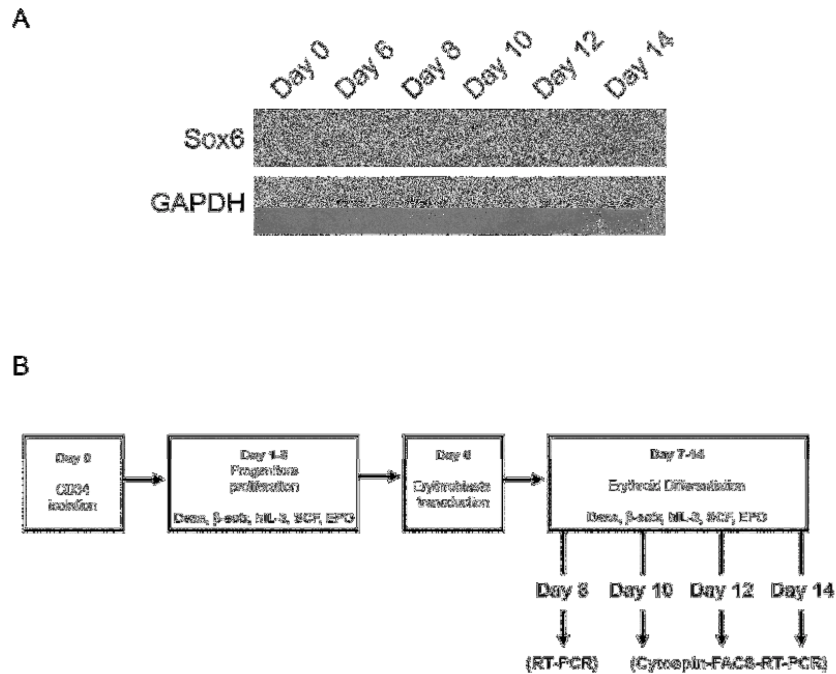


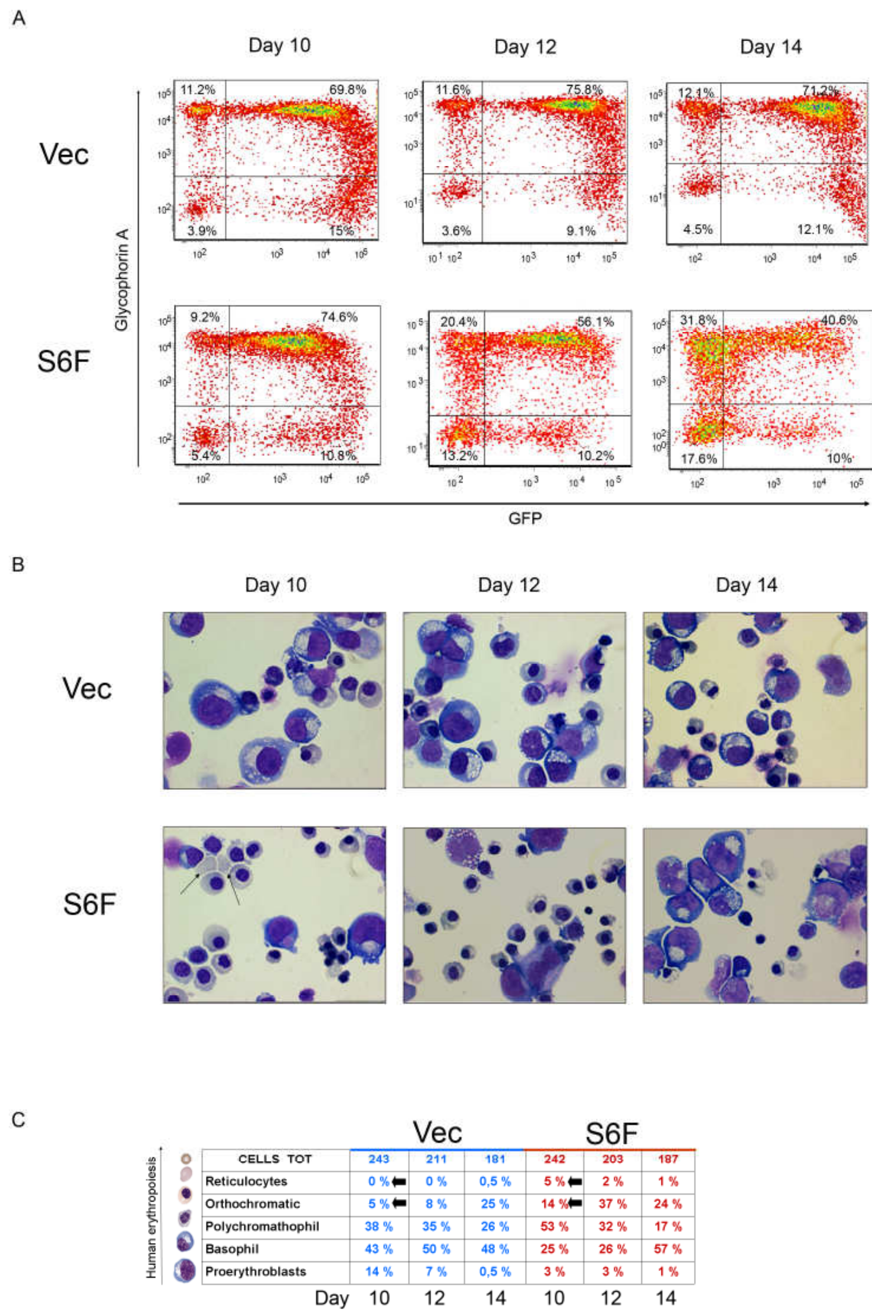
Figure 5: (A) Semi quantitative RT-PCR on erythroid cultures from CB at the different days of culture indicated above the figure. Upper panel: Sox6 expression; Lower panel: GAPDH. The progression toward erythroid differentiation, evident from the increasing percentage of GpA positivity at the different days (FACS analysis) is indicated below the figure.
(B) Experimental outline. The cytokines used in the medium are indicated.

The evaluation of the percentage of GFP positive cells by FACS analysis, show that cells are transduced with similar efficiency by both constructs (Vec, empty vector: 84,9 %; S6F, Sox6 overexpressing vector: 85,4%). The analysis was subsequently carried out on samples taken at days 10, 12 and 14, according to the flow chart on fig. and data from a representative culture are shown in the figures below.

We first followed erythroid maturation of the culture by evaluating the number of GFP and Glycophorin A (CD235) double positive cells by FACS analysis (Fig. 6A). In a S6F-transduced representative culture, this value reaches its peak at day 10 (74,6%) and then declines to 56.1% at day 12 and to 40.6% at day14, suggesting a progressive loss of transduced cells. In contrast, in the empty vector transduced culture, $GPA^{+}GFP^{+}$ double positive population increases from day 10 (69.9%) to day 12 (75.8%) and finally slightly decreases to 71.2% at day14.

To better correlate these data with erythroid maturation, cells from the same samples as above, were cytospun and differential cells counts were performed to score the relative number of cells at different stages of erythroid maturation. At day 10, in S6F-transduced cells, many polychromathophil and orthocromatic erythroblasts, usually appearing at later days in control cultures, are already present (Fig. 6B and cells count in table 6C). More strickingly, 3 to 5% (depending on the experiment) of reticulocytes are scored, while in the control culture a maximum of 0,5-1% of reticulocytes is observed at the end of the culture (day 14). In parallel with the increased number of more mature cells (poly,ortho,ret), a decrease of more immature (basophils + proeryhtroblasts) cells is observed in Sox6 overexpressing cells (28% versus 57%). The prevalence of more mature cells with respect to the control is further maintained at day 12.

Finally, the distribution of the different cell types between the two cultures returns equal at day 14, when the pool of GFP+ Sox6 overexpressing cells reaches the minimum and thus presumably fails to significantly contribute to the culture.



	Vec			S6F		
CELLS TOT	243	211	181	242	203	187
Reticulocytes	0 % ⬅	0 %	0,5 %	5 % ⬅	2 %	1 %
Orthochromatic	5 % ⬅	8 %	25 %	14 % ⬅	37 %	24 %
Polychromathophil	38 %	35 %	26 %	53 %	32 %	17 %
Basophil	43 %	50 %	48 %	25 %	26 %	57 %
Proerythroblasts	14 %	7 %	0,5 %	3 %	3 %	1 %
Day	10	12	14	10	12	14

Figure 6: Sox6 enhances erythroid differentiation in Cord Blood-derived cell cultures.

(A) FACS analysis on cells transduced either with the Empty Vector (Vec) or with the Sox6 overexpressing vector (Sox6): x axis GFP expression; Y axis GPA expression

(B) MayGrünwald-Giemsa on cytospin preparations of the same samples as above

C) Differential counts on cells from the same cytospins as in D. More than 200 cells were scored for each samples.

We then analyzed the expression of globin genes at the same time points by Real time PCR (Fig 7). At day 10, α, γ and β globin expression is strongly increased in S6F-transduced cells, when compared to the control culture, whereas ε expression, although barely detectable in these perinatal "adult type" cultures (note the scale on the y axis), does not change significantly, suggesting a repressive role of Sox6 on the human ε-globin transcription (Fig. 7 and Fig. 4D).

At day 12, the globins genes expression become similar in the two culture. Finally, at day 14, (when the Vec-transduced cells reach the peak of erythroid differentiation) whereas the control culture maintains a plateau level of globins expression, the S6F colture has already declined in both GpA and GFP double positivity and globin expression levels.

Taken together these data suggest that Sox6 overexpression enhances (increased number of reticulocytes) and anticipates (differentiation peak at day10 instead of 14) the normal erythroid differentiation kinetic of the culture.

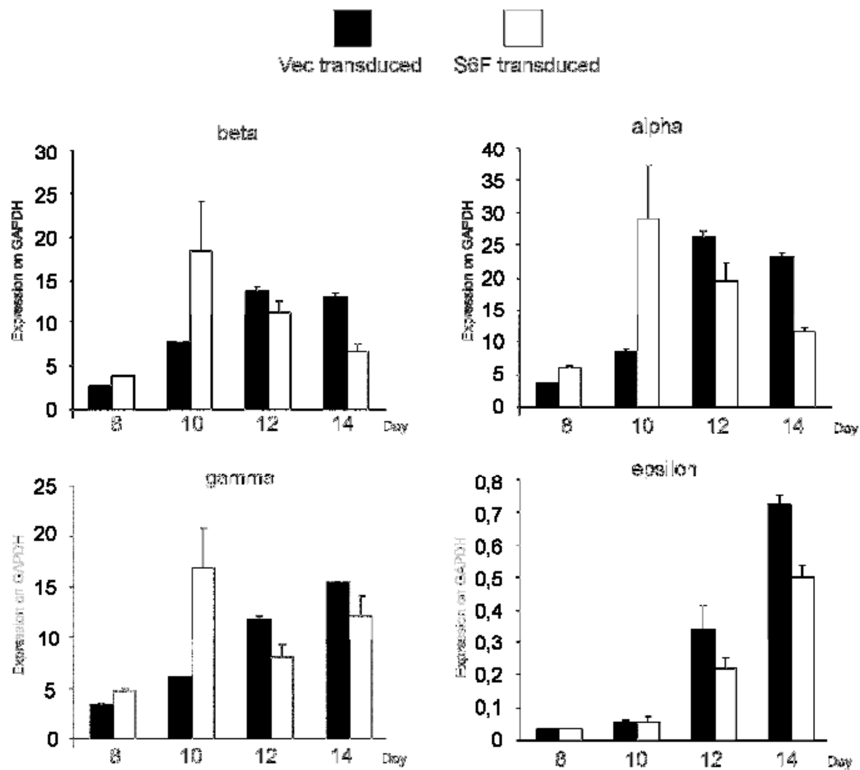


Figure 7: Histograms represent the quantitation by Real time RCR of globin mRNAs, obtained at days 10, 12, 14 of the culture. Sox6 anticipates the expression of the alpha, beta and gamma, but not of the epsilon globin gene.

In a second set of experiments we directly transduced $CD34^+$ cells with either the empty or the S6F vectors as above, and we subsequently analysed the transduced cultures, starting from day 6, to day14. However, while in the case of the empty vector the percentage of GFP positive cells is about 90% at day 6 and this value remains constant in the following days during the following erythroid maturation, in the S6F-tranduced culture the number of GFP positive cells is about 25% at day 6 and progressively decline to 3% by day14.

The same transduced cells were also seeded in parallel in methylcellulose. In this clonigenic assay we observed a reduced number of BFU-E in the S6Ftransduced cells, in parallel with a reduced number of total CFU observed, and of the percentage of GFP positive colonies (scored by fluorescence microscope observation). PCR analysis was carried out to detect the presence of the integrated lentivirus on DNA extracted from a representative number of colonies at the end of the culture and confirmed the percentage of GFP positive scored by visual observation.

Together these data suggest that S6 overexpression at the CD34+ stage decrease progenitors survival leading to a depletion of their pool.

The search of Sox6 targets on DNA

The effect of Sox6 overexpression observed in both K562 an in primary $CD34^+$ progenitor cells, prompted us to identify the pathways responsible of the enhanced erythropoiesis upon Sox6 induction. Since Sox6 is a transcription factor, we looked for its hypothetical target genes using an *in-silico* approach.

We took advantage of the web tool TFBS cluster (Transcription Factor Binding Site - http://hscl.cimr.cam.ac.uk/TFBScluster -) which allows to identify conserved patterns of binding sites present in evolutionary conserved regulatory regions in the mammalian genomes (Donaldson, I. J. and Gottgens, B., 2007).

The only known erythroid Sox6 target site is the εy globin promoter in mouse, and it is composed by two Sox consensi, with an opposite orientation, and spaced by 8 nucletides (Fig. 8A). We use this

sequence as model to find similar consensi through the human genome: two binding sites, possibly with an opposite orientation, but very close each other (in a cluster of 30-35 nucleotides). The software found 875 double Sox consensus sequences, in mouse-man conserved regions, interspearsed in the human genome. Only a minority of them (6%) were found near the transcription start site of known genes, while the others were located in introns (56%), or downstream the 3' end of known genes (38%) (Fig 8B). Of course, every single target found is of potential interest, but we started to analyse the 56 consensi found within 10 Kbp in 5'position of a known gene, where generally reside most of the regulatory sequences that regulates the transcription of a gene. Since our idea was to study genes potentially regulated by Sox6, which are important for erythroid differentiation, to further narrow the range of our research, we matched these 56 target genes potentially regulated by Sox6, with the data of our microarray analysis discussed in chapter 2.

This approach revealed us 7 genes of extreme interest: among them, interestingly, we found SOCS3 and Sox6 itself (Fig. 8C).

