

# Sox6 Directly Silences Epsilon Globin Expression in Definitive Erythropoiesis

Zanhua Yi<sup>1</sup>✉, Orit Cohen-Barak<sup>1</sup>✉, Nobuko Hagiwara<sup>2</sup>, Paul D. Kingsley<sup>3</sup>, Deborah A. Fuchs<sup>4</sup>, Drew T. Erickson<sup>1</sup>, Elliot M. Epner<sup>5</sup>, James Palis<sup>3</sup>, Murray H. Brilliant<sup>1\*</sup>

**1** Department of Pediatrics, University of Arizona, College of Medicine, Tucson, Arizona, United States of America, **2** Department of Internal Medicine, Division of Cardiovascular Medicine, University of California Davis, Davis, California, United States of America, **3** Department of Pediatrics, Center for Pediatric Biomedical Research, University of Rochester Medical Center, Rochester, New York, United States of America, **4** Department of Pathology, University of Arizona, College of Medicine, Tucson, Arizona, United States of America, **5** Department of Hematology & Oncology, Oregon Health & Science University, Portland, Oregon, United States of America

**Sox6 is a member of the Sox transcription factor family that is defined by the conserved high mobility group (HMG) DNA binding domain, first described in the testis determining gene, Sry. Previous studies have suggested that Sox6 plays a role in the development of the central nervous system, cartilage, and muscle. In the Sox6-deficient mouse,  $p^{100H}$ ,  $\epsilon\gamma$  globin is persistently expressed, and increased numbers of nucleated red cells are present in the fetal circulation. Transfection assays in GM979 (erythroleukemic) cells define a 36-base pair region of the  $\epsilon\gamma$  proximal promoter that is critical for Sox6 mediated repression. Electrophoretic mobility shift assay (EMSA) and chromatin immunoprecipitation (ChIP) assays demonstrate that Sox6 acts as a repressor by directly binding to the  $\epsilon\gamma$  promoter. The normal expression of Sox6 in wild-type fetal liver and the ectopic expression of  $\epsilon\gamma$  in  $p^{100H}$  homozygous fetal liver demonstrate that Sox6 functions in definitive erythropoiesis. The present study shows that Sox6 is required for silencing of  $\epsilon\gamma$  globin in definitive erythropoiesis and suggests a role for Sox6 in erythroid cell maturation. Thus, Sox6 regulation of  $\epsilon\gamma$  globin might provide a novel therapeutic target in the treatment of hemoglobinopathies such as sickle cell anemia and thalassemia.**

Citation: Yi Z, Cohen-Barak O, Hagiwara N, Kingsley PD, Fuchs DA, et al. (2006) Sox6 directly silences epsilon globin expression in definitive erythropoiesis. PLoS Genet 2(2): e14.

## Introduction

Sry type HMG box (Sox6) is a member of the Sox transcription factor family characterized by the conserved high mobility group (HMG) domain, consisting of 79 amino acids involved in DNA recognition and binding [1]. Sox transcription factors bind to the minor groove of DNA and cause a 70°–85° bend of the DNA that leads to local conformational changes [2,3], while most other transcription factors target the major groove of DNA [4]. Therefore, Sox proteins may perform part of their function as architectural proteins by organizing local chromatin structure and assembling other DNA-bound transcription factors into biologically active, sterically defined multiprotein complexes. Sox6 has been reported to be able to act as either an activator or a repressor, depending on its interactors and its target promoter context [5,6]. Intriguingly, Sox6 has also been shown to act as a general splicing factor that participates in pre-mRNA splicing [7]. Depletion of Sox6 in HeLa cell extracts blocked splicing of multiple substrates, and expression of the HMG domain of either Sox6, Sox9, or Sry in the extracts restored splicing, indicating functional overlap of these proteins [7]. Regardless of how Sox6 functions in regulating gene expression, previous studies have demonstrated that Sox6 is an important regulatory molecule that plays a role in the development of the central nervous system [8–11], cartilage [6,12,13], and muscle [14,15]. A Sox6-null mutant mouse ( $p^{100H}$ ) has previously been identified in our laboratory [14]. Mice homozygous for  $p^{100H}$  show delayed growth, develop myopathy and arterioventricular heart block, and die within 2 wk after birth [14]. The  $p^{100H}$  mutant allele is associated with a Chromosome 7 inversion that disrupts both the  $p$  gene and the Sox6 gene (and no other gene within 50,000

nucleotides of the chromosomal breakpoints) [14]. Because the  $p$  gene functions solely in pigmentation [16], the Sox6 transcription factor is implicated in all other phenotypes.

Among the HMG box proteins distantly related to Sry (the first member identified of the Sox transcription factor family) that similarly bind to the minor groove and bend DNA, but without sequence specificity, are the ubiquitously expressed HMG1 and HMG2 proteins [17]. Modulation of DNA structure by these and other HMG proteins can mediate long-range enhancer function on both DNA and chromatin-assembled genes by bringing together distant regions of DNA and associated factors. Specifically, HMG proteins have been shown to modulate  $\beta$ -globin genes [18–21].

The mouse  $\beta$ -globin genes { $\epsilon\gamma$ ,  $\beta$ h1,  $\beta$ -major, and  $\beta$ -minor} are clustered on Chromosome 7 and they are highly homologous to their human counterparts in organizational

**Editor:** Gregory Barsh, Stanford University, United States of America

**Received** September 28, 2005; **Accepted** December 20, 2005; **Published** February 3, 2006

A previous version of this article appeared as an Early Online Release on December 21, 2005 (DOI: 10.1371/journal.pgen.0020014.eor).

**DOI:** 10.1371/journal.pgen.0020014

**Copyright:** © 2006 Yi et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Abbreviations:** bp, base pair; ChIP, chromatin immunoprecipitation; dpc, days post coitus; EMSA, electrophoretic mobility shift assay; HMG, high mobility group; PCR, polymerase chain reaction; Sox6, Sry type HMG box; WT, wild-type

\* To whom correspondence should be addressed. E-mail: mhb@peds.arizona.edu

✉ These authors contributed equally to this work.

## Synopsis

Beta-globin gene switching—the transition from embryonic to fetal to adult synthesis of specific globin chains—results in hemoglobins with different affinity for oxygen. This system is a longstanding paradigm for developmental biology and is directly relevant to human disease, since small amounts of normal embryonic or fetal beta-globins can “balance” the detrimental effect of abnormal or missing adult globins in diseases such as sickle cell anemia and beta-thalassemia.

In the current study, the transcription factor Sox6 was identified as a novel and crucial silencing factor of epsilon (embryonic) globin through a somewhat serendipitous pathway. The authors had previously identified a chromosomal inversion,  $p^{100H}$ , by virtue of its effect on the pink-eyed dilution gene and found that the same inversion also disrupts the Sox6 gene. Using  $p^{100H}$  mutant mice as a tool for identifying downstream targets of Sox6, the authors discovered that epsilon-globin levels were dramatically elevated, paving the way for a series of molecular genetic experiments demonstrating that Sox6 directly binds to and normally inhibits transcription from the epsilon-globin gene promoter. This work provides fundamental new insights into regulation of globin gene transcription during development, and provides new clues for manipulating globin gene transcription as an approach to treat human blood diseases.

structure and function [22]. High-level expression of these genes requires a regulatory element, the locus control region that is characterized by a set of nuclease hypersensitive sites spread over 25 kb located 5' of the  $\epsilon$  gene [23]. The  $\beta$ -globin genes are expressed in a tissue- and development-specific fashion. In mice, erythropoiesis originates in the embryonic yolk sac where primitive erythroid cells express  $\epsilon$  and  $\beta$ h-1 globins [22]. At 11.5 d post coitus (dpc), erythropoiesis shifts to the fetal liver where definitive erythroid cells express adult  $\beta$  globins ( $\beta$  major and minor) [22]. The  $\epsilon$  gene is silenced in definitive erythroid cells. The mechanism of silencing of its human counterpart,  $\epsilon$  globin, has been studied extensively. In definitive erythropoiesis,  $\epsilon$  is activated and silenced autonomously [24,25], although in primitive erythropoiesis  $\epsilon$  also appears to be regulated competitively [26]. The  $\gamma$ -globin to adult  $\beta$ -globin switch is controlled by promoter competition for the LCR [24,25].

All the elements responsible for silencing the  $\epsilon$  globin gene are within the  $\epsilon$  gene or in adjacent sequences [27], suggesting that silencing is primarily gene autonomous. Using promoter deletion analyses in transgenic mouse models and cell transfection assays, multiple DNA elements important to the silencing process have been previously identified in both the proximal and the distal  $\epsilon$  gene promoter [27]. Their corresponding transcription factors, such as GATA-1, YY-1, COUP-TF, and DRED have been identified and shown to directly bind to these DNA elements (as part of protein complexes) to regulate  $\epsilon$  silencing [27]. Thus, it appears that the silencing of the  $\epsilon$  gene involves a complicated network of multiple *cis* elements and transacting proteins.

In addition to playing an important role in the development of the central nervous system [8–11], cartilage [6,12,13], and muscle [14,15], it was shown that Sox6 is upregulated in long-term hematopoiesis stem cells (LT-HSC) compared with multipotent progenitors of adult mouse bone marrow lineage [28]. In this study, we describe that Sox6 also exerts

pleiotropic effects on erythropoiesis. These effects include delayed maturation of erythrocytes (that normally enucleate prior to entering the bloodstream [27]) and higher expression of embryonic globin genes. The most extreme effect is the persistence of high expression of the embryonic  $\epsilon$  globin gene. Here we describe and characterize the effects of Sox6 on the  $\epsilon$  globin gene. We show that Sox6 binds to the proximal promoter of  $\epsilon$  globin and represses its transcription. In wild-type (WT) mice, Sox6 is not expressed in yolk sac blood islands, but is expressed in fetal liver, the opposite expression pattern of  $\epsilon$  globin. In the absence of Sox6,  $\epsilon$  globin is ectopically expressed in the fetal liver, demonstrating that Sox6 functions in definitive erythropoiesis.

