

# The *HBS1L-MYB* intergenic interval associated with elevated HbF levels shows characteristics of a distal regulatory region in erythroid cells

Karin Wahlberg,<sup>1</sup> Jie Jiang,<sup>1</sup> Helen Rooks,<sup>1</sup> Kiran Jawaid,<sup>1</sup> Fumihiko Matsuda,<sup>2,3</sup> Masao Yamaguchi,<sup>2,3</sup> Mark Lathrop,<sup>3</sup> \*Swee Lay Thein,<sup>1,4</sup> and \*Steve Best<sup>1</sup>

<sup>1</sup>Division of Gene and Cell Based Therapy, King's College London School of Medicine, London, United Kingdom; <sup>2</sup>Center for Genomic Medicine, Kyoto University Graduate School of Medicine, Yoshida, Japan; <sup>3</sup>Centre National de Génotypage, Institut Génomique, Commissariat à l'Energie Atomique, Evry, France; and <sup>4</sup>Department of Haematological Medicine, King's College Hospital, London, United Kingdom

***HBS1L-MYB* intergenic polymorphism (*HMIP*) on chromosome 6q23 is associated with elevated fetal hemoglobin levels and has pleiotropic effects on several hematologic parameters. To investigate potential regulatory activity in the region, we have measured sensitivity of the sequences to DNase I cleavage that identified 3 tissue-specific DNase I hypersensitive sites in the core intergenic interval. Chromatin immunoprecipitation with microarray (ChIP-chip) analysis showed**

**strong histone acetylation in a defined interval of 65 kb corresponding to the core *HBS1L-MYB* intergenic region in primary human erythroid cells but not in non-*MYB*-expressing HeLa cells. ChIP-chip analysis also identified several potential *cis*-regulatory elements as strong GATA-1 signals that coincided with the DNase I hypersensitive sites present in *MYB*-expressing erythroid cells. We suggest that *HMIP* contains regulatory sequences that could be important in hema-**

**topoiesis by controlling *MYB* expression. This study provides the functional link between genetic association of *HMIP* with control of fetal hemoglobin and other hematologic parameters. We also present a large-scale analysis of histone acetylation as well as RNA polymerase II and GATA-1 interactions on chromosome 6q, and  $\alpha$  and  $\beta$  globin gene loci. The data suggest that GATA-1 regulates numerous genes of various functions on chromosome 6q. (Blood. 2009;114:1254-1262)**

## Introduction

Variable levels of fetal hemoglobin (HbF,  $\alpha_2\gamma_2$ ) persist into adulthood, and although they have no clinical consequences in otherwise healthy individuals, high HbF levels have a major impact on the principal  $\beta$  hemoglobin disorders— $\beta$  thalassemia and sickle cell disease. Increased HbF production mitigates the severity of both diseases.<sup>1-3</sup> The level of HbF in adults is inherited as a quantitative trait, and is largely genetically controlled with a heritability of 0.89.<sup>4</sup>

Three loci—*HBS1L-MYB* intergenic region on chromosome 6q23, *BCL11A* on chromosome 2p16, and the  $\beta$  globin cluster on chromosome 11—account for up to 50% of the variation in HbF levels in patients with sickle cell anemia or  $\beta$  thalassemia and in healthy European whites.<sup>5-7</sup> The *HBS1L-MYB* intergenic region alone contributes approximately 20% of the overall trait variance in healthy European whites,<sup>5,8</sup> and 3% to 7% of the trait variance in African-American and Brazilian patients with sickle cell anemia.<sup>6</sup>

The panel of single nucleotide polymorphisms in the *HBS1L-MYB* region that account for the effects of the 6q locus<sup>6,8-10</sup> reside in a nearly contiguous segment of 79 kb distributed in 3 linkage disequilibrium blocks, referred to as *HBS1L-MYB* intergenic polymorphism (*HMIP*) blocks 1, 2, and 3.<sup>8</sup> Genetic variants that show the strongest effects are concentrated in 24 kb of *HMIP* block 2, located 33 kb upstream of *HBS1L* and 65 kb upstream of *MYB*.<sup>8</sup> The mechanisms through which these variants operate to increase HbF are still not clear, but studies suggest that the biologic effects are likely to involve regulation of the flanking genes—*HBS1L* and *MYB*. *MYB* and *HBS1L* expression was significantly reduced in