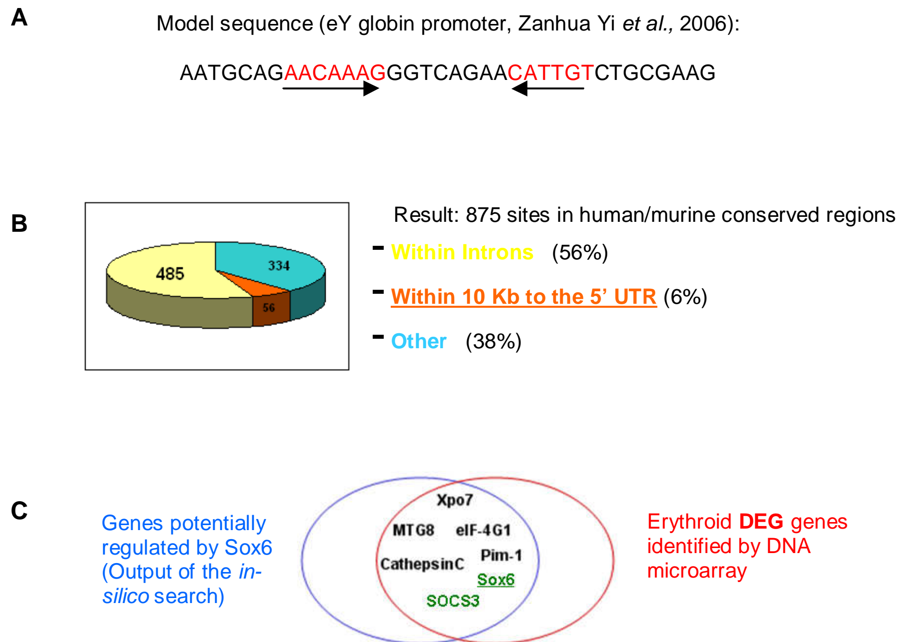


Figure 8: (A) The Sox6 binding sites on the εy globin promoter, used as a model sequences in our in-silico search. (B) The output of the search using TFBScluster. (C) Genes lying near potential Sox6 targets sites, merged with DEGs of erythroid system (see microarray data in chapter 2).

For both SOCS3 and Sox6, the double Sox consensus is very close to the transcription start site, and within an highly conserved region among mammalian genomes (Fig. 9).

In the case of SOCS3, this region is 2.7 Kbp upstream to the promoter, while in the case of Sox6 itself, the consensus resides in the hypothetical promoter itself. We deeply studied this region, and its regulation by Sox6, as I'll show in the following chapter.

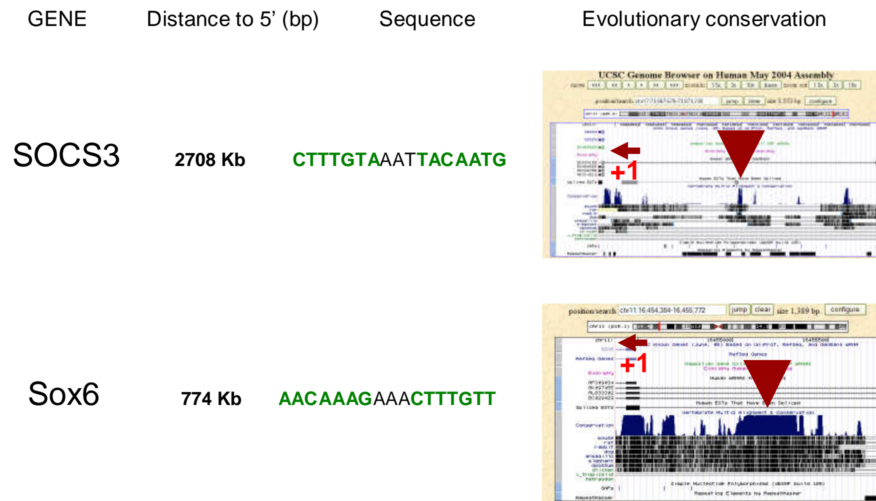


Figure 9: The distance from the transcription start site (+1) and evolutionary conservation of the two consensi we focused on, within SOCS3 and Sox6 itself loci, are schematized. On the right a cartoon taken from UCSC database (http://genome.ucsc.edu/): the transcription is from the right to the left (red arrows), while the position of the binding sites found is indicated by the red arrowheads.

SOCS3 expression, as shown in figure 2, is highly upregulated upon Sox6 overexpression, 3 hours already after Sox6 transduction. This observation disclose the possibility that SOCS3 is a really early Sox6 target gene. Of course the direct transcriptional regulation of SOCS3 by Sox6 needs to be validated from further studies.

SOCS3 increased expression might explain, at least in part, the consequences of Sox6 overexpression: the block in the progression through the cell cycle and in the growth of K562 cell line, and in human primary erythroblasts. Its role might be to withdraw the

progenitor cells from the proliferation phase, leading them toward the final stages of erythtoid differentiation.

DISCUSSION AND FUTURE AIMS

Sox6 enhances erythoid maturation

In this chapter we described the phenotypic changes occuring upon Sox6 overexpression in human cells.

The human K562 erythroleukemic cell line is a well described human model, and grows in culture indefinitely. The overexpression of Sox6 alone blocks its progression through the cell cycle, stopping its growth and leading to the culture exhaustion 10 days after Sox6 transduction. This strong effect on the cell divisions, is accompanied by a partial erythroid differentiation of these cells. While K562 express normally low levels of globin transcripts, Sox6 overexpression upregulates them: the pellet of these cells after centrifugation is red by hemoglobin accumulation, and several benzidine-positive cells appear. This is a very impressing result, considering the erythroleukemic nature of this cell line, generally capable of differentiating toward an erythroid destiny, only by the addition of some chemical inducers such as the Hemin (an Heme molecule precursor) in high concentration. This differentiation toward erythroid maturation is very clear also in a human primary culture of erythroid progenitors, extracted and isolated from the cord blood. Sox6 overexpression enhances and accelerate the terminal maturation of this culture, as clearly shown by differential cell counting: more mature orthochromatic cells, and reticulocytes appear in Sox6 transduced

culture compared with the control culture (transduced with the empty vector), accompanied with higher globins expression levels. While the normal control culture reaches its peak of erythroid marker expression, GlycophorinA, at the end of the culture at day 14, Sox6 transduced one shows an anticipate and clearly more intense peak at day 10. The peak of maturation at day 10 in Sox6 transduced culture, is also evident by the analysis on globins' transcripts, which are highly stimulated if compared with control culture starting from day 10.

Sox6 starts to be expressed in erythroid progenitors, and its expression rises while cells are maturing. Its role in the erythroid differentiation may be the activation of different downstream erythroid specific genes, and the downregulation of other genes that need to be switched off. Anticipating, and augmenting its expression, we obtained an enhancement and an anticipation of the differentiation, demonstrating its capability to regulate (almost in part) different aspects of erythroid maturation.

Sox6 molecular mechanism of action

During erythroid differentiation, the terminal maturation is strictly linked to the cell cycle withdrawal. Sox6 might regulates this two connected processes, since its expression level seems critical for their balance. In fact, no changes in the expression levels of other major regulators of erythropoiesis (GATA1, GATA2, EKLF, NF-E2p45, SCL/Tal1) are present, suggesting the possibility that Sox6 acts on other pathways.

One hypothesis is the interference by Sox6 in the Epo-Stat5 signalling pathway. Epo is one of the major cytokine regulating the erythroid

system, allowing the survival of erythroid committed progenitors and protecting them from apoptosis. After erythopoietin binding to its receptor (EpoR), the signal is transduced through a phosphorilation cascade to Stat5A and Stat5B (Signal Transduction and Activation of transcription), that dimerize and translocate to the nucleus, where they activate several target genes. Epo sustains the survival of erythroid progenitors, while in the last stages of maturation, erythroid committed progenitors must undergo only few cell cycle divisions, before completing the maturation process.

3 hours after Sox6 transduction, the transcription of SOCS3 gene (Suppressor Of Cytokine Signalling) is strongly increased, suggesting it as an early Sox6 target. SOCS3 is indeed able to block the Epo-Stat5 signalling transduction, by binding Janus Kinase 2 (Jak2), preventing its phosphorilation and the further Stat5 activation (Ingley E. Et al., 2004).

The direct activation of SOCS3 by Sox6 might be an exhaustive explanation of the observation that Sox6, if transduced in high not-committed progenitors leads to their premature death, but it is able to pushes the already committed progenitors toward terminal maturation.

Sox6 target genes and future aims

To identify which are the genes Sox6 is able to directly regulate, we performed a bioinformatic analysis using TFBS cluster web software, that allowed us to find potentially regulatory regions bearing highly conserved Sox6 putative binding sites, within evolutionary conserved regions between murine and human genomes. We used as model the

only know erythroid Sox6 binding site published, that resides in the murine εy-globin promoter, and is composed by two adjacent sox consensi, spaced by 8 nucleotide, and with an opposite orientation.
Interestingly, among the great number of data this analysis gave us, we could identify a very well conserved double Sox consensus within a conserved regulatory region, 2.7 Kbp upstream the transcription start site of the SOCS3 gene. This genomic region includes two very close Sox consensi, spaced by 4 nucleotides, and with an opposite orientation. It greatly resembles the Sox6 consensus found in the εy-globin promoter.
A direct SOCS3 transcriptional regulation by Sox6, comes out as a very likely hypothesis. To demonstrate the capability of Sox6 to bind this region in vitro, we are setting up an EMSA (Electrophoretic mobility shift assay) experiment, in which a radio labelled probe, corresponding to the double Sox consensus upstream the SOCS3 gene will be incubated with a protein extract containing Sox6: the binding between Sox6 and the probe, would be visualized as a shift in the migration of the radio labelled probe, in a gel electrophoresis run. Then, to demonstrate Sox6 binding in vivo on this region, a ChIP (Chromatin Immuno-Precipitation) will be performed. The whole chromatin, is fixated with formaldehyde, fragmented by sonication, and immunoprecipitated by an anti-SOX6 specific antibody: the immunoprecipiation of this specific genomic region, will be tested by PCR amplification using sequence-specific primers.
All these data, with our observation described in this chapter, would provide a simple but clear molecular mechanism by which a single gene, Sox6, is able to control and determine the fate of a cell, easing

out the cell from the cell cycle, and switching on a lineage specific genetic program of transcription.

REFERENCES

Cohen-Barak, O., N. Hagiwara, M. F. Arlt, J. P. Horton, and M. H. Brilliant. 2001. Cloning, characterization and chromosome mapping of the human SOX6 gene. Gene **265**:157-164.

Cohen-Barak, O., D. T. Erickson, M. S. Badowski, D. A. Fuchs, C. L. Klassen, D. T. Harris, and M. H. Brilliant. 2007. Stem cell transplantation demonstrates that Sox6 represses epsilon y globin expression in definitive erythropoiesis of adult mice. Exp.Hematol. **35**:358-367.

Connor, F., E. Wright, P. Denny, P. Koopman, and A. Ashworth. 1995. The Sry-related HMG box-containing gene Sox6 is expressed in the adult testis and developing nervous system of the mouse. Nucleic Acids Res. **23**:3365-3372.

Demaison, C., K. Parsley, G. Brouns, M. Scherr, K. Battmer, C. Kinnon, M. Grez, and A. J. Thrasher. 2002. High-level transduction and gene expression in hematopoietic repopulating cells using a human immunodeficiency [correction of imunodeficiency] virus type 1-based lentiviral vector containing an internal spleen focus forming virus promoter. Hum.Gene Ther. **13**:803-813.

Donaldson, I. J. and B. Gottgens. 2007. CoMoDis: composite motif discovery in mammalian genomes. Nucleic Acids Res. **35**:e1.

Dumitriu, B., M. R. Patrick, J. P. Petschek, S. Cherukuri, U. Klingmuller, P. L. Fox, and V. Lefebvre. 2006. Sox6 cell-autonomously stimulates erythroid cell survival, proliferation, and terminal maturation and is thereby an important enhancer of definitive erythropoiesis during mouse development. Blood **108**:1198-1207.

Haan, S., S. Wuller, J. Kaczor, C. Rolvering, T. Nocker, I. Behrmann, and C. Haan. 2009. SOCS-mediated downregulation of mutant Jak2 (V617F, T875N and K539L) counteracts cytokine-independent signaling. Oncogene **28**:3069-3080.

Hagiwara, N., B. Ma, and A. Ly. 2005. Slow and fast fiber isoform gene expression is systematically altered in skeletal muscle of the Sox6 mutant, p100H. Dev.Dyn. **234**:301-311.

Hamada-Kanazawa, M., K. Ishikawa, K. Nomoto, T. Uozumi, Y. Kawai, M. Narahara, and M. Miyake. 2004. Sox6 overexpression causes cellular aggregation and the neuronal differentiation of P19 embryonic carcinoma cells in the absence of retinoic acid. FEBS Lett. **560**:192-198.

Hamada-Kanazawa, M., K. Ishikawa, D. Ogawa, M. Kanai, Y. Kawai, M. Narahara, and M. Miyake. 2004. Suppression of Sox6 in P19 cells leads to failure of neuronal differentiation by retinoic acid and induces retinoic acid-dependent apoptosis. FEBS Lett. **577**:60-66.

Han, Y. and V. Lefebvre. 2008. L-Sox5 and Sox6 drive expression of the aggrecan gene in cartilage by securing binding of Sox9 to a far-upstream enhancer. Mol.Cell Biol. **28**:4999-5013.

Ikeda, T., S. Kamekura, A. Mabuchi, I. Kou, S. Seki, T. Takato, K. Nakamura, H. Kawaguchi, S. Ikegawa, and U. I. Chung. 2004. The combination of SOX5, SOX6, and SOX9 (the SOX trio) provides signals sufficient for induction of permanent cartilage. Arthritis Rheum. **50**:3561-3573.

Ingley, E., P. A. Tilbrook, and S. P. Klinken. 2004. New insights into the regulation of erythroid cells. IUBMB.Life **56**:177-184.

Lefebvre, V. 2009. The SoxD transcription factors - Sox5, Sox6, and Sox13 - are key cell fate modulators. Int.J.Biochem.Cell Biol.

Lozzio, B. B. and C. B. Lozzio . 1977. Properties of the K562 cell line derived from a patient with chronic myeloid leukemia. Int.J.Cancer **19**:136.

Migliaccio, G., R. Di Pietro, G. di, V, A. Di Baldassarre, A. R. Migliaccio, L. Maccioni, R. Galanello, and T. Papayannopoulou. 2002. In vitro mass production of human erythroid cells from the blood of normal donors and of thalassemic patients. Blood Cells Mol.Dis. **28**:169-180.

