Mir-144 selectively regulates embryonic $\alpha$-hemoglobin synthesis
during primitive erythropoiesis

Yan-Fang Fu$^{1,5}$, Ting-Ting Du$^{1,5}$, Mei Dong$^1$, Kang-Yong Zhu$^1$, Chang-Bin Jing$^1$,
Yong Zhang$^1$, Lei Wang$^1$, Hong-Bo Fan$^1$, Yi Chen$^1$, Yi Jin$^1$, Gui-Ping Yue$^4$,
Sai-Juan Chen$^1$, Zhu Chen$^1$, Qiu-Hua Huang$^1$, Qing Jing$^4$, Min Deng$^{1,3}$, Ting Xi Liu$^{1,2,3,*}$

1 Laboratory of Development and Diseases and State Key Laboratory for Medical Genomics and Key Laboratory of Stem Cell Biology, Institute of Health Sciences, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences & Shanghai Jiao Tong University School of Medicine, and Shanghai Institute of Hematology, RuiJin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, 200025

2 Model Organism Division, E-Institutes of Shanghai Universities, Shanghai, 200025

3 Shanghai Stem Cell Institute, Shanghai Jiao Tong University School of Medicine, Shanghai, 200025

4 Laboratory of Nucleic Acid and Molecular Biomedicine, Institute of Health Sciences, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences and Shanghai Jiao Tong University School of Medicine, Shanghai, 200025

5 These authors contribute equally to this work

* Correspondence: Ting Xi Liu, Laboratory of Development and Diseases, Institute of Health Sciences (IHS), Room 408, Building 1, 225 South Chong Qing Road, Shanghai 200025. Tel: 86-21-63857025; Fax: 86-21-63857029; E-mail: txliu@sibs.ac.cn

Key words: microRNA; miR-144; Krüppel-like transcription factor; $\alpha$-hemoglobin; Erythropoiesis; Zebrafish
Abstract

Precise transcriptional control of developmental stage-specific expression and switching of \( \alpha \) and \( \beta \)-globin genes is significantly important to understand the general principles controlling gene expression and the pathogenesis of thalassemia. Although transcription factors regulating \( \beta \)-globin genes have been identified, little is known about the microRNAs and trans-acting mechanism controlling \( \alpha \)-globin genes transcription. Here, we show that an erythroid lineage-specific microRNA gene, \( \text{miR-144} \), expressed at specific developmental stages during zebrafish embryogenesis, negatively regulates the embryonic \( \alpha \)-globin, but not embryonic \( \beta \)-globin genes expression, through physiologically targeting \( \text{klfd} \), an erythroid-specific Krüppel-like transcription factor. Klfd selectively binds to the CACCC boxes in the promoters of both \( \alpha \)-globin and \( \text{miR-144} \) genes to activate their transcriptions, thus forming a negative feedback circuitry to fine-tune the expression of embryonic \( \alpha \)-globin gene. The selective effect of \( \text{miR-144-Klfd} \) pathway on \( \text{globin} \) gene regulation may thereby constitute a novel therapeutic target for improving the clinical outcome of patients with thalassemia.
Introduction

Over the past 50 years, studies on the hemoglobin synthesis during erythropoiesis have led to considerable understanding of not only the general principle controlling eukaryotic gene expression, but also the regulatory mechanism underlying the pathogenesis of human genetic diseases such as thalassemia. All vertebrates studied to date possess multiple globin genes that are expressed in different developmental stages, to generate different hemoglobins with distinct oxygen affinities during development. In human, α-globin cluster contains three functional α-globin genes (ζ, α1 and α2) that are differentially expressed during embryonic or yolk sac (ζ), and adult (α1 and α2) erythropoiesis. The β-globin gene cluster consists of five genes that are also developmentally regulated: ε expressed in yolk sac, Gγ and Aγ in fetal liver, and δ and β in bone marrow. Although the erythroid- and developmental stage-specific expressions of globin genes have been known for decades, the precise molecular mechanisms directing the exquisite regulation and switching remain largely unknown.

Several types of regulatory mechanisms have been implicated in the embryonic and adult globins expression. Structural removal of genetic materials containing either α- or β-globin genes, deletion of their respective upstream cis-regulatory elements (e.g., HS40 for α-globins and locus control region for β-globins, respectively), and gain-of-function mutation situated between the α-globin genes and their cis-regulatory elements have been shown to contribute to the lack and reduced expression of corresponding hemoglobins and accounts for a high percentage of patients with α- and β-thalassemia.

Trans-acting mechanisms involving both erythroid-specific and general transcriptional regulators are also crucial for the developmental control of β-globin
Transcription factors selectively regulating stage-specific expression of \( \beta \)-\textit{globin} genes have been identified. EKLF, an erythroid Krüppel-like factor, specifically interacts with the CACCC boxes of \( \beta \)-, but \( \gamma \) or \( \epsilon \)-genes.\(^6\) In EKLF-deficient mice, the adult \( \beta \)-\textit{globin} is markedly decreased, without affecting embryonic \( \epsilon \gamma \)-, \( \beta h1 \)- and \( \alpha \)-\textit{globins} expression.\(^7\) FKLF (KLF11) interacts with the CACCC box of \( \gamma \)gene to activate its transcription,\(^8\) while LKLF (KLF2) is required for \( \epsilon \) expression \textit{in vivo}.\(^9\) It has been proposed that erythroid-specific transcription factors (EKLF, GATA-1 and FOG-1) bind to the CACCC box of \( \beta \)-\textit{globin} gene promoters, providing individual \textit{globin} gene family members with a competitive advantage for the interaction with the \textit{cis-locus control region} to regulate their expression.\(^10\) Although these advances in understanding \( \beta \)-\textit{globin} gene regulation at the transcriptional level, little is known about \textit{trans}-acting factors regulating \( \alpha \)-\textit{globin} expression.\(^1\)

MicroRNAs (miRNAs) are endogenous ~22 nt RNAs that play important regulatory roles at the posttranscriptional level in animals and plants by targeting mRNAs for cleavage or translational repression.\(^11\) Currently, the target genes of miRNAs are mainly identified by a combination of bioinformatic searches for potential miRNA recognition element (MRE) in the 3’UTR of the target gene and subsequent experimental validations of predicated miRNA-target interactions conducted usually with luciferase reporter assay in cultured cells \textit{in vitro}.\(^11,12\) Although these progresses, \textit{in vivo} validations of miRNA targets in a physiological context with more reliable approaches represents a significant challenge in dissecting the miRNA-mediated regulatory pathways in development and diseases.\(^12\) Three miRNAs (\textit{miR-221, miR-222, miR-24}) required for the development of
mature erythrocytes through targeting c-kit, activin type I receptor ALK4 in vitro, respectively, have been identified, and a recent report show that miR-451 is a direct target of the critical hematopoietic transcription factor GATA-1. However, the miRNA-mediated signaling posttranscriptionally regulating hemoglobin gene expression in vivo remains largely unknown.

