The role of OMICS research in understanding phenotype variation in thalassaemia: the THALAMOSS project

Roberto Gambari

Department of Life Sciences & Biotechnology, Ferrara University, Ferrara, Italy

Abstract

The β-thalassaemias are a group of severe and rare anaemias with monogenic inheritance, a complex systemic phenotype and several treatment-related complications, caused by more than 300 mutations of the β-globin gene. Novel therapeutic protocols, most of which are based on still experimental treatments, show great promise but significant variability of success between patients. These strategies include chemical/molecular induction of the endogenous β-like γ-globin gene or the restoration of clinically relevant β-globin levels by gene therapy. A small number of modifiers with significant impact on disease penetrance, severity and efficacy of treatments are known, but most remain elusive. Improvements of existing treatment regimens and optimization and application of novel treatments will critically depend on the characterization of additional disease modifiers and the stratification of patients for customized treatment regimens. This requires extensive analyses based on "OMICS", an English-language neologism which refer to different but connected fields in molecular biology and biochemistry, such as genomics, transcriptomics, exomics, proteomics, metabolomics. The major objective of OMICS is a collective characterization of pools of biological molecules (gene sequences, transcripts, proteins and protein domains) controlling biological structures, functions and dynamics, including several involved in pathological conditions. One of the most interesting observations of genomics in β-thalassaemias is the association between genomic sequences and high fetal haemoglobin (HbF) levels, in consideration of the fact that high HbF levels are usually associated with milder forms of β-thalassaemia. Related to this issue, is the possibility to predict response to different therapeutic protocols on the basis of genomic analyses. For instance, three major loci (Xmn1-HBG2 single nucleotide polymorphism, HBS1L-MYB intergenic region on chromosome 6q, and BCL11A) contribute to high HbF production. Pharmacogenomic analysis of the effects of hydroxyurea (HU) on HbF production in a collection of β-thalassemia and sickle cell disease (SCD) patients allowed the identification of genomic signatures associated with high HbF. Therefore, it can hypothesized that genomic studies might predict the response of patients to treatments based on hydroxyurea, which is at present the most used HbF inducer in pharmacological therapy of β-thalassaemia. Transcriptomic/proteomic studies allowed to identify the zinc finger transcription factor B-cell lymphoma/leukemia 11A (BCL11A) as the major repressor of HbF expression. The field of research on g-globin gene repressors (including BCL11A) is of top interest, since several approaches can lead to pharmacologically-mediated inhibition of the expression of g-globin gene repressors, leading to g-globin gene activation. Among these strategies, we underline direct targeting of the transcription factors by aptamers or decoy molecules, as well as inhibition of the mRNA coding g-globin gene repressors with shRNAs, antisense molecules, peptide nucleic acids (PNAs) and microRNAs. In this respect, the THALAMOSS FP7 Project (THALAssaemia MOdular Stratification System for personalized therapy of β-thalassemia, www.thalamoss.eu) aims develop a universal sets of markers and techniques for stratification of β-thalassaemia patients into treatment subgroups for (a) onset and frequency of blood transfusions, (b) choice of iron chelation, (c) induction of fetal hemoglobin, (d) prospective efficacy of gene-therapy. The impact of THALAMOSS is the provision of novel biomarkers for distinct treatment subgroups in β-thalassaemia (500–1000 samples from participating medical centres), identified by combined genomics, proteomics, transcriptomics and tissue culture assays, the development of new or improved products for the cell isolation, characterization and treatment of β-thalassaemia patients and the establishment of routine techniques for detection of these markers and stratification of patients into treatment groups. Translation of these activities into the product portfolio and R&D methodology of participating SMEs will be a major boost for them as well as for the field. THALAMOSS tools and technologies will (a) facilitate identification of novel diagnostic tests, drugs and treatments specific to patient subgroups and (b) guide conventional and novel therapeutic approaches for β-thalassaemia, including personalized medical treatments.

Correspondence: Roberto Gambari, Department of Life Sciences & Biotechnology, Ferrara University, Ferrara, Italy
E-mail: gam@unife.it

Key words: β-thalassaemias; genomics; transcriptomics; proteomics; HbF induction; gene therapy.