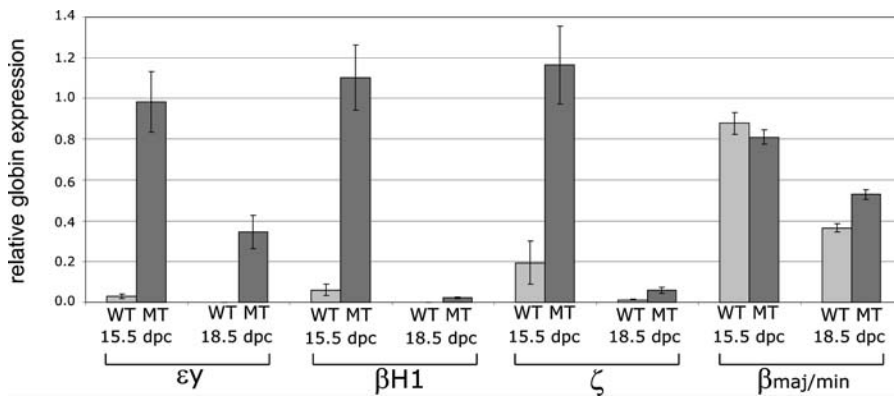
## Results

### Persistent Expression of the Embryonic Globin, $\epsilon$ , in Sox6-Deficient Mice

The  $\epsilon$  globin gene was initially identified as an upregulated transcript in the  $p^{100H}$  mouse using subtractive hybridization to identify Sox6 downstream targets. This initial observation was confirmed in an independent knock-out allele of Sox6 [13] using real-time polymerase chain reaction (PCR) (unpublished data). Real-time PCR was used to quantitate the expression levels of other globin genes in  $p^{100H}$  mutant and WT mouse livers at two developmental stages, 15.5 dpc and 18.5 dpc. As shown in Figure 1, the  $\epsilon$  gene is expressed at high levels at both time points in mutant mice, in contrast to the decline in expression observed in WT mice. No difference was seen between WT and heterozygous mice (unpublished data). Interestingly, the expression levels of the other two embryonic globin genes ( $\zeta$  and  $\beta$ h1) are also higher in  $p^{100H}$  homozygous mice, compared with WT mice, but to a much lesser extent than seen for  $\epsilon$  globin at 18.5 dpc (Figure 1). Moreover, the expression level of adult  $\beta$  globin is also somewhat higher in  $p^{100H}$  homozygous mice than in WT mice at 18.5 dpc (Figure 1). Perinatal lethality of mutant mice (presumably from the heart defect [14]) precludes us from evaluating postnatal globin expression. The graphs in Figure 1 illustrate real time PCR results that were performed in triplicate (standard deviation of the data is shown by error bars). Because all of the assays were performed at the same time with the same internal control, the levels shown are relative levels and are thus comparable across all samples and are in agreement with previously published results for WT fetal mice [29]. We note that the level of  $\epsilon$  expression in the livers of 15.5 dpc and 18.5 dpc homozygous mutant mice is statistically equivalent to the level of  $\beta$ maj/min expression in the livers of 15.5 dpc and 18.5 dpc homozygous WT mice.

### Transfection Studies Using GM979 Cells Indicate That Sox6 Directly Represses the $\epsilon$ Gene Promoter at the Transcriptional Level

Real-time PCR assays (Figure 1) measure steady-state levels of  $\epsilon$  mRNA, not transcriptional activity of the  $\epsilon$  promoter. To investigate whether Sox6 directly acts on the  $\epsilon$  gene promoter at the transcriptional level, we used an *in vitro* transient transfection assay and GM979 cells, a murine erythroleukemic cell line that expresses both  $\epsilon$  and adult beta globins [30]. We generated an  $\epsilon$  promoter reporter construct (E-Luc) by fusing a micro-LCR ( $\mu$ LCR) element (2.5 kb) [31] to the  $\epsilon$  proximal promoter (2.2 kb), followed by the



**Figure 1.** Real-Time PCR of Globin Genes

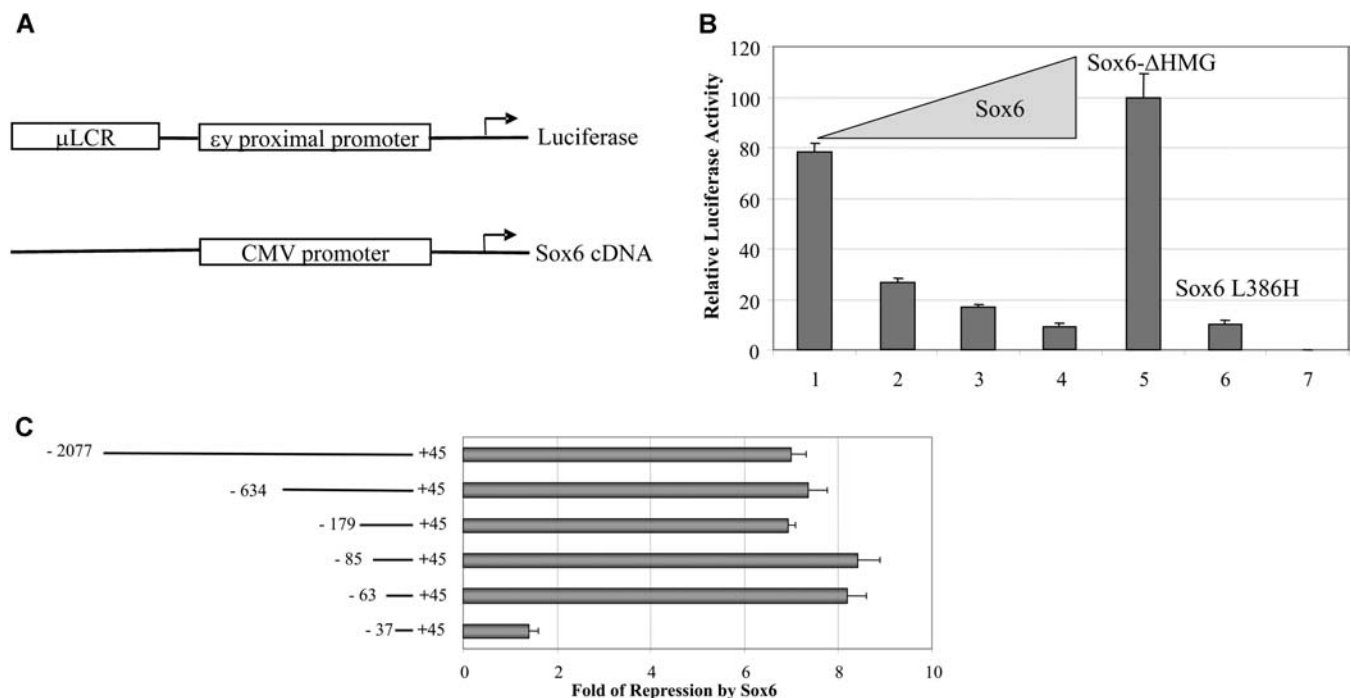
The levels of expression of  $\epsilon\gamma$ ,  $\beta\text{H1}$ , zeta, and  $\beta\text{maj/min}$  were measured at 15.5 dpc and 18.5 dpc in homozygous WT and  $p^{100H}$  mutant littermates by real-time PCR (see Materials and Methods). Relative expression levels in the livers of each genotype are graphed for each globin gene (performed in triplicate and normalized with GAPDH). Standard deviation is indicated by bars.

DOI: 10.1371/journal.pgen.0020014.g001

luciferase reporter gene, as shown in Figure 2A (detailed in Materials and Methods). Overexpression of Sox6 in GM979 cells by transient transfection leads to a dosage-dependent repression of E-Luc reporter activity (Figure 2B). In contrast, overexpression of a truncated Sox6 protein that lacks its HMG domain [32] (similar to the  $p^{100H}$  mouse allele) fails to repress

E-Luc activity (Figure 2B). These data indicate that Sox6 acts to repress the  $\epsilon\gamma$  promoter at the transcriptional level.

Sox6 has been shown to act as a repressor and to interact with a widely expressed co-repressor, CtBP2, on the *fgf-3* promoter [5]. CtBP2 is expressed in GM979 cells (unpublished data). To investigate whether the interaction with CtBP2 is