Zebrafish is an ideal system for the study of embryonic erythropoiesis. Embryonic hematopoiesis in the zebrafish begins in two anatomically and functionally distinct regions: the anterior- and posterior-lateral plate mesoderm (A- and P-LPM). Cell fate mapping experiments at the shield stage of development show that the P-LPM arises from the most ventral region opposite the embryonic organizer (or shield). Cells in the P-LPM migrate medially in an anterior-to-posterior wave to form the intermediate cell mass (ICM). The ICM is the major site of primitive erythropoiesis (equivalent of mammalian yolk sac), with gata-1 being exclusively expressed in this region during somitogenesis. Five zebrafish embryonic α- and β-globins (α-E1, α-E3, β-E1, β-E2 and β-E3), and several adult α- and β-globin genes (α-A1, β-A1 and β-A2) have been characterized, and adult globins do not appear to be expressed at high level during embryogenesis.

In this study, we show that an erythroid-specific miRNA, miR-144, is expressed at the specific developmental stages between 20-32 and 72-120 hours postfertilization (hpf), to negatively regulate embryonic α-hemoglobin (α-E1), but not β-hemoglobin genes expression, through targeting the 3’UTR of Krüppel-like factor D (klfd) gene in a physiological context. The Klfd interacts with the CACCC sites in the promoters of either α-E1 globin or miR-144 gene, forming an intricate feedback circuitry to selectively
regulate embryonic \(\alpha\)-globin expression during primitive erythropoiesis.

Materials and methods

Fish care
Zebrfish maintenance, breeding, and staging were performed as described previously. The zebrafish facility and zebrafish study were approved by the institutional review board (IRB) of Institute of Health Sciences, Shanghai Institutes of Biological Sciences, Chinese academy of Sciences.

Cloning and plasmid construction
All primer sequences and plasmid constructions are available in the Table S3. Target sequences for 2\times PT, 2\times MRE and 3\times MRE were synthesized as DNA oligonucleotides (Invitrogen) and cloned into the 3’UTR of pCS2\(+\)-EGFP plasmid.

Locked nucleic acid (LNA), miRNA duplexes, morpholinos, capped mRNA synthesis
The \(miR-144\) LNA-modified oligonucleotides labeled with or without digoxigenin were purchased from Takara and Exiqon. The \(miR-144\) duplex and its 2-base pair mismatch control were synthesized from Invitrogen. The \(kfl\) morpholino antisense oligonucleotides and its 5-mismatch control were purchased from GENE TOOLS. All sequences were available in Table S3. Capped mRNAs were synthesized with the Massage machine kit according to manufacturer’s instructions (Ambion).

Whole-Mount miRNA and mRNA In Situ Hybridization and Apoptosis Assay
Whole-mount in situ hybridization with Dig-labeled LNA and mRNA antisense probes were performed as described previously.\textsuperscript{22,23} Dig-labeled miR-144 LNA probe and fluorescein-labeled gata-1 or myo D probe were used for two-color in situ hybridization\textsuperscript{24} at 56 °C. The purple color was developed first with NBT/BCIP as substrate (Vector Laboratories), followed by the development of red color with Fast Red tablets (Roche). Whole-mount TUNEL assays was performed as described previously.\textsuperscript{24,25}

**Generation of anti-Klfd antibody**

A rabbit polyclonal antiserum against zebrafish Klfd was generated using a His-Klfd (N-terminal) fusion protein as an antigen source. A 5’ fragment was amplified by PCR from the full-length klfd cDNA. The resulting PCR product of 810 bp encoding the amino-terminal 270 amino acids of the Klfd polypeptide was digested with BamH1/Xho1 and subcloned into the BamH1/Xho1 sites of the plasmid pET-28a (+) vector. A His-Klfd (N-terminal) fusion protein of the expected size (34 kDa) was expressed in E. coli strain DH5α, purified with Ni-NTA Agarose (Qiagen), and used to immune a rabbit (AbMART).

**Immmunochemistry and Western blot**

Whole-mount immunostaining was performed as described previously,\textsuperscript{26} with a rabbit polyclonal antibody against phosphorylated histone 3 (pH3). For Western blot, embryos at 16 and 22 hpf were deyolked as described previously.\textsuperscript{27} Embryos were homogenized in lysis buffer (20 mM Tris HCl [pH 7.4], 150 mM NaCl, 5 mM EDTA, 10% glycerol and 0.1% Triton X-100) containing protease inhibitor cocktail and phosphatase inhibitor
Signals were detected with mouse anti-actin antibody (1:10,000) and rabbit anti-zebrafish Klfd (1:2,500) for 2 hours at room temperature or overnight at 4°C, followed by incubation with appropriate horseradish peroxidase conjugated-secondary antibody (1:10,000) and enhanced chemiluminescence kit (Cell Signaling). Band densities were quantified with Image-Pro Plus software.

**Northern analysis**

Embryos were collected at 18, 24 and 30 hpf during development. For LNA-mediated miR-144 knockdown, one-cell stage embryos were injected with 0.3 ng of either miR-144 LNA or 5-base pair mismatch control. Total RNA from 160-300 embryos was extracted with Trizol (Invitrogen), which was subsequently enriched with mirVana miRNA isolation kit (Ambion) and separated on 15% of Urea-PAGE gel. Northern blot were probed with Dig-labeled miR-144 or miR-1 LNA probe as described previously, and was visualized using DIG luminescent detection kit for nucleic acids (Roche).

