Minireview

A Melanesian α-thalassemia mutation suggests a novel mechanism for regulating gene expression

Qiliang Li

Address: Division of Medical Genetics, Department of Medicine, University of Washington, Seattle, WA 98195, USA.
Email: li111640@u.washington.edu

Published: 24 October 2006

Genome ***Biology*** 2006, **7:**238 (doi:10.1186/gb-2006-7-10-238)

The electronic version of this article is the complete one and can be found online at http://genomebiology.com/2006/7/10/238

Abstract

A Melanesian variant of the genetic disease α-thalassemia has recently been shown to be due to a single-nucleotide polymorphism located between the adult α-globin genes and their enhancers. The finding that this mutation creates a novel promoter provides support for a mechanism of gene regulation by facilitated chromatin looping.

Approximately 0.1% of the human genome sequence is responsible for the variation among individuals, and the majority of these differences are single-nucleotide polymorphisms (SNPs). Although most SNPs are stable and have no deleterious effects, others are likely to contribute to individuality, disease susceptibility, and individual responses to therapeutic drugs. At present, such 'functional' SNPs have mostly been identified in diseases that are caused by defects in a single gene (monogenic diseases), although SNPs have also been linked to complex diseases such as hypertension, diabetes, heart disease, and cancer, as well as to responses to drugs. SNPs are important not only in medicine but also in basic molecular biology as they represent a natural library of variations that can be used to elucidate and validate mechanisms of gene expression *in vivo*.

In this regard, a SNP in the gene for myostatin - a transcription factor inhibiting muscle development - has recently been shown to contribute to the muscular hypertrophy typical of the Texel breed of sheep [1]. A G→A transition in the 3′ untranslated region of the gene for myostatin creates a target site for a microRNA, which results in translational inhibition of myostatin and consequent muscle growth. Another SNP has recently been shown by De Gobbi *et al.* [2] to be responsible for a severe form of α-thalassemia found in Melanesia. This SNP disrupts gene expression by another novel mechanism, creating an illegitimate promoter site between the globin enhancers and the adult α-globin gene cluster, which has the effect of downregulating expression of the cluster. The molecular mechanism underlying this downregulation is not yet established, but the new findings provide the basis for some interesting speculations.

Diseases caused by defects in the structure and expression of the globin genes, and thus of hemoglobin, represent the most complete repertoires of monogenic defects known to date. In nearly all cases, the molecular basis of these hemoglobinopathies has been identified [3]. Defects have been identified in protein structure, gene expression, and chromatin organization. The underlying genetic defect in individuals from Melanesia with a particular form of α-thalassemia has, however, eluded researchers for a long time. The disease is characterized by severe anemia, consequent on a marked downregulation of α-globin and the production of excess tetramers of β-globin (β4), also known as hemoglobin H (HbH), which precipitate in the red blood cells. Extensive analysis of the α-globin cluster and the surrounding 300 kb of DNA, however, revealed no deletions or chromosomal rearrangements.

The human α-globin locus is located within the telomeric region of chromosome 16 (Figure 1a). The locus contains a ζ-globin gene, which is expressed at the embryonic stage of development, two α-globin genes (α_1 and α_2), which are expressed at the fetal and adult stages, and several minor genes that are expressed at a low level, including the α^D-globin gene. The physiological levels of α-like globin gene expression depend on a group of enhancers that lie distal to the 5′ side of the ζ gene, with each enhancer being