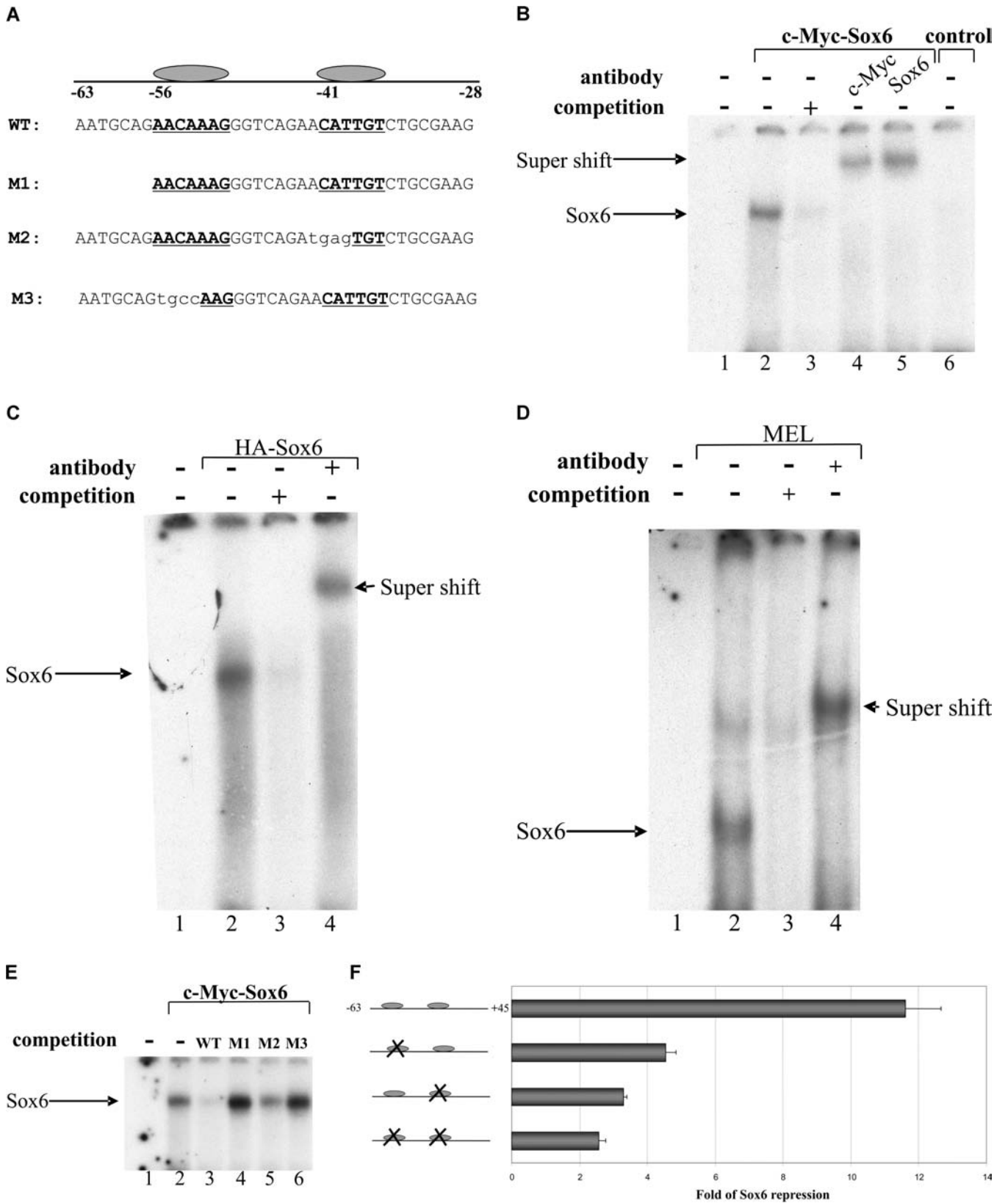
**Figure 2.** The Effect of Sox6 on the  $\epsilon\gamma$  Promoter

(A) Constructs of the  $\epsilon\gamma$  promoter reporter (E-luc) and Sox6 overexpression vector. The E-luc reporter construct consists of a 2.5-kb  $\mu\text{LCR}$  element, a 2.2-kb  $\epsilon\gamma$  proximal promoter, and the luciferase reporter in the pGL-3 basic plasmid (see Materials and Methods). Sox6 expression is driven by the CMV promoter.

(B) Sox6 represses  $\epsilon\gamma$  promoter activity in a dosage-dependent manner. In GM979 cells, the E-Luc  $\epsilon\gamma$  promoter reporter construct was co-transfected (1) without overexpression of Sox6; (2–4) with increasing amounts of CMV-Sox6 overexpression vector; (5) with a truncated version of Sox6 that lacks its HMG domain; (6) with a mutant version of Sox6 (L386H) that has previously been shown to abolish interaction with CtBP2; or (7) with an empty reporter plasmid (without  $\epsilon\gamma$  promoter and  $\mu\text{LCR}$  element).

(C) Promoter deletion analyses to delimit the critical sequence. The 2.2-kb proximal promoter or deletions of it, as indicated on the left (numbering relative to +1 = the transcription start site of  $\epsilon\gamma$  globin, see Materials and Methods), were engineered in reporter constructs as in (A) and were transfected along with CMV driven Sox6 to GM979 cells (see Materials and Methods for details). The relative repression by Sox6 on the activity of the different reporter constructs is shown. All experiments were done in triplicate.

DOI: 10.1371/journal.pgen.0020014.g002



**Figure 3.** Analysis of the Minimal Region (36 bp) of the Proximal  $\epsilon\gamma$  Promoter Responsive to Sox6

(A) The sequence of the 36-bp fragment and its mutant versions used in EMSA. The WT 36-bp DNA sequence (–63 to –28) of the  $\epsilon\gamma$  globin proximal promoter contains two Sox/Sox6 consensus binding sites, shown in bold underline. Versions with truncation of this sequence (M1) or mutation of one of the two consensus binding sites (M2 and M3) are also shown.

(B) EMSA with c-Myc-tagged Sox6. EMSA was performed using the 36-bp radio-labeled WT probe (as shown in (A)) and c-Myc tagged Sox6 translated in vitro using reticulocyte lysate (see Materials and Methods). Lane 1: radio-labeled free probe (run out of the gel); Lane 2: no competition, no antibody; Lane 3: competition with 200-fold excess cold probe, no antibody; Lane 4: no competition, c-Myc antibody (producing a supershift); Lane 5: no competition, Sox6 antibody (producing a supershift); Lane 6: no competition, no antibody using in vitro translated vector containing c-Myc, but not Sox6.

(C) EMSA with HA-tagged Sox6. EMSA was performed similarly as in (B) using HA-tagged Sox6 translated in vitro. Lane 1: radio-labeled free probe (run out of the gel); Lane 2: no competition, no antibody; Lane 3: competition with 200-fold excess cold probe, no antibody; Lane 4: no competition, HA antibody (producing a supershift).

(D) EMSA using MEL cell nuclear extracts and the 36-bp WT probe. Lane 1: radio-labeled free probe (run out of the gel); Lane 2: no competition, no antibody; Lane 3: competition with 200-fold excess cold probe, no antibody; Lane 4: no competition, Sox6 antibody (producing a supershift).

(E) EMSA with c-Myc-tagged Sox6, WT and mutant versions of the 36-bp fragment in competition. EMSA was performed using the radio-labeled 36-bp WT probe and the c-Myc tagged Sox6 translated in vitro. Lane 1: radio-labeled free probe; Lane 2: no competition, no antibody. Competition was performed using 200-fold excess cold probes corresponding to WT (Lane 3), M1 (Lane 4), M2 (Lane 5), and M3 (Lane 6).

(F) Both consensus Sox/Sox6 binding sites are required for Sox6 responsiveness. GM979 cells were transfected with a reporter construct (Figure 2A) containing –63 to +45 of the  $\epsilon\gamma$  proximal promoter together with the CMV-Sox6 overexpression vector. Mutations of the consensus binding sites were also tested (M3, M2, M2 plus M3, see (A)). The fold repression of Sox6 with the WT or mutant constructs is shown. The baseline activities of the mutagenized reporter constructs are comparable to the WT construct.

DOI: 10.1371/journal.pgen.0020014.g003

required for Sox6 repression of the  $\epsilon\gamma$  promoter, we introduced a point mutation (L386H) in the Sox6 protein that has been previously reported to be sufficient to abolish Sox6-CtBP2 interaction [5]. This amino acid change is not in the HMG DNA binding domain. However, this mutant version of Sox6 retains the ability to repress the  $\epsilon\gamma$  promoter in the transfection assay (Figure 2B), indicating that Sox6 represses the  $\epsilon\gamma$  promoter in a CtBP2-independent manner. Deletion analysis of the  $\epsilon\gamma$  promoter, as shown in Figure 2C, defined a region (–63 to –37) within the  $\epsilon\gamma$  proximal promoter that is critical for Sox6 repression. Analysis of this short region reveals two Sox/Sox6 consensus binding sites [5] (Figure 3A).

**EMSA and ChIP Assays Show that Sox6 Directly Binds to the  $\epsilon\gamma$  Promoter**

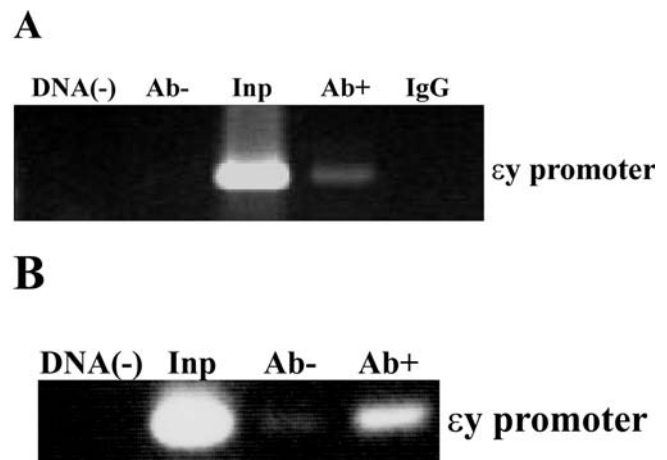
Sox6 might repress the  $\epsilon\gamma$  promoter, either through direct physical contact with the promoter or by regulating intermediates affecting the  $\epsilon\gamma$  promoter. To investigate whether Sox6 is directly associated with the  $\epsilon\gamma$  promoter, we first performed electrophoretic mobility shift assays (EMSA) using a c-Myc-tagged Sox6 in a reticulocyte lysate-based transcription/translation in vitro system. The probes used are listed in Figure 3A. The 36-base pair (bp) WT probe corresponds to the critical region of the  $\epsilon\gamma$  promoter defined in our promoter deletion analyses. This probe contains two consensus Sox/Sox6 binding sites. Also included in our EMSA are three mutated probes that are, either truncated (M1), or mutated (M2 and M3) in Sox/Sox6 binding sites (Figure 3A). Sox6 is able to physically associate with the 36-bp region (Figure 3B) within the  $\epsilon\gamma$  promoter defined by the deletion analysis experiments (Figure 2C). The 36-bp probe was shifted by the tagged Sox6 protein. Moreover, both c-Myc and Sox6 antibodies supershift the band, indicating that the binding is Sox6-specific. To rule out the possibility that the c-Myc tag itself binds to the probe, an HA-tagged Sox6 was used in another EMSA that confirmed these results (Figure 3C).