**Establishment of Tg(zgata-1:EGFP), Tg(zgata-1:EGFP-2×PT) and Tg(zgata-1:EGFP-klfd 3’UTR) transgenic lines**

Transgene plasmids flanked by I-Sce1 sites were prepared with endotoxin-free kit (Promega). Microinjection was performed at one-cell stage embryos with 2 nl of injection solution containing 50 pg/µl of DNA, 0.5 × I-Sce1 buffer and 0.5 units/µl I-Sce1 meganuclease (New England Biolabs). Injected embryos were raised to sexual maturity (F0 founders) and crossed to wild-type zebrafish to generate F1 progeny, which were
screened for EGFP expression in the P-LPM and ICM at 16 and 24 hpf, respectively. The EGFP+ F1 embryos were raised to adults to establish the stable transgenic lines.

Single-embryo semi-quantitative and real-time quantitative RT-PCR

Single-embryo RT-PCR in multiple individual embryos was performed as described.25 PCR conditions were as follows: 50°C, 30 min; 95°C, 10 min; 30 cycles of 94°C, 30 s; 58°C, 30 s; 72°C, 1 min; then 72°C, 10 min. PCR products were separated on a 1.5% agarose gel. Real-time quantitative RT-PCR was performed using SYBR Green (TOYOBO Co). Zebrafish gapdh housekeeping gene was used as an internal control. Primer sequences are available in Table S3.

Bioinformatics

3’UTR of 118 ICM-expressing genes (http://zfin.org) were retrieved from RefSeq database (http://www.ncbi.nlm.nih.gov/RefSeq/). Bioinformatics analyses of miRNA binding sites were performed with miRanda v1.0b30 and RNAhybrid.31 The free energy (ΔG) of sequences flanking the miR-144 MRE was determined using mFold.32 Syntenic analysis was performed as described previously.25

Embryonic chromatin immunoprecipitation (E-ChiP)

E-ChiP was performed as previously described with minor modifications.33 Please see details in Supplemental Material and Methods.

Results
*miR-144* is evolutionarily conserved and expressed in erythroid progenitors at specific developmental stages

Previous study shows that *miR-144* is expressed in the ICM of 24 hpf zebrafish embryos. To determine the role of *miR-144* during erythropoiesis, we examined the spatio-temporal expression of *miR-144* by whole-mount miRNA *in situ* hybridization (WISH) using a digoxigenin (Dig)-labeled locked nucleic acid (LNA)-modified DNA oligonucleotides as a probe. While *miR-144* was not expressed maternally (Figure S1), the earliest zygotic transcription of *miR-144* could only be detected in the ICM of 20 hpf embryos, then became undetectable at 36 and 48 hpf, and redetected in the posterior blood island at 72 and 120 hpf (Figure 1A and Figure S1, arrows). The *miR-144* was undetectable in the homozygous *cloche* (*clom39*) embryo (Figure 1B, arrows, comparing with its sibling), a mutant that lacks hematopoiesis and vasculogenesis due to an as-yet-unidentified mutation upstream of the early hematopoietic transcription factor *scl*, and was co-localized with the erythroid progenitor marker *gata-1* in the ICM (Figure 1C and inset). The zebrafish *miR-144* shares the same nucleotide acid identities with other species, and six zebrafish genes, along with *miR-144* and *miR-451*, define a ~180 Kb genomic region on chromosome 5 (Figure 1D right), which is syntenic to the human *MiR-144* genomic locus on chromosome 17 (Figure 1D left). Taken together, the results indicate that zebrafish *miR-144* is an evolutionarily conserved orthologue of human *MiR-144*, expressed in erythroid progenitors under tight developmental regulation.

**Endogenous miR-144 is functionally active in vivo**

To determine whether the *miR-144* is functional for gene silencing, we first examined
whether the mature miR-144 is expressed in vivo. Northern blot analysis showed that the endogenous mature miR-144 (22 nt in length) was not detected in embryos at 18 hpf, but readily detected at 24 hpf and decreased at 30 hpf (Figure 1E), which was consistent with the WISH result (Figure 1A). To test whether the mature miR-144 is functional in physiological context, we established two stable “reporter” transgenic lines under the erythroid-specific gata-1 promoter.35 The two lines were designated as Tg(zgata-1:EGFP) and Tg(zgata-1:EGFP-2×PT) (Figure 2A,B; Table S1). The Tg(zgata-1:EGFP-2×PT) carried two tandem perfectly complementary target sites (2×PT) for miR-144 binding in the 3’UTR of EGFP reporter gene (Figure 2B red boxes).

The EGFP transcripts and EGFP protein expression appeared normal in the P-LPM of 16 hpf F2 embryos derived from both Tg(zgata-1:EGFP) and Tg(zgata-1:EGFP-2×PT) lines (Figure 2A,B, left panels, red arrows). By 24 hpf, however, a dramatic reduction of EGFP expression was observed only in the ICM region of the Tg(zgata-1:EGFP-2×PT), but not in the Tg(zgata-1:EGFP) embryos (Figure 2A,B right panels). In contrast, the EGFP expression was not affected in the dorsal neurons of both transgenic lines (Figure 2A,B, black arrowheads), The observation was consistent with the facts that endogenous miR-144 was only expressed in the ICM after 20 hpf (Figure 1A,E), and that a miRNA will direct mRNA cleavage if the target transcript is perfectly complementary in sequence.11 Together, these results indicate that the mature miR-144 is functionally active in the developing erythroid progenitors in vivo.