Acknowledgements: RG is granted by *Fondazione Cariparo* (*Cassa di Risparmio di Padova e Rovigo*), UE FP7 THALAMOSS Project (THALAssaemia MOdular Stratification System for personalized therapy of beta-thalassemia) and by Telethon GGP10124 and by AIRC. This research is also supported by *Associazione Veneta per la Lotta alla Talassemia* (AVLT), Rovigo.

Thalassemia Reports 2014; 4:4877
doi:10.4081/thal.2014.4877

Introduction

The β-thalassaemias are a group of severe and rare anaemias with monogenic inheritance, a complex systemic phenotype and treatment-related complications.[1,2] Novel and mostly experimental treatments, such as the chemical induction of the endogenous β-like γ-globin gene or the

pagepress

restoration of β-globin levels by gene therapy, show great promise but significant variability of success between patients.[2] A small number of modifiers with significant impact on disease penetrance, severity and efficacy of treatments are known, but most remain elusive. Improvements of existing treatment regimens and optimization and application of novel treatments will critically depend on the characterization of additional disease modifiers and the stratification of patients for customized treatment regimens.[2-4] This requires extensive analyses based on "OMICS".

OMICS

The word "OMICS" is an English-language neologism which refer to different but connected fields in molecular biology and biochemistry ending in *-omics*, such as genomics, transcriptomics, exomics, proteomics, metabolomics. The major objective of OMICS is a collective characterization of pools of biological molecules (gene sequences, transcripts, proteins and protein domains) controlling biological structures, functions and dynamics, including several involved in pathological conditions.

In vitro experimental systems: erythroid precursor cells from β-thalassemia patients

The OMICS approach needs suitable experimental model systems from β-thalassaemia patients, which recapitulate erythroid differentiation in vitro. These experimental model systems are required when transcriptomic and proteomic studies have to be conducted on erythroid cells, but are also very important to study epigenetics. In several published reports, the most used experimental model system is constituted by erythroid precursor cells (ErPCs) from β-thalassaemia patients.[1] This experimental system was demonstrated to be very useful in assessing toxicity and efficacy of any therapeutic intervention, as well as predicting therapeutic response in clinical management of β-thalassaemia patients. ErPCs from peripheral blood (PB) are widely used, while access to bone-marrow or mobilized blood samples, which incidentally are the cells preferentially used in clinical applications, is more restricted.[2] Using PB-derived ErPCs, it is possible to obtain large cultures of relatively pure and synchronized erythroid cell population in which compounds can be added at specific stages of maturation. In the procedure developed by Fibach *et al.*,[5] the culture is divided into two phases: first, an erythropoietin (EPO)-independent proliferation phase, in which peripheral blood cells are first cultured in the presence of a combination of growth factors, but in the absence of EPO, and second, a differentiation phase when the culture, supplemented with EPO, generates orthochromatic normoblasts and enucleated erythrocytes, with cells decreasing in size and accumulating haemoglobin and large cellular clusters assuming a reddish colour and give brown-red pellets upon centrifugation.[1,5] This system recapitulates many aspects of *in vivo* erythropoiesis, including globin RNA metabolism, cell cycle kinetics, expression of cell surface antigens, iron and ferritin metabolism and recruitment of transcription factors,[1-5] and allows analysis of Hb content by a variety of techniques, such as alkaline denaturation, benzidine staining, capillary electrophoresis, cation-exchange HPLC for haemoglobins and reversed-phase HPLC for globin chains.[6]