erythroid cultures of individuals with high HbF levels, whereas overexpression of *MYB* in K562 cells inhibited  $\gamma$ -globin expression supporting *MYB*'s role in HbF regulation.<sup>11</sup> Further, *HBS1L* and *MYB* expression was positively correlated in erythroid progenitor cells, and *HBS1L* expression correlated with the genetic variants associated with HbF.<sup>8</sup> Variability in *HMIP* block 2 was subsequently shown to have a pleiotropic effect on erythrocyte count and volume, and platelet and monocyte counts in healthy Europeans.<sup>12</sup> These observations suggest that the *HBS1L-MYB* intergenic region is functionally active, containing distal regulatory sequences for the flanking genes—*HBS1L* and *MYB*. The function of *HBS1L* is unknown but it encodes a protein with apparent GTP-binding activity, involved in the regulation of a variety of critical cellular processes,<sup>13</sup> and *MYB* encodes a transcription factor involved in oncogenesis and with an essential role in erythropoiesis.<sup>14-16</sup>

Initially, we investigated the regulatory potential of *HMIP* block 2 by measuring sensitivity of the sequences in this region to DNase I cleavage that identified multiple DNase I hypersensitive sites in the region in K562 cells. We then proceeded to a comprehensive analysis of the regulatory potential of a large region of chromosome 6q using chromatin immunoprecipitation (ChIP) and microarray (ChIP-chip) analysis on primary human erythroid progenitor cells. We identified strong signals of histones H3 and H4 acetylation in the *HBS1L-MYB* intergenic region (indicative of active chromatin) especially concentrated in block 2, in basophilic erythroblasts when the globin genes and *MYB* are fully active. Tissue specificity of the regulatory activity in the intergenic region

Submitted March 12, 2009; accepted June 7, 2009. Prepublished online as *Blood* First Edition paper, June 15, 2009; DOI 10.1182/blood-2009-03-210146.

\*S.L.T. and S.B. contributed equally to this work.

The online version of this article contains a data supplement.

The publication costs of this article were defrayed in part by page charge payment. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 USC section 1734.

© 2009 by The American Society of Hematology

was demonstrated by minimal histone acetylation in HeLa cells. ChIP-chip also demonstrated interactions of the erythroid-specific transcription factor GATA-1 with several sites in the *HBSIL-MYB* interval. GATA-1 signals in coincidence with the DNase I hypersensitive sites in *HMIP* block 2 strongly suggest the presence of regulatory elements. Regulatory activity of the intergenic region was validated by presence of intergenic transcripts in erythroid precursor cells in a tiling microarray. We postulate that the regulatory elements distally control *MYB* expression, which in turn influences erythroid differentiation and, indirectly, the control of HbF levels.

The use of microarrays also allowed us to compare patterns of activity in the candidate interval with other regions encompassing widely expressed genes across 70 Mb of chromosome 6q. We further provide a large-scale analysis of GATA-1 occupancy in erythroid cells that includes the entire  $\beta$  and  $\alpha$  globin gene clusters.

## Methods

### Cells and cell cultures

Cell lines were maintained in RPMI-1620 medium (Sigma-Aldrich) with the addition of 10% of fetal calf serum (FCS; PAA-laboratories), 2 mM L-glutamine (Sigma-Aldrich), 0.1 mg/mL streptomycin, and 18 units/mL penicillin (Sigma-Aldrich). Concentrations were kept at 0.5 to  $1.0 \times 10^6$  cells/mL. K562 cells were treated with 40  $\mu$ M hemin (Sigma-Aldrich) for 24 hours to induce differentiation.

Primary human erythroid cells were cultured from peripheral blood using a 2-phase liquid system as previously described.<sup>11,17</sup> Cytospins of erythroid progenitors from different days of culture were stained using a Giemsa staining set (Hema "Gurr"; VWR) according to the manufacturer's protocol.