Efficient and specific ablation of miR-144 function in vivo by locked nucleic acid

To examine the function of miR-144 in erythroid development, we took advantage of
LNA technology, which has been demonstrated to efficiently and specifically silence miRNA functions in vitro and in vivo. We first established a transient reporter system in which the synthetic miR-144 duplex was co-injected with either EGFP-2×PT or EGFP-2×PTmut capped mRNA sensors (mut: the first 2-8 nt at the 5’ seed matching region of miR-144 has been changed from 5’-ACAGUAU-3’ to 5’-ACAUGAU-3’) into one-cell stage embryos. The EGFP fluorescence was dramatically reduced in the 22 hpf embryos injected with miR-144 duplex plus EGFP-2×PT (Figure S2A,B), but remained unaffected in the embryos injected with miR-144 duplex plus EGFP-2×PTmut mRNAs (Figure S2C). Consistently, injection of EGFP-2×PT mRNAs into wild-type embryos resulted in a disappearance of EGFP transcripts specifically in the ICM of 22 hpf embryos (Figure S2D arrow); while the appropriate EGFP expression was observed in the ICM of the embryos injected with EGFP-2×PTmut sensor (Figure S2E arrow). We then assessed the specificity and dosage of the antisense LNA-modified oligonucleotides that could effectively block the miR-144 function in vivo. Injection of 0.3 ng miR-144 LNA, but not its 5-base pair mismatch LNA control into one-cell stage embryos, was able to completely rescue the EGFP expression suppressed by co-injection of the capped EGFP-2×PT mRNAs and miR-144 duplex (Figure S2F,G). Furthermore, the miR-144 LNA, but not its LNA control, was able to completely rescue the EGFP transcripts and protein fluorescence suppressed by endogenous miR-144 gene in the ICM of F2 embryos derived from the Tg(zgata-1:EGFP-2×PT) transgenic line at 24 hpf (Figure 2C,D arrows), due to an efficient and specific knockdown of the endogenous mature miR-144 as evidenced by Northern blot analysis (Figure S2H). Thus, we injected miR-144 LNA at the dosage of 0.3 ng per embryo to investigate the phenotypes resulting from a knockdown of
the miR-144 gene.

Selective increase of embryonic α-globin expression in miR-144-deficient embryos during primitive erythropoiesis

The miR-144 LNA-injected embryos appeared morphologically normal, compared with LNA control-injected embryos. The earliest phenotype can be observed at 22 hpf, when 32% of miR-144-knockdown embryos showed a slight cellular accumulation in the posterior ICM (Figure 3A, middle panel, arrow, n=460/1439). In approximately 8% of miR-144 knockdown embryos, a more obvious ICM expansion was also observed (Figure 3A, right panel, arrow, n=115/1439). To molecularly characterize the phenotypes, we performed WISH analyses with twelve hematopoietic- and vascular-specific genes for different stages of primitive hematopoiesis (scl for stem cells; gata-1, fog-1 and alas2 for erythroid progenitors; α-E1, α-E3, β-E1 and β-E2 hemoglobin for mature red cells; pu.1, mpo and l-plastin for granulocytes; and flk-1 for vascular endothelial cells) at 22, 28, 33 and 48 hpf. Surprisingly, only significant increase (43%, 119/275) (Figure 3B, middle panel, arrow), and massive induction (8%, 23/275) (Figure 3B, right panel, arrow) of α-E1 globin expression were observed in the miR-144-knockdown embryos at 22 hpf and later stages (Figure S3,S4). The alas2 transcripts increased slightly (29%, 14/49) (Figure 3C, middle panel, arrow), with about 27% (13/49) having an ectopic expression observed in the posterior blood island (Figure 3C, right panel, arrowhead). A slight increase was also detected for α-E3 globin (Figure 3D arrow). No obvious changes in the levels of β-E1 and β-E2 globins were noted at 22 hpf and later stages (Figure 3E,F and Figure S4). Other genes gata-1, fog-1 and scl (Figure
3G,H and Figure S5), *pu.1*, *mpo*, *l-plastin* and *flk-1* (Figure 3I and Figure S6) all expressed appropriately at 22 hpf and later stages after the onset of circulation. Semi-quantitative single-embryo and real-time RT-PCR confirmed the unbalanced expression that the α-E1 transcripts were significantly increased relative to other genes such as β-E2 and *mpo* (Figure 3J and Figure S7). The possibilities that abnormal dorsoventral patterning, increased proliferation or decreased apoptosis of erythroid progenitors appeared unlikely to account for the up-regulation of α-E1globin, because other hematopoietic-specific genes expressed normally, and no obvious changes in the levels of the mitotic marker phosphorylated histone 3 (pH3), and apoptotic TUNEL staining were observed (Figure 3K,L).

*miR-144 physiologically targets klfd, an erythroid-specific krüppel-like transcription factor*

Individual miRNAs appear to regulate the expression of numerous targets posttranscriptionally, and paired expression profiles of miRNAs and mRNAs can be used to identify functional miRNA-target relationships with high precision. To identify the potential physiological targets for *miR-144*, we first screened 3′UTR of 118 ICM-expressing genes registered at ZFIN database (http://zfin.org/) for potential *miR-144* recognition elements (MRE), using RNAhybrid and miRanda computational algorithms. We identified 11 genes as putative *miR-144* targets (*alas2*, *cebpα*, *gata-1*, *lmo2*, *mcam*, *pfkfb4*, *runx1*, *scl*, *tfe3a*, *klf12* and *klfd*), most of which carried a putative MRE (Table S2). We also used mFold to analyze the local mRNA secondary structure and the free energy (ΔG) of 70 bp 5’ and 3’ sequences flanking the putative MRE as
described previously, and found that most of the putative targets had a relatively high ΔG (Table S2). To test whether these putative MREs could be responsible for silencing their expression by miR-144, we placed multimers of about 40-70 bp containing the predicted miR-144 MRE (2-3 copies) from the lmo2, alas2, klf12 and klf3 3’UTRs into the 3’UTR of an EGFP reporter plasmid. We microinjected the in vitro synthesized capped EGFP, EGFP-2×PT and EGFP-2 or 3×MREs mRNAs individually with either control or miR-144 duplex into one-cell stage zebrafish embryos (DsRed mRNAs were co-injected as an injection control), and measured the level of EGFP fluorescence to determine the effect of miR-144 on EGFP translation at 22 hpf. While no changes in EGFP fluorescence were observed for the EGFP and EGFP-2 or 3×MREs of lmo2, alas2 and klf12 (Figure S8A and S8C-E), a dramatic suppression of EGFP expression was obtained in the embryos injected with control EGFP-2×PT and EGFP-2×MRE-klfd mRNAs (Figure S8B,F), suggesting that the MRE of klf3 3’UTR can be efficiently targeted by miR-144. Interestingly, the miR-144 MRE was located in a 3’UTR region with highest free energy ΔG (5’ ΔG: -6.90 and 3’ ΔG: -3.80 kcal/mol⁻¹), compared with other predicated target genes (Table S2), suggesting the local accessibility of the klf3 MRE and a strong predictive value of the ΔG index.40