Nicola, N. A. and C. J. Greenhalgh. 2000. The suppressors of cytokine signaling (SOCS) proteins: important feedback inhibitors of cytokine action. Exp.Hematol. **28**:1105-1112.

Ogata, H., T. Chinen, T. Yoshida, I. Kinjyo, G. Takaesu, H. Shiraishi, M. Iida, T. Kobayashi, and A. Yoshimura . 2006. Loss of SOCS3 in the liver promotes fibrosis by enhancing STAT3-mediated TGF-beta1 production. Oncogene **25**:2520-2530.

Ogata, H., T. Kobayashi, T. Chinen, H. Takaki, T. Sanada, Y. Minoda, K. Koga, G. Takaesu, Y. Maehara, M. Iida, and A. Yoshimura. 2006. Deletion of the SOCS3 gene in liver parenchymal cells promotes hepatitis-induced hepatocarcinogenesis. Gastroenterology **131**:179-193.

Rigby, R. J., J. G. Simmons, C. J. Greenhalgh, W. S. Alexander, and P. K. Lund. 2007. Suppressor of cytokine signaling 3 (SOCS3) limits damage-induced crypt hyper-proliferation and inflammation-associated tumorigenesis in the colon. Oncogene **26** :4833-4841.

Smits, P., P. Li, J. Mandel, Z. Zhang, J. M. Deng, R. R. Behringer, B. de Crombrugghe, and V. Lefebvre. 2001. The transcription factors L-Sox5 and Sox6 are essential for cartilage formation. Dev.Cell **1**:277-290.

Werner, M. H. and S. K. Burley . 1997. Architectural transcription factors: proteins that remodel DNA. Cell **88**:733-736.

Wolffe, A. P. 1994. Architectural transcription factors. Science **264**:1100-1101.

Yi, Z., O. Cohen-Barak, N. Hagiwara, P. D. Kingsley, D. A. Fuchs, D. T. Erickson, E. M. Epner, J. Palis, and M. H. Brilliant. 2006. Sox6 directly silences epsilon globin expression in definitive erythropoiesis. PLoS.Genet. **2**:e14.

Chapter 4

AUTOREGULATION OF SOX6 EXPRESSION DURING HUMAN ERYTHROID DIFFERENTIATION

(Manuscript submitted to Journal of Biological Chemistry)

AUTOREGULATION OF SOX6 EXPRESSION DURING HUMAN ERYTHROID DIFFERENTIATION

Claudio Cantu' [1], Vito Grande [1], Letizia Cassinelli [1], Ilaria Alborelli [1], Rossella Ierardi [2], Luisa Ronzoni [3], Maria Domenica Cappellini [3], Giuliana Ferrari [2,4], Sergio Ottolenghi [1], Antonella Ronchi[1]

From Dipartimento di Biotecnologie e Bioscienze, Università di Milano-Bicocca, Italy[1]. San Raffaele Telethon Institute for Gene Therapy (HSR-TIGET), San Raffaele Scientific Institute, 20132 Milan, Italy[2]. Dipartimento di Medicina Interna, Università di Milano, Fondazione Policlinico Mangiagalli, Regina Elena, IRCCS, Milano, Italy[3]. Vita-Salute San Raffaele University, Milan, Italy[4]

Running head: Sox6 autoregulation during erythroid differentiation

Address correspondence to: Antonella Ronchi, Dipartimento di Biotecnologie e Bioscienze, Università di Milano-Bicocca, P.za della Scienza 2 , 20126, Italy. Phone: +39-02-6448-3337, Fax: +39-02-6448-3314. E-mail: antonella.ronchi@unimib.it

ABSTRACT

Sox6 is a member of the Sox family of transcription factors, which play critical roles in determining cell fate and differentiation. Sox6 has been demonstrated to control multiple aspects of definitive erythropoiesis in mouse: Sox6 deficiency is associated to anemia, poor maturation of erythroid cells and increased levels of εy-globin in adult cells.

In this work we show that, in human CD34+-derived erythroid cultures and in mouse bone marrow cell populations, Sox6 expression peaks at the erythroblast stage and then decreases along with erythroid differentiation.

The overexpression of exogenous Sox6 by lentiviral transduction in both human primary hematopoietic cells and in the human erythoid K562 cell line, results in a marked reduction of the endogenous transcript, suggesting a negative feedback regulation of Sox6 expression. The downregulation of the endogenous Sox6 gene is mediated by a double evolutionarily conserved Sox6 binding site within the Sox6 promoter. This sequence is directly bound *in vitro* (Electrophoresis Mobility Shift Assay, EMSA) and *in vivo* (Chromatin immunoprecipitation, ChIP) by Sox6, that represses its own transcription. We propose that this negative autoregulation feedback might be relevant to control the downregulation of Sox6 expression observed in late erythroid maturation.

INTRODUCTION

Sox6 is a member of the Sox (Sry-type HMG box) family of transcription factors, characterized by the presence of an HMG box, a domain that recognizes the minor groove on DNA. The binding of Sox proteins to DNA forces it to bend at about 75°, introducing local conformational changes. The ability of Sox proteins to bind in close proximity to other transcription factors and to distort DNA suggests that they can act as "architectural proteins", possibly by promoting the assembly of biologically active multiprotein complexes. In turn, modulation of interactions between distant chromatin domains has been proposed to play a crucial role in gene regulation by bringing together promoter/enhancer regions. According to this general model, Sox6 has been suggested to act both as an activator or as a repressor, depending on its interactions and on its target sequences (1, 2, 3).

The Sox domain recognizes a very degenerate (A/T)(A/T)CAA(A/T)G consensus, making very difficult to identify its *in vivo* targets: the best characterized and validated Sox6 target sequence on the regulatory elements of the chondrocyte gene Col2a1 is, for example, composed by four sites each having different mismatches relative to the HMG box consensus (4). Moreover, the presence of adjacent pairs of Sox sites on different known targets, suggests that double Sox sites might likely be the preferred Sox6 targets, although the relative arrangement and orientation of the two sites is not yet clearly defined (1, 2, 3 and refs ref. therein).

Murine Sox6 null mutants (p100) show delayed growth, myopathy, arterioventricular heart block and die within 2 weeks after birth (5).

Sox6 is indeed required for proper formation of heart, nervous system (5, 6, 7), cartilage (4, 8), cardiac and skeletal muscle (9, 10). Recent reports indicate that complete Sox6 ablation causes a perturbation of erythropoiesis resulting in the presence of increased numbers of nucleated and misshaped red cells in the fetal circulation and in a strong relative increase of embryonic (εy) globin gene expression (11, 12, 13). In particular, Sox6 directly silences εy globin expression in murine definitive erythropoiesis by binding to a double Sox6 site lying within a 36bp region on the εy proximal promoter (13). Moreover, embryonic liver stem cells from Sox6 null mice engrafted into lethally irradiated wt adult mice, show levels of εy expression in spleen and bone marrow higher than those observed in control mice transplanted with wild type cells (12).

The emerging critical role of Sox6 in erythropoiesis prompted us to study in more detail its expression and regulation in human erythropoiesis using, as model systems, primary CD34+ cells undergoing in vitro erythroid differentiation and the human erythroleukemic cell line K562.

We demonstrate that Sox6 expression peaks in early erythroblasts and decreases along with erythroid differentiation. Further, Sox6 is able to repress its own transcription by direct binding to an evolutionarily conserved double Sox6 site lying within its proximal promoter.

EXPERIMENTAL PROCEDURES

Cell cultures CD34+ cells were immunopurified from Human Cord Blood (HCB), cultured for two weeks in StemSpan (Stem Cell Technologies) containing 20% of fetal bovine serum (FBS, Hyclone) and supplemented with 10 ng/ml of human stem cell factor (hSCF), 1 U/ml of human erythropoietin (EPO), 1 ng/ml of human interleukin-3 (hIL-3), 10^{-6} M dexamethasone (Sigma), and 10^{-6} M β-estradiol (Sigma). $CD34^{+}$ cells were seeded at a concentration of 10^{5} cells/ml and diluted over time to maintain the cell concentration in the range of 1–2x10^{6} cells/ml. Cells were collected and analysed on days 6, 8, 10, 12 and 14.

CD34+ cells from human peripheral adult blood (HPB) were immunopurified after buffy coat isolation and erythroid cultures were done as in ref. 14: cells were cultured at a density of 10^{5} cells/mL in alpha-minimal essential medium in the presence of SCF, IL-3, Epo. 1 μg/mL of cyclosporine A was added to inhibit lymphocytes–monocytes growth. After 7 days of culture, cells were replated in erythroid differentiation medium. Cell samples were collected and analysed on days 7 and 14. For both cultures, erythroid differentiation was monitored by FACS analysis for the expression of GlycophorinA (CD235), and by cell morphology analysis on cytocentrifuged samples stained with May–Grunwald–Giemsa (not shown).Primary human cells were obtained according to Institutional guidelines.K562 cells were cultured in RPMI medium supplemented by 10% Fetal Bovine Serum, PenStrep and L-glutamine.

Electrophoretic mobility shift assay (EMSA) ^{32}P-labeled DNA probes were incubated with 10 µgrs of nuclear or total extracts, for 20 min at 15°C in a buffer containing 5% glycerol, 50 mM NaCl, 20 mM Tris, pH 7.9, 0.5 mM EDTA, 5 mM MgCl, 1 mM dithiothreitol (DTT), 500 ng/µl poly(dG-dC), and 50 ng/µl bovine serum albumin (BSA) in a 15-µl final reaction mixture. The reaction mixture was then loaded onto a 8% polyacrylamide gel (29:1 acrylamide-bisacrylamide ratio) and run at 4°C at 150 V for 3 h. Nuclear extracts were prepared according to standard protocols (29, 2). The antibodies used for the supershift analysis were: anti-FLAG, (Sigma F7425); anti-GATA1 (SantaCruz N6, sc-265). Oligonucleotide probes:

WT probe: Fw: 5’ CCTCTGTAACAAAGTTTCTTTGTTTTAATGG 3’

Rev: 5’ CCATTAAAACAAAGAAACTTTGTTACAGAGG 3’

Mutated probe: Fw: 5’ CCTCTGTGGCAGAGTGTCTGTGTGTGAATGG 3’

Rev: 5’ CCATTCACACACAGACACTCTGCCACAGAGG 3’

Sox6 overexpression vectors The Sox6 murine cDNA (17) was kindly provided by Prof. Michiko Hamada-Kanazawa, Kobe-Gakuin University, Japan. The Sox6 cDNA was transferred into the pCMV-Tag 4B plasmid (Stratagene), in frame with a 3’ FLAG epitope, (EcoRI - EcoRV restriction sites), to produce the Sox6FLAG expression vector used in transfection assays. The Sox6 recombinant protein lacks the 49 C-ter aminoacids: this shorter molecule fully retains Sox6 biological proprieties (17) and allows to discriminate between endogenous and exogenous Sox6. The Sox6-FLAG cassette (EcoRI - KpnI) was blunted and cloned immediately upstream to the IRES-Emerald GFP cassette into the blunted BamHI site of the pHR SIN BX IR/EMW (derived from pHR SIN CSGW, (18)) lentiviral

vector. The two packaging plasmids psPAX2 and pMD-VSVG were used to produce Lentiviral pseudo-particles in 293T cells (www.lentiweb.com).

The expression level of the exogenous Sox6FLAG vs the endogenous transcript was estimated in K562 cells and in cultures of primary Cord Blood (at day 10, peak of sox6 expression) by Real Time PCR, using GAPDH mRNA as a standard for comparison. (Supplementary Fig. 1)

Lentiviral harvesting protocol Exponentially growing HEK 293T cells were transfected using Lipofectamine 2000 (Invitrogen) with the three vectors lentiviral system. About 48 h after transfection, the supernatant containing recombinant viruses was harvested, filtered (0.45 µm), centrifuged at 20.000g for 8 hours, and then frozen at –80 °C. Lentiviruses were titrated on HEK 293T cells by measuring the percentage of GFP positive cells by FACS analysis.

Luciferase reporter plasmids The human Sox6 promoter region from nucleotide -1116 to nucleotide –1 was obtained by direct amplification from genomic DNA with Phusion High-fidelity DNA Polymerase (Finnzymes) using the following primers :

Fw:5’ATCGGTACCGGGCTGAGTTAGATATTTATTTC 3’; Rev:5’ ATCTCGAGAGATCTGAATTCATGAAAGTGACCTG 3’, containing a KpnI and XhoI restriction site (underlined), respectively, for further cloning into the corresponding sites of the pGL3 luciferase reporter vector (Promega), either containing or not a downstream TK minimal promoter. The DNA region of 234 bp (from nt. –991 to nt.-759) containing the double Sox6 site, was amplified with the following primers:

Fw: 5’ ACGTGGTACCGATCCATTGTTTTTCAGAAGG 3’

Rev :5’ ACGTCTCGAGAACAAAGAAACTTTGTTACAGAGGC 3’

containing the KpnI and XhoI sites, as above, for cloning into the pGL2 luciferase reporter vector (Promega), upstream to the minimal Gata-1 promoter and Gata-1 erythroid specific enhancer region HS2 (16). To mutate the double Sox consensus, the same region was amplified with a Reverse primer containing the same mutations proved to abolish Sox6 binding in EMSA assay:

Rev: 5’ ACGTCTCGAGCACACAGACACTCTGCCACAGAGGC 3’.

All the amplified DNA regions were sequenced to avoid undesired mutations.

Transfection experiments K562 human erythroleukemic cells were grown in RPMI 1640 medium supplemented with PenStrep, L-glutamine and 10% fetal bovine serum. 1,5 x 10^5 exponentially growing K562 cells were transfected in 0,5 ml of Opti-MEM medium (Invitrogen) as final volume, using 2µl of Lipofectamine 2000 (Invitrogen), with 900 ng of the reporter plasmid, and increasing amount (from 200 ng to 1,6 µgrs) of the Sox6 expression plasmid (pCMVSox6Tag4B), per well. The pCMV-Tag 4B empty vector was added to each transfection at the concentration required to equalize the total amount of DNA transfected in each reaction. After 48 h, total cellular extracts were prepared and the Luciferase activity was measured according to the Promega Luciferase reporter system protocol. All experiments were repeated in triplicate with at least three independent plasmid preparations.

Chromatin Immunoprecipitation (ChIP) assay Briefly, K562 (10^6 cells for each Immunoprecipitation reaction) were fixed with 0.4%

formaldehyde for 10 minutes at room temperature, and chromatin was sonicated to a size of about 500 bp. Immunoprecipitation was performed after overnight incubation with anti-FLAG antibody (Sigma F-7425), and subsequent incubation with protein A agarose (Upstate biotechnology). Immunoprecipitated DNA was then analysed by amplifying an equivalent of 10^4 cells DNA with the following oligonucleotides:

Sox6 promoter:

Fw: 5' TTTGAAAGAATACAGCCTCTG 3'

Rev: 5' ATGCATTAAGGTGGTTTGGTA 3'

GAPDH:

Fw: 5' CGGAGTCAACGGATTTGGTCGTAT 3'

Rev: 5' AGCCTTCTCCATGGTGGTGAAGAC 3'

Immunoprecipitation was repeated 3 times on independent Chromatin preparations.