Flow cytometry of primary erythroid cells was performed with anti-human CD71 monoclonal antibodies (FITC conjugated, 555536; BD Biosciences) or anti-human glycophorin A (GPA, phycoerythrin [PE] conjugated, R7078; DAKO) as previously described.<sup>11</sup>

### DNase I hypersensitivity analyses

DNase I hypersensitivity analysis of *HMIP-2* was performed on 2 biologic replicates of induced and uninduced K562 cells, and Jurkat cells. Nuclei were treated with 70 units of DNase I at 37°C for 3 minutes as determined by optimization experiments (supplemental methods and supplemental Figure 1, available on the *Blood* website; see the Supplemental Materials link at the top of the online article). Real-time quantitative polymerase chain reaction (PCR)<sup>18,19</sup> was performed in triplicate on 20 ng DNA samples using SYBR Green PCR Mastermix (Applied Biosystems) and the ABI Prism 7900HT Sequence Detection system (Applied Biosystems).

The *HMIP-2* region could be covered in 68 overlapping fragments of approximately 500 bp (PCR primer sequences available on request). Relative sensitivity to DNase I for each target was calculated by converting delta  $C_T$  (difference in  $C_T$  values between treated and untreated DNAs) to a linear scale and plotted as a function of primer position. Values were normalized to the negative control *NEFM* (supplemental methods and supplemental Figure 1) to account for differences in treatment conditions.

### Chromatin immunoprecipitation

Antibodies used for ChIP experiments included anti-diacetylated histone H3 (K9 and 14; no. 06-599; Millipore), anti-tetra-acetylated histone H4 (K5, 8, 12, and 16; no. 06-866; Millipore), anti-RNA polymerase II (no. 05-623; Millipore), and anti-GATA-1 (no. sc-1234; Santa Cruz Biotechnology).

ChIP experiments for acetylated histones were performed using the EZ-ChIP protein G kit (no. 17-371; Millipore) according to the manufac-

er's protocol with minor modifications (Dr David Garrick, Weatherall Institute of Molecular Medicine, Oxford, United Kingdom). ChIP for RNA polymerase II and GATA-1 was performed using the Magna ChIP protein G kit (no. 17-611; Millipore) according to the manufacturer's protocol with minor modifications.

ChIP assays were performed on cultured primary human erythroid cells (phase II, day 10) and HeLa cells. Cells ( $5 \times 10^7$  per experiment) were cross-linked in 10 mL growth medium with 1% formaldehyde (Sigma-Aldrich) for 10 minutes at room temperature, and the chromatin was sonicated ( $10 \times 15$  seconds, Sonic VibraCell at 40% efficiency) to a size of approximately 500 base pairs (bp; range 200-1000 bp). Immunoprecipitations were performed after an overnight incubation with 5 to 10  $\mu$ g of the appropriate antibody, with protein G agarose beads, or with protein G magnetic beads. A sample containing no antibody was used as a negative control.

ChIP material was validated by SYBR Green quantitative PCR before microarray analyses using different positive control targets. Enrichment of a specific target sequence in ChIP material was calculated relative to input DNA, and the results were normalized to a control sequence in exon 1 of the neurofilament gene (*NEFM*) representing an inactive gene (data not shown).

### Microarray analysis of ChIP material

Two microarrays, both from Roche-NimbleGen, were used in these experiments as described in the supplemental methods. The first array encompassed 70 Mb (positions 93 424 310 to 165 905 673; hg18) on chromosome 6, including the 6q23 HbF locus. The second was a custom array and included the globin loci as controls. The microarray data have been deposited with Gene Expression Omnibus (GEO) under accession number GSE16541.<sup>20</sup>