Next, nuclear extracts from MEL cells were used in EMSA employing the same 36-bp probe. MEL cells, a murine erythroleukemic cell line, express adult  $\beta$  globins, but not  $\epsilon\gamma$  [33]. Sox6 directly binds to this DNA sequence in MEL cells (Figure 3D). The intact consensus Sox/Sox6 binding sites of the DNA probe are required for the binding, as shown in the competition assay (Figure 3E). Ablation of putative Sox/Sox6

binding sites (M1 and M3) abolish their ability to compete in EMSA (Figure 3E). The M2 mutant probe may compete partially with WT binding.

To investigate the functional significance of the intact Sox/Sox6 binding sites, the  $\epsilon\gamma$  promoter reporter constructs with mutagenized Sox/Sox6 binding sites were co-transfected with the Sox6 overexpression vector into GM979 cells. Consistent with the EMSA results, the mutant  $\epsilon\gamma$  promoter reporter constructs (with either one or both Sox/Sox6 binding sites mutagenized) do not result in significant promoter repression in transfection studies (Figure 3F). Thus, both sites are required for maximal repression of  $\epsilon\gamma$  by Sox6, but not to the same degree.

We also tested whether Sox6 binds to the  $\epsilon\gamma$  promoter in vivo using chromatin immunoprecipitation (ChIP) (Figure 4).

**Figure 4.** ChIP Assay

MEL cells (A) and 15.5-dpc fetal liver cells (B) were treated as detailed in Materials and Methods. 10% of the sample was saved as total input (Inp); remaining samples were divided: plus Sox6 antibody (Ab+), minus Sox6 antibody (Ab–), as well as no DNA (DNA–) and normal rabbit IgG (IgG) that served as negative controls. Other controls for these experiments included PCR within the promoter of the  $\alpha$ -globin gene and intron 24 of the  $p$  gene. Both were negative (unpublished data). PCR was carried out using primer pairs flanking the Sox/Sox6 binding sites (see Material and Methods) of the  $\epsilon\gamma$  proximal promoter. For all reactions, we used 2  $\mu$ l of immuno-precipitated DNA and 2  $\mu$ l of 1/100 total input. Semiquantitative PCR was done within the exponential range. Multiple independent experiments were done.

DOI: 10.1371/journal.pgen.0020014.g004

The Sox6-containing complex was immunoprecipitated from MEL cells or from liver cells of 15.5 dpc WT mice using Sox6 antibody. Figure 4 shows that the  $\epsilon\gamma$  proximal promoter is readily immunoprecipitated with Sox6 antibody in both MEL cells and liver cells. Normal IgG was used as a negative control (Figure 4A). The above data (Figures 3 and 4) clearly indicate that Sox6 acts as a repressor of the  $\epsilon\gamma$  gene by directly binding to the  $\epsilon\gamma$  promoter, probably as a dimer.

### The Persistent Expression of $\epsilon\gamma$ Globin in Sox6-Deficient Mice Is Due to a Defect in the $\epsilon\gamma$ -Gene-Silencing Mechanism in Definitive Erythroid Cells

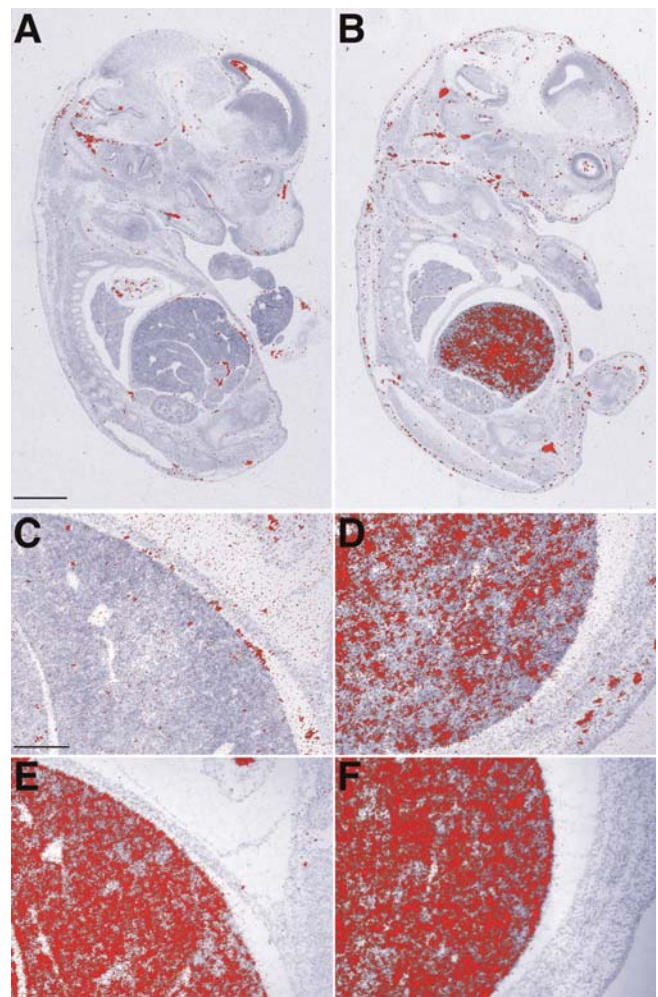
Normally, the  $\epsilon\gamma$  globin gene is exclusively expressed in primitive erythrocytes and silenced in definitive erythrocytes. To determine whether the persistent expression of  $\epsilon\gamma$  globin is due to residual primitive erythrocytes or is due to ectopic expression of  $\epsilon\gamma$  globin in definitive erythrocytes, we examined the spatial pattern of  $\epsilon\gamma$  globin transcripts in mouse embryos by in situ hybridization (Figure 5). As expected,  $\epsilon\gamma$  globin is not expressed in the WT 14.5-dpc liver, the site of definitive erythropoiesis in the fetus. In contrast, abundant ectopic  $\epsilon\gamma$  mRNA expression is seen in the liver of 14.5-dpc mutants (Figure 5 A–D). However, the expression of  $\beta$ major/min globin is equally abundant in both WT and  $p^{100H}$  mutant mice (Figure 5 E and F). These data demonstrate that the persistent high levels of  $\epsilon\gamma$  are due to ectopic expression in the definitive erythroid cells that mature in the fetal liver, suggesting that there is an intrinsic defect of the  $\epsilon\gamma$  silencing mechanism in Sox6-null mice.

### The Expression Pattern of Sox6 Suggests a Role in Definitive Erythropoiesis

The observation that Sox6-deficient mice ectopically express  $\epsilon\gamma$  globin in liver, where definitive erythroid cells mature, suggests that Sox6 is an important regulator in definitive erythropoiesis. To determine the temporal and spatial expression pattern of Sox6, Northern blot and in situ hybridization assays were employed. As shown in Figure 6A, Sox6 is detectable by Northern blot beginning at 10.5 dpc, coincident with the temporal onset of definitive erythropoiesis in the liver. Furthermore, in situ hybridization shows that Sox6 is highly transcribed in 12.5-dpc liver, but not in yolk sac blood islands at 7.5 dpc (Figure 6B). Therefore, Sox6 expression is temporally and spatially coincident with definitive, but not primitive, erythropoiesis. These data, taken together with the observation that  $\epsilon\gamma$  globin is highly expressed in the liver cells of 14.5-dpc  $p^{100H}$  mutant mice (Figure 5), demonstrate that Sox6 functions in definitive erythropoiesis.

### Mutant $p^{100H}$ Mice Have Higher Numbers of Nucleated Red Cells

Among the other Sox6 effects in erythropoiesis, we have noticed that there are more nucleated red blood cells circulating in  $p^{100H}$  mutant mice than in WT mice at 14.5 dpc and 18.5 dpc (Figure 7A). However, at postnatal day 10.5, we do not see circulating nucleated red cells in either WT or mutant mice, suggesting that this may be a transient effect. In addition, the mutant liver shows a significant increase in hematopoietic precursor cells including nucleated erythrocytes at 18.5 dpc (Figure 7B). This alteration is noted as early



**Figure 5.** In Situ Hybridization of  $\epsilon\gamma$  and  $\beta$ major/min transcripts in WT and  $p^{100H}$  Mutant Mice at 14.5 dpc

Expression of  $\epsilon\gamma$  globin (A–D) and  $\beta$  globin (E and F) in sagittal sections of 14.5-dpc fetuses is shown in pseudocolor (red).

(A) WT fetus.

(B)  $p^{100H}$  Homozygous mutant fetus.

(C and D) Inset boxes are shown. Prominent expression of  $\epsilon\gamma$  globin is only detected in the liver of mutant mice.

(E and F) In contrast, both WT and mutant mouse livers express adult  $\beta$ major/min, respectively. No signals were detected above background using sense probes (unpublished data).

The size bars represent 1 mm for (A) and (B); 100  $\mu$ m for (C–F).