Genomic analysis for β-thalassemia

This field of investigation has as major objective the finding of correlation between genomic variations and expression of particular phenotypes in thalassemia. One of the most interesting observations in the last years is the association between genomic sequences and high fetal haemoglobin (HbF) levels, in consideration of the fact that high HbF levels are usually associated with milder forms of β-thalassaemia. Related to this issue, is the possibility to predict response to different therapeutic protocols on the basis of genomic analyses. The issue of the relationship between g-globin gene repressors and levels of HbF in erythroid cells was also the subject of a recent review paper by Thein and Menzel,[7] reporting the progress in the understanding of the persistence of HbF in adults. Three major loci (Xmn1-HBG2 single nucleotide polymorphism, HBS1L-MYB intergenic region on chromosome 6q, and BCL11A) contribute to high HbF production. As far as the intergenic region HBS1L-MYB it was recently found by Stadhouders *et al.*[8] that several HBS1L-MYB intergenic variants affect regulatory elements that are occupied by key erythroid transcription factors within this region. These elements interact with MYB, a critical regulator of erythroid development and HbF levels. They found that several HBS1L-MYB intergenic variants reduce transcription factor binding, affecting long-range interactions with MYB and MYB expression levels. These data provide a functional explanation for the genetic association of HBS1L-MYB intergenic polymorphisms with human erythroid traits and HbF levels. In conclusion, according with the review by Thein and Menzel,[7] and in agreement with several additional studies, putative repressors of g-globin gene transcription are Oct-1,[9] MYB[10] and BCL-11A.[11] For instance, the zinc finger transcription factor B-cell lymphoma/leukemia 11A (BCL11A) was recently shown as the major repressor of HbF expression by genome-wide association studies (GWAS) leading to the identification of a new HbF-associated locus on chromosome 2, located within the gene BCL11A.[11] It has been in fact reported that transgenic deactivation of BCL11A reactivates HbF and corrects a humanized sickle-haemoglobin mouse model, knockdown of BCL11A leads to significant HbF induction in human cells, similar to knockdown of its positive regulator KLF1.[2] In order to move from general concept to therapeutic application, controlled and stable shRNA-mediated HbF induction has achieved an efficiency of potential clinical relevance.[2,12,13] This field of research is of top interest, since several approaches can lead to pharmacologically-mediated inhibition of the expression of g-globin gene repressors, leading to g-globin gene activation. Among these strategies, we underline, in addition to the already mentioned delivery of shRNAs, direct targeting of the transcription factors by aptamers or decoy molecules, as well as inhibition of the mRNA coding g-globin gene repressors with antisense molecules, peptide nucleic acids (PNAs) and microRNAs.[4]

A second examples linking genomic/pharmacogenomics analyses has been reported by Borg *et al.*,[14] who performed a very interesting pharmacogenomic analysis of the effects of hydroxyurea (HU) on HbF production in a collection of β-thalassemia and sickle cell disease (SCD) compound heterozygotes and a collection of healthy and KLF1-haploinsufficient adults. This extensive study was undertaken to identify genomic signatures that follow high HbF, and allowed to identify KLF10 as a possible candidate. The research effort of these investigators, therefore, was focused on genotype analysis of β-thalassemia major and intermedia patients as well as on cohorts of β-thalassemia/SCD compound heterozygous patients that do or do not respond to HU treatment. These analyses showed that a mutant state of the KLF10 3'UTR is not present in β-thalassemia intermedia patients and is underrepresented in β-thalassemia/SCD compound heterozygous patients that respond well to HU treatment. Therefore, it can hypothesized that genomic studies might predict the response of patients to hydroxyurea, which is at present the most used HbF inducer in pharmacological therapy of β-thalassaemia.[14]

Transcriptomic/proteomic analysis for β-thalassemia

Erythroid precursor cells (ErPC) from β-thalassaemia patients are excellen model systems for transcriptomic and proteomic analyses (Figure 1). Several important issues can be answered following transcriptomic and proteomic studies. For instance studies on different cohorts of ErPCs separated on the basis of (a) endogenous levels of HbF or (b) response in vitro to HbF inducers might help in identifying the expression of genes which are associated to high-HbF phenotype or efficient response to HbF inducers. This is particularly important since increasing number of evidences allow to hypothesize that in vitro studies on HbF inducers predict in vivo response.[15,16]