RNA isolation and RT-PCR Total RNA from 10^5 cells (both K562 or differentiating human CD34+ cells at the different time points) was purified with TRI Reagent (Applied Biosystem AM9738). Before cDNA synthesis, RNA was treated with RQ1 DNase (Promega) for 30 min at 37°C. cDNA was prepared using the High Capacity cDNA Reverse Transcription Kit (Applied Biosystem cat n°4368814) and then diluted 1:50 for amplification. Negative control reactions (without Reverse Transcriptase) gave no signal. Real time analysis was performed using ABI Prism 7500, (PE Applied Biosystems). Primers were designed to amplify 100- to 150-bp amplicons and were based on sequences from the Ensembl database (http://www.ensembl.org/Homo_sapiens/Info/Index). Samples from

three or more independent experiments were analyzed in triplicate. Specific PCR product accumulation was monitored by SYBR Green dye fluorescence in 25-µl reaction volume. Dissociation curves confirmed the homogeneity of PCR products.

The same forward primer was used to amplify both endogenous and exogenous Sox6 cDNA:

Fw:5' GAGGCAGTTCTTTACTGTGG 3'. To discriminate between endogenous and transduced RNA variants, two different reverse primers were used:

Rev:1 5' CCGCCATCTGTCTTCATA 3' complementary to the extreme 3' of Sox6 transcript, and

Rev:2 5' CTTATCGTCGTCATCCTTGTA 3', that matches with the FLAG coding region.

Primers for GAPDH were:

Fw: 5' ACGGATTTGGTCGTATTGGG 3'

Rev: 5' TGATTTTGGAGGGATCTCGC 3'

Mouse Sox6 primers

Fw: 5' TGCGACAGTTCTTCACTGTGG 3'

Rev: 5' CGTCCATCTTCATACCATACG 3'

Mouse HPRT primers:

Fw: 5' CCTGCTGGATTACATTAAAGCACTG 3'

Rev: 5' GTCAAGGGCATATCCAACAACAAAC 3'

Western blot K562 total and nuclear extracts were prepared according to standard protocols (15, 16) and proteins were subjected to SDS-PAGE separation and blotting.

The endogenous Sox6 protein was detected by the anti Sox6 (c-20) Sta Cruz sc-17332 antibody raised against the 50aa C-ter epitope

(lacking in the exogenous Sox6FLAG protein). The Sox6FLAG protein was detected by the use of the anti-FLAG antibody Sigma F7425. Protein loading was checked by reprobing filters with a monoclonal anti hnrnp-C1/C2 antibody (Sigma R5028). Antibodies binding was detected by using appropriate horseradish peroxidise-conjugated IgG and revealed by ECL (LiteAblot, Euroclone).

Cell sorting Freshly extracted Mouse Bone Marrow cells were disaggregated in Phosphate-Buffered-Saline, and incubated with the following labelled antibodies: allophycocyanin (APC) anti-mouse CD117 (c-Kit); PE anti-mouse CD71, and FITC anti mouse TER119, all from Becton Dickinson, RD. The sorting was performed by a MoFlo (DAKO-Cytomation) cell sorter and purity of the obtained cells populations was >95%.

RESULTS

Sox6 expression during erythoid maturation

Sox6 expression was analysed in primary human erythroid liquid cultures undergoing *in vitro* erythroid differentiation, starting from CD34+ cells purified either from adult peripheral (HPB) or cord blood (HCB).

In these cultures CD34+ cells are amplified to erythroblasts and then induced to terminal differentiation into mature erythrocytes. In cord blood cells-derived cultures (19), RT-PCR on Sox6 transcript was performed at day 0 (purified CD34+ cells), during erythroblast amplification (day6), at the end of the erythroid precursors amplification stage (day 8) and then at days 10, 12 and 14 during erythroid induced differentiation. As shown in Fig. 1A, Sox6 expression is absent in CD34+ cells, starts to be detectable at day 8 (72% of Glycophorin A positive cells estimated by FACS analysis, not shown), reaches a peak around day 12 (85% of Glycophorin A positive cells), to finally decrease at day 14, corresponding to the end of the culture and to the maximum extent of erythroid differentiation (nearly 90% of Glycophorin A positive cells).

Sox6 expression shows a similar pattern in the two phase erythroid culture from CD34+ positive cells purified from adult peripheral blood (14), Fig.1B: Sox6 mRNA is absent in CD34+ positive cells (day 0), accumulates at the end of the erythroblasts amplification stage (day 7, corresponding to 67% of Glycophorin A positive cells) and then decreases at day 14 of the culture, when erythroid differentiation is completed (97% of GlycophorinA positivity). Finally, we analysed

Sox6 expression in mouse Bone Marrow FACS sorted cells populations representing progressive stages of erythroid maturation. As shown in Fig.1C, Sox6 is modestly expressed in the more immature (Kit+CD71-) cell population, increases together with erythoid differentiation (CD71high Ter119low and CD71highTer119high cells) and then falls in the most mature cells (CD71low TER119high).

The latter data, confirm in an in vivo system, the observation made in the ex vivo primary cultures.

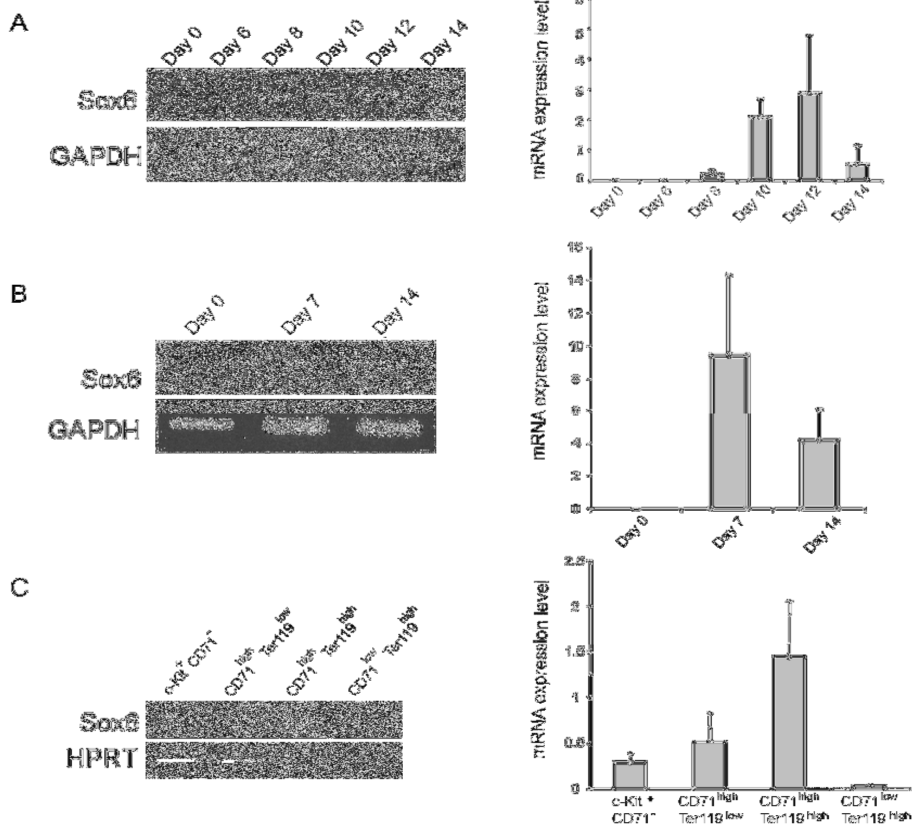


Fig. 1 Sox 6 expression during erythroid maturation.

Semi-quantitative (left) and RealTime (right) PCRs were performed on cDNAs obtained by retrotranscription of RNA samples from cells indicated above the figures. The Sox6 expression level is relative to the expression of the GAPDH gene (human cells) and HPRT (mouse cells).

A. Erythroid liquid cultures from CD34+ cells purified from cord blood. Day 0: purified CD34+ cells; day 8: beginning of the erythroid differentiation phase. Day 14: end of the culture, when 95% of cells are differentiated to Glycophorin A positive cells.

B. Erythroid liquid cultures from CD34+ cells purified from adult peripheral blood. Day 0 corresponds to the freshly purified CD34+ cells, day 7 to the end of the erythroblastic expansion and day 14 to the end of erythroid differentiation.

C. Mouse Bone Marrow cells sorted according to their erythroid maturation, from more immature kit+CD71-, to progressively more mature stages CD71highTer119low, CD71highTer119high, CD71low Ter119high.

The Sox6 gene is repressed by high levels of exogenously overexpressed Sox6

The downregulation of Sox6 transcription in late differentiation described above might be due either to an indirect effect mediated by other transcription factors or to a direct repressive action of Sox6 on its own regulatory sequences.

To answer this question, we overexpressed Sox6 as a SOX6FLAG fusion protein by lentiviral transduction of both HCB cells and of K562 erythroleukemic cells, a well characterized model for the study of the initial stages of erythroid differentiation. K562 cells can be induced to partially differentiate upon hemin addition; under these conditions, K562 cells express low levels of Sox6, similar to those detected prior to hemin addition (Fig.2A)

In both K562 and HCB cells, the overexpression of the exogenous Sox6 gene is mirrored by a dramatic reduction of both the

endogenous Sox6 transcript and protein (Fig 2B, 2C), suggesting that high levels of Sox6 repress the transcription of the endogenous gene.

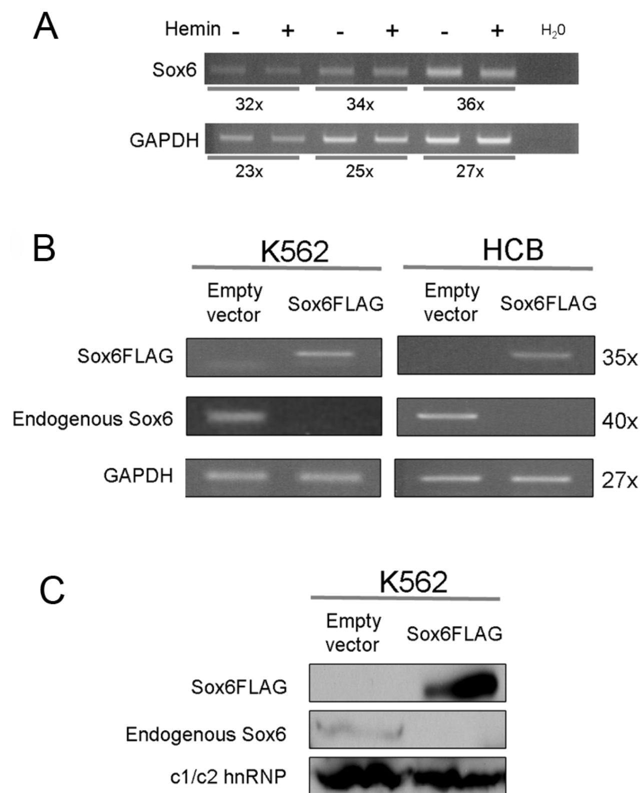


Fig.2 Sox6 exogenous expression represses the endogenous Sox6 gene

A. Semi-quantitative RT-PCR on cDNA from K562 cells, treated (+) or not (-) with hemin (50µM for 4 days). Upper panel: Sox6 expression; Lower panel: GAPDH expression. PCR amplification cycles are indicated below the lanes.

B. Semiquantitative RT-PCR on cDNA from K562 (left) and HCB (right) cells transduced with either the Sox6 overexpressing vector or the corresponding empty vector. GAPDH was used to normalize cDNAs. The number of amplification cycles used for each set of primers is indicated on the right of the figure. In Real time quantification, the endogenous Sox6 transcript -in K562 and HCB cells overexpressing the Sox6FLAG transcript - was scored as undetectable (not shown).

C. Western blot analysis. Nuclear extracts from $7x10^5$ cells were loaded in each lane. An anti-FLAG antibody was used to detect the exogenous Sox6, while the endogenous protein was revealed by anti-Sox6 antibody raised again the C-ter portion of the protein, absent in the exogenous protein. The anti c1/c2 hnRNP antibody was used to normalize for protein loading.

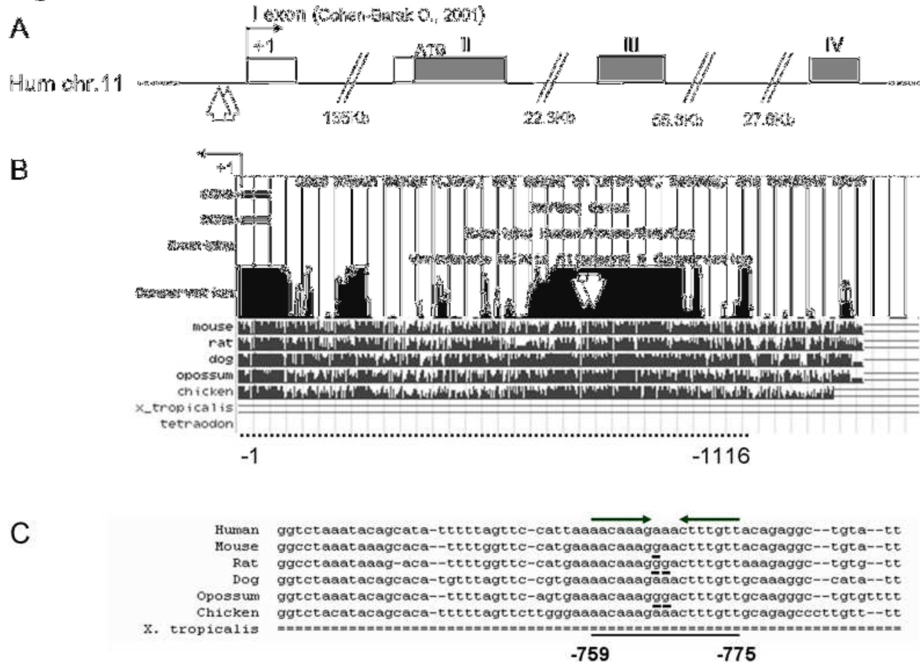
The Sox6 gene contains Sox6 binding sites within its regulatory regions.

We then looked for potential Sox6 binding sites within the regulatory regions of the Sox6 gene (Fig.3).

