



Mini-Review

Regulation of human alpha-globin gene expression and alpha-thalassemia

D.M. Ribeiro and M.F. Sonati

Departamento de Patologia Clínica, Faculdade de Ciências Médicas,
Universidade Estadual de Campinas, Campinas, SP, Brasil

Corresponding author: M.F. Sonati
E-mail: sonati@fcm.unicamp.br

Genet. Mol. Res. 7 (4): 1045-1053 (2008)

Received May 9, 2008

Accepted July 1, 2008

Published October 14, 2008

ABSTRACT. Hemoglobin and globin genes are important models for studying protein and gene structure, function and regulation. We reviewed the main aspects of regulation of human α -globin synthesis, encoded by two adjacent genes (α_2 and α_1) clustered on chromosome 16. Their expression is controlled mainly by a regulatory element located 40 kb upstream on the same chromosome, the α -major regulatory element, whose activity is restricted to a core fragment of 350 bp, within which several regulatory protein binding sites have been found. Natural deletions involving α -major regulatory element constitute a particular category of α -thalassemia determinants in which the α -globin genes are physically intact but functionally inactive.

Key words: Hemoglobin; Globin genes; HS-40; Thalassemia;
Gene expression; Alpha-major regulatory element

loci depends on local cis-acting sequences, such as the promoter sequences, and remote cis-acting sequences, such as the locus control region (LCR) of the β -globin gene cluster and the major regulatory element (MRE) of the α -globin gene cluster (Forrester et al., 1987; Grosveld et al., 1987; Higgs et al., 1990; Jarman et al., 1991).

In the β -cluster, the regulatory sequences are associated with five erythroid-specific DNase I hypersensitive sites (named HS-1 to HS-5) distributed over a region 4-20 kb upstream of the ϵ -gene. The entire segment of DNA that includes these sites is referred to as the β -LCR (Forrester et al., 1987; Grosveld et al., 1987) (Figure 1A).

Similarly, in the α -globin cluster, four erythroid-specific DNase I hypersensitive sites, located 10 (HS-10), 33 (HS-33), 40 (HS-40), and 48 (HS-48) kb upstream of the ζ -globin mRNA cap site, have been identified (Higgs et al., 1998). Characterization of natural deletions, analysis of interspecific hybrids and stable transfectants, and studies of transgenic mice indicate that only the regulatory sequence located 40 kb upstream of the α -cluster (HS-40) has a significant effect on α -gene expression (Hatton et al., 1990; Liebhaber et al., 1990; Wilkie et al., 1990; Romao et al., 1991, 1992; Sharpe et al., 1992; Flint et al., 1994, 1996; Higgs et al., 1998; Anguita et al., 2002; Hartevelde et al., 2005; Viprakasit et al., 2003, 2006). This sequence was later termed the α -MRE (Jarman et al., 1991; Figure 1B).

Because the α - and β -clusters have a common ancestry, coordinated expression and similar organization, it was assumed that regulation of globin gene expression by β -LCR and α -MRE was similar. However, experimental evidence from *in vivo* and *in vitro* studies suggests different roles for the β -LCR and α -MRE in β -globin and α -globin gene expression, respectively (Vyas et al., 1992; Craddock et al., 1995). Transgenic mice carrying constructs with the β -globin gene under the control of β -LCR show expression levels proportional to the number of gene copies integrated, irrespective of the positions at which they are integrated into the mouse genome (Grosveld et al., 1987). In contrast to β -LCR, α -MRE does not confer position independence or copy-number dependence on transgenic α -globin gene expression (Sharpe et al., 1992; Gourdon et al., 1994).

These differences in the regulation of gene expression could be explained by chromatin effects at the integration site. It has been suggested that β -LCR, in addition to its activity as an enhancer, mediates a transition from “closed” chromatin to “open” transcriptionally active chromatin in a tissue-specific manner. Even though α -MRE has some ability to form “open” chromatin in transgenic mice, it does not appear to be required for organization of the chromatin structure of the α -globin locus, which is located in a constitutively “open” chromatin environment (Vyas et al., 1992; Craddock et al., 1995).

Studies on the chromatin structure of the β -globin cluster showed that β -LCR interacts with promoters of active β -like genes through looping, forming a nuclear compartment dedicated to RNA polymerase II transcription, termed active chromatin hub (ACH). The formation of this erythroid-specific structure is also developmentally regulated (Tolhuis et al., 2002; Patrinos et al., 2004). Transcription factors seem to be essential for its spatial organization. Drissen et al. (2004), using chromatin conformation capture (3C) technology to investigate the mouse β -globin locus, demonstrated that erythroid Krüppel-like transcription factor, required for adult β -globin gene expression, is also necessary for complete β -ACH formation.