Several publications that appeared in the last three years have confirmed that the g-globin gene expression is under a strong negative transcriptional control. Apart the theoretical importance, this conclusion indicates the potential therapeutic use of targeting these transcription factors to treat hemoglobinopathies. The involvement of BCL11A was already mentioned. Interstingly, associated with BCL11A expression is the erythroid Kruppel-like factor 1 (EKLF1, KLF1), an adult β-globin gene-specific zinc finger transcription factor which is a positive regulator of BCL11A transcription. Interestingly, when KLF1 was knocked down in erythroid progenitor CD34+ cells, g-globin expression was induced.[2,4] In another set of studies, DRED (direct repeat erythroid definitive) was shown to be a repressor complex that binds to the direct repeat (DR) elements in the - and g-globin gene promoter. Two of the components in this complex are the orphan nuclear receptors TR2 and TR4. Enforced expression of TR2/TR4 increased fetal g-globin gene expression in adult erythroid cells from β-YAC transgenic mice and also in adult erythroid cells from the humanized SCD mice. These studies clearly demonstrate that manipulation of transcription factors (including the already mentioned Oct-1 and MYB) efficiently reactivates g-globin gene expression during adult definitive erythropoiesis.

The miRNA-transcription factor network

MicroRNAs (miRNAs, miRs) are a family of small non-coding RNAs that regulate gene expression by targeting mRNAs in a sequence specific manner, inducing translational repression or mRNA degradation at the level of the RNA-induced silencing complex (RISC). This complex is responsible for the gene silencing. MicroRNAs have been found deeply involved in the control of erythroid differentiation.[17] The number of relevant studies on the possible effects of miRNAs on HbF production by erythroid cells is growing.[2] The first very intriguing observation in this field of investigation was reported by Sankaran *et al.*, who observed that in human trisomy 13, there is delayed switching and persistence of HbF and elevation of embryonic hemoglobin in newborns.[18]