The Sox6 consensus binding sequence on DNA is very poorly defined: Sox6 is known to recognize the minor groove of DNA, similarly to other members of the HMG proteins family, but very few Sox6 target sites have been validated so far. Among them, the only defined Sox6 target in erythropoietic cells is the 36bp sequence containing two Sox6 binding sites lying within the εy globin proximal promoter and mediating the εy gene repression in adult erythroid cells. This sequence consists of two AACAA(A/T)G sites, in opposite orientation, spaced by 8 nucleotides (13).

We used this sequence as a model to search for putative double Sox6 sites within the Sox6 locus, taking advantage of the USCS database (http://genome.ucsc.edu/) and of the TESS Transcription Element Search System (http://www.cbil.upenn.edu/cgi-bin/tess/tess) software. This approach revealed the presence of an evolutionarily conserved putative double Sox6 binding site -775 nt. upstream to the Sox6 transcription start site (Fig 3). This site, mapping in a block of 31 conserved nucleotides, is composed of 2 single sites in opposite

orientation, spaced by 3 bases. The two Sox6 sites composing the double consensus are fully conserved from chicken to man whereas few substitutions are present within the spacing nucleotides. This same sequence was also identified as a potential Sox binding site by the TFBScluster software (http://hscl.cimr.cam.ac.uk/TFBScluster (20), which allows to identify evolutionarily conserved complex patterns of binding sites present in regions of high regulatory potential defined on the basis of their location and on their conservation among species.



<u>Fig. 3</u> Mapping of the double Sox6 binding site on the human Sox6 promoter

A. Schematic representation of the 5’ region of the human Sox6 locus on chromosome 11. The thick arrows indicate the position of the double Sox6 binding site.

B. UCSC genome browser (http://genome.ucsc.edu/) graphical map of the human Sox6 promoter region. Note that while in panel A the transcription orientation in

from left to right, in panel B it is from right to left. The thick arrows indicate the Sox6 binding sites (same ones as in A). The -1116nt. region of the promoter studied in the transfection experiments described in this paper is indicated with a dotted line under the panel.

C. Nucleotide conservation of the region (nt. -759 to nt. -775) containing the two Sox6 binding sites (black arrows). The few substitutions in the spacer region between the two sites are underlined.

Sox6 binds in vitro and in vivo to the putative Sox6 binding sequences identified within the Sox6 promoter

To test the ability of Sox6 to bind to the above identified consensus, we set up Electrophoretic Mobility Shift Assay (EMSA) experiments using as a probe either an oligonucleotide encompassing the double Sox6 site or a corresponding oligonucleotide mutated in the Sox6 consensus (Fig. 4A). As shown in Fig. 4B the wt probe, when incubated with K562 nuclear extracts (lanes1 to 6), gives a weak retarded band (lane1). When nuclear extracts from K562 cells overexpressing the Sox6FLAG protein are used (lanes 2-6), a stronger band is observed (lane 2). This band is specifically supershifted by increasing amounts of an anti-FLAG antibody (lanes 3-4), but not by an unrelated anti-GATA-1 (lanes 5-6) antibody. Similarly, when total cell extracts are used, a barely detectable band is observed in K562 (lane 7). A strong specific band appears in total extracts from K562 overexpressing cells (lane 8) and is properly supershifted by the anti-FLAG antibody (lane9).

As expected, in K562 Sox6FLAG cells, the overexpression of exogenous Sox6 is accompanied by a decrease of the band corresponding to the endogenous Sox6 protein, (compare lane 1 with

lanes 3 and 4, where the exogenous Sox6 is supershifted by the anti-FLAG antibody and no endogenous Sox6 band is visible), in agreement with the reduction of the endogenous Sox6 mRNA and protein observed in Sox6 transduced cells (see above, Fig.2).
When tested in the same conditions, the probe mutated in the Sox6 consensus (lanes 10 to18) fails to give any retarded band, confirming that the binding observed on the wt probe is specifically due to the double Sox6 site.
To assess the *in vivo* ability of Sox6 to bind to the -775 double site, we performed Chromatin Immunoprecipitation on K562 cells transduced with either the construct overexpressing the Sox6FLAG protein or the corresponding empty vector. The anti-FLAG antibody was used in the chromatin immunoprecipitation reaction. As shown in Fig. 4C, the anti-FLAG antibody efficiently immunoprepipitates the Sox6 double site region in the presence of exogenous Sox6FLAG protein (lane 6). The same sequence is not immunoprecipitated by the same anti-FLAG antibody when chromatin from K562 cells transduced with the empty vector (lane 5) is used. As a further negative control, normal Rabbit-IgG fail to immunoprecipitate the Sox6 promoter region in both Sox6FLAG and empty vector-K562 chromatin samples (lanes 3-4). Moreover, primers designed on the GAPDH gene were used as a control (lower panel) and no amplification was observed on the same immunoprecipitated samples as above.
Together, these data confirm that Sox6 is indeed able to bind both in vitro and in vivo to its own promoter.

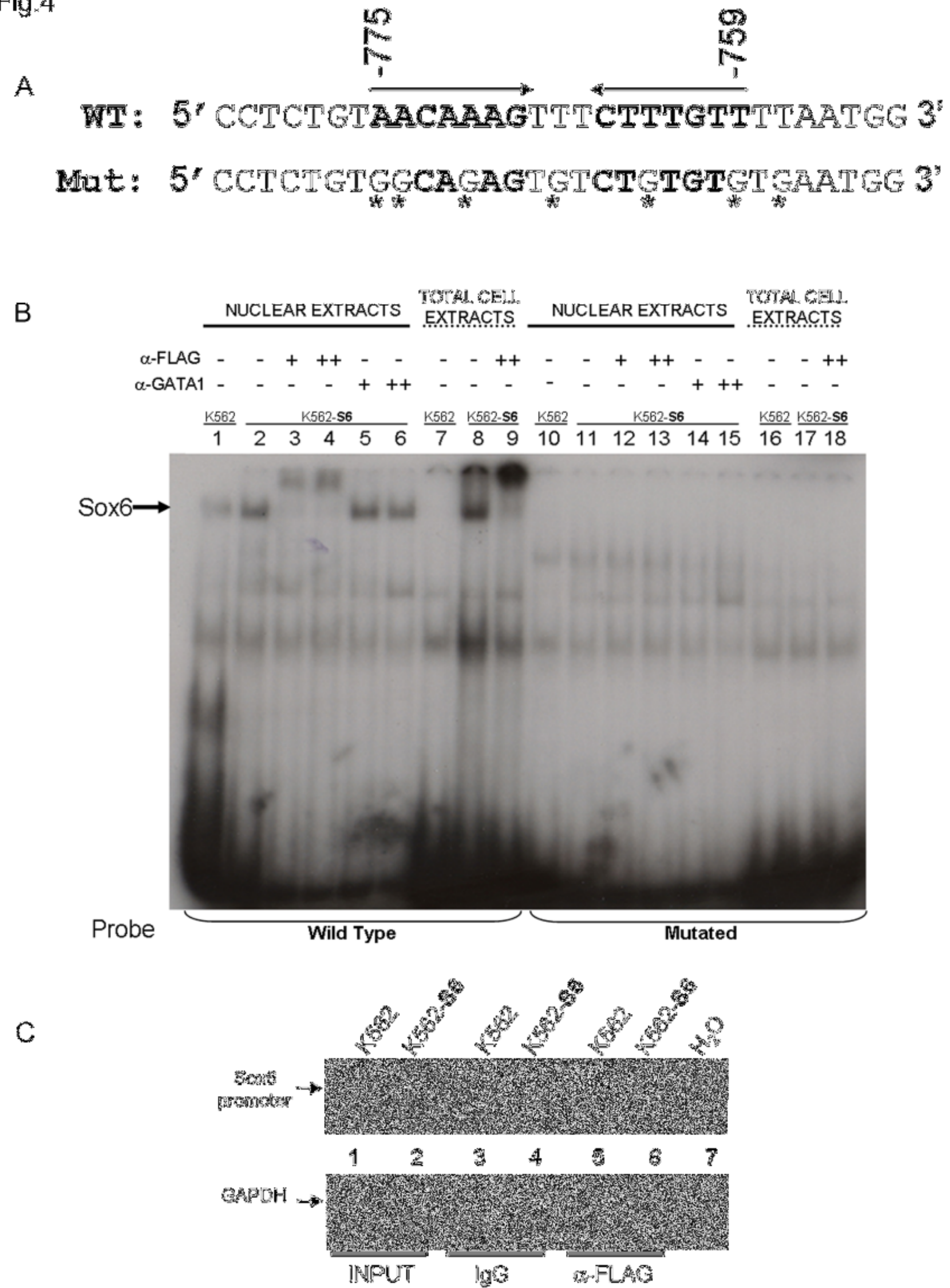


Fig. 4 Sox6 binds *in vitro* and *in vivo* to the -775-759 double Sox6 binding site on its promoter

A. Nucleotide sequence of the probes used in EMSA experiments. The Sox6 sites are indicated by the arrows and the mutations destroying the Sox6 consensus are marked by asterisks.
B. EMSA experiments: the probes in panel A were ^{32}P labelled and incubated with either nuclear (lanes 1 to 6 and 10 to 16) or total cell extracts from K562 cells or K562 cells overexpressing Sox6FLAG (K562-S6). The probe mutated in the Sox6 consensus sites, fails to give any retarded band when tested in the same conditions as for the wild type probe (lane 10-18).
C. Chromatin immunoprecipitation experiments. The anti-FLAG or rabbit IgG antibodies were used to immunoprecipitate chromatin from K562 or K562 overexpressing Sox6FLAG (K562-S6) cells and the region encompassing the double Sox6 binding site was amplified. Lanes 1 and 2: input chromatins. Lanes 3 and 4: normal rabbit IgG. Lane 5 and 6: anti-FLAG antibody (that recognizes the Sox6FLAG transduced protein) Lane 7: water. Upper panel: primers designed on the Sox6 promoter region. Lower panel: primers on the GAPDH gene as a negative control.

The -775 Sox6 binding sites within the Sox6 promoter mediate transcriptional repression in K562 cells.

To test the contribution of the -775bp Sox6 binding sites in regulating Sox6 transcription, we prepared a series of Luciferase reporter constructs that we assayed by transfection experiments in K562 cells. Each experiment described below was done with at least two independent DNA plasmid preparations, each of them transfected at least in triplicate.

We first cloned the 1116 nt upstream to the Sox6 gene transcriptional start site (1116S6), corresponding to the promoter region conserved between man and mouse (Fig 3A) immediately upstream to the Luciferase reporter gene. This construct, transfected in K562 cells,

behaves as a weak promoter, when compared to the promoterless pGL3 plasmid activity and to that of a minimal tk promoter (Fig 5A). Since Sox6 overexpression leads to silencing of the endogenous gene, we reasoned that to observe the potential repression of Sox6 on its own promoter, via its direct binding to the -775 sites, we should work in the context of a highly active erythroid element. We then restricted the Sox6 5' flanking region to 234nt. containing the upstream double Sox6 site and we cloned it 5' to an erythroid regulatory cassette (consisting of the core of the HS2 Gata-1 element, regulated by the GATA-1 and CP2 transcription factors, linked to the 330nt. minimal Gata-1 promoter, S6GATApGL2) (Fig. 5). We also made a similar construct (S6MGATApGL2) in which the double Sox6 site was mutated by the introduction of the same point mutations previously shown to completely abolish Sox6 binding in EMSA (Fig. 4A).

These constructs were cotransfected in K562 cells together with increasing amounts of a Sox6 overexpressing plasmid.

As shown in Fig. 5B, the S6GATApGL2 plasmid is progressively repressed in a dose dependent manner by the simultaneous cotransfection of increasing amounts of the Sox6 expressing vector (7, 11, 15 dark columns, in the histogram). On the other hand, the corresponding S6MGATApGL2 construct, mutated in the Sox6 consensus (columns 8, 12, 16) is insensitive to Sox6 repression. Moreover, the effect of Sox6 on the double Sox6 binding site is highly specific, since the cotransfection of a Sox4 overexpressing plasmid (at the highest concentration used for Sox6) fails to modify the activity of the S6GATApGL2 reporter (column 19).

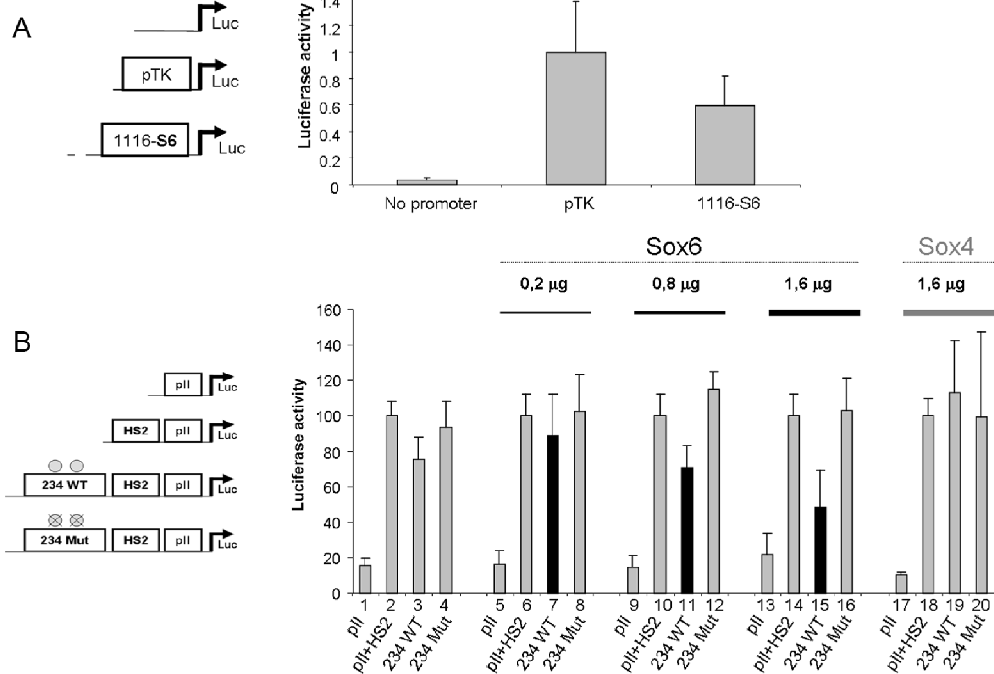


Fig. 5 Sox6 represses its own promoter

A. Transfection experiments in K562 cells. The constructs used are schematically represented on the left of the figure. The Luciferase activity is given in arbitrary units, standard deviations are represented on the top of each column. The region corresponding to nt. -1 to -1116 of the Sox6 promoter (1116-S6) was cloned upstream to the Luciferase reporter gene and its activity was compared to that of a promoterless pGl3 plasmid and to that of a Tk minimal promoter. In these conditions, the 1116-S6 element behaves as a weak promoter.