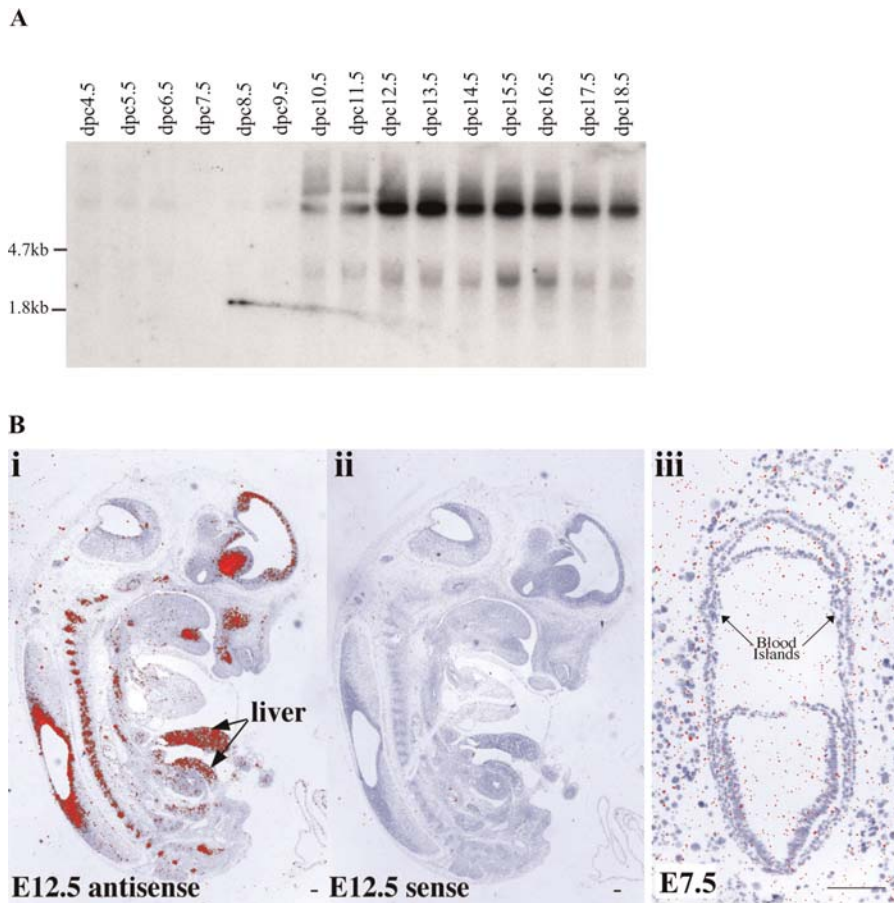
DOI: 10.1371/journal.pgen.0020014.g005

as 14.5 dpc. These observations suggest that besides silencing the  $\epsilon\gamma$  globin gene, Sox6 may affect red cell maturation.

### Discussion

In this report, we show that Sox6 is a novel factor in the complicated regulation mechanism of globin genes. In the Sox6 null mouse, there is a transient effect on the embryonic globin genes,  $\zeta$  and  $\beta$ H1, and a persistent upregulation of the  $\epsilon\gamma$  globin gene. Sox6 directly regulates and binds to the proximal promoter of  $\epsilon\gamma$  gene and represses the  $\epsilon\gamma$ -globin gene in definitive erythropoiesis.

Sox6 belongs to group D of the Sox family of proteins that includes Sox5, 12, 13, and 23 [34]. Group D Sox proteins contain a coiled-coiled domain that mediates homo- and



**Figure 6.** Expression Pattern of Sox6 by Northern Blot and In Situ Assays

(A) Sox6 expression during embryonic development shown by Northern blot. Each lane contains 20  $\mu$ g of total RNA from embryos whose ages are listed above each lane as dpc. The filter was hybridized with a  $^{32}$ P-labeled 575-bp mouse Sox6 cDNA fragment (nucleotides 1353–1927). Numbers on the left are sizes of standard marker fragments in kb.

(B) Sox6 expression shown by in situ hybridization. Panel i: Sagittal section through an E12.5 mouse embryo using antisense Sox6. mRNA distribution is represented by pseudocolored red signal superimposed on the counterstained specimen. Sox6 transcripts are detected primarily in the fetal liver, developing nervous system, chondrocytes and craniofacial area. Panel ii: The sense control probe shows no signal above background. Panel iii: E7.5 embryo hybridized to antisense probe for Sox6. No signal is detected above background specifically in blood islands (or with the sense probe, unpublished data).

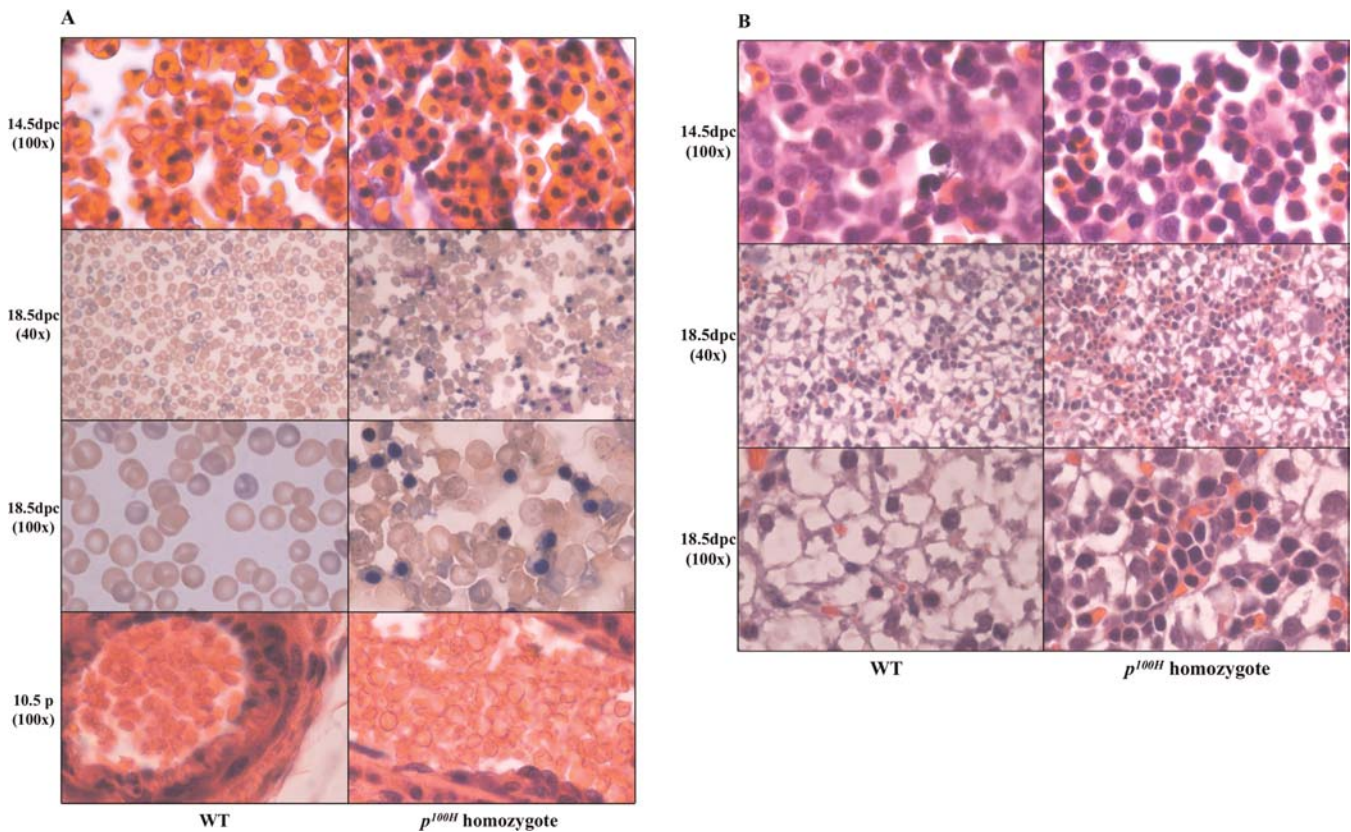
The size bars represent 100  $\mu$ m.

DOI: 10.1371/journal.pgen.0020014.g006

heterodimerization [6,35]. Functionally, dimerization of Sox5 and Sox6 has been shown to greatly increase the binding efficiency of the two Sox proteins to DNA that contains adjacent Sox sites [6]. In addition, Sox6 binds more strongly to an HMG-box dimer motif than to a single HMG-box motif [5]. Therefore, it appears that target genes for group D Sox proteins, such as Sox6, probably harbor pairs of HMG binding sites with a configuration compatible with binding of D-Sox protein dimers. Indeed, in the present study, the defined Sox6 target sequence of the  $\epsilon$ y promoter contains two Sox/Sox6 consensus sites (Figure 3A). Functionally, both sites are essential for Sox6 binding to the  $\epsilon$ y promoter and repression of its activity (Figure 3E and 3F). These observations suggest that Sox6 binds to this sequence of the  $\epsilon$ y promoter either as a homodimer or as a heterodimer with other Sox proteins. Because Sox proteins recognize a short 6-bp core-binding sequence that allows for considerable degeneracy, the specificity of their actions is thought to rely upon interactions with other transcription factors [36]. In our EMSAs, we had to run the electrophoresis on a 4%–6% gel

for at least 4–8 h to detect the Sox6-associated band, suggesting that Sox6 is part of a high molecular weight complex. A few other  $\epsilon$ y globin repressors have been reported to bind to DNA sequences near the Sox/Sox6 consensus sites, including the DRED complex [37] and COUP-TF [38]. Sox6 might interact with these factors and form a large repression complex. Identification of other components of the Sox6-containing complex associated with the  $\epsilon$ y promoter will shed light on its mechanism of repression.

Sox proteins bind and bend linear DNA by partial intercalation in the minor groove, and can also bind to four-way junctions [2–4]. Therefore, one attractive model to explain how Sox6 proteins control gene expression is that they function as architectural factors bound to DNA, influencing local chromatin structure by bending DNA and by assembling multiprotein transcriptional complexes. By changing the local chromatin structure, Sox6 could either interfere with binding of other activators to the promoter or facilitate binding of other repressors. Another example of a repressor that interferes with an activator on the  $\epsilon$ y



**Figure 7.** The Blood and Liver Phenotype of WT and  $p^{100H}$  Homozygous Mice

(A) Red blood cells of WT mice (left panels) and  $p^{100H}$  homozygous mice (right panels) are shown at the indicated ages.

(B) Liver cells of WT mice (left panels) and  $p^{100H}$  homozygous mice (right panels) are shown at the indicated ages.

DOI: 10.1371/journal.pgen.0020014.g007

promoter is DRED. DRED interferes with EKLF, an activator, in binding to the  $\epsilon$  promoter [39]. Two HMG architectural proteins (distantly related to the Sox family of transcription factors), HMG-I and HMG-Y, were demonstrated to bind to the human adult  $\beta$  globin silencers (silencers I and II) and cause bending of the DNA, facilitating the binding of other repressors [40].