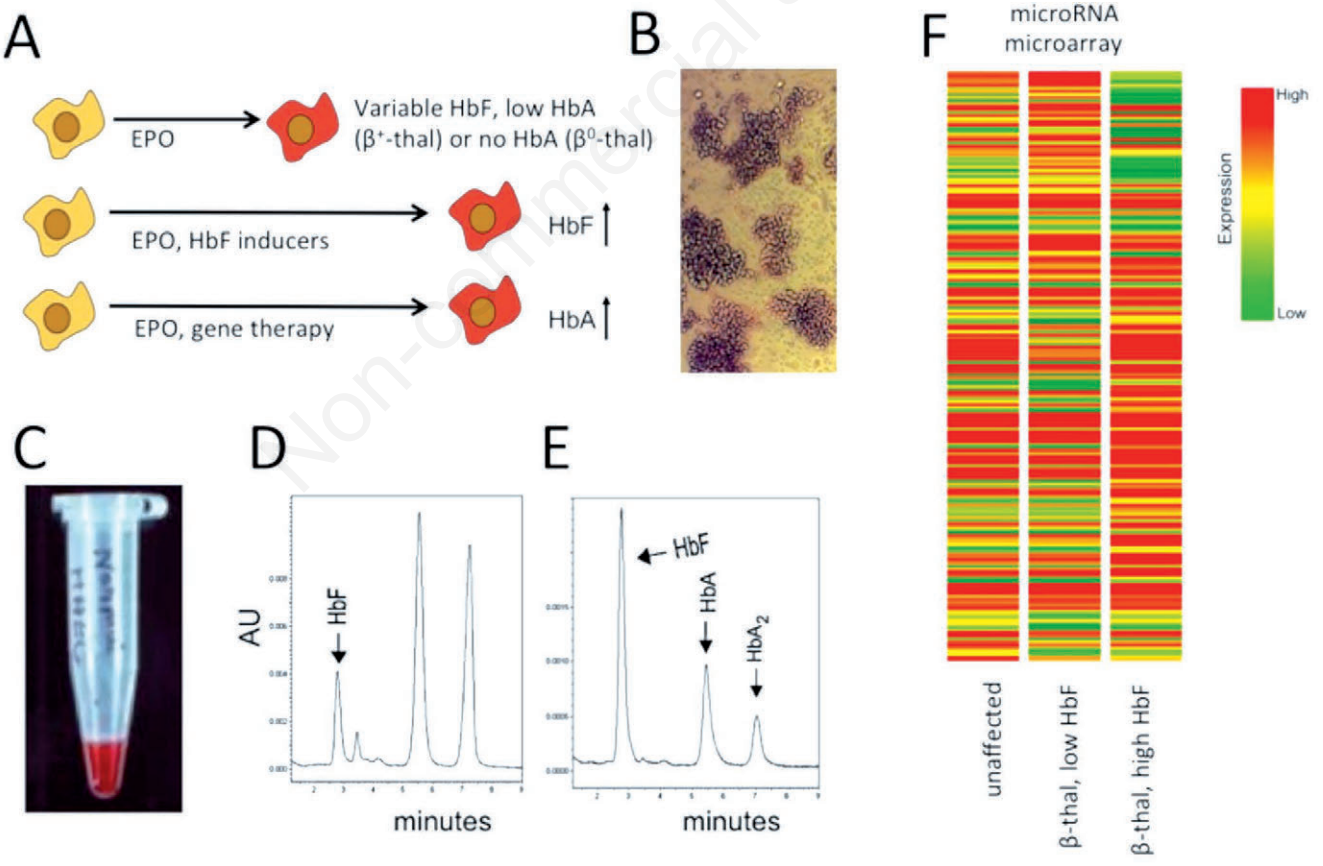

Figure 1. (A) Scheme describing the erythroid precursor cells (ErPCs) system used for assessing the activity of HbF inducers and gene-therapy vectors. B,C. Microscopic analysis (B) and red-pellet (C) of ErPCs, showing high level of Hb accumulation. D,E. HPLC analysed of EPO-induced ErPC from two b-thalassemia patients, one exhibiting low HbF production (D), the other exhibiting high HbF production (E). (F) Comparison of microRNA analysis using microarray. RNA was isolated from an unaffected donor, a b-thalassemia patient producing low HbF, or a b-thalassemia patient producing high levels of HbF, as indicated. Modified from Gambari[3] and Bianchi *et al.*[20]

pagepress

In partial trisomy cases, this trait maps to chromosomal band 13q14; by examining the genes in this region, two microRNAs, miR-15a and miR-16-1, appear as top candidates for the elevated HbF levels. Indeed, increased expression of these microRNAs in primary human erythroid progenitor cells results in elevated fetal and embryonic hemoglobin gene expression. Moreover, this group showed that MYB mRNA is a direct target of these microRNAs; interestingly, MYB, as already discussed, plays an important role in silencing the fetal and embryonic hemoglobin genes. The list of microRNAs proposed to be involved in down-regulation of g-globin gene repressors and, consequently, in up-regulation of g-globin gene transcription is increased in the last few years. For instance, Lulli *et al.* showed that miR-486-3p regulates BCL11A expression by binding to the extra-long isoform of BCL11A 3'UTR.[19] Overexpression of miR-486-3p in erythroid cells resulted in reduced BCL11A protein levels, associated to increased expression of g-globin gene, whereas inhibition of physiological miR-486-3p levels increased BCL11A and, consequently, reduced g-globin expression. The data obtained indicate that BCL-11A, one of the major repressor of g-globin gene expression, is a molecular target of miR-486-3p; accordingly, pharmacological mediated up-regulation of miR-486-3p might lead to BCL-11A down-regulation and, consequently activation of the g-globin gene expression. If the research in this field of investigation will confirm that microRNAs up-regulated in HbF expressing erythroid cells recognize mRNA coding TF repressors of g-globin gene expression, the strategy of design molecules able to mimic activity of those microRNAs for reactivation of the g-globin genes could be very appealing. Other microRNAs found up-regulated in association with g-globin gene expression were miR-210[20] and miR-23a/27a.[21] In conclusion, the findings that microRNAs are involved in g-globin anticipate the possibility that their pharmacological alteration might be a key strategy for increase HbF in erythroid cells.

THALAMOSS: THALAssaemia MOdular Stratification System for personalized therapy of β-thalassemia

Aims

THALAMOSS (www.thalamoss.eu, started 11/2012) aims develop a universal sets of markers and techniques for stratification of β-thalassaemia patients into treatment subgroups for (a) onset and frequency of blood transfusions, (b) choice of iron chelation, (c) induction of fetal hemoglobin, (d) prospective efficacy of gene-therapy.