Investigating the α -globin cluster in mouse erythroid cells, Zhou et al. (2006) suggested that it could also be regulated through recruitment of active α -like gene promoters and regulatory elements to an ACH-like structure, occupied by flanking co-localized housekeeping genes. In nonerythroid cells, the housekeeping genes would be still co-localized, but

α -globin genes would be excluded from α -ACH. More recently, Vernimmen et al. (2007), employing a quantitative form of the 3C approach and chromatin immunoprecipitation (ChIP) assays, proposed a detailed model to explain mouse α -globin locus activation during erythroid differentiation. In erythroid precursors that do not yet express the globin genes, multiprotein complexes containing transcription factors initially bind the remote regulatory sequences, progressing along the chromosome until all elements, including the α -globin promoters, are bound and the associated chromatin is modified (Anguita et al., 2004). At the next stage of erythropoiesis, as proerythroblasts undergo terminal differentiation to form intermediate and late erythroblasts, the α -globin genes are activated. The general transcriptional machinery, including RNA Polymerase II, is first recruited to the upstream regulatory elements, while the α -globin promoters are occupied by the SP/X-Krüppel-like transcription factors. Only late in the differentiation process, the transcriptional machinery is recruited to the α -globin promoters, with looping occurring between the regulatory elements and the α -globin genes (Vernimmen et al., 2007). There seem to be two purposes for looping between regulatory elements and α -globin promoters: to deliver the general transcriptional machinery (Johnson et al., 2003; Szutorisz et al., 2005a,b) and to relocate the promoter into the ACH structure (Vernimmen et al., 2007).

De Gobbi et al. (2007) investigated gene activation in human erythropoiesis. Using ChIP-chip technology, they analyzed the pattern of transcription factor binding and the epigenetic modifications across a 220-kb telomeric region of chromosome 16 in primary erythroid and nonerythroid cells. All known *cis*-acting regulatory elements in the human α -globin cluster were identified and no additional erythroid-specific regulatory elements were found. The authors then constructed a humanized mouse model containing the human and mouse α -cluster orthologous regions coexpressing in the same animal. Some significant differences were observed in transcription factor binding and histone modification in the human and mouse clusters, showing that they are regulated in different ways. These orthologous sequences play different roles in these two species, and an additional species-specific element (HS-12) seems to be recruited in the mouse cluster, emphasizing that some caution is required when the mouse is used as a model to investigate human gene regulation.

THE HUMAN α -MAJOR REGULATORY ELEMENT

The regulatory element α -MRE behaves as a classical enhancer; its main function in the normal chromosomal environment is to activate and enhance expression from the ζ -globin and the α -globin promoters (Zhang et al., 1993).

The functional domain of this element is restricted to a 350-bp core fragment, in which several well-conserved nuclear protein binding sites have been identified. These include four potential binding sites for the erythroid-specific factor GATA-1, four CACC boxes and two binding sites for the erythroid-factor NF-E2 (Jarman et al., 1991) that are occupied *in vivo* in an erythroid lineage-specific and developmental stage-specific manner. *In vivo*, three of four GATA-1 sites, both NF-E2 sites and one CACC box are occupied in erythroid cells, but not in nonerythroid cells (Strauss et al., 1992). De Gobbi et al. (2007), using ChIP-chip technology, demonstrated the binding pattern of the GATA-1/SCL complex and the entire pentameric erythroid complex (GATA-1, SCL, E2A, LMO2 and Ldb-1), along with both p45 and p18 NF-E2 subunits across the α -MRE.

Table 1. Alpha-major regulatory element (α -MRE) haplotypes.

Positions of the α -MRE polymorphic sites						
in the 16p13.3 region (UCSC Genome Browser)	103,564	103,597	103,625	103,666	103,675	103,678
in the cloned α -MRE fragment (Harteveld et al., 2002)	+96	+130	+158	+199	+209	+212
Haplotypes						
A	C	G	C	G	G	G
B	C	A	C	A	G	G
C	C	A	C	A	A	G
D	C	G	T	G	G	G
E	C	G	C	G	G	C
F	A	G	C	G	G	G

Bold letters indicate the nucleotide substitutions in relation to the A haplotype.

Table 2. Alpha-major regulatory element (α -MRE) haplotype frequencies in the population groups studied.