Sox6 expression is temporally and spatially coincident with definitive (but not primitive) erythropoiesis (Figure 6), and Sox6 represses  $\epsilon$  globin expression both in vivo (Figure 1) and in vitro (Figure 2). Moreover, in situ hybridization clearly shows that the persistent expression of  $\epsilon$  globin in  $p^{100H}$  mutant mice is due to defects in the silencing mechanism of definitive erythropoiesis that takes place in the liver (Figure 5). Taken together, these data demonstrate that Sox6 functions in definitive erythropoiesis to silence  $\epsilon$  globin expression.

The expression level of  $\epsilon$  globin in homozygous Sox6 null mice at 15.5 dpc and 18.5 dpc is statistically equivalent to the level of  $\beta_{maj}/\min$  expression in the livers of 15.5-dpc and 18.5-dpc homozygous WT mice (Figure 1). This demonstrates that ectopic expression of the  $\epsilon$  globin gene is quite robust in homozygous mutant mice. The expression levels of two other embryonic globin genes ( $\zeta$  and  $\beta_{H1}$ ) are also higher in  $p^{100H}$  homozygotes, compared with WT. Like  $\epsilon$ , levels of  $\zeta$  and  $\beta_{H1}$  are dramatically higher in mutant mice at 15.5 dpc. However, unlike  $\epsilon$  globin,  $\zeta$  and  $\beta_{H1}$

decline in expression by day 18.5 dpc (Figure 1), suggesting that  $\epsilon$  is regulated differently than  $\zeta$  and  $\beta_{H1}$ . It is possible that Sox6 has a general effect on embryonic globin genes (and erythrocyte maturation) in addition to a specific role in silencing  $\epsilon$ .

Although most  $p^{100H}$  mutant mice die just after being born, a rare few survive longer. None have been observed to live longer than 2 wk after birth [14]. We examined a single archived sample of liver RNA from a mutant mouse on postnatal day 13.5 for globin gene expression and detected high levels of  $\epsilon$  globin in this RNA sample, compared with undetectable  $\epsilon$  RNA in WT control mice. At this point in development, the levels of  $\zeta$  and  $\beta_{H1}$  RNA were undetectable both in mutant and WT; however, adult  $\beta$ -like globin RNA levels were moderately elevated in the mutant RNA compared with WT (unpublished data), similar to what we observe at 18.5 dpc (Figure 1). These findings suggest that Sox6 continues to function postnatally to silence  $\epsilon$  globin expression and has a unique function in the regulation of  $\epsilon$ -globin. The mechanism by which Sox6 regulates the other embryonic globin genes remains to be elucidated.

Sox6 has other effects in erythropoiesis, including a delay in enucleation/maturation in  $p^{100H}$  mutant mice. This may be the result of indirect effects, such as stress-induced proliferation (resulting from cardiac defects) and/or anemia. Severe anemia can lead to rapid premature release of red cells, prior



to their complete maturation. However, the hematocrit of 18.5-dpc mutant mice is only 20% lower than that of WT (unpublished data), and this mild anemia is probably not sufficient to explain the extent of nucleated red cells. Alternatively, Sox6 itself may play a role in red cell terminal differentiation, as it has been shown to be an important factor in cardiac [15], neuronal [10], astrocytic [11], and cartilage differentiation programs [32,41–44].

The restoration of normal enucleation of red cells in Sox6-deficient mouse by postnatal day 10.5 may result from functional compensation of other Sox proteins (expressed at later developmental stages), since functional redundancy is a recurring theme with Sox proteins [13,45,46]. Moreover, erythropoiesis has already shifted from fetal liver to bone marrow by postnatal day 10.5. The accompanying change in the microenvironment of red cell production may permit normal enucleation. Identification of Sox6 downstream target genes and its interacting proteins will shed light on the role of Sox6 in red cell terminal differentiation and the enucleation process.

Recently, *in vivo* and *in vitro* analyses suggest that reactivation of human  $\epsilon$ -globin would be therapeutically beneficial to adults with sickle cell disease [47], providing a rationale for detailed investigations into the molecular basis of  $\epsilon$ -globin gene silencing. The present study identifies a novel repressor, Sox6, which binds to the  $\epsilon$ y proximal promoter, potentially as part of a larger repression complex. Because murine Sox6 and its human counterpart are 94% identical at the amino acid level [48], it is possible that human Sox6 may also be important in human  $\epsilon$  globin silencing. There is significant sequence homology between the human and mouse  $\epsilon$  promoter regions, and the human promoter contains at least two potential Sox6 binding sites. Indeed, the existence of a silencer of the human  $\epsilon$  globin gene has been proposed [49,50]. Thus, elucidation of the Sox6 repression mechanism and identification of other components of the Sox6-containing complex may further our understanding of  $\epsilon$  globin regulation and potentially reveal additional molecular targets for the treatment of sickle cell anemia and  $\beta$  thalassemias.

## Materials and Methods

**Plasmid construction.** The  $\epsilon$ y promoter deletion reporter plasmid (E-luc) was generated by PCR amplification of the  $\epsilon$ y proximal promoter. A 2.2-kb fragment upstream of the  $\epsilon$ y globin initiation codon (ATG) was used, because it has been shown that all sequences required for  $\epsilon$  gene silencing are located within a 3.7-kb EcoRI fragment containing about 2 kb of sequence upstream of the  $\epsilon$  globin gene cap site [50]. Nucleotide numbering is relative to the transcription start site. The transcriptional start site is based on the longest cDNA in the Phantom (Functional Annotation of Mouse) database. The PCR primers contained XhoI and HindIII sites that were used to clone the  $\epsilon$ y promoter fragment upstream of the firefly luciferase gene in pGL3 Basic (Promega, Madison, Wisconsin, United States). A 2.5-kb SstI/XhoI fragment of  $\mu$ L $\beta$  9.3 (micro LCR; [31]) was then inserted upstream of the  $\epsilon$ y promoter in the above pGL3 Basic plasmid, resulting in a reporter construct in which luciferase expression is driven by the  $\epsilon$ y promoter. A series of deletion constructs of the  $\epsilon$ y promoter were generated similarly. Forward primers with an XhoI site include: MHB1457, 5'CCGCTCGAGTCTAGGCAAACACTCA3' (–2077 to –2052); MHB1503, 5'CCGCTCGAGTCTCTACTGTCACTCCCTG3' (–634 to –605); MHB1505, 5'CCGCTCGAGGGAGCCAAAAAAGAAATGC3' (–197 to –169); MHB1506, 5'CCGCTCGAGCTGACCAATGGCTTCAAG3' (–85 to –58); MHB1532, 5'CCGCTCGAGAATGCAGAACAAAGGTCAGA3' (–63 to –34); and MHB1507, 5'CCGCTCGAGGTCG

CGAAGAATAAAAAGGC 3' (–37 to –9). All forward primers were used in combination with the reverse primer HindIII site: MHB1477, 5'CGGAAGCTTGGGAGGTTGCTGGTGA3' (+45 to +20).

Sox6-pcDNA3.1 [15] was used to overexpress Sox6. A truncated version of the Sox6 overexpression construct (Sox6- $\Delta$ HMG-pcDNA3.1) that lacks the HMG domain was generated, as described by others [32]. Mutagenesis of Sox/Sox6 consensus binding sites of the  $\epsilon$ y promoter were done by PCR. Forward primers used to generate these mutagenized  $\epsilon$ y promoter reporter constructs include: MHB1661, 5'CCGCTCGAGAATGCAGTCCCAAGGGTCAGAACATTGTCTGCGAAG3' (–63 to –19); MHB1662, 5'CCGCTCGAGAAATGCAGAACAAAGGGTCAGATGCTGCGAAGAA3' (–63 to –16); and MHB1663, 5'CCGCTCGAGAATGCAGTCCCAAGGGTCAGATGAGTGTCTGCGAAGAA 3' (–63 to –18).

**Quantitation of globin mRNA.** RNA was first reverse transcribed to cDNA. Primers for cDNA PCR amplification of globin genes were obtained from Primerbank [51]. All primers were searched against the NCBI database to confirm specificity. For  $\epsilon$ y globin: MHB1666, 5'TGGCCTGTGGAGTAAAGGTCAA3'; and MHB1667, 5'GAAGCAGAGGACAAGTCCCA3'. For  $\zeta$  globin: MHB1668, 5'CTACCCCCAGACGAAGACCTA3'; and MHB1669, 5'CTTAACCGCATCCCCTACGG3'. For  $\beta$ H1 globin: MHB1672, 5'TGGACAACCTCAAGGAGACC3'; and MHB1673, 5'ACCTCTGGGGTGAATTCCT3'. For  $\beta$ maj/min globin: MHB1674, 5'ATGGCCTGAATCACTTGGAC3'; and MHB1675, 5'ACGATCATATTGCCAGGAG3'. Using the SYBR green supermix kit with ROX (Bio-Rad, Hercules, California, United States), PCR amplification was run on an ABI7000 (Applied Biosystems, Foster City, California, United States) at the University of Arizona core facility. All PCR was performed in a 25- $\mu$ l reaction with 12.5  $\mu$ l SYBR green supermix. GAPDH mRNA levels were used as control for input RNA. Standard curve analyses were performed to test the efficiency of the amplifications. Triplicates were done for each PCR reaction. Relative quantitative values were calculated in the ABI Prism 7000 SDS Software (Applied Biosystems) and normalized to GAPDH in Microsoft Excel (Redmond, Washington, United States).