Workpackages

THALAMOSS is organized in the following Workpackages: WP1. Recruitment, patient characterization and development of culture technologies for erythroid precursor cells; WP2. Omics analyses; WP3. Novel therapeutic approaches; WP4. Data management and analysis; WP5. Dissemination and exploitation; WP6. Regulatory and ethical issues; WP7. Program management.

Impact

The impact of THALAMOSS is the provision of novel biomarkers for distinct treatment subgroups in β-thalassaemia (500–1000 samples from participating medical centres), identified by combined genomics, proteomics, transcriptomics and tissue culture assays, the development of new or improved products for the cell isolation, characterisation and treatment of β-thalassaemia patients and the establishment of routine techniques for detection of these markers and stratification of patients into treatment groups.

Expected products

Translation of these activities into the product portfolio and R&D methodology of participating SMEs will be a major boost for them as well as for the field. THALAMOSS tools and technologies will (a) facilitate identification of novel diagnostic tests, drugs and treatments specific to patient subgroups and (b) guide conventional and novel therapeutic approaches for β-thalassaemia, including personalised medical treatments.

Key researchers

Key researchers of THALAMOSS are R. Gambari (Ferrara University, Italy, UNIFE), M. Kleanthous (The Cyprus Foundation for Muscular Dystrophy Research, Cyprus, CING), S. Philipsen (Erasmus Universitair Medisch Centrum Rotterdam, The Netherlands, EMC), E. Katsantoni (Biomedical Research Foundation, Academy of Athens, Greece, BRFF), S. Rivella (Weill Cornell Medical College, NY, USA, CU), P. Holub (Masaryk University, Czech Republic, MU), R. Galanello (Cagliari University, Italy, UNICA), SL. Thein (King's College Hospital, UK, KCL), E. Voskaridou (Laiko General Hospital, Greece, LGHA). Participating SMEs are Biocep (Israel), NovaMechanics Ltd. (Cyprus) and IRBM (Italy). Industrial activities are also provided by Harbour Antibodies (HA, The Netherlands). ThalaMoSS is financed through the FP7-HEALTH-2012-INNOVATION-1 call, project number 306201.

Conclusions: from OMICS to personalized therapy of thalassemia

As elsewhere summarized and reviewed, β-thalassemia are caused by more than 300 different mutations; moreover, thalassemia patients can be stratified in accordance with several markers, such as polymorphisms associated with high levels of HbF production (for instance XmnI, rs1427407, rs10189857, rs9399137). It is expected, therefore, that the management of β-thalassemia patients might depend on stratified classes of β-thalassemia patients. For instance, in the case of gene therapy, the protocols might be different when β^0- (for example $\beta^0$39/$\beta^0$39, β^0-IVSI-1/β^0-IVSI-1 or $\beta^0$39/β^0-IVSI-1) or β^+- (for example β^+-IVSI-110/ β^+-IVSI-110, β^+-IVSI-6/ β^+-IVSI-6 or β^+-IVSI-110/ β^+-IVSI-6) genotypes are considered. Moreover, β^0/β^+ genotypes should be carefully studied with respect to their response of gene therapy based on vectors carrying a normal β-globin gene. On the other hand, we expect that gene therapy based on vectors carrying a g-globin genes or expressing shRNAs against mRNAs coding repressor of g-globin gene transcription (such as BCL-11A and KLF-1) should be carefully discussed in the case the patients involved are already expressing high endogenous HbF. Pharmacogenomic-based studies have clearly demonstrated that several genomic variations (not restricted to the human β-globin gene cluster, are significantly associated with response to treatment in β-type hemoglobinopathies patients, with chemical HbF inducers, such as hydroxyurea. In this specific case, personalized treatments are expected to be considered after genomic/transcriptomic analysis able to predict the in vivo response.