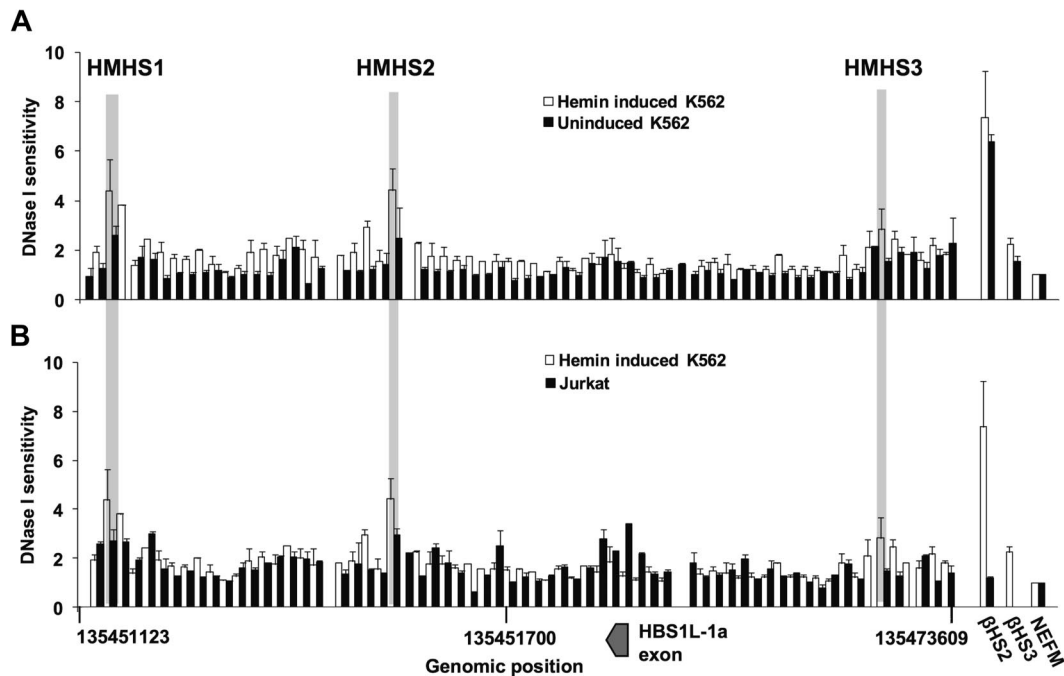
Input and ChIP DNA were amplified using a whole genome amplification kit (WGA1; Sigma-Aldrich) applying a previously described protocol adapted for ChIP material.<sup>21</sup> Amplified DNA was purified using a QIAquick PCR purification kit (QIAGEN) with buffer PBI substituted for buffer PB.

Arrays were hybridized and washed using Roche-NimbleGen kits according to the manufacturer's protocol. Scanning was performed using a GenePix 4000B Scanner (Molecular Devices). Detailed protocol, including data extraction and analysis, is shown in supplemental methods.

### Transcript Tiling Array

A customized Affymetrix GeneChip Tiling Array was designed to identify novel transcripts<sup>22</sup> between positions 135 323 209 and 135 582 003 (NCBI build 36<sup>23</sup>) that encompass the entire *MYB* and *HBSIL* genes as well as the intergenic region; 116 858 bases containing human repetitive sequence were excluded from probe design. The remaining 141 937 bases were tiled by overlapping oligonucleotides. The starting position of each 25mer oligo was shifted by 2 bases within *HBSIL* and *MYB*, and by one base within the intergenic region. Oligonucleotides were designed for both DNA strands. A total of 224 258 oligonucleotide probes was included in each array.

Total RNA was extracted from erythroid cells (liquid culture phase II, days 3 and 5)<sup>11,17</sup> using Tri-Reagent (Sigma-Aldrich). Five micrograms was treated with the RiboMinus Transcriptome Isolation Kit (Invitrogen) to remove ribosomal RNA before cDNA synthesis according to the manufacturer's instructions. Synthesis of cDNA, fragmentation, and end labeling was performed using the GeneChip Whole Transcript Double-Stranded Target Assay kit (Affymetrix) according to the manufacturer's instructions. After hybridization for 16 hours, the arrays were washed and stained using a Fluidics Station 450 (Affymetrix). Scanning was performed by GeneChip Scanner 3000 7G (Affymetrix). Data analysis was performed using Affymetrix Tiling Analysis Software Version 1.1 and data were visualized using the Affymetrix Integrated Genome Browser. Because double-stranded cDNA was used, the transcript maps correspond to the signals from both strands.



**Figure 1.** DNase I digestion profiles of the *HBS1L-MYB* intergenic region in *HMIP* block 2. DNase I sensitivity was analyzed by real-time PCR using 68 primer pairs spanning the 24-kb *HMIP-2* region. Control primers targeting the  $\beta$ -globin LCR HS2 and HS3 ( $\beta$ HS2 and  $\beta$ HS3) and the *NEFM* gene were also included in experiments. Relative sensitivity to DNase I for each target was calculated from delta  $C_T$  values between treated and untreated samples and normalized to *NEFM*. DNase I sensitivity was plotted as a function of primer position. Error bars represent differences between 2 biologic repeats. (A) DNase I sensitivity in *HMIP-2* compared between hemin-induced and uninduced K562 cells. Three DNase I hypersensitive sites (HMHS1, HMHS2, and HMHS3; gray bars) were identified in hemin-induced K562 cells. (B) DNase I sensitivity in *HMIP-2* compared between hemin-induced K562 cells and Jurkat cells. The 3 hypersensitive sites in hemin-induced K562 cells showed no or less sensitivity to DNase I in Jurkat cells. Hypersensitivity was instead observed in Jurkat cells at the site of the *HBS1L-1a* exon, which showed no hypersensitivity in K562 cells (induced and uninduced). The  $\beta$ HS3 control was not included for DNase I sensitivity analyses in Jurkat cells.