**In situ hybridization.** Antisense probes were designed to murine  $\epsilon$ y globin nucleotides 509–584;  $\beta$ maj globin nucleotides 458–549; and mouse Sox6 nucleotides 1353–1927. Embryos were fixed overnight by immersion in 4% paraformaldehyde, embedded in paraffin, sectioned at 5  $\mu$ m, and adhered to charge modified slides (VWR, West Chester, Pennsylvania, United States). Slides were processed for *in situ* hybridization as described [52] using *in vitro* transcribed RNA probes labeled with <sup>33</sup>P. Darkfield and brightfield images were obtained with a Nikon Optiphot microscope (Nikon, Melville, New York, United States) and SPOT RT-Slider digital camera (Diagnostic Instruments, Sterling Heights, Michigan, United States). Objectives used were 1 $\times$  (NA = 0.04) and 10 $\times$  (NA = 0.5). Images were processed, pseudocolored, and combined using Photoshop (Adobe, San Jose, California, United States) software with Fovea Pro (Reindeer Graphics, Asheville, North Carolina, United States) plugins. Original images are available.

**Histology.** 18.5-dpc embryos were exsanguinated and peripheral blood smears were prepared from both mutant and WT mice. The slides were Wright-stained and read by DAF. For whole mount analysis, 14.5-dpc WT and mutant embryos, and postnatal day–10.5 mice were fixed in 10% formalin, paraffin-embedded, sectioned at 5  $\mu$ m, and stained with hematoxylin and eosin. Liver samples (at 14.5 dpc and 18.5 dpc) were prepared in a similar manner. Images were obtained with Nikon Labophot-2 microscope. Objectives used were E Plan 40/0.65 160/0.17 Nikon (40 $\times$  objective), E Plan 100/1.25 oil 160/0.17 Nikon (100 $\times$  objective). The camera was a Nikon Coolpix 4300. Original images are available.

**Northern blot.** A mouse embryonic tissue Northern blot filter (Seegene, Rockville, Maryland, United States) was hybridized with a Sox6 probe generated by RT-PCR (nucleotides 1353–1927) and labeled with [ $\alpha$ -<sup>32</sup>P]dCTP, by random primer labeling (RediprimeII; Amersham Biosciences, Buckinghamshire, England, United Kingdom). The hybridization was performed in phosphate buffered 7% SDS hybridization solution. Blots were washed with 0.2 $\times$  SSC, 1% SDS at 60  $^{\circ}$ C prior to exposure to X-ray film (Kodak, Rochester, New York, United States) at –80  $^{\circ}$ C for 6 d.

**Cell culture and transfection.** GM979 cells (Coriell Cell Repositories, Camden, New Jersey, United States) were cultured in Ham's F12 with 2 mM L-glutamine supplemented with heat-inactivated 10% fetal calf serum (Ivitrogen, Carlsbad, California, United States), penicillin (100 units/ml), streptomycin 100  $\mu$ g/ml, and L-glutamine (2 mM). MEL cells were cultured in DMEM supplemented as above (without heat inactivating the serum). GM979 cells ( $4 \times 10^5$ ) in log phase of growth were transfected with plasmids by FuGENE6 (Roche,

Indianapolis, Indiana, United States). Cells were transfected with  $\epsilon$  promoter reporter constructs (500 ng) along with either empty vector or Sox6 overexpression vector (1000 ng). In assays of dosage effect, we used 200 ng, 500 ng, and 1000 ng. pRL-CMV 15ng (Promega) was used as a control for transfection efficiency.

**Nuclear protein extract and in vitro translation of Sox6.** Nuclear extracts were prepared from MEL cells ( $2 \times 10^7$ ) using a kit (Active Motif, Carlsbad, California, United States). The Sox6 in vitro translation expression vector, tagged with c-Myc and HA, was described before [15]. The translation was performed in a reticulocyte lysate based in vitro translational system (TNT<sup>®</sup> Quick Coupled Transcription/Translation Systems, Promega). A vector without the Sox6 coding sequence was also translated as a negative control.

**Antibodies.** Sox6 antibodies used in this study were either kindly provided by Dr. Enzo Lalli (Université Louis Pasteur, France [7]) or commercially obtained (Catalog No. sc-17332 X, Santa Cruz Biotechnology, Santa Cruz, California, United States). All Sox6 antibodies generated similar results. c-Myc antibody was purchased from Invitrogen. Normal rabbit IgG antibody was obtained from Upstate Biotechnology (Lake Placid, New York, United States).

**EMSA.** Single-stranded complementary oligonucleotides were annealed and end-labeled with [ $\gamma$ -<sup>32</sup>P] ATP with T4 polynucleotide kinase. EMSA was performed with 5  $\mu$ g of nuclear proteins from MEL cells or 3  $\mu$ l of in vitro-translated Sox6 along with the reticulocyte lysate in binding buffer: 100 mM NaCl, 10% glycerol, 200 ng/ $\mu$ l BSA, 50 ng/ $\mu$ l poly (dI-dC) or poly (dG-dC), 10 mM HEPES (pH 7), 0.1 mM EDTA, 0.25 mM DTT, 0.6 mM MgCl<sub>2</sub>. For competition or supershift assays, the indicated unlabelled oligonucleotide competitor (200-fold molar excess) or antibody (3  $\mu$ l) was added 30 min to 60 min prior to addition of radiolabeled probe. Following addition of the radiolabeled probe, the samples were incubated for 30 min or 60 min at room temperature and loaded on a 4% or 6% (w/v) polyacrylamide gel. Electrophoresis was performed at a constant 19 mA for 4–8 h at room temperature, and the gels were dried prior to autoradiography. Antibodies used for supershift analyses included c-Myc, HA, and Sox6 antibodies (described as above). The DNA sequences of the oligonucleotides are as follows (only forward oligos are listed): For the 36-bp WT probe: 5'AATGCAGAACAAGGGTCAGAACATTGTCTGCGAAG3' (MHB1556); for mutant probe 1 (M1): 5'AACAAAGGGTCAGAACATTGTCTGCGAAG3' (MHB1644); for mutant probe 2 (M2): 5'AATGCAGAACAAGGGTCAGAtgagTGTCTGCGAAG3' (MHB1648); for mutant probe 3 (M3): 5'AATGCAGTgcccAAGGGTCAGAACATTGTCTGCGAAG3' (MHB1650).

**ChIP assay.** As described by Nouzova [53], in brief: Cells from MEL cells ( $4 \times 10^7$ ) or fetal liver cells from three 15.5-dpc WT mice were treated with 1% formaldehyde for 10 min at 37 °C, rinsed in ice-cold 1× Hanks' balanced salt solution with 0.1% EDTA containing protease inhibitors, collected by centrifugation at 4 °C, resuspended in a SDS lysis buffer containing protease inhibitors, and incubated on

ice for 10 min. DNA-protein complexes were sonicated to 200 and 600 bp. One-tenth of the sample was set aside for input control, and the remaining sample was precleared with protein A-Sepharose (Amersham Biosciences, Piscataway, New Jersey, United States). Following preclearing, the samples were split into thirds: one sample treated with anti-Sox6, a second treated with normal rabbit IgG, and the third sample without Ab. The last two were used as negative controls. The chromatin-antibody complexes were eluted, and the DNA protein cross-links were reversed with 5 M NaCl at 65 °C for 4 h. Input DNA or immunoprecipitated DNA was used as a template in the PCR reaction. PCR amplification of the  $\epsilon$  promoter was performed and yielded a 172-bp amplicon, corresponding to nucleotides -31 to +140 of the  $\epsilon$  promoter (primers MHB1688, 5'CGAAGAATAAAAGGCCACCA3'; and MHB1689, 5'GCTTCACCACCAACCCTCTC3'). PCR was performed under the following conditions: 95 °C for 15 min followed by 30 cycles at 95 °C for 30 s, 60 °C for 30 s, and 72 °C for 45 s, ending with a final extension at 72 °C for 5 min.

## Supporting Information

### Accession Numbers

Accession numbers for the genes and gene products discussed in this paper are  $\beta$ H1 globin (GenBank NM\_008219),  $\beta$ maj globin (J00413) (in situ),  $\beta$ maj/min globin (NM\_008220) (real time PCR),  $\epsilon$  globin cDNA (NM\_008221) from the Fantom (Functional Annotation of Mouse) database (<http://fantom2.gsc.riken.jp>),  $\zeta$  globin (GenBank NM\_010405), and mouse Sox6 (U32614).

## Acknowledgments

The authors thank Dr. Bernard Futscher for help with the ChIP assays, Dr. Veronique Lefebvre for tissue samples from Sox6 KO mice, and Dr. Dominick DeLuca and Dr. Michael Bulger for helpful discussions. We also thank Rongji Chen and Enrique Claudio for excellent technical support. ZY was a PhD candidate in the Graduate Interdisciplinary Program in Genetics at The University of Arizona and this work was submitted in partial fulfillment of the requirements for the PhD.

**Author contributions.** ZY and OCB conceived, designed, and performed the experiments. ZY, OCB, and MHB analyzed the data. NH, PDK, DAF, DTE, EME, JP, and MHB contributed reagents/materials/analysis tools. JP contributed to the discussion. ZY and OCB wrote the paper.