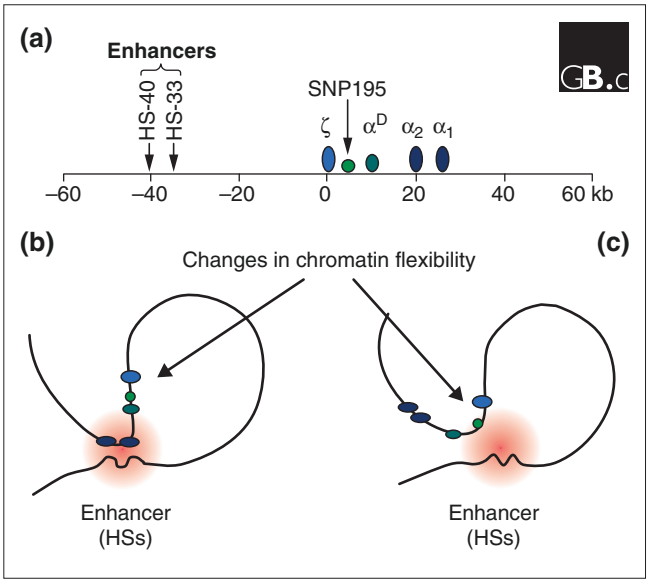

Figure 1
Possible mechanism for the downregulation of α-globin gene expression in Melanesian hemoglobin H (HbH) disease. **(a)** Schematic diagram of the human α-globin locus. The ζ-globin gene (the light-blue oval) is expressed in the embryonic stage of development and is silenced at around 6 to 8 weeks of gestation. The α-globin genes (the dark-blue ovals) are activated in fetal liver, and then in bone marrow in the adult. The physiological levels of α-like globin gene expression depend on the actions of upstream enhancers (HS-33 and HS-40) - mainly HS-40, which is located 40 kb 5′ of the ζ-gene. The scale on the figure indicates distances in kilobases from the start of the α-gene cluster. The single nucleotide polymorphism (SNP) 195 is shown as a green circle. **(b,c)** A possible explanation for the SNP195 promoter-induced downregulation of the α-globin genes. Effective interaction between proteins bound by the enhancer (depicted schematically as a red circle indicating the range of influence of the enhancer) and the α-globin genes is essential for their high-level expression, and is accomplished by chromatin looping. (b) In the normal locus the SNP195 region is lightly acetylated and chromatin flexibility favors interaction between the enhancer and the α_1- and α_2-genes. (c) When the SNP195 promoter site (green circle) is activated in Melanesian HbH disease, histone acetylation is increased and the chromatin becomes more flexible as a consequence, resulting in a change in loop size. This change means that the enhancer now preferentially interacts with the new promoter, and no longer influences expression of the globin genes.

characterized by an erythroid-specific DNase I hypersensitive site (HS). Of these, HS-40 has the most powerful enhancer activity [4].

De Gobbi *et al.* [2] have now characterized the mutation responsible for Melanesian HbH disease as a gain-of-function allele of SNP 195 that creates a new promoter in an intergenic region just upstream of the α^D-globin gene, and approximately 13 kb upstream of the adult α-globin genes (Figure 1a). This promoter separates the adult α-globin genes from their upstream enhancers and has the effect of severely downregulating their expression, revealing a novel means of disrupting α-globin gene expression. The disease α-thalassemia results when production of the α-globin and β-globin chains that make up hemoglobin is unbalanced. Hemoglobin is a tetramer of two α-like and two β-like globin chains and the two kinds of globins are normally synthesized at equal levels. Downregulation of one copy of the α-globin gene causes anemia with mild changes in red blood cells, the so-called α-thalassemia trait, but when α-globin gene expression is reduced to less than 50% of normal, the excess β-globin chains form tetramers that precipitate in the red blood cell, causing a more severe anemia called HbH disease.

The mutation identified by De Gobbi *et al.* [2] is an A→G transition lying between the ζ gene and the α^D gene. Remarkably, they found that transcription from the region around the mutant was increased 1,000 times compared to the wild-type chromatin, as analyzed by the tilted array expression assay [2], but that expression of the α^D gene, located immediately downstream of SNP195, was reduced by around 80-fold. Reverse transcription-coupled PCR (RT-PCR) of the RNA isolated from a Melanesian patient showed that transcription of the α_1- and α_2-globin genes, which are located approximately 13 kb downstream of the mutant, were also decreased in the mutant allele, as expected from the phenotype. The A→G transition created a known binding site (TAATAA→TGATAA) for the erythroid-specific *trans*-acting factor GATA1. This altered binding site also nucleates the binding of a pentameric erythroid complex, including the transcription factors SCL, E2A, LMO2, and Ldb-1, as analyzed by chromatin immunoprecipitation (ChIP) studies [2]. Unlike the wild-type SNP allele, the mutant allele binds RNA polymerase II, suggesting that a new promoter has been created by the mutation. In addition, De Gobbi *et al.* [2] carried out a ChIP assay that showed that the mutation resulted in an increase in acetylated histones H3. In summary, they found that SNP195 creates a new promoter-like element between the upstream regulatory element and its cognate promoters. This element, when activated, causes significant downregulation of the α^D, α2, and α1 genes that lie downstream, thus causing the thalassemia.

Trapped enhancers

This Melanesian form of HbH disease is the first natural example of a mutation that causes the function of an enhancer to be 'trapped' by an intervening promoter. Similar observations have been reported previously in several transgenic studies. The insertion of a gene for hygromycin B resistance between the DNase-hypersensitive sites HS-1 and HS-2 in the locus control region for the murine β-globin locus resulted in the inactivation of the downstream β_m-globin gene, located approximately 40 kb downstream of the locus control region [5]. In the granzyme B locus the insertion of the PGK-*neo* gene (a neomycin phosphotransferase gene driven by the phosphoglycerate kinase I promoter) into the furthest downstream gene in the cluster severely reduced the normal expression of the downstream genes within the locus, even those at a distance greater than 100 kb from the mutation [6], suggesting that the enhancing activity of an

upstream regulatory element is disrupted by the inserted gene. External genes (a neomycin-resistance gene or α-globin gene) have been inserted by homologous recombination into the human α-globin gene locus in a hybrid MEL cell line, which harbors one copy of the human chromosome 16, in both possible orientations either upstream or downstream of the HS-40 region. In this case, each insertion led to a severe decrease in HS-40-dependent transcription of the α-globin genes approximately 50 kb downstream [7]. The common feature in these experiments is that when an active promoter is placed between a distal enhancer and its cognate gene, the enhancer activity is caught by the inserted promoter, resulting in downregulation of the distal gene(s).