## Results

### *HBS1L-MYB* intergenic region is sensitive to DNase I in erythroid cells

Variants in block 2 of the *HBS1L-MYB* intergenic region account for the majority of the genetic association with HbF in the 6q QTL region. We performed DNase I hypersensitivity analysis along the entire 24 kb of *HMIP-2* in K562 cells as a screen for genetic regulatory elements. K562 cells treated with hemin were induced to differentiate, resulting in strong up-regulation of globin gene expression and simultaneous down-regulation of *MYB* expression.<sup>24,25</sup> After 24 hours of hemin treatment (at 40  $\mu$ M),  $\beta$  globin expression increased 9.5-fold, that of  $\gamma$  globin, 5-fold, whereas *MYB* expression decreased by 7-fold (data not shown). DNase I hypersensitivity profiles were studied in uninduced and hemin-induced K562 cells, and in Jurkat cells (T-cell leukemia), representing a nonerythroid cell line. Two biologic repeats were performed for each cell type. Analysis of both  $\beta$ HS2 and  $\beta$ HS3 (positive controls) and *NEFM* (negative control) were included for K562 cells and  $\beta$ HS2 and *NEFM*, for Jurkat cells.

In uninduced K562 cells, several sites within *HMIP-2* showed sensitivity to DNase I above background levels, indicating an open chromatin structure (Figure 1A). When cells were induced to differentiate, the region showed a general increase in DNase I sensitivity; 3 sites referred to here as *HBS1L-MYB* (HM) HS1, HS2, and HS3, in particular, showed a marked increase in sensitivity compared with background levels (Figure 1A). DNase I sensitivity also increased for  $\beta$ HS2 and  $\beta$ HS3 controls, consistent with the induction in globin gene expression. HMHS1, HMHS2, and HMHS3 also showed stronger sensitivity to DNase I than the

$\beta$ HS3 control, thereby reaching a threshold level for hypersensitivity (Figure 1A). There was no difference in DNase I sensitivity for *NEFM* in uninduced and induced K562 cells.

As expected,  $\beta$ HS2 was not sensitive to DNase I in Jurkat cells, a nonerythroid cell line (Figure 1B). In the *HMIP* block 2 region, Jurkat cells show similar background levels and generally, a similar DNase I sensitivity profile to induced K562 cells, but with much less sensitivity at HMHS1, HMHS2, and HMHS3, indicating that DNase I sensitivity at these sites is tissue specific (Figure 1B). The strongest sensitivity to DNase I in Jurkat cells coincided with the putative promoter region of the alternative *HBS1L* exon (exon 1a), which showed a low degree of sensitivity in K562 cells (both uninduced and induced). This is consistent with expression of the *HBS1L-1a* transcript in Jurkat cells but not in K562 cells.<sup>8</sup> To validate the promoter prediction at *HBS1L-1a*, we examined its functional property in a reporter assay; the region showed activity in Jurkat cells but not in K562 cells (supplemental Figure 2). The *HBS1L* exon 1a promoter therefore served as a positive internal control for DNase I hypersensitivity at an active regulatory element within the *HMIP* block 2 region.

### Characterization of erythroid and nonerythroid cells for use in chromatin immunoprecipitation

The identification of DNase I hypersensitive sites in *HMIP* block 2 in hemin-induced K562 cells suggested that the intergenic region contained regulatory elements active in erythroid lineages and encouraged further functional analysis of the interval. We assessed the chromatin activity and transcription factor binding throughout the *HBS1L-MYB* and flanking regions of chromosome 6q, and tissue-specific activity of these profiles. Primary human erythroid progenitors cultured in a 2-phase liquid system were analyzed to

select optimum time of harvest when *HBSIL*, *MYB*, and the globin genes were fully expressed. HeLa cells served as examples of a *MYB*-negative cell line.