Populations	Haplotype					
	A	B	C	D	E	F
Dutch (35) ^a	0.43	0.57				
Italian (53)	0.56	0.43	0.01			
Indian (39)	0.67	0.32	0.01			
Chinese (23)	0.74	0.26				
Indonesian (34)	0.78	0.22				
African (32)	0.73	0.11		0.16		
Pygmy (60)	0.76	0.16		0.04	0.02	0.02
Parakanã (70)	0.70	0.30				
Xikrin (95)	0.87	0.13				

Source: Harteveld et al., 2002; Ribeiro et al., 2003.

^aNumber of individuals studied.

The α -MRE is genetically polymorphic; this polymorphism was first studied by Harteveld et al. (2002) in seven population groups from Africa, Europe and Asia. Six different α -MRE haplotypes, named A to F, were found (Table 1). These haplotypes were also studied in native Indians from two non-miscegenated tribes (the Parakanã and Xikrin) in Amazonia, in northern Brazil (Ribeiro et al., 2003). The α -MRE haplotype frequencies in these population groups are shown in Table 2.

Most of these polymorphisms are not expected to interfere with regulation of α -gene expression, since they are between binding sites for nuclear factors or in sites considered not to be active *in vivo*, except for haplotype D, in which the base substitution at position 158 changes the first consensus binding site of factor NF-E2 (Andrews et al., 1993). However, no experimental studies have been made to evaluate whether the polymorphisms responsible for the different α -MRE haplotypes are able to influence its enhancer activity.

ALPHA-THALASSEMIA

Thalassemias are inherited hemoglobin disorders characterized by a quantitative reduction of the α - or β -globin chains (Weatherall and Clegg, 1981). Alpha-thalassemias result from under-

production of the α -globin chains of fetal and adult hemoglobin. Down-regulation of one or two of the four α -globin genes leads to mild alterations in red blood cells. When three α -genes are involved, excess β -globin chains form tetramers (β_4 - Hb H), causing a moderate to severe hemolytic anemia called Hb H disease. Inactivation of the four α -genes results in tetramerization of γ -chains (γ_4 - Hb Bart's) and in the Hb Bart's hydrops fetalis syndrome (Weatherall and Clegg, 1981). Most α -thalassemia determinants are deletions that remove the α -genes; inactivating point mutations are a less common cause of this type of thalassemia, but they may occur at high frequencies in certain areas under selective pressure by malaria (Higgs, 1993).

Deletions involving the regulatory element α -MRE are sporadic but characterize a particular category of α -thalassemia determinants in which the α -globin genes are physically intact but functionally inactive (Romao et al., 1991). At present, about 15 deletions that remove the α -MRE and cause α -thalassemia have been found, 11 of which have been fully characterized (Hatton et al., 1990; Liebhaber et al., 1990; Wilkie et al., 1990; Romao et al., 1991, 1992; Flint et al., 1994, 1996; Hartevelde et al., 2005; Viprakasit et al., 2003, 2006) (Figure 2). These patients inherited chromosomes with deletions that removed from 35 to 160 kb of the region upstream of the α -cluster, while the linked α -genes remained intact, with entirely normal DNA sequences. All these patients had α -thalassemia and their phenotypes were consistent with severe down-regulation (<1-2%) of α -gene expression from the affected chromosome (Higgs et al., 1998). The hematological findings in the patients showed that simple heterozygotes for an α -MRE deletion [$(\alpha\alpha)/\alpha\alpha$] have low mean corpuscular volume (<80 fentoliters) and low mean corpuscular hemoglobin (<25 picograms) and are indistinguishable from patients with only two α -genes ($--/\alpha\alpha$). Compound heterozygotes [$(\alpha\alpha)/-\alpha$] have the clinical phenotype of Hb H disease and are indistinguishable from patients who inherit a single functional α gene ($--/-\alpha$) (Viprakasit et al., 2003, 2006).