B. Cotransfection experiments. A 234 nt. fragment containing either the wt or the mutated Sox6 double site was cloned upstream to a highly active erythroid cassette composed of the Gata-1 gene minimal promoter (pII) and the Gata-1 HS2 erythroid enhancer (HS2). The mutations in the Sox6 consensus are the same as in Fig 4A, already shown to abolish Sox6 binding in EMSA (Fig. 4A). These constructs were cotransfected in K562 cells together with increasing amounts (from 0.2 mgrs to 1,6 mgrs) of a Sox6 expressing plasmid. The 234wt fragment is repressed in a dose dependent manner by the addition of the cotransfected Sox6 plasmid (black bars,

columns 7, 11, 15), while the corresponding mutated (234mut) element is insensitive to Sox6 cotransfection (columns 8, 12, 16). A Sox4 overexpressing plasmid (as the highest concentration of 1,6 mgrs used for Sox6 overexpression) fails to repress the Sox6-derived 234 element, suggesting a Sox6 highly specific effect (lane 19).

DISCUSSION

The observation (Fig.1) that Sox6 is progressively induced during erythroid differentiation, peaks at the intermediate stages, and then declines at the latest stages of maturation, prompted a study of the effects of Sox6 overexpression in erythroid differentiation. In the course of this work we showed that exogenous Sox6 greatly accelerates erythroid maturation and hemoglobin synthesis (to be reported elsewhere), while strongly repressing its own expression (Fig. 2). Here, we investigate in detail this effect, showing that Sox6 is a repressor acting on an upstream Sox6 enhancer. We suggest that this negative autoregulatory loop is an important component of the Sox6 downregulation observed during erythroid differentiation.

A repressor role for Sox6

Recent literature demonstrated the role of Sox6 in mouse definitive erythropoiesis, although the molecular mechanisms of its action are still unclear. In particular, despite its profound effect on red cell survival, proliferation and differentiation (11), the only Sox6 target known in erythropoietic cells is the 36bp sequence within the ey globin promoter (13). The direct binding of Sox6 to this sequence has been proposed to repress εy expression in adult erythropoietic cells, thus suggesting the Sox6 requirement for proper embryonic-adult globin switching to occur. The εy site is composed of two single Sox consensus sites in opposite orientation spaced by 8 nucleotides. The prevalence of paired Sox sites on the few known Sox6 target genes (1, 2, 3) suggests that two sites are likely required for Sox6 function, although very little is known about their relative configuration and

their spacing. The double binding site that we identified within the proximal promoter of the human Sox6 gene itself is composed, according to the "paired sites module", by two sites in opposite orientation and identical in their sequence (5'AACAAAG 3' and 5'CTTTGTT 3', respectively), spaced by 3 nucleotides. As in the case of the Sox6 sites on the εy globin promoter, this newly identified Sox6 target site is repressed by Sox6 binding. Moreover, in recent literature, Sox6 has been proposed to work as a repressor, through interactions with a variety of partner factors, in several cell types (10, 21, 22, 23).
On the other hand, Sox6, together with the highly related Sox5 and Sox9 genes, was originally identified as a master gene in chondrogenic differentiation, where this "Sox trio" activates the expression of chondrogenic specific genes, such as type II collagen (*Col2a1*), aggrecan (*Agc*), cartilage link protein (*Crtl1*) and matrilin (3, 4, 24, 25). On these targets, Sox5, Sox6 and Sox9 are thought to secure each other to their binding sites to activate transcription (note that, while Sox6 and Sox5 do not have a transactivation domain, Sox9 possesses it, thus working as a typical transcription factor (26)).
Altogether, these data suggest that interactions with various protein partners and/or different DNA arrangements of Sox6 target sites might underlie the different transcriptional outputs observed (activation or repression) in different systems. To clarify this issue, it will be of great interest to understand which factors are recruited by Sox6 on its own gene and on its other targets in erythroid cells.
Sox6 downregulates its own expression by direct binding to an upstream negative element

Our data suggest that the expression level of Sox6 must be downregulated during late erythropoiesis and that this downregulation is at least in part mediated by a feedback regulatory loop through the direct binding of Sox6 to its own promoter. Sox 6 autoregulation is likely to be dependent on a relatively high level of Sox6 expression. In fact, partial erythroid differentiation of K562 cells, upon hemin treatment, is not associated per se with Sox6 downregulation (Fig.2A). However, when Sox6 is overexpressed in the same cells (K562) or in primary erythroid cultures, endogenous Sox6 downregulation is readily apparent Fig.2B).

We have identified a double Sox consensus binding site 775nt. upstream to the transcriptional start site that is directly bound by Sox6 and is required for Sox6-dependent repression (Fig. 3). Recently, Ikeda et al (27) identified an alternative Sox6 promoter active in embryonic tissues and located 128kb upstream to the sequence investigated in the present study.

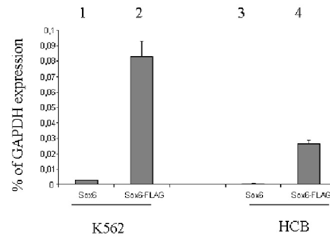
The corresponding transcript is far less abundant in K562 cells than the main transcript studied in this paper, and decreases, as the main transcript, upon Sox6 overexpression (see supplementary Fig.2). Whether this transcript is also directly repressed by Sox6 in erythroid cells is presently not known.

Genetic circuits of autoregulation of key transcription factors, affecting the rate of their synthesis by influencing the rate of transcription (either by autoactivation or autorepression), are known to govern many processes involving progressive commitment of pluripotent progenitors into lineage restricted cells (28, 29) where they are thought to reinforce cell fate decisions. Many examples are known

within the hematopoietic system: Gata-1, for example, is expressed at low levels in multipotent progenitors, but becomes abundant in committed erythroid precursors where its transcription is sustained by a positive feedback loop (30, 31 and refs therein). Finally, Gata-1 expression declines at late stages of erythroid differentiation (32, 33). Interestingly, the initial upregulation of Gata-1 in erythroid differentiation is also linked to repression of another member of the family, Gata-2, which predominates in early progenitors, and may bind to a subset of Gata1-binding sites, often eliciting opposite effects to those of Gata-1 (34). In this regard, it would be of great interest to test whether other Sox family members can share common targets with Sox6, either at different stages of erythroid differentiation or in other tissues in which Sox6 is important for cell differentiation. A candidate for such a role in hematopoiesis is Sox4, which is expressed in early progenitors prior to Sox6 and is progressively repressed during differentiation (unpublished results), with a kinetic opposite to that of Sox6 induction.

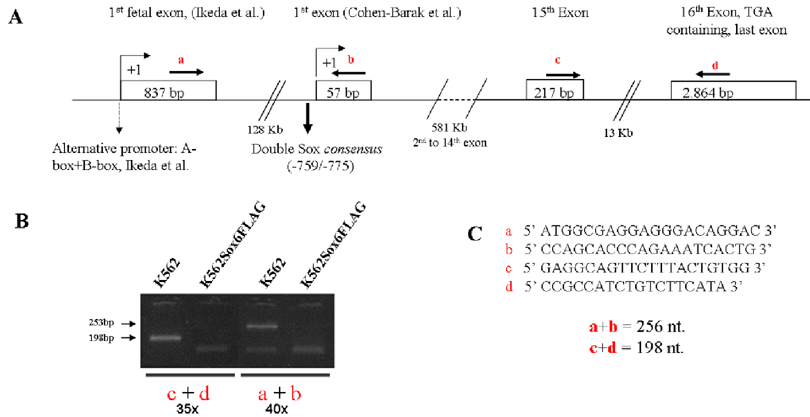
Supplementary Fig.1

The expression level of the exogenous Sox6FLAG vs the endogenous transcript was estimated by Real Time PCR in K562 cells and in primary Human Cord Blood cells at day 10 of the culture, the central day of Sox6 expression, by comparing their expression with GAPDH expression, used as internal standard



Columns hight represents the level of expression of endogenous Sox6 (1,3) and Sox6FLAG (2,4) relatively to GAPDH, used as internal standard.
Columns 1 and 3: untransduced cells;
columns 2 and 4: cells transduced with the Sox6FLAG vector.
Primer pairs were designed to discriminate between endogenous and exogenous Sox6-FLAG transcripts (see experimental procedures). In both case the ratio is about 30 fold (K562: 29, CB: 30).
Standard deviations refer to threee independent amplifications.

Supplementary Fig. 2



A) Schematic representation of the human Sox6 locus on Chromosome 11. The two published transcription start sites are indicated by the arrows at the beginning of the corresponding exons (+1). The position of the double Sox6 binding site studied in this paper is indicated by the thick arrow. The relative position of the primers used for RT-PCR (a+b and c+d) are indicated above the exons.

B) RT-PCR on cDNA from K562 and K562Sox6FLAG cells. Note that while the c+d primers pair detects both transcripts, the a+b pair only amplifies the upstream "fetal" transcript described by Ikeda et al. and that the number of PCR cycles shown are a+b=35x ; c+d=40x, respectively. This indicates that the transcript published by Cohen-Barak is largely the most abundant in K562 cells.

C) Sequence of primers used in the RT-PCR reactions and expected length of the PCR products.

REFERENCES

1) Guth, S. I. and Wegner, M. (2008) *Cell Mol.Life Sci.* **65,** 3000-3018

2) Kiefer, J. C. (2007) *Dev.Dyn.* **236,** 2356-2366

3)Lefebvre, V., Dumitriu, B., Penzo-Mendez, A., Han, Y., and Pallavi, B. (2007) *Int.J.Biochem.Cell Biol.* **39,** 2195-2214

4) Lefebvre, V., Li, P., and de Crombrugghe, B. (1998) *EMBO J.* **17,** 5718-5733

5)Hagiwara, N., Klewer, S. E., Samson, R. A., Erickson, D. T., Lyon, M. F., and Brilliant, M. H. (2000) *Proc.Natl.Acad.Sci.U.S.A* **97,** 4180-4185

6)Connor, F., Wright, E., Denny, P., Koopman, P., and Ashworth, A. (1995) *Nucleic Acids Res.* **23,** 3365-3372

7)Stolt, C. C., Schlierf, A., Lommes, P., Hillgartner, S., Werner, T., Kosian, T., Sock, E., Kessaris, N., Richardson, W. D., Lefebvre, V., and Wegner, M. (2006) *Dev.Cell* **11,** 697-709

8) Ikeda, T., Kamekura, S., Mabuchi, A., Kou, I., Seki, S., Takato, T., Nakamura, K., Kawaguchi, H., Ikegawa, S., and Chung, U. I. (2004) *Arthritis Rheum.* **50,** 3561-3573

9) Cohen-Barak, O., Yi, Z., Hagiwara, N., Monzen, K., Komuro, I., and Brilliant, M. H. (2003) *Nucleic Acids Res.* **31,** 5941-5948

10) Hagiwara, N., Yeh, M., and Liu, A. (2007) *Dev.Dyn.* **236,** 2062-2076

11) Dumitriu, B., Patrick, M. R., Petschek, J. P., Cherukuri, S., Klingmuller, U., Fox, P. L., and Lefebvre, V. (2006) *Blood* **108,** 1198-1207

12) Cohen-Barak, O., Erickson, D. T., Badowski, M. S., Fuchs, D. A., Klassen, C. L., Harris, D. T., and Brilliant, M. H. (2007) *Exp.Hematol.* **35,** 358-367

13) Yi, Z., Cohen-Barak, O., Hagiwara, N., Kingsley, P. D., Fuchs, D. A., Erickson, D. T., Epner, E. M., Palis, J., and Brilliant, M. H. (2006) *PLoS.Genet.* **2,** e14

14) Ronzoni, L., Bonara, P., Rusconi, D., Frugoni, C., Libani, I., and Cappellini, M. D. (2008) *Blood Cells Mol.Dis.* **40,** 148-155

15) Schreiber, E., Matthias, P., Muller, M. M., and Schaffner, W. (1989) *Nucleic Acids Res.* **17,** 6419

16)Bose, F., Fugazza, C., Casalgrandi, M., Capelli, A., Cunningham, J. M., Zhao, Q., Jane, S. M., Ottolenghi, S., and Ronchi, A. (2006) *Mol.Cell Biol.* **26,** 3942-3954

17) Hamada-Kanazawa, M., Ishikawa, K., Nomoto, K., Uozumi, T., Kawai, Y., Narahara, M., and Miyake, M. (2004) *FEBS Lett.* **560,** 192-198

18) Demaison, C., Parsley, K., Brouns, G., Scherr, M., Battmer, K., Kinnon, C., Grez, M., and Thrasher, A. J. (2002) *Hum.Gene Ther.* **13,** 803-813

19) Migliaccio, G., Di Pietro, R., di, G., V, Di Baldassarre, A., Migliaccio, A. R., Maccioni, L., Galanello, R., and Papayannopoulou, T. (2002) *Blood Cells Mol.Dis.* **28,** 169-180

20) Donaldson, I. J. and Gottgens, B. (2006) *Nucleic Acids Res.* **34,** 524-528

21) Murakami, A., Ishida, S., Thurlow, J., Revest, J. M., and Dickson, C. (2001) *Nucleic Acids Res.* **29,** 3347-3355

22) Iguchi, H., Urashima, Y., Inagaki, Y., Ikeda, Y., Okamura, M., Tanaka, T., Uchida, A., Yamamoto, T. T., Kodama, T., and Sakai, J. (2007) *J.Biol.Chem.* **282,** 19052-19061

23) von Hofsten, J., Elworthy, S., Gilchrist, M. J., Smith, J. C., Wardle, F. C., and Ingham, P. W. (2008) *EMBO Rep.* **9,** 683-689

24) Rentsendorj, O., Nagy, A., Sinko, I., Daraba, A., Barta, E., and Kiss, I. (2005) *Biochem.J.* **389,** 705-716

25) Lefebvre, V. (2002) *J.Bone Miner.Metab* **20,** 121-130

26) Han, Y. and Lefebvre, V. (2008) *Mol.Cell Biol.* **28,** 4999-5013

27) Ikeda, T., Saito, T., Ushita, M., Yano, F., Kan, A., Itaka, K., Moro, T., Nakamura, K., Kawaguchi, H., and Chung, U. I. (2007) *Biochem.Biophys.Res.Commun.* **357,** 383-390

28) Akashi, K., D. Traver, T. Miyamoto, and I. L. Weissman. (2000). *Nature* **404**, 193-197

29) Orkin, S. H. (2000) *Nat.Rev.Genet.* **1,** 57-64

30) Ferreira, R., Ohneda, K., Yamamoto, M., and Philipsen, S. (2005) *Mol.Cell Biol.* **25,** 1215-1227

31) Shimizu, R. and Yamamoto, M. (2005) *Semin.Cell Dev.Biol.* **16,** 129-136

32) Tanabe, O., Shen, Y., Liu, Q., Campbell, A. D., Kuroha, T., Yamamoto, M., and Engel, J. D. (2007) *Genes Dev.* **21,** 2832-2844

33) Whyatt, D., Lindeboom, F., Karis, A., Ferreira, R., Milot, E., Hendriks, R., de Bruijn, M., Langeveld, A., Gribnau, J., Grosveld, F., and Philipsen, S. (2000) *Nature* **406,** 519-524

34) Bresnick, E. H., Martowicz, M. L., Pal, S., and Johnson, K. D. (2005) *J.Cell Physiol* **205,** 1-9

FOOTNOTES

Abbreviations used: HPB: human peripheral blood, HCB: human cord blood.