It has been shown that an ectopic expression of 3’UTR and miRNA may not be reliable and does not guarantee a functional interaction. To validate that the klf3 3’UTR is a bona fide endogenous target for miR-144 in physiological context, we placed the full-length of klf3 3’UTR into the 3’UTR of EGFP reporter gene under the erythroid-specific gata-1 promoter. We screened 138 microinjected zebrafish adults and established 10 transgenic F0 founder lines, which were referred to as Tg(zgata-1:EGFP-klf3
F1 and F2 embryos derived from these transgenic lines all expressed EGFP specifically in the ICM, with some ectopic EGFP expression observed in the dorsal neurons, skin and hatching gland (Table S1). While expression of both EGFP mRNA (by WISH) and EGFP protein (by Western blot) were normally detected in the control Tg(zgata-1:EGFP) embryos at 16 and 24 hpf (Figure 4A, arrows; Figure 4C, left), we observed a marked decrease in the levels of both EGFP mRNA and EGFP protein in the ICM of 24 hpf, but not in the P-LPM of 16 hpf embryos derived from the line 5 of Tg(zgata-1:EGFP-klfd 3’UTR)line 5 (Figure 4B, arrows; Figure 4C, right). The reduction of EGFP expression can be rescued by microinjection of miR-144 LNA (Figure 4D), but not by its 5-base pair mismatch LNA control (Figure 4E). The same results were also obtained from F2 embryos derived from additional lines 1, 6 and 8 of Tg(zgata-1:EGFP-klfd 3’UTR)lines 1, 6, 8 (Figure S9). No changes were observed for EGFP expression in the dorsal neurons, where the miR-144 is absent (Figure 4A,B arrowheads).

Furthermore, western blot analysis using a rabbit polyclonal antibody specifically against zebrafish Klfd protein (see below) showed that the endogenous Klfd protein level was significantly increased in miR-144 knockdown embryos at 22 hpf (Figure 4F lanes 1-2). It was noted that the EGFP mRNAs appeared to be degraded by miR-144, a phenomenon that has been observed in the zebrafish primordial germ cells.\(^{41}\) Taken together, we concluded that the transcription factor klfd is a physiological target of miR-144.

Klfd selectively binds to \(\alpha-E1\) hemoglobin promoter at specific stage to activate its transcription
Consistent with the observation of klfd as a target gene for miR-144, WISH analyses showed that the spatio-temporal expression of klfd is similar to that of miR-144. The earliest expression of klfd was detected at 18 hpf (Figure 5A arrow), but not at 16 hpf (data not shown), and remained the expression throughout 48 hpf (Figure 5A) and later stages as observed previously.\(^{42}\) Injection of klfd mRNA (20 pg) into one-cell stage embryos significantly increased \(\alpha\)-E1 expression (48%, 17/35), with a slight increase, if any, in the levels of alas2 transcripts (Figure 5B and Figure S10). The expression of \(\beta\)-E2, gata-1, mpo, and mesodermal myo D and ectodermal krox 20 genes was not altered by klfd overexpression at 22 hpf (Figure 5B and Figure S10), consistent with the phenotypes observed in miR-144-deficient embryos. Western blot analysis using a rabbit polyclonal antibody specifically against zebrafish Klfd showed that injection of a klfd-specific morpholino oligonucleotides (Figure S11A,B) reduced endogenous Klfd protein levels in a dose-dependent manner (Figure 5C lanes 3-5), with 90% loss of Klfd protein being observed after injection of 12 ng per embryo of the same morpholino (Figure 5C lane 5). By contrast, injection 12 ng of the 5-base pair mismatch control morpholino (5-MM) had no effect on the level of Klfd protein (Figure 5C lane 2). As a result, a significant decrease in the levels of \(\alpha\)-E1 (25%, 12/49), but not \(\beta\)-E2, gata-1 and mpo transcripts was observed in the 22 and 36 hpf of Klfd-knockdown embryos (Figure 5D and data not shown). Furthermore, O-dianisidine staining showed that the embryonic hemoglobin syntheses were decreased in the Klfd-deficient embryos at 36 (96%, 24/25) and 54 hpf (94%, 33/35), compared with 5-MM morpholino-injected embryos (Figure 5E arrows). Consistently, knockdown of Klfd is also able to significantly rescue the overproduction and ectopic expression of \(\alpha\)-E1 globin induced by miR-144 deficiency. More than 95% of miR-144 and Klfd double-knockdown embryos demonstrated normal levels of \(\alpha\)-
E1 globin (Figure 5F, right panel, n=73/78), while about 38% of miR-144-knockdown embryos (n=27/71) demonstrated a significant increase of α-E1 globin in the ICM and posterior blood island (Figure 5F, left panel, arrow, n=27/71). These results strongly suggest that α-E1 globin is one of the target genes downstream of Klfd transcription factor.

Previous studies show that CACCC box is essentially elements for binding of KLF family transcription factors.6,42 Because six consecutive CACCC boxes occupied ~3.0 kb promoter upstream of the α-E1 globin translational start site (Figure 5G), we therefore test whether Klfd is able to directly bind to these consensus elements by whole-embryo chromatin immunoprecipitation (E-ChiP) assay.33 Chromatin fragments were extracted from 16 and 22 hpf wild-type embryos, and were immunoprecipitated with the same rabbit polyclonal anti-Klfd antibody. DNA from the immunoprecipitates was amplified by PCR using primers located to the indicated genomic regions (I-III) (Figure 5G horizontal arrows). The results showed that the Klfd was richly associated with the promoter region “I” at 22 hpf, but not at 16 hpf (Figure 5G), consistent with the observation that the earliest klfd expression was detected at 18 hpf (Figure 5A). No binding activity was measurable for regions “II” and “III” at either 16 or 22 hpf (Figure 5G). With the same chromatin preparations, no binding reactivity was observed in the CACCC boxes of β-E1 (Figure 5G regions IV and V), alas2 (Figure 5H regions I and II) and β-E3 globin promoters (Figure 5I regions III and IV) in 16 and 22 hpf embryos analyzed. The results suggest that the Klfd only interacts with the distal CACCC sites on the α-E1 globin promoter to positively regulate its expression in developing erythrocytes.