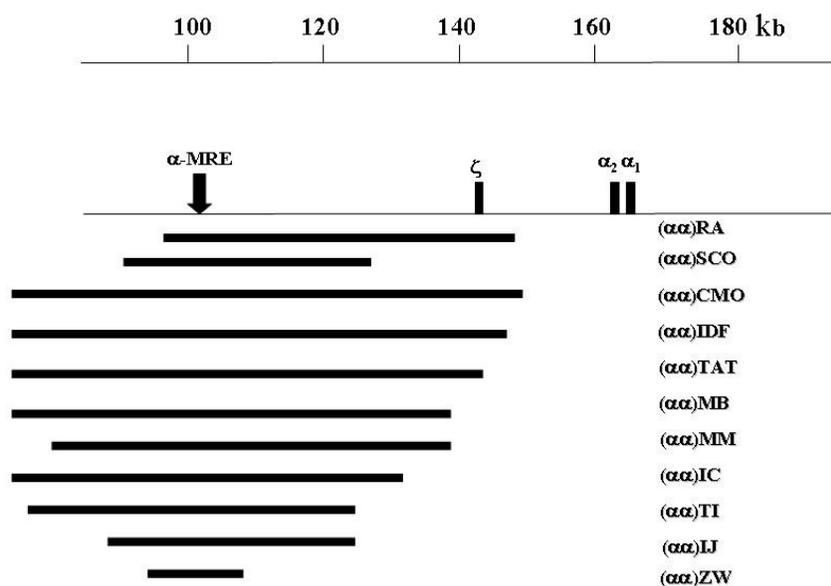


Figure 2. Diagram of upstream deletions of the α -cluster. The black boxes above the line represent the genes, and the α -major regulatory element (α -MRE) is shown as an arrow. The black boxes below the line represent the deletions (adapted from Viprakasit et al., 2003 and Hartevelde et al., 2005).

A recent study demonstrated the existence of an additional mechanism causing α -thalassemia, namely, a single nucleotide polymorphism in a non-genic region between the α -globin genes and their regulatory element. This polymorphism creates a new promoter-like element that interferes with normal activation of all downstream α -like genes and disrupts α -gene expression, probably as a result of its preferential interaction with the regulatory element α -MRE, outcompeting the endogenous α -globin promoters (De Gobbi et al., 2006).

Deletions of α -MRE have also been found in the Brazilian population. In 2002, Wenning et al. described a patient with Hb H disease resulting from a combination of the $-\alpha^{3.7}$ rightward deletion (the most common α -thalassemic determinant) with the $(\alpha\alpha)^{MM}$ deletion, which removes the regulatory element α -MRE, abolishing α -gene expression from the affected chromosome. The $(\alpha\alpha)^{MM}$ deletion was first characterized by Romao et al., in 1991, in a patient from the Azores Islands, Portugal. The Brazilian case is the first description of this type of α -thalassemia in the Latin-American continent. A number of other cases have yet to be characterized. Awareness concerning these deletions in populations is important because heterozygotes present reduced red cell indices in the absence of iron deficiency, β -thalassemia or α -globin gene alterations.

Regulation of α -globin gene expression by α -MRE is an important model for understanding the control systems of other eukaryotic genes and for investigating the interplay between gene transcription and modifications to chromatin structure related to gene function (Higgs and Wood, 2008). Future studies should focus on achieving a better understanding of human α -cluster regulation during erythroid differentiation, since most of the research in this field has been carried out in mice; it is also important to establish exactly how this cis-activating element interacts with the α -gene promoters, and to evaluate whether α -MRE variability influences enhancer activity.

ACKNOWLEDGMENTS

Research supported by FAPESP/Brazil (#02/13801-7; fellowship #03/07412-0) and CNPq/Brazil (grant #475481/2006-2).

REFERENCES

- Andrews NC, Erdjument-Bromage H, Davidson MB, Tempst P, et al. (1993). Erythroid transcription factor NF-E2 is a haematopoietic-specific basic-leucine zipper protein. *Nature* 362: 722-728.
- Anguita E, Sharpe JA, Sloane-Stanley JA, Tufarelli C, et al. (2002). Deletion of the mouse alpha-globin regulatory element (HS -26) has an unexpectedly mild phenotype. *Blood* 100: 3450-3456.
- Anguita E, Hughes J, Heyworth C, Blobel GA, et al. (2004). Globin gene activation during haemopoiesis is driven by protein complexes nucleated by GATA-1 and GATA-2. *EMBO J.* 23: 2841-2852.
- Beutler E, Lichtman MA, Coller BS, Kipps TJ, et al. (2001). *Williams Hematology*. McGraw-Hill Companies, NewYork.
- Bunn HF and Forget BG (1986). *Hemoglobin: Molecular, Genetics, and Clinical Aspects*. 3rd edn. W.B. Saunders Company, Philadelphia.
- Craddock CF, Vyas P, Sharpe JA, Ayyub H, et al. (1995). Contrasting effects of alpha and beta globin regulatory elements on chromatin structure may be related to their different chromosomal environments. *EMBO J.* 14: 1718-1726.
- De Gobbi M, Viprakasit V, Hughes JR, Fisher C, et al. (2006). A regulatory SNP causes a human genetic disease by creating a new transcriptional promoter. *Science* 312: 1215-1217.
- De Gobbi M, Anguita E, Hughes J, Sloane-Stanley JA, et al. (2007). Tissue-specific histone modification and transcription factor binding in alpha globin gene expression. *Blood* 110: 4503-4510.
- Drissen R, Palstra RJ, Gillemans N, Splinter E, et al. (2004). The active spatial organization of the β -globin locus requires