How does a promoter trap the activity of a nearby enhancer and downregulate the expression of a distal gene? So far there is no consensus. One proposal is that a *trans*-activating complex recruited by the enhancer is able to track along the chromatin fiber and is captured by the first promoter it encounters, leading to the inactivation of downstream genes [8]. This model assumes the directionality of enhancer activity, and thus is able to explain the preference for the proximal promoter over the distal one. However, the tracking model has difficulty in explaining why the expression of the α-globin gene was reduced to similar levels when a *neo* gene was placed either 5′ or 3′ to the HS-40 enhancer [7]. An alternative and widely accepted model for the interaction of distal enhancers with their promoters is chromatin looping [9,10]. This proposes that transcriptional enhancement by an enhancer distal to the cognate genes is mediated by the formation of a chromatin loop that brings the two elements into proximity (Figure 1b). Because a 30 nm chromatin fiber has a certain stiffness, chromatin flexibility will determine the loop size [11,12]. The facilitated chromatin looping model hypothesizes that chromatin flexibility, and thus the preferential looping profile, is modulated by histone acetylation and other modifications [11]. Greater flexibility favors the formation of a smaller loop. On the basis of this model, an event that can increase the level of histone acetylation (for instance, the presence of an active promoter) at a region near an enhancer will downregulate expression of distal genes on either side of the enhancer (Figure 1c).

The Melanesian HbH disease is the first natural *in vivo* example showing that the generation of a promoter, which increases the level of histone acetylation in the region between the distal enhancer and the α-globin genes, damages the expression of the downstream genes. If facilitated chromatin looping proves to be the mechanism in this case, it might turn out to be a more general regulatory process in multigene clusters that require a remote enhancer for high-level transcription. The discovery by De Gobbi *et al.* [2] of such a mutation in an apparently functionless intergenic region suggests that intergenic regions can participate in gene regulation in eukaryotes, and that searches for functional SNPs should include such regions.

Acknowledgements

The author's laboratory is funded by NIH grant HL73439.

References

1. Clop A, Marcq F, Takeda H, Pirottin D, Tordoir X, Bibe B, Bouix J, Caiment F, Elsen JM, Eychenne F, *et al.*: **A mutation creating a potential illegitimate microRNA target site in the myostatin gene affects muscularity in sheep.** *Nat Genet* 2006, **38:** 813-818.
2. De Gobbi M, Viprakasit V, Hughes JR, Fisher C, Buckle VJ, Ayyub H, Gibbons RJ, Vernimmen D, Yoshinaga Y, de Jong P, *et al.* **A regulatory SNP causes a human genetic disease by creating a new transcriptional promoter.** *Science* 2006, **312:**1215-1217.
3. Steinberg MH: *Disorders of Hemoglobin: Genetics, Pathophysiology and Clinical Management.* Cambridge: Cambridge University Press; 2001.
4. Higgs DR, Wood WG, Jarman AP, Sharpe J, Lida J, Pretorius IM, Ayyub H: **A major positive regulatory region located far upstream of the human alpha-globin gene locus.** *Genes Dev* 1990, **4:**1588-1601.
5. Kim CG, Epner EM, Forrester WC, Groudine M: **Inactivation of the human beta-globin gene by targeted insertion into the beta-globin locus control region.** *Genes Dev* 1992, **6:**928-938.
6. Pham CT, MacIvor DM, Hug BA, Heusel JW, Ley TJ: **Long-range disruption of gene expression by a selectable marker cassette.** *Proc Natl Acad Sci USA* 1996, **93:**13090-13095.
7. Esperet C, Sabatier S, Deville MA, Ouazana R, Bouhassira EE, Godet J, Morle F, Bernet A: **Non-erythroid genes inserted on either side of human HS-40 impair the activation of its natural alpha-globin gene targets without being themselves preferentially activated.** *J Biol Chem* 2000, **275:**25831-25839.
8. Blackwood EM, Kadonaga JT: **Going the distance: a current view of enhancer action.** *Science* 1998, **281:**60-63.
9. Tolhuis B, Palstra RJ, Splinter E, Grosveld F, de Laat W: **Looping and interaction between hypersensitive sites in the active beta-globin locus.** *Mol Cell* 2002, **10:**1453-1465.
10. Carter D, Chakalova L, Osborne CS, Dai YF, Fraser P: **Long-range chromatin regulatory interactions *in vivo*.** *Nat Genet* 2002, **32:** 623-626.
11. Li Q, Barkess G, Qian H: **Chromatin looping and the probability of transcription.** *Trends Genet* 2006, **22:**197-202.
12. Bon M, Marenduzzo D, Cook PR: **Modeling a self-avoiding chromatin loop: relation to the packing problem, action-at-a-distance, and nuclear context.** *Structure* 2006, **14:**197-204.