Basophilic erythroblasts from phase II, day 10, were chosen for ChIP analysis (supplemental methods and supplemental Figure 3). At this stage, expression profiles indicated that the cells were in a state of high transcriptional activity; the candidate genes (*HBSIL* and *MYB*) as well as *GATA1* and the globin genes were expressed at high levels (data not shown).

To investigate tissue specificity of activity in the intergenic region and the relation between activity and candidate gene expression, we compared histone acetylation patterns in the intergenic region between erythroid precursor cells and a cell line lacking *HBSIL* and *MYB* expression. Eight cell lines, which included Jurkat, HL60 (promyelocytic leukemia), U937 (monocytic leukemia), HeLa (cervical cancer), HEK293 (kidney), HuH-7 (liver carcinoma), U2OS (osteosarcoma), and HKC-8 (renal epithelial), were screened, together with K562 cells and primary erythroid cells for *HBSIL* and *MYB* expression using TaqMan reverse transcription PCR. *HBSIL* was highly expressed in all lines analyzed (real-time PCR  $C_T$  values of 23-26), which indicates a housekeeping function (supplemental Figure 4A). In contrast, *MYB* expression varied dramatically; it was highly expressed in all hematopoietic cell-related lines ( $C_T$  values of 24-26) but minimally expressed in other lines ( $C_T$  values of 31-35; supplemental Figure 4B). HeLa, which showed relatively low *HBSIL* expression (10% of expression in erythroid precursors) and insignificant *MYB* expression (0.01% of expression observed in erythroid precursors), was chosen to represent a *MYB*-negative cell line in ChIP experiments.

### Overview of histone acetylation, GATA-1, and RNA polymerase II interactions across chromosome 6q

On viewing the ChIP-chip data from erythroid precursors across 70 Mb of chromosome 6q represented on the array, it was evident that all antibodies showed similar patterns of activity. The strongest peaks were found in gene-rich areas, whereas large intergenic sequences lacked signal (supplemental Figure 5). Well-defined areas of high levels of AcH3 were seen at transcriptional start sites (TSSs) with a less defined pattern for AcH4 as previously described.<sup>26</sup> Interestingly, abundant GATA-1 signals were found over the entire 70-Mb region in coincidence with RNAP II signals at active genes. The gene-free *HBSIL-MYB* interval showed strong signals for all antibodies in erythroid cells, indicating a high level of activity in this region.

Publicly available expression data from the UCSC Genome Browser (<http://genome.ucsc.edu>)<sup>27</sup> revealed that a majority of the genes that showed strong signals for histone acetylation, RNAP II, and GATA-1 were hematopoietic specific, but a few ubiquitous genes were also included. The genes represented a wide range of functions including transcription factors, adhesion receptors, and signaling proteins, reflecting a broad target repertoire for GATA-1 as a transcription factor (supplemental Table 1).

### The *HBSIL-MYB* intergenic region is highly active in erythroid precursors

ChIP-chip data of GATA-1, AcH3, AcH4, and RNAP II in erythroid precursors and AcH3 in HeLa cells was analyzed across a 2.5-Mb region of the 6q23 HbF locus, encompassing the 5 protein coding genes (including the *HBSIL-MYB* intergenic region), and flanking sequences (Figure 2). Biologic ChIP-chip replicates using primary

human erythroblasts (phase II day 10 of culture) from 2 individuals provided very similar profiles.