We are grateful to Drs. Thamar van Dijk, Sjaak Philipsen, Michiko Hamada-Kanazawa, Cristina Fugazza and Tariq Enver for reagents and advice for lentiviral work.

This work was supported by PRIN 2007 to M.D.C., by the Italian Telethon Foundation to G.F., Fondazione CARIPLO 2005 to S.O. and PRIN 2006 to A.R.

Chapter 5

FINAL DISCUSSION: MOLECULAR AND CLINICAL CONSIDERATIONS

Summary

The aim of my thesis was to find new genes whose function is crucial for erythrtoid differentiation and/or the regulation of the globin switching process. With this purpose we analysed the gene expression profiles of mouse hemopoietic cells extracted from fetal livers (between days E11.5 and E13.5, period in which the globin switching occurs) at different stages of erythroid differentiation. Among genes whose expression level rose the most, we chose Sox6 for further functional studies. We demonstrated that Sox6 enforced expression induces terminal erythroid maturation in the erythroleukemic cell line K562 and in primary cultures of $CD34^+$ progenitors extracted from human cord blood. We then aimed to unravel the molecular mechanisms underlying Sox6 function, by searching its direct target binding sites on DNA. Among them we found a possible direct regulation of Sox6 on SOCS3 transcription, a gene involved in the regulation of the progression through the cell cycle. Finally, we demonstrated that Sox6, when overexpressed, negatively auto-regulates its transcription by directly binding on its own promoter.

Identification of new genes important for erythropoiesis

Hemopoietic cells were purified from livers of mouse embryos at day E11.5, E12.5 and E13.5 (corresponding to the period during which the globin switching takes place), and FACS sorted on the basis of their expression of c-kit and Ter119 (Chapter 2, figure 1). The transcriptional analysis of these highly purified and homogeneous cell populations, identifies differentially expressed genes (DEGs) between progenitors and more differentiated cells, that are presumably required

for erythroid maturation. Some genes known to vary in erythroid differentiation, for example the globin genes, were used as internal control of our experiment (Chapter 2, figure 2). Many other genes whose level of expression significantly rise during erythroid maturation, but for which no function is already known in the erythroid system, are considered.

On the other side, we could also identify DEGs between cell populations at the same degree of differentiation, but isolated at different days of mouse embryonic development. At day E11.5 of mouse development, erythrocytes express embryonic globins (εy and βH1) and at E13.5 the vast majority of them already express adult ones (βmajor and βminor) (McGrath K. and Palis J., 2008): thus, DEGs between E11.5 and E13.5 are possibly involved in the repression of ε and βH1, and/or the activation of β globin.

The genes found in this way were then divided in functional categories. We focused on the category of transcription factors, with the goal to study genes potentially capable of activating or inhibiting the transcription of many other erythroid specific genes.

We chose for further analyses the transcription factor Sox6, whose expression level rises at all three days of development analysed. It is worth noting that another member of the same family of transcription factors, Sox4, has an opposite behaviour if compared with Sox6 (see figure 6 of chapter 2): it is expressed in early progenitors and progressively declines during differentiation. Since Sox proteins bind similar consensus sequences on the DNA, this observation suggests that Sox6 and Sox4 may regulate the same target genes in different stages of erythroid differentiation, and in opposite fashion.

Sox6 in the human erithroid system

Sox6 is highly upregulated within differentiating erythroid progenitors, in all three days of development studied. It is a member of Sry related HMG box family of transcription factors, which are known to be fundamental for the proper development of several tissues; in particular Sox6 is known to play important functions in the developing central nervous system (Hamada-Kanazawa M., et al., 2004a, Hamada-Kanazawa M., et al., 2004b), in cartilage and in muscle (Lefebvre V., 2009, Han Y. And Lefebvre V., 2008, Ikeda T. et al., 2004, Smits P. et al., 2001; Hagiwara N., et al., 2000). Sox6 absence in mice development leads to a strong impairment of definitive erythropoiesis (Dumitriu B., et al., 2006). Despite these reports, no data are available so far about its molecular mechanisms of action. Moreover, no insights into Sox6 role in human erythroid differentiation and human globin genes regulation are reported.

Sox6 overexpression in erytholeukemic K562 cells (Lozzio and Lozzio, 1977), has a strong effect: they start to differentiate and transcribe several erythroid specific genes, including heme-synthesis enzymes and globin chains (Chapter 3, figure 3). Megakaryocytic genes such as GpIIa and GpIIIb, which are normally upregulated in K562 by TPA addition (Alitalo R., 1990), are downregualated in K562 overexpressing Sox6, indicating that their differentiation potential is restricted to the erythroid lineage. Since K562 cells are normally able to differentiate only upon induction with high concentration of chemical inducers, such as hemin (ferriprotoporphyrin IX) or chemotherapeutic drugs (Rowley PT. et

al., 1992; Charnay P., Maniatis T., 1983), it interesting that a single gene, Sox6, is able to activate many aspects of erythroid differentiation in K562, in the total absence of any other chemical stimuli.

Similar results were obtained in an in vitro system, using primary CD34+ human cord blood cells induced to differentiate into the erythroid lineage (Chapter 3, figure 6.)

Sox6 normally starts to be expressed in erythroid progenitors, at the erythroblast stage (day 8 of this culture), and peaks in more differentiated precursors (day 10 and 12). By producing an artificial early peak of Sox6 expression by Sox6 lentiviral transduction at day 6, we obtained a strong acceleration of differentiation, which was evident by the increased number of more differentiated cells (orthochromatic erythroblasts) and, even more striking, by the appearance of completely enucleated cells at day 10 of the culture; note that enucleated cells are almost absent, at this stage, in the control culture (Chapter 3, figure 6). This result confirmed in a more physiologic model, the ability of Sox6 to accelerate and boost the normal erythroid differentiation we observed in K562, suggesting that Sox6 is a major regulator of the erythropoietic differentiation program.

Sox6 and the globin switching

Sox6 was recently reported as a repressor of the ε-globin gene in mouse (Yi Z. et al., 2006). As reported in figure 4 of chapter 3, Sox6 strongly stimulates the transcription of all the globins that K562 cells are normally able to produce, including the ε-globin. This result is

partially in conflict to the repression reported by Yi et al., and might reflect a different regulation of transcription between murine and human homolog globins. However, Sox6, in a context of general globins transcriptional activation, induces an increase in the ratio γ/ε and α/ζ, suggesting that the embryonic genes are "repressed" as compared to γ and α genes. Moreover, the exogenous transduced Sox6 is able to bind the human ε-globin promoter, although its consensus does not seem to be completely conserved (Figure 4D and 4C respectively, chapter 3), thus suggesting a repressive role of Sox6 on ε-globin transcription also in human cells.

The strong upregulation of globin genes transcription upon Sox6 overexpression is also confirmed by the RT-PCR analysis done on the primary culture of erythroid progenitors, transduced with the same Sox6 containing lentiviral vector (see figure7, chapter 3). All globins are upregulated, especially at day 10, but the levels of ε transcript (barely detectable in these "adult type" cultures) are unchanged.

Taken together, these results suggest that Sox6 is a general activator of the erythroid specific transcriptional program, but has a marginal role in the human globin switching process: it could be of interest to test whether Sox6 really plays a fundamental role into the ε- to γ-globin switch, in the early phase of human embryonic development.

Sox6 targets on DNA

Sox6 activation of globin transcription does not seem to be a direct effect: it happens few days after Sox6 lentiviral transduction, and no evident conserved Sox-consensus binding sites in globin regulatory regions were found (see below). Neither Sox6 effect is mediated by

already known major transcriptional regulators of eryhtopoiesis GATA1, GATA2, EKLF of NF-E2p45, since their expression level is unchanged upon Sox6 overexpression (Figure 3,chapter 3). We thus searched genome-wide Sox6 targets on DNA, using an in-silico approach exploiting the web tool TFBScluster software (Donaldson IL., Gottgens B., 2007 - http://hscl.cimr.cam.ac.uk/TFBScluster -). Sox6, as a member of group D of Sox factors, does not have a trans-activation domain, and exploits its transcriptional function by interacting with other partners (as Sox5 and Sox9 in the cartilage system (Lefebvre V. et al., 2007)), or by the HDAC recruitment on the cyclin D1 promoter (Iguchi H., et al., 2007). Moreover, the majority of described Sox6 binding sites are paired HMG box consensi (Guth S.I. and Wegner M., 2008; Kiefer JC., 2007; Lefebvre V., 2007). The εy promoter, the only known example of erythroid Sox6 target before our work, is perfectly consistent with this observations, and is composed of two binding sequences, with an opposite orientation (Yi Z. et al., 2006).

On these basis, using TFBScluster, we searched for double Sox-consensi, lying within a cluster of no more than 35 nucleotides considering evolutionary conserved sequences of the human genome. Although the exclusion of human genomic regions not conserved in other mammals represents a potential loss of information, this was necessary to restrict the output data of our first search. Despite these constraints, the number of potential Sox6 targets found by TFBScluster was outstanding: more than 800 potential Sox6 target regions in different positions with respect to known genes. We then further narrowed our analysis to those regions found nearby the

transcription initiation site of genes expressed in erythropoiesis (DEGs in our experiment previously described in chapter 2), and we focused on two of them (figure 9, chapter 3): SOCS3 (Suppressor of cytokines signalling) and Sox6 itself.

Sox6 blocks the progression through the cell cycle, and activates SOCS3

The profound effect on differentiation upon Sox6 overexpression in erythroid progenitors cells, is accompanied with a strong reduction in their proliferation rate: K562 cells, despite their leukemic nature stop growing and die in culture 10 days after Sox6 transduction (See figure 1, chapter 3); if transduced in $CD34^+$ progenitors, Sox6 causes early cell death, probably because of a block in the cell divisions program similar to that observed in K562 (not shown).

Among genes involved in cell cycle regulation I tested, SOCS3 is the only one whose expression levels significantly change upon Sox6 transduction: its increase in K562 cells is evident already 3 hours after transduction (Chapter 3, figure 2).

The Sox consensus found within the SOCS3 locus is located 2.7 Kbp upstream to the transcription start site in a very well conserved short region of nearly 120 nucleotides (Figure 9, chapter 3). Preliminary results using the luciferase reporter gene assay (not shown), suggest that Sox6 directly regulates SOCS3 transcription in a dose dependent manner. In parallel, ChIP (Chromatin Immunoprecipitation) will be performed to determine whether Sox6 binds this region in vivo.

SOCS3 is involved in the downregulation of STAT5B activity, which activates the IGF (Insuline growth factor) gene transcription (Usenko T. et al., 2007). It is known that the downregulation of the IGF gene interferes with the cell cycle progression of BCR-ABL positive cells, driven by the constitutive activation of Abl receptor (Lakshmikuttyamma A. et al., 2009).

Our working hypothesis is the direct SOCS3 activation by Sox6, which blocks blocks the proliferation of K562 cells and hemopoietic progenitors in culture, by inhibiting the phosphorilation of STAT5B, thus downregulating the secretion of IGF, which is in turn responsible of the continued divisions of BCR-ABL cells.

We are now planning to check the phosphorilation level of STAT5B after Sox6 tranduction, and to confirm the validity of the mechanism we hypothesized, the ability of SOCS3overexpression to mimic Sox6 effect in the erythroid progenitors would be necessary.

Sox6 regulates its own transcription

Since Sox6 is able to activate molecular mechanisms controlling cell proliferation and differentiation, it is of crucial importance to understand how Sox6 expression is regulated during development. The identification of the conserved double Sox consensus in the Sox6 locus itself, suggests its own auto-regulation.

Very little is known about Sox6 transcriptional regulation. A “fetal” promoter has been recently described (Ikeda T. et al., 2007), lying 135Kbp upstream to the previously discovered transcription start site (Cohen-Baràk O. et al., 2001).

The consensus we found maps approximately 1.1 Kbp upstream of the main transcription start site identified by Cohen-Baràk et al. in position -759 / -775, and it is well conserved in several mammalian species (Chapter 4, figure 3).

This consensus is composed by two sites in opposite orientation, but identical in their sequence (5'AACAAAG 3' and 5'CTTTGTT 3'), spaced by 3 nucleotides. Few substitutions are present in the 3 nucleotides between the two consensi, suggesting their stringent evolutionary importance. We demonstrated that Sox6 is able to specifically bind this sequence with high affinity in vitro (EMSA) and in vivo (ChIP) (see chapter 4, figure 4B and 4C respectively).

As in the case of the Sox6 consensus site on the εy globin promoter, this newly identified target site is repressed by Sox6 binding: the activity of this region is downregulated by Sox6, in a dose-dependent manner. Moreover, point mutations in the Sox consensus completely abolish this repression. This repression activity seems to be Sox6 specific, since Sox4 protein is not able to give any effect (Chapter 4, figure 5). Finally, lentiviral transduction of Sox6 switches off the transcription of the endogenous Sox6 mRNA, that virtually disappears (Chapter 4, figure 2).

A repressive role of Sox6 has been recently proposed to occur in several other cell types (Han Y. and Lefebvre V., 2008; Iguchi H. et al., 2007; Murakami A. et al., 2001; von Hofsten J. et al., 2008).

The prevalence of paired Sox sites on Sox6 target genes described so far (Guth S.I. and Wegner M., 2008; Kiefer JC., 2007; Lefebvre V., 2007) suggests that two sites are required for Sox6 function, although very little is known about their relative configuration and spacing.

Sox6 levels peek in vivo (in fetal liver and erythroid adult progenitors from bone marrow) in relatively mature progenitors (Chapter 4, figure 1), but sharlply declines at subsequent stages. Based on our transfection and ChIP experiments (Chapter4, figure 5 and 4C respectively), we propose that Sox6 may participate in its own transcriptional repression.