**Funding.** The authors received no specific funding for this study.

**Competing interests.** The authors have declared that no competing interests exist. ■

## References

- Pevny L, Placzek M (2005) SOX genes and neural progenitor identity. *Curr Opin Neurobiol* 15: 7–13.
- Ferrari S, Harley VR, Pontiggia A, Goodfellow PN, Lovell-Badge R, et al. (1992) SRY, like HMG1, recognizes sharp angles in DNA. *EMBO J* 11: 4497–4506.
- Connor F, Cary PD, Read CM, Preston NS, Driscoll PC, et al. (1994) DNA binding and bending properties of the post-meiotically expressed Sry-related protein Sox-5. *Nucleic Acids Res* 22: 3339–3346.
- van de Wetering M, Oosterwegel M, van Norren K, Clevers H (1993) Sox-4, a Sry-like HMG box protein, is a transcriptional activator in lymphocytes. *EMBO J* 12: 3847–3854.
- Murakami A, Ishida S, Thurlow J, Revest JM, Dickson C (2001) SOX6 binds CtBP2 to repress transcription from the Fgf-3 promoter. *Nucleic Acids Res* 29: 3347–3355.
- Lefebvre V, Li P, de Crombrughe B (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.
- Ohe K, Lalli E, Sassone-Corsi P (2002) A direct role of SRY and SOX proteins in pre-mRNA splicing. *Proc Natl Acad Sci U S A* 99: 1146–1151.
- Connor F, Wright E, Denny P, Koopman P, Ashworth A (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.
- Hamada-Kanazawa M, Ishikawa K, Nomoto K, Uozumi T, Kawai Y, et al. (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, Ishikawa K, Ogawa D, Kanai M, Kawai Y, et al. (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.
- Scheel JR, Ray J, Gage FH, Barlow C (2005) Quantitative analysis of gene expression in living adult neural stem cells by gene trapping. *Nat Methods* 2: 363–370.
- Lefebvre V, Behringer RR, de Crombrughe B (2001) L-Sox5, Sox6 and Sox9 control essential steps of the chondrocyte differentiation pathway. *Osteoarthritis Cartilage* 9 (Suppl A): S69–S75.
- Smits P, Li P, Mandel J, Zhang Z, Deng JM, et al. (2001) The transcription factors L-Sox5 and Sox6 are essential for cartilage formation. *Dev Cell* 1: 277–290.
- Hagiwara N, Klewer SE, Samson RA, Erickson DT, Lyon MF, et al. (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.
- Cohen-Barak O, Yi Z, Hagiwara N, Monzen K, Komuro I, et al. (2003) Sox6 regulation of cardiac myocyte development. *Nucleic Acids Res* 31: 5941–5948.
- Brilliant MH (2001) The mouse p (pink-eyed dilution) and human P genes, oculocutaneous albinism type 2 (OCA2), and melanosomal pH. *Pigment Cell Res* 14: 86–93.
- Thomas JO (2001) HMG1 and 2: Architectural DNA-binding proteins. *Biochem Soc Trans* 29: 395–401.
- Bagga R, Michalowski S, Sabnis R, Griffith JD, Emerson BM (2000) HMG I/Y regulates long-range enhancer-dependent transcription on DNA and chromatin by changes in DNA topology. *Nucleic Acids Res* 28: 2541–2550.
- Dyer MA, Naidoo R, Hayes RJ, Larson CJ, Verdine GL, et al. (1996) A DNA-bending protein interacts with an essential upstream regulatory element of the human embryonic beta-like globin gene. *Mol Cell Biol* 16: 829–838.
- Dyer MA, Hayes PJ, Baron MH (1998) The HMG domain protein SSRP1/

- PREIBF is involved in activation of the human embryonic beta-like globin gene. *Mol Cell Biol* 18: 2617–2628.
21. Drew LR, Tang DC, Berg PE, Rodgers GP (2000) The role of trans-acting factors and DNA-bending in the silencing of human beta-globin gene expression. *Nucleic Acids Res* 28: 2823–2830.
  22. Trimborn T, Gribnau J, Grosveld F, Fraser P (1999) Mechanisms of developmental control of transcription in the murine alpha- and beta-globin loci. *Genes Dev* 13: 112–124.
  23. Li Q, Peterson KR, Fang X, Stamatoyannopoulos G (2002) Locus control regions. *Blood* 100: 3077–3086.
  24. Raich N, Enver T, Nakamoto B, Josephson B, Papayannopoulou T, et al. (1990) Autonomous developmental control of human embryonic globin gene switching in transgenic mice. *Science* 250: 1147–1149.
  25. Harju S, McQueen KJ, Peterson KR (2002) Chromatin structure and control of beta-like globin gene switching. *Exp Biol Med (Maywood)* 227: 683–700.
  26. Tanimoto K, Liu Q, Bungert J, Engel JD (1999) Effects of altered gene order or orientation of the locus control region on human beta-globin gene expression in mice. *Nature* 398: 344–348.
  27. Stamatoyannopoulos G (2005) Control of globin gene expression during development and erythroid differentiation. *Exp Hematol* 33: 259–271.
  28. Forsberg EC, Prohaska SS, Katzman S, Heffner GC, Stuart JM, et al. (2005) Differential expression of novel potential regulators in hematopoietic stem cells. *PLoS Genet* 1: e28. DOI: 10.1371/journal.pgen.0010028
  29. Whitelaw E, Tsai SF, Hogben P, Orkin SH (1990) Regulated expression of globin chains and the erythroid transcription factor GATA-1 during erythropoiesis in the developing mouse. *Mol Cell Biol* 10: 6596–6606.
  30. Zitnik G, Hines P, Stamatoyannopoulos G, Papayannopoulou T (1991) Murine erythroleukemia cell line GM979 contains factors that can activate silent chromosomal human gamma-globin genes. *Proc Natl Acad Sci U S A* 88: 2530–2534.
  31. Forrester WC, Novak U, Gelinas R, Groudine M (1989) Molecular analysis of the human beta-globin locus activation region. *Proc Natl Acad Sci U S A* 86: 5439–5443.
  32. Fernandez-Lloris R, Vinals F, Lopez-Rovira T, Harley V, Bartrons R, et al. (2003) Induction of the Sry-related factor SOX6 contributes to bone morphogenetic protein-2-induced chondroblastic differentiation of C3H10T1/2 cells. *Mol Endocrinol* 17: 1332–1343.
  33. Willing MC, Nienhuis AW, Anderson WF (1979) Selective activation of human beta—but not gamma—globin gene in human fibroblast x mouse erythroleukemia cell hybrids. *Nature* 277: 534–538.
  34. Bowles J, Schepers G, Koopman P (2000) Phylogeny of the SOX family of developmental transcription factors based on sequence and structural indicators. *Dev Biol* 227: 239–255.
  35. Takamatsu N, Kanda H, Tsuchiya I, Yamada S, Ito M, et al. (1995) A gene that is related to SRY and is expressed in the testes encodes a leucine zipper-containing protein. *Mol Cell Biol* 15: 3759–3766.
  36. Kamachi Y, Uchikawa M, Kondoh H (2000) Pairing SOX off: With partners in the regulation of embryonic development. *Trends Genet* 16: 182–187.
  37. Tanabe O, Katsuoka F, Campbell AD, Song W, Yamamoto M, et al. (2002) An embryonic/fetal beta-type globin gene repressor contains a nuclear receptor TR2/TR4 heterodimer. *EMBO J* 21: 3434–3442.
  38. Filipe A, Li Q, Deveaux S, Godin I, Romeo PH, et al. (1999) Regulation of embryonic/fetal globin genes by nuclear hormone receptors: A novel perspective on hemoglobin switching. *EMBO J* 18: 687–697.
  39. Tanimoto K, Liu Q, Grosveld F, Bungert J, Engel JD (2000) Context-dependent EKLf responsiveness defines the developmental specificity of the human epsilon-globin gene in erythroid cells of YAC transgenic mice. *Genes Dev* 14: 2778–2794.
  40. Chase MB, Haga SB, Hankins WD, Williams DM, Bi Z, et al. (1999) Binding of HMG-I(Y) elicits structural changes in a silencer of the human beta-globin gene. *Am J Hematol* 60: 27–35.
  41. Lefebvre V (2002) Toward understanding the functions of the two highly related Sox5 and Sox6 genes. *J Bone Miner Metab* 20: 121–130.
  42. Chimal-Monroy J, Rodriguez-Leon J, Montero JA, Ganan Y, Macias D, et al. (2003) Analysis of the molecular cascade responsible for mesodermal limb chondrogenesis: Sox genes and BMP signaling. *Dev Biol* 257: 292–301.
  43. Ikeda T, Kamekura S, Mabuchi A, Kou I, Seki S, et al. (2004) The combination of SOX5, SOX6, and SOX9 (the SOX trio) provides signals sufficient for induction of permanent cartilage. *Arthritis Rheum* 50: 3561–3573.
  44. Smits P, Dy P, Mitra S, Lefebvre V (2004) Sox5 and Sox6 are needed to develop and maintain source, columnar, and hypertrophic chondrocytes in the cartilage growth plate. *J Cell Biol* 164: 747–758.
  45. Wegner M (1999) From head to toes: The multiple facets of Sox proteins. *Nucleic Acids Res* 27: 1409–1420.
  46. Stolt CC, Lommes P, Friedrich RP, Wegner M (2004) Transcription factors Sox8 and Sox10 perform non-equivalent roles during oligodendrocyte development despite functional redundancy. *Development* 131: 2349–2358.
  47. He Z, Russell JE (2002) A human embryonic hemoglobin inhibits Hb S polymerization in vitro and restores a normal phenotype to mouse models of sickle cell disease. *Proc Natl Acad Sci U S A* 99: 10635–10640.
  48. Cohen-Barak O, Hagiwara N, Arlt MF, Horton JP, Brilliant MH (2001) Cloning, characterization and chromosome mapping of the human SOX6 gene. *Gene* 265: 157–164.
  49. Li Q, Blau CA, Clegg CH, Rohde A, Stamatoyannopoulos G (1998) Multiple epsilon-promoter elements participate in the developmental control of epsilon-globin genes in transgenic mice. *J Biol Chem* 273: 17361–17367.
  50. Raich N, Papayannopoulou T, Stamatoyannopoulos G, Enver T (1992) Demonstration of a human epsilon-globin gene silencer with studies in transgenic mice. *Blood* 79: 861–864.
  51. Wang X, Seed B (2003) A PCR primer bank for quantitative gene expression analysis. *Nucleic Acids Res* 31: e154.
  52. Palis J, Kingsley PD (1995) Differential gene expression during early murine yolk sac development. *Mol Reprod Dev* 42: 19–27.
  53. Nouzova M, Holtan N, Oshiro MM, Isett RB, Munoz-Rodriguez JL, et al. (2004) Epigenomic changes during leukemia cell differentiation: Analysis of histone acetylation and cytosine methylation using CpG island microarrays. *J Pharmacol Exp Ther* 311: 968–981.