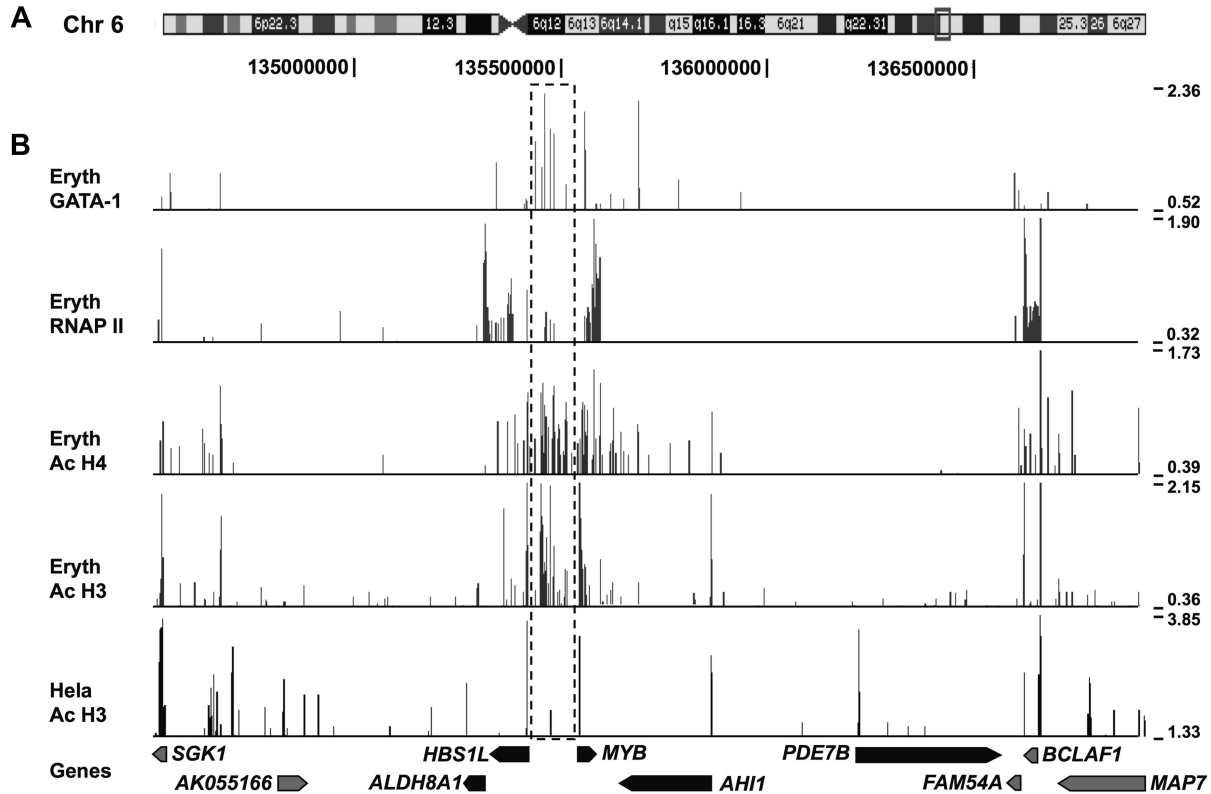
Histone acetylation patterns and RNAP II signal in the 6q23 locus in erythroid precursors were consistent with previous expression analysis of genes in the locus<sup>11</sup> (Figure 2). The *PDE7B* and *ALDH8A1* genes are not expressed in erythroid precursors and consistently showed no signal of acetylation or RNAP II interaction. *AH11* that is expressed at low levels showed some histone acetylation signal at the promoter regions. In contrast, the highly expressed *HBSIL* and *MYB* genes were associated with strong RNAP II signal and histone acetylation around the promoter regions as well as coding regions. The RNAP II antibody used detects both nonphosphorylated inactive RNAP II as well as phosphorylated actively elongating forms. Interestingly, although *MYB* is the most highly expressed gene in the region, we see no RNAPII interaction at the immediate promoter-proximal 5' end, but instead, high levels in the body of the gene and a large accumulation toward the 3' end and beyond (Figures 2-3). We speculate that the lack of RNAPII at the promoter is a result of rapid transcription leaving no inactive RNAPII stalling at the initiation complex. The phenomenon of RNAPII accumulation at the 3' end of actively transcribed genes has previously been observed and is likely to reflect pausing and dephosphorylation of the polymerase before release.<sup>28</sup> With nonexpressed genes, RNAP II was often seen as a strong signal but only at the 5' preinitiation complex. Strikingly, the AcH3 and AcH4 signals were equally strong in the *HBSIL-MYB* intergenic region as around the *HBSIL* and *MYB* promoters and exons, indicating that this region is highly active in erythroid precursors. In the context of the whole of 6q region analyzed, there was little evidence of such high intergenic activity elsewhere.

In contrast to the strong signals observed in erythroid cells, HeLa cells showed minimal histone acetylation, consistent with the low/absent candidate gene expression in these cells.