Many examples of key transcription factors, able to affect their rate of transcription, by autoactivation or by autorepression, were already known. GATA1, for example, probably the most important erythroid transcription factor, is able to positively regulate its own transcription in erythroid progenitors (Shimizu R. and Yamamoto M., 2005; Ferreira R. et al., 2005), but at the end of the differentiation process, it needs to be downregulated (Tanabe O. et al., 2007). In fact its overexpression in erythroid cells inhibits their terminal maturation, and leads to letal anemia (Whyatt D. et al., 2000).

It would be of great interest to test whether other proteins, perhaps belonging to the Sox family, can activate Sox6 expression in the early phases of erythroid differentiation, possibly by binding the Sox consensus we found. A possible candidate is Sox4, which has an opposite behaviour if compared with Sox6 (see figure 6 of chapter 2), since it is expressed in early progenitors and progressively drops during differentiation.

The clinical relevance

Mutation of coding or intron sequences of the β-globin gene or, more rarely, within their regulatory regions (promoters, enhancers etc.), is the cause of diseases such as sickle cell anemia and β-thalassemia,

affecting millions of people in the world (Vichinsky E., 2007). Even though recent progresses in medical treatment of hemoglobin disorders (iron chelation, transfusion, bone marrow transplantation) have extended life expectancy of these patients, a definitive cure is not yet available (Cunningham MJ., 2008); however, an improved understanding of the molecular basis of the disease, has provided clues for potential molecular targets (Quek L. and Thein SL., 2007).
Several medical observations started few decades ago, demonstrate that the maintenance of γ-globin expression through adult life, is able to significantly ameliorate the clinical conditions of thalassemic patients (Forget B., 1998).
Our investigation to find genes differentially expressed in the murine and human erythroid system, may highlight new hypothetical targets for molecular treatment of hemoglobinopaties. In fact, if the molecular events in hemoglobin switching or γ globin gene reactivation are better understood, HbF could be reactivated to some extent in adult cells, leading to a substantial amelioration for these genetic blood disorders (Bank A., 2006). In this respect, our investigation has not yet been successful. However it might be useful in principle to ameliorate another major defect of thalassemia, ineffective erythropoiesis (i.e. the failure of erythroblasts to complete their maturation and their premature death in the bone marrow) (Libani IV. et al., 2008). The strong induction of differentiation by Sox6 on erythroid precursor cells, now encourages us to test whether Sox6 would be able to compensate at least partially the maturation defect of thalassemic cells, by transduction of hemopoietic progenitors derived from patients with different forms of thalassemia. Although Sox6

introduction directly in the human bone marrow would not be possible, the elucidation of the molecular mechanisms by which Sox6 induces erythroid differentiation could better explain the physiologic defect present in patients with these diseases, and might provide clues to new molecular treatments. In this regard, the activation of SOCS3 brought about by Sox6, and the resulting signal transduction changes, may be interesting and amenable to pharmacological treatments.

The possibility of a direct activation of SOCS3 by Sox6 in the hemopoietic tissue, and probably, in other tissues where Sox6 is expressed (Cohen-Baràk O. et al., 2001), indicates that Sox6 has a role in the regulation of the normal or aberrant cell cycle progression.

Further, molecular elucidation of SOCS3 regulation mechanism can in fact have a profound impact in the research concerning tumors from different tissues treatment: its mis-regulation is the cause of numerous proliferative disorders in humans: of lung and liver (Baltayiannis G. et al., 2008) and of the digestive tract (Isomoto H., 2009).

Of interest, in the hemopoietic system, SOCS3 is involved in the aberrant signaling transduction of myeloproliferative neoplastic disorders (Kota J. et al., 2008), and in the molecular mechanism causing Polycythemia Vera (Chen G., Prchal JT., 2006). It has been very recently shown how the overexpression of SOCS3 is indeed able to revert the molecular defect of Polycythemia Vera cells, correcting the hypersensitivity to IGF molecule, that lead to erythocytes precursors overgrowth (Usenko T. et al., 2007). In this regard, the strong augment of SOCS3 transcript in Sox6 overexpressing cells, whether direct or indirect, motivates further studies.

REFERENCES

Alitalo, R. 1990. Induced differentiation of K562 leukemia cells: a model for studies of gene expression in early megakaryoblasts. Leuk.Res. **14**:501-514.

Baltayiannis, G., N. Baltayiannis, and E. V. Tsianos. 2008. Suppressors of cytokine signaling as tumor repressors. Silencing of SOCS3 facilitates tumor formation and growth in lung and liver. J.BUON. **13**:263-265.

Bank, A. 2006. Regulation of human fetal hemoglobin: new players, new complexities. Blood **107**:435-443.

Bernard, P. and V. R. Harley. 2009. Acquisition of SOX transcription factor specificity through protein-protein interaction, modulation of Wnt signalling and post-translational modification. Int.J.Biochem.Cell Biol.

Charnay, P. and T. Maniatis. 1983. Transcriptional regulation of globin gene expression in the human erythroid cell line K562. Science **220**:1281-1283.

Chen, G. and J. T. Prchal. 2006. Polycythemia vera and its molecular basis: an update. Best.Pract.Res.Clin.Haematol. **19**:387-397.

Cohen-Barak, O., N. Hagiwara, M. F. Arlt, J. P. Horton, and M. H. Brilliant. 2001. Cloning, characterization and chromosome mapping of the human SOX6 gene. Gene **265**:157-164.

Connor, F., E. Wright, P. Denny, P. Koopman, and A. Ashworth. 1995. The Sry-related HMG box-containing gene Sox6 is expressed in the adult testis and developing nervous system of the mouse. Nucleic Acids Res. **23**:3365-3372.

Cunningham, M. J. 2008. Update on thalassemia: clinical care and complications. Pediatr.Clin.North Am. **55**:447-60, ix.

Donaldson, I. J. and B. Gottgens. 2007. CoMoDis: composite motif discovery in mammalian genomes. Nucleic Acids Res. **35**:e1.

Dumitriu, B., M. R. Patrick, J. P. Petschek, S. Cherukuri, U. Klingmuller, P. L. Fox, and V. Lefebvre. 2006. Sox6 cell-autonomously stimulates erythroid cell survival, proliferation, and terminal maturation and is thereby an important enhancer of definitive erythropoiesis during mouse development. Blood **108**:1198-1207.

Ferreira, R., K. Ohneda, M. Yamamoto, and S. Philipsen. 2005. GATA1 function, a paradigm for transcription factors in hematopoiesis. Mol.Cell Biol. **25**:1215-1227.

Forget, B. G. 1998. Molecular basis of hereditary persistence of fetal hemoglobin. Ann.N.Y Acad.Sci. **850**:38-44.

Guth, S. I. and M. Wegner. 2008. Having it both ways: Sox protein function between conservation and innovation. Cell Mol.Life Sci. **65**:3000-3018.

Hagiwara, N., B. Ma, and A. Ly . 2005. Slow and fast fiber isoform gene expression is systematically altered in skeletal muscle of the Sox6 mutant, p100H. Dev.Dyn. **234**:301-311.

Hamada-Kanazawa, M., K. Ishikawa, K. Nomoto, T. Uozumi, Y. Kawai, M. Narahara, and M. Miyake. 2004. Sox6 overexpression causes cellular aggregation and the neuronal differentiation of P19 embryonic carcinoma cells in the absence of retinoic acid HAMADAKANAZAWA2004A. FEBS Lett. **560**:192-198.

Hamada-Kanazawa, M., K. Ishikawa, D. Ogawa, M. Kanai, Y. Kawai, M. Narahara, and M. Miyake. 2004. Suppression of Sox6 in P19 cells leads to failure of neuronal differentiation by retinoic acid and induces retinoic acid-dependent apoptosis. FEBS Lett. **577**:60-66.

Han, Y. and V. Lefebvre. 2008. L-Sox5 and Sox6 drive expression of the aggrecan gene in cartilage by securing binding of Sox9 to a far-upstream enhancer. Mol.Cell Biol. **28**:4999-5013.

Iguchi, H., Y. Urashima, Y. Inagaki, Y. Ikeda, M. Okamura, T. Tanaka, A. Uchida, T. T. Yamamoto, T. Kodama, and J. Sakai. 2007. SOX6 suppresses cyclin D1 promoter activity by interacting with beta-catenin and histone deacetylase 1, and its down-regulation induces pancreatic beta-cell proliferation. J.Biol.Chem. **282**:19052-19061.

Ikeda, T., S. Kamekura, A. Mabuchi, I. Kou, S. Seki, T. Takato, K. Nakamura, H. Kawaguchi, S. Ikegawa, and U. I. Chung. 2004. The combination of SOX5, SOX6, and SOX9 (the SOX trio) provides signals sufficient for induction of permanent cartilage. Arthritis Rheum. **50**:3561-3573.

Ikeda, T., T. Saito, M. Ushita, F. Yano, A. Kan, K. Itaka, T. Moro, K. Nakamura, H. Kawaguchi, and U. I. Chung . 2007. Identification and characterization of the human SOX6 promoter. Biochem.Biophys.Res.Commun. **357**:383-390.

Ingley, E., P. A. Tilbrook, and S. P. Klinken. 2004. New insights into the regulation of erythroid cells. IUBMB.Life **56**:177-184.

Isomoto, H. 2009. Epigenetic alterations in cholangiocarcinoma-sustained IL-6/STAT3 signaling in cholangio- carcinoma due to SOCS3 epigenetic silencing. Digestion **79 Suppl 1**:2-8.

Kiefer, J. C. 2007. Back to basics: Sox genes. Dev.Dyn. **236**:2356-2366.

Lefebvre, V. and P. Smits. 2005. Transcriptional control of chondrocyte fate and differentiation. Birth Defects Res.C Embryo.Today **75**:200-212.

Lefebvre, V., B. Dumitriu, A. Penzo-Mendez, Y. Han, and B. Pallavi. 2007. Control of cell fate and differentiation by Sry-related high-mobility-group box (Sox) transcription factors. Int.J.Biochem.Cell Biol. **39**:2195-2214.

Lefebvre, V. 2009. The SoxD transcription factors - Sox5, Sox6, and Sox13 - are key cell fate modulators. Int.J.Biochem.Cell Biol.

Lakshmikuttyamma A., Pastural E., Takahashi N., Sawada K., Sheridan DP., DeCoteau JF., Geyer CR. 2009. Bcr-Abl induces autocrine IGF-1 signaling. Oncogene. **27**:3831-3844.

Libani IV., Guy EC., Melchiori L., Schiro R., Ramos P., Breda L., Scholzen T., Chadburn A., Liu Y, Kernbach M., Baron-Lühr B., Porotto M., de Sousa M., Rachmilewitz EA., Hood JD., Cappellini MD., Giardina PJ., Grady RW., Gerdes J., Rivella S.. 2008.

Decreased differentiation of erythroid cells exacerbates ineffective erythropoiesis in beta-thalassemia. Blood. **3**:875-885.

Lozzio, B. B. and C. B. Lozzio . 1977. Properties of the K562 cell line derived from a patient with chronic myeloid leukemia. Int.J.Cancer **19**:136.

McGrath, K. and J. Palis. 2008. Ontogeny of erythropoiesis in the mammalian embryo. Curr.Top.Dev.Biol. **82**:1-22.

Migliaccio, G., R. Di Pietro, G. di, V, A. Di Baldassarre, A. R. Migliaccio, L. Maccioni, R. Galanello, and T. Papayannopoulou. 2002. In vitro mass production of human erythroid cells from the blood of normal donors and of thalassemic patients. Blood Cells Mol.Dis. **28**:169-180.

Murakami, A., S. Ishida, J. Thurlow, J. M. Revest, and C. Dickson. 2001. SOX6 binds CtBP2 to repress transcription from the Fgf-3 promoter. Nucleic Acids Res. **29**:3347-3355.

Rowley PT., Farley BA., LaBella S., Giuliano R., Leary JF.. 1992. Single K562 human leukemia cells express and are inducible for both erythroid and megakaryocytic antigens. Int J Cell Cloning. **10**:232-240.

Quek, L. and S. L. Thein. 2007. Molecular therapies in beta-thalassaemia. Br.J.Haematol. **136**:353-365.

Shimizu, R. and M. Yamamoto. 2005. Gene expression regulation and domain function of hematopoietic GATA factors. Semin.Cell Dev.Biol. **16**:129-136.

Smits, P., P. Li, J. Mandel, Z. Zhang, J. M. Deng, R. R. Behringer, B. de Crombrugghe, and V. Lefebvre. 2001. The transcription factors L-Sox5 and Sox6 are essential for cartilage formation. Dev.Cell **1**:277-290.

Tanabe, O., Y. Shen, Q. Liu, A. D. Campbell, T. Kuroha, M. Yamamoto, and J. D. Engel. 2007. The TR2 and TR4 orphan nuclear receptors repress Gata1 transcription. Genes Dev. **21**:2832-2844.

Usenko, T., D. Eskinazi, P. N. Correa, D. Amato, Y. Ben David, and A. A. Axelrad. 2007. Overexpression of SOCS-2 and SOCS-3 genes reverses erythroid overgrowth and IGF-I hypersensitivity of primary polycythemia vera (PV) cells. Leuk.Lymphoma **48** :134-146.

Vichinsky, E. 2007. Hemoglobin e syndromes. Hematology Am.Soc.Hematol.Educ.Program.79-83.

von Hofsten, J., S. Elworthy, M. J. Gilchrist, J. C. Smith, F. C. Wardle, and P. W. Ingham. 2008. Prdm1- and Sox6-mediated transcriptional repression specifies muscle fibre type in the zebrafish embryo. EMBO Rep. **9**:683-689.

Whyatt, D., F. Lindeboom, A. Karis, R. Ferreira, E. Milot, R. Hendriks, M. de Bruijn, A. Langeveld, J. Gribnau, F. Grosveld, and S. Philipsen. 2000. An intrinsic but cell-nonautonomous defect in GATA-1-overexpressing mouse erythroid cells. Nature **406**:519-524.

Yi, Z., O. Cohen-Barak, N. Hagiwara, P. D. Kingsley, D. A. Fuchs, D. T. Erickson, E. M. Epner, J. Palis, and M. H. Brilliant. 2006. Sox6 directly silences epsilon globin expression in definitive erythropoiesis. PLoS.Genet. **2**:e14.

Università degli Studi di Milano-Bicocca

CONSULTAZIONE TESI DI DOTTORATO DI RICERCA

Il sottoscritto Claudio Cantu', n° matricola 036251,
nato a Vaprio d'Adda, il 06/12/1981,
autore della tesi di DOTTORATO dal titolo:

"THE SOX6 TRANSCRIPTION FACTOR:

ITS ROLE IN HUMAN AND MURINE

ERYTHROID DIFFERENTIATION

AND MECHANISMS FOR ITS REGULATION"

AUTORIZZA

La consultazione della tesi stessa, fatto divieto di riprodurre, in tutto o in parte, quanto in essa contenuto.

Data, 12/11/2009 Firma