- the transcription factor EKLF. *Genes Dev.* 18: 2485-2490.
- Flint J, Craddock CF, Villegas A, Bentley DP, et al. (1994). Healing of broken human chromosomes by the addition of telomeric repeats. *Am. J. Hum. Genet.* 55: 505-512.
- Flint J, Rochette J, Craddock CF, Dode C, et al. (1996). Chromosomal stabilisation by a subtelomeric rearrangement involving two closely related *Alu* elements. *Hum. Mol. Genet.* 5: 1163-1169.
- Forrester WC, Takegawa S, Papayannopoulou T, Stamatoyannopoulos G, et al. (1987). Evidence for a locus activation region: the formation of developmentally stable hypersensitive sites in globin-expressing hybrids. *Nucleic Acids Res.* 15: 10159-10177.
- Fritsch EF, Lawn RM and Maniatis T (1980). Molecular cloning and characterization of the human β -like globin gene cluster. *Cell* 19: 959-972.
- Gourdon G, Sharpe JA, Wells D, Wood WG, et al. (1994). Analysis of a 70 kb segment of DNA containing the human ζ and alpha-globin genes linked to their regulatory element (HS-40) in transgenic mice. *Nucleic Acids Res.* 22: 4139-4147.
- Grosveld F, van Assendelft GB, Greaves DR and Kollias G (1987). Position-independent, high-level expression of the human β -globin gene in transgenic mice. *Cell* 51: 975-985.
- Hartevelde CL, Muglia M, Passarino G, Kielman MF, et al. (2002). Genetic polymorphism of the major regulatory element HS-40 upstream of the human α -globin gene cluster. *Br. J. Haematol.* 119: 848-854.
- Hartevelde CL, Voskamp A, Phylipsen M, Akkermans N, et al. (2005). Nine unknown rearrangements in 16p13.3 and 11p15.4 causing alpha- and beta-thalassaemia characterised by high resolution multiplex ligation-dependent probe amplification. *J. Med. Genet.* 42: 922-931.
- Hatton CS, Wilkie AO, Drysdale HC, Wood WG, et al. (1990). Alpha-thalassaemia caused by a large (62 kb) deletion upstream of the human α globin gene cluster. *Blood* 76: 221-227.
- Higgs DR (1993). α -Thalassaemia. *Baillieres Clin. Haematol.* 6: 117-150.
- Higgs DR and Wood WG (2008). Long-range regulation of alpha globin gene expression during erythropoiesis. *Curr. Opin. Hematol.* 15: 176-183.
- Higgs DR, Wood WG, Jarman AP, Sharpe J, et al. (1990). A major positive regulatory region located far upstream of the human α -globin gene locus. *Genes Dev.* 4: 1588-1601.
- Higgs DR, Sharpe JA and Wood WG (1998). Understanding alpha globin gene expression: a step towards effective gene therapy. *Semin. Hematol.* 35: 93-104.
- Hughes JR, Cheng JF, Ventress N, Prabhakar S, et al. (2005). Annotation of cis-regulatory elements by identification, subclassification, and functional assessment of multispecies conserved sequences. *Proc. Natl. Acad. Sci. U. S. A.* 102: 9830-9835.
- Jarman AP, Wood WG, Sharpe JA, Gourdon G, et al. (1991). Characterization of the major regulatory element upstream of the human alpha-globin gene cluster. *Mol. Cell. Biol.* 11: 4679-4689.
- Johnson KD, Grass JA, Park C, Im H, et al. (2003). Highly restricted localization of RNA polymerase II within a locus control region of a tissue-specific chromatin domain. *Mol. Cell. Biol.* 23: 6484-6493.
- Liebhaber SA, Griese EU, Weiss I, Cash FE, et al. (1990). Inactivation of human α -globin gene expression by a *de novo* deletion located upstream of the α -globin gene cluster. *Proc. Natl. Acad. Sci. U. S. A.* 87: 9431-9435.
- Patrinos GP, de Krom M, de Boer E, Langeveld A, et al. (2004). Multiple interactions between regulatory regions are required to stabilize an active chromatin hub. *Genes Dev.* 18: 1495-1509.
- Perutz MF, Rossmann MG, Cullis AF, Muirhead H, et al. (1960). Structure of myoglobin: A three-dimensional Fourier synthesis at 5.5 Angstrom resolution, obtained by X-ray analysis. *Nature* 185: 416-422.
- Ribeiro DM, Figueiredo MS, Costa FF and Sonati MF (2003). Haplotypes of alpha-globin gene regulatory element in two Brazilian native populations. *Am. J. Phys. Anthropol.* 121: 58-62.
- Romao L, Osorio-Almeida L, Higgs DR, Lavinha J, et al. (1991). Alpha-thalassaemia resulting from deletion of regulatory sequences far upstream of the alpha-globin structural genes. *Blood* 78: 1589-1595.
- Romao L, Cash F, Weiss I, Liebhaber S, et al. (1992). Human α -globin gene expression is silenced by terminal truncation of chromosome 16p beginning immediately 3' of the ζ -globin gene. *Hum. Genet.* 89: 323-328.
- Sharpe JA, Chan-Thomas PS, Lida J, Ayyub H, et al. (1992). Analysis of the human alpha globin upstream regulatory element (HS-40) in transgenic mice. *EMBO J.* 11: 4565-4572.
- Strauss EC, Andrews NC, Higgs DR and Orkin SH (1992). *In vivo* footprinting of the human α -globin locus upstream regulatory element by guanine and adenine ligation-mediated polymerase chain reaction. *Mol. Cell. Biol.* 12: 2135-2142.
- Szutorisz H, Canzonetta C, Georgiou A, Chow CM, et al. (2005a). Formation of an active tissue-specific chromatin domain initiated by epigenetic marking at the embryonic stem cell stage. *Mol. Cell. Biol.* 25: 1804-1820.
- Szutorisz H, Dillon N and Tora L (2005b). The role of enhancers as centres for general transcription factor recruitment. *Trends Biochem. Sci.* 30: 593-599.