Transactivation of miR-144 promoter by Klfd
We hypothesized that Klfd may bind and activate the miR-144 promoter based on three observations: firstly, the earliest expression of klfd was detected at 18 hpf (Figure 5A), two hours before the appearance of miR-144 transcripts in the ICM (Figure 1A). Secondly, the miR-144 expression was abolished in Klfd-deficient embryos (52%, 26/50) (Figure 6A arrow), but significantly elevated in the embryos over-expressing 20 pg of klfd mRNAs (82%, 34/41) (Figure 6B arrows). Finally, promoter analysis found that there were 12 consecutive CACCC sites occupying ~5.0 Kb upstream of the miR-144 transcriptional start site (TSS) (Figure 6C). To test the hypothesis, we performed E-ChiP again with the same preparations of immunoprecipitated chromatin fragments used in the Figure 5G-I, and amplified by PCR using primers located to the indicated regions (I-V) (Figure 6C). The Klfd was found predominantly enriched in the proximal promoter region “V”, but not in the middle and distal regions represented by I-IV at 22 hpf (Figure 6D right). No detectable binding activity was observed in 16 hpf embryos (Figure 6D left), when neither klfd nor miR-144 was expressed (Figure 1A and Figure 5A). Taken together, the results strongly suggested that the interaction of Klfd with the proximal miR-144 promoter contributed to the stage-specific expression of miR-144.

Discussion

In this study, we show that an erythroid-specific miR-144 is expressed at specific developmental stages between 20-32 and 72-120 hpf during zebrafish embryogenesis. The precise onset of miR-144 expression at 20 hpf appears to be dependent on Krüppel-like factor D (klfd), a zinc finger transcription factor that is first expressed in erythroid ICM at 18 hpf. The Klfd binds to the proximal CACCC boxes on the miR-144 promoter.
to activate its transcription. The *miR-144* targets Klfd physiologically, hence forming a feedback circuitry that acts selectively to regulate the embryonic *α-E1*, but not *β-globins* expression during the late stage of primitive erythropoiesis (Figure 6E). Interestingly, a slight increase of *alas2* (a speed-limited enzyme in the heme biosynthesis), *α-E3* and *β-E2*, are also observed upon *miR-144* deficiency. Because *miR-144* does not target the 3’UTR of *alas2* (Figure S8) and *α-E3* and *β-E2*, and no evidence indicates that Klfd binds to the CACCC boxes of *alas2* promoter (Figure 5H), the possibility that the *miR-144* might target additional gene(s) other than *klfd*, or a compensatory mechanism secondary to the upregulation of *α-E1* globin, cannot be excluded. Notably, about 8% of *miR-144* knockdown embryos show drastically expansion of *α-E1* expression outside of the expression domain of *miR-144*. Because no increase was observed for the mitotic marker phosphorylated histone 3 (pH3) in the ICM, it appears to be unlikely that the expansion is due to the hyper-proliferation of erythropoietic cells. It is possible that the expression of *α-E1* might be stimulated in non-blood cells through a cell non-autonomous mechanism, or very weak expression of *miR-144* in non-blood cells which cannot be detected by ISH. Furthermore, the possibility of a local lysis of red blood cells due to the unbalanced synthesis of *α- and β-globin*, and a consequent release of *α-E1* transcript from the lytic red blood cells into the intercellular environment cannot be excluded.

Previous study suggests that unidentified CACCC box-interacting factors distinct from Sp1-4, EKLF (KLF1) and BKLF (KLF3), may contribute to the developmental expression of embryonic *ζ-globin*, the mammalian homolog of zebrafish *α-E1 globin*. In zebrafish, Klf4 (also known as Biklf) has been shown to preferentially interact with the CACCC box in the embryonic *β-globin*, but not the embryonic *α-globin* gene promoter,
and knockdown of Klf4 by morpholino results in decreased expression of embryonic $\beta$-E3 globin.\textsuperscript{44,45} Thus, the Klfd identified in this study represents the first transcription factor selectively regulating embryonic $\alpha$-globin expression. The results suggest that members of Krüppel-like factor family appear to be conserved regulators of both $\alpha$- and $\beta$-globins expression during vertebrate evolution. Because $\text{miR-144}$ is also detected at 3 and 5 dpf in the posterior blood island, a site for definitive hematopoiesis\textsuperscript{46}, it can be anticipated that the $\text{miR-144-Klfd}$ signaling is also involved in the adult $\alpha$-globin regulation, but can not be addressed in the present study due to the short lifespan of $\text{miR-144}$ LNA and $\text{klfd}$ morpholino oligonucleotides in zebrafish embryos. A zebrafish or murine line with stable $\text{miR-144}$ inactivation or erythroid-specific overexpression should help to further clarify this issue.

The zebrafish $\text{klfd}$ gene was originally cloned in the year of 2001,\textsuperscript{42} and its mammalian orthologue was not yet identified.\textsuperscript{42,44} Protein sequence alignment and synteny analyses suggest that although zebrafish Klfd does not have a known ortholog in mammalian genome, it is most closed to mammalian KLF1 and KLF2 by sequence and syntenic analyses (Figure S12),\textsuperscript{42} and may be related to an ancestor of these genes. In addition, like murine $\text{KLF1}$, zebrafish $\text{klfd}$ is expressed within both embryonic erythrocytes and definitive erythrocytes (Figure 5A).\textsuperscript{42} Although it is initially reported that KLF1 does not affect embryonic $\zeta$-globin gene expression,\textsuperscript{7} a recent study shows that $\text{KLF1}^{-/-}$ embryos have about 3-fold lower $\zeta$-globin gene expression than wild-type.\textsuperscript{47} The similar effects of zebrafish Klfd and mammalian KLF1 on the regulation of zebrafish embryonic $\alpha$-E1 globin and mammalian embryonic $\zeta$-globin, respectively, suggest that zebrafish $\text{klfd}$ might be the functional ortholog of mammalian $\text{KLF1}$, which warrants
further study.

The evolutionary conservation and close genomic position between miR-451 and miR-144 (92 bp in between) also suggests that the two miRNAs might share the same set of transcriptional regulatory machinery. Interestingly, miR-451 has been found to be dramatically induced with erythroid differentiation. Given the potential of microRNA as a therapeutic target, and reactivation of the embryonic α- and β-globin chains has been shown to rescue the lethality of mice with α- and β-thalassemia, the identification of miR-144 signaling in the embryonic α-globin regulation in zebrafish embryos might provide novel insights into the pathogenesis of humans disorders associated with deregulated hemoglobin synthesis.