References

1. Gambari R, Fibach E. Medicinal chemistry of fetal hemoglobin inducers for treatment of β-thalassemia. Curr Med Chem 2007;14:199-212.
2. Finotti A, Breda L, Lederer CW, Bianchi N, Zuccato C, Kleanthous

M, Rivella S, Gambari R. Recent trends in the gene therapy of β-Thalassemia. Journal of Blood Medicine, 2014 [In press].
3. Gambari R. Alternative options for DNA-based experimental therapy of β-thalassemia. Expert Opin Biol Ther 2012;12:443-62.
4. Finotti A, Gambari R. Recent trends for novel options in experimental biological therapy of β-thalassemia. Expert Opin Biol Ther 2014;14:1443-54.
5. Pope SH, Fibach E, Sun J, Chin K, Rodgers GP. Two-phase liquid culture system models normal human adult erythropoiesis at the molecular level. Eur J Haematol 2000;64:292-303.
6. Fibach E, Bianchi N, Borgatti M, Prus E, Gambari R. Mithramycin induces fetal hemoglobin production in normal and thalassemic human erythroid precursor cells. Blood 2003;102:1276-81.
7. Thein SL, Menzel S. Discovering the genetics underlying foetal haemoglobin production in adults. Br J Haematol 2009;145:455-67.
8. Stadhouders R, Aktuna S, Thongjuea S, et al. HBS1L-MYB intergenic variants modulate fetal hemoglobin via long-range MYB enhancers. J Clin Invest 2014;124:1699-710.
9. Xu XS, Hong X, Wang G. Induction of endogenous γ-globin gene expression with decoy oligonucleotide targeting Oct-1 transcription factor consensus sequence. J Hematol Oncol 2009;2:15.
10. Jiang J, Best S, Menzel S, et al. cMYB is involved in the regulation of fetal hemoglobin production in adults. Blood 2006;108:1077-83.
11. Sankaran VG, Menne TF, Xu J, et al. Human fetal hemoglobin expression is regulated by the developmental stage-specific repressor BCL11A. Science 2008;322:1839-42.
12. Sankaran VG. Targeted therapeutic strategies for fetal hemoglobin induction. Hematology Am Soc Hematol Educ Program 2011;2011: 459-65.
13. Wilber A, Hargrove PW, Kim YS, et al. Therapeutic levels of fetal hemoglobin in erythroid progeny of β-thalassemic CD34+ cells after lentiviral vector-mediated gene transfer. Blood 2011;117:2817-26.
14. Borg J, Phylactides M, Bartsakoulia M, et al. KLF10 gene expression is associated with high fetal hemoglobin levels and with response to hydroxyurea treatment in β-hemoglobinopathy patients. Pharmacogenomics 2012;13:1487-500.
15. Italia K, Jijina F, Merchant R, et al. Comparison of in-vitro and in-vivo response to fetal hemoglobin production and γ-mRNA expression by hydroxyurea in Hemoglobinopathies. Indian J Hum Genet 2013;19:251-8.
16. Pecoraro A, Rigano P, Troia A, et al. Quantification of HBG mRNA in primary erythroid cultures: prediction of the response to hydroxyurea in sickle cell and β-thalassemia. Eur J Haematol 2014;92:66-72.
17. He L, Hannon GJ. MicroRNAs: small RNAs with a big role in gene regulation. Nat Rev Genet 2004;5:522-31.
18. Sankaran VG, Menne TF, Š epanovi D, et al. MicroRNA-15a and -16-1 act via MYB to elevate fetal hemoglobin expression in human trisomy 13. Proc Natl Acad Sci USA 2001;108:1519-24.
19. Lulli V, Romania P, Morsilli O, et al. MicroRNA-486-3p regulates g-globin expression in human erythroid cells by directly modulating BCL11A. PLoS One 2013;8:e60436.
20. Bianchi N, Zuccato C, Lampronti I, Borgatti M and Gambari R. Expression of miR-210 during erythroid differentiation and induction of γ-globin gene expression. BMB reports 2009;42:493-9.
21. Ma Y, Wang B, Jiang F, et al. A feedback loop consisting of microRNA 23a/27a and the β-like globin suppressors KLF3 and SP1 regulates globin gene expression. Mol Cell Biol 2013;33:3994-4007.