- Tolhuis B, Palstra RJ, Splinter E, Grosveld F, et al. (2002). Looping and interaction between hypersensitive sites in the active beta-globin locus. *Mol. Cell* 10: 1453-1465.
- Vernimmen D, De Gobbi M, Sloane-Stanley JA, Wood WG, et al. (2007). Long-range chromosomal interactions regulate the timing of the transition between poised and active gene expression. *EMBO J.* 26: 2041-2051.
- Viprakasit V, Kidd AM, Ayyub H, Horsley S, et al. (2003). *De novo* deletion within the telomeric region flanking the human alpha globin locus as a cause of α thalassaemia. *Br. J. Haematol.* 120: 867-875.
- Viprakasit V, Harteveld CL, Ayyub H, Stanley JS, et al. (2006). A novel deletion causing α thalassemia clarifies the importance of the major human alpha globin regulatory element. *Blood* 107: 3811-3812.
- Vyas P, Vickers MA, Simmons DL, Ayyub H, et al. (1992). Cis-acting sequences regulating expression of the human alpha-globin cluster lie within constitutively open chromatin. *Cell* 69: 781-793.
- Weatherall DJ and Clegg JB (1981). *The Thalassemia Syndromes*. 3rd edn. Blackwell Scientific Publications, Oxford.
- Wenning MR, Harteveld CL, Giordano PC, Kimura EM, et al. (2002). Hemoglobin H disease resulting from the association of the - α 3.7 rightward deletion and the ($\alpha\alpha$)MM deletion in a Brazilian patient. *Eur. J. Haematol.* 69: 179-181.
- Wilkie AO, Lamb J, Harris PC, Finney RD, et al. (1990). A truncated human chromosome 16 associated with alpha thalassaemia is stabilized by addition of telomeric repeat (TTAGGG)_n. *Nature* 346: 868-871.
- Zhang Q, Reddy PM, Yu CY, Bastiani C, et al. (1993). Transcriptional activation of human ζ 2 globin promoter by the α globin regulatory element (HS-40): functional role of specific nuclear factor-DNA complexes. *Mol. Cell. Biol.* 13: 2298-2308.
- Zhou GL, Xin L, Song W, Di LJ, et al. (2006). Active chromatin hub of the mouse α -globin locus forms in a transcription factory of clustered housekeeping genes. *Mol. Cell. Biol.* 26: 5096-5105.