Acknowledgments

We thank Drs. Yi Zhou and Len Zon at Children’s Hospital in Boston for providing Cloche mutant, and Shuo Lin at UCLA for gata-1 promoter plasmid. The work was supported in part by National Basic Research Program of China (2007CB947003), National Natural Science Foundation of China (30525019 to T.X.L and 30771185 to M.D), Hundred Scholars Award of Chinese Academy of Sciences, Shanghai PUJIANG project (06PJ14104), Knowledge Innovation Program of the Chinese Academy of Sciences (KSCX2-YW-R-03), Chinese National High Tech Program 863 (2006AA02A405 to S.J.C) and Key Discipline Program of Shanghai Municipal Education Commission (Y0201 to S.J.C and 06BZ013 to M.D). We thank Dr. Dang Sheng Li and all members of the laboratory for helpful discussions.
Authorship: YF Fu, TT Du, M Dong, KY Zhu, CB Jing performed experiments and analyzed data. Y Zhang, L Wang, HB Fan, Y Chen, Y Jin, GP Yue, QH Huang and M Deng assisted with experiments. YF Fu, TT Du and TX Liu designed research plan and wrote the paper. SJ Chen, Z Chen and Q Jing critically read the paper. The authors have no conflicts of interest to declare.

References


Figure legends

Figure 1. Erythroid- and developmental stage-specific expression of an
evolutionarily conserved miR-144 gene. (A) Whole-mount miRNA in situ hybridization (WISH) was performed at the indicated stages. Note that the expression of miR-144 is first detected in the ICM of 20 hpf embryos and become undetectable at 36 hpf (arrows). Embryos are lateral views with the head to the left. (B) Two-color WISH analysis of miR-144 (black) and myo D (red) in the cloche\textsuperscript{m39} mutant and wild-type sibling embryo at 24 hpf. Note that the miR-144 expression is undetectable in the cloche mutant (right panel, arrow). (C) Co-localization of miR-144 (black) and gata-1 (red) at 26 hpf. Inset is the amplified view of the ICM. (D) The miR-144 and miR-451 genes reside within a region of zebrafish chromosome 5 that is syntenic with the human miR-144 and miR-451 locus. (Left) Eight genes including miR-144 and miR-451 are located within a genomic region on human chromosome 17, according to the latest version of the human genome draft (http://genome.ucsc.edu/). (Right) Eight zebrafish homologues of these genes are listed according to their map positions on chromosome 5 (http://www.ensembl.org). (E) Northern blot analysis of mature miR-144 expression at 18, 24 and 30 hpf. miR-144 duplex and 5S RNA were used as positive and loading control, respectively.

Figure 2. miR-144 is functional active in vivo. (A,B) Expression of EGFP mRNAs and EGFP fluorescence in 16 and 24 hpf F1 embryos derived from Tg(zgata-1:EGFP) transgenic line 1 (A) or Tg(zgata-1:EGFP-2xPT) transgenic line 2 (B). Note that both EGFP mRNAs and green fluorescence expressed normally in the P-LPM of 16 hpf, but became undetectable in the ICM of 24 hpf embryos from Tg(zgata-1:EGFP-2xPT) transgenic line. Amplified views of fluorescent ICM were shown in the insets. Arrowhead indicates non-specific EGFP expression in the dorsal neurons. (C,D) Rescue of both
EGFP mRNAs and protein fluorescence by microinjection of miR-144 LNA (C), but not its 5-base pair mismatch control (D), into one-cell stage F1 embryos from Tg(zgata-1:EGFP-2xPT) transgenic line 2.

**Figure 3.** miR-144 selectively inhibits embryonic α-E1 globin expression. (A) Morphology of 22 hpf embryos injected with either miR-144 LNA control (left) or miR-144 LNA oligonucleotides (middle and right). (B) WISH assays for α-E1 globin. Note that a significant increase of α-E1 globin mRNAs was observed in the miR-144 knockdown embryos (middle, arrows), with approximately 8% of miR-144 knockdown embryos showing massive induction (right, arrow). (C,D) WISH assays for alas2 (C) and α-E3 globin (D). A slight increase (alas2 and α-E3) and an ectopic expression (alas2, right panel, arrowhead) were noted. (E-I) WISH assays for β-E1 hemoglobin (E), β-E2 hemoglobin (F), gata-1 (G), scl (H) and l-plastin (I, arrows). No obvious changes were detected. (J) Representative results of genes analyzed by semi-quantitative, single-embryo RT-PCR in miR-144 LNA and control-injected embryos (n=3). (K) Whole-mount immunohistochemistry staining for mitotic phosphorylated histone 3 (pH3). Brown dots denote cells undergoing mitosis. (L) Apoptotic assay. Arrows denote red dots and TUNEL-positive cells.

**Figure 4.** miR-144 targets klfd in physiological context. (A,B) Expression of EGFP mRNAs and EGFP fluorescence in 16 and 24 hpf F2 embryos derived from either Tg(zgata-1:EGFP) transgenic line 2 (A), or Tg(zgata-1:EGFP-klfd 3’UTR) transgenic line 5 (B). Arrowheads indicate non-specific expression of EGFP mRNAs in the dorsal
neurons. (C) Western blot analyses of EGFP protein in 16 and 24 hpf embryos from either Tg(zgata-1:EGFP) transgenic line 2 or Tg(zgata-1:EGFP-klf6 3’UTR) transgenic line 5. (D,E) Rescue of EGFP mRNAs by microinjection of either miR-144 or control LNA into one-cell stage F2 embryos from Tg(zgata-1:EGFP-3’UTR) transgenic line 5. (F) Western blot analysis of endogenous Klfd protein levels in miR-144-deficient embryos at 22 hpf.

**Figure 5. Klfd selectively transactivates α-E1 globin promoter.** (A) WISH analysis of klf6 mRNAs at the indicated stages. Note that the klf6 is first detected in the ICM of 18 hpf embryos. Embryos are lateral views with the head to the left. (B) Forced expression of klf6 mRNAs (20 pg) increases α-E1 globin expression at 24 hpf (arrows). (C) Knockdown of Klfd protein by morpholino in vivo. Western blot analysis of embryos injected with 5-MM control (Ctrl) or klf6-specific morpholino oligonucleotides at the indicated doses. The proteins extract from 20 embryos at 22 hpf were loaded into each lane. WT: wild type. Actin was used as a loading control. (D) Knockdown of Klfd reduces α-E1 globin expression at 22 hpf (arrows). (E) Decreased hemoglobin staining by O-dianisidine at 36 and 54 hpf (arrows). (F) Knockdown of Klfd is able to significantly rescue the overproduction and ectopic expression of α-E1 globin induced by miR-144 deficiency (arrow). The mpo expression was not affected. (G-I) The Klfd selectively binds to the distal CACCC elements of α-E1 gene promoter. E-ChiP analysis of chromatin extracted from 16 and 22 hpf embryos. PCR was performed using the primers located to the indicated promoter regions for α-E1 (G, regions I-III), β-E1 (G, regions IV and V), alas2 (H, regions I and II) and β-E3 (I, regions III and IV). Note that
the Klf6d was significantly enriched at the CACCC boxes (region I) of α-E1 globin promoter at 22 hpf (g), but not at 16 hpf. The results were repeated four times with separate batches of chromatin preparations.

Figure 6. Klf6d binds to miR-144 promoter to activate its transcription. (A,B) Markedly decreased expression of miR-144 in Klf6d-deficient 22 hpf embryos (A arrows), while increased expression of miR-144 was observed in 22 hpf embryos injected with 20 pg of capped klf6d mRNAs (B arrows). (C) Schematic representation of 5.0 Kb miR-144 promoter upstream of the transcriptional start sites (+1). The positions of CACCC boxes were indicated by the vertical arrows and arabic numbers. Horizontal arrows indicate the primers located to the promoter regions (I-V). EH453145 is the GenBank accession number of cDNA encoding zebrafish miR-144/miR-451 precursor sequence. (D) The Klf6d selectively interacts with the proximal CACCC boxes (region V) of miR-144 promoter. Note that the Klf6d only binds to the proximal miR-144 promoter at 22 hpf, but not at 16 hpf. (E) A proposed model of miR-144-Klf6d-mediated trans-acting mechanism in the regulation of embryonic α-E1 globin.
Figure 1

A

![Images of developing embryos at different time points (18 h to 36 h).]

B

**myoD / miR-144**

- sibling 24 h
- cloche 24 h

C

**gata-1 / miR-144**

- 26 h

D

**Human chr.17**

<table>
<thead>
<tr>
<th>Mb</th>
<th>Gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.458</td>
<td>CARKL</td>
</tr>
<tr>
<td>3.518</td>
<td>TMEM93</td>
</tr>
<tr>
<td>3.523</td>
<td>P2RX5</td>
</tr>
</tbody>
</table>

**zfish chr.5**

- mir-144
- mir-451
- eral1
- C17orf63
- ERAL1
- mir-451
- mir-144
- TBX2

E

**Northern Blot**

- 18 h
- 24 h
- 30 h

- 22 nt

- SS rRNA
Figure 2
Figure 3

<table>
<thead>
<tr>
<th>miR-144 LNA Ctrl</th>
<th>miR-144 LNA knockdown</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A</strong> Brightfield</td>
<td></td>
</tr>
<tr>
<td>22 h</td>
<td></td>
</tr>
<tr>
<td><strong>B</strong> α-E1 globin</td>
<td></td>
</tr>
<tr>
<td>22 h</td>
<td></td>
</tr>
<tr>
<td><strong>C</strong> alas2</td>
<td></td>
</tr>
<tr>
<td>22 h</td>
<td></td>
</tr>
<tr>
<td><strong>D</strong> β-E1 globin</td>
<td></td>
</tr>
<tr>
<td>22 h</td>
<td></td>
</tr>
<tr>
<td><strong>E</strong> α-E3 globin</td>
<td></td>
</tr>
<tr>
<td>22 h</td>
<td></td>
</tr>
<tr>
<td><strong>F</strong> β-E2 globin</td>
<td></td>
</tr>
<tr>
<td>22 h</td>
<td></td>
</tr>
<tr>
<td><strong>G</strong> gata-1</td>
<td></td>
</tr>
<tr>
<td>22 h</td>
<td></td>
</tr>
<tr>
<td><strong>H</strong> scl</td>
<td></td>
</tr>
<tr>
<td>22 h</td>
<td></td>
</tr>
<tr>
<td><strong>I</strong> l-plastin</td>
<td></td>
</tr>
<tr>
<td>22 h</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>J</th>
<th>RT-PCR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ctrl</td>
</tr>
<tr>
<td></td>
<td>LNA</td>
</tr>
<tr>
<td>α-E1</td>
<td></td>
</tr>
<tr>
<td>alas2</td>
<td></td>
</tr>
<tr>
<td>β-E1</td>
<td></td>
</tr>
<tr>
<td>gata-1</td>
<td></td>
</tr>
<tr>
<td>mpo</td>
<td></td>
</tr>
<tr>
<td>gapdh</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>K</th>
<th>Ctrl</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LNA</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>L</th>
<th>Ctrl</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LNA</td>
</tr>
</tbody>
</table>
Figure 5

A

B

C

D

E

F

G

H

I
Figure 6

A Ctrl  
22 h  
klfd mo

B WT  
22 h  
klfd mRNA

C

-4315
-4040
-3944
-3879
-3553
-3353
-3287
-190
-147
-133
-1753

CACCC
GGGTT

+1 TSS
pre-mir-144
pri-mir-144; EH453145

D

16 h  
22 h  
Klfd
no Ab
Input

E

mir-144
mature mir-144
Klfd
alpha-E1 hemoglobin transcription
Mir-144 selectively regulates embryonic \(\alpha\)-hemoglobin synthesis during primitive erythropoiesis

Yan-Fang Fu, Ting-Ting Du, Mei Dong, Kang-Yong Zhu, Chang-Bin Jing, Yong Zhang, Lei Wang, Hong-Bo Fan, Yi Chen, Yi Jin, Gui-Ping Yue, Sai-Juan Chen, Zhu Chen, Qiu-Hua Huang, Qing Jing, Min Deng and Ting Xi Liu