GATA-1 signal in the 6q23 locus was concentrated around the *MYB* and *HBSIL* area, with the strongest peaks in the core intergenic region. In fact, the GATA-1 signal in the *HBSIL-MYB* intergenic region represented the most significant GATA-1 peaks in the entire 70-Mb region covered on the 6q array.

### The *HMIP-2* and *-3* regions show characteristics of a distal regulatory region in erythroid precursors

A closer view of the *HBSIL-MYB* intergenic region revealed that the histone H3 and H4 acetylation in erythroid precursor cells was found in a defined 65-kb interval encompassing the *HMIP-2* and *-3* regions (Figure 3). Within this interval, the strongest AcH3 and GATA-1 signals were concentrated in the *HMIP-2* region. The intergenic region included 7 peaks of GATA-1 signal, 3 of which were within *HMIP-2* and 1 was in *HMIP-3*. The GATA-1 signals in the *HMIP-2* region all coincided with DNase I hypersensitive sites identified in induced K562 cells as a further indication of these sites being functional regulatory elements. In addition to a GATA-1 signal approximately 7 kb 5' of *MYB*, strong signals were seen in intron 5 of *MYB* (just upstream of exon 6) and in intron 8 (just upstream of exon 9). No GATA-1 signal, however, was detected at the immediate *MYB* promoter. GATA-1 signal was also observed at the *HBSIL* promoter region. There was a prominent coincidence of GATA-1 signal on the array with conserved GATA-1 motifs (human/mouse/rat alignment available from the UCSC Genome Browser; <http://genome.ucsc.edu>),<sup>27</sup> which support a functional relevance of these sites. Some weak RNAP II signals were also

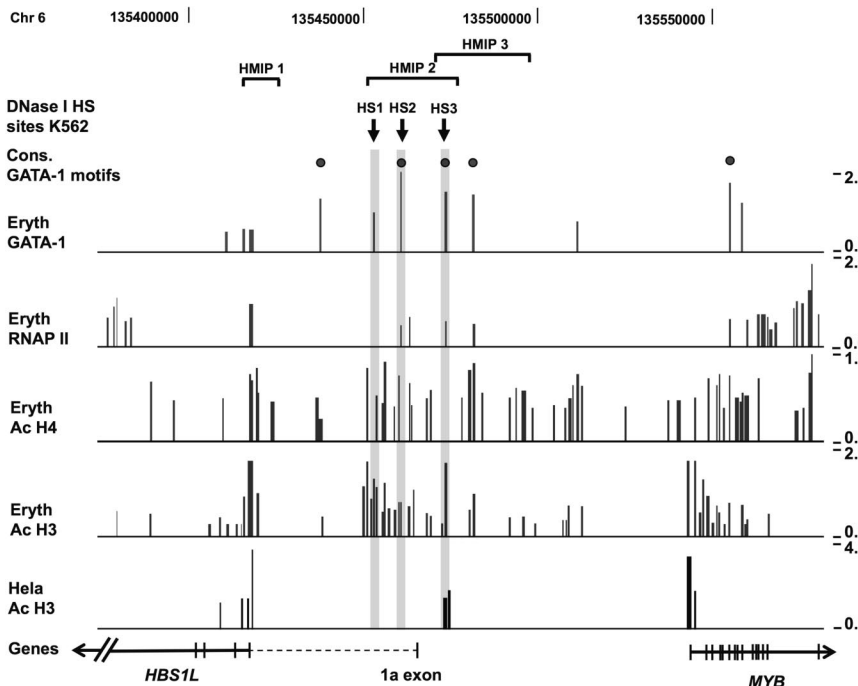


**Figure 2.** ChIP-chip data for the 6q23 locus in erythroid precursor and HeLa cells. (A) The box on chromosome 6 represent a 2.5-Mb region (position chr6: 134 503 000-136 928 000), which is covered in this figure and includes the 6q23 locus. (B) Results from ChIP-chip experiments for GATA-1, RNAP II, AcH4, and AcH3 in erythroid precursors and AcH3 in HeLa cells. Genes in the 6q23 locus are shown below as black arrows and other genes in gray. The *HBS1L*-*MYB* intergenic region is indicated by the boxed area marked with dashed lines to highlight the strong histone acetylation and GATA-1 binding signals in erythroid precursor cells. Data are displayed in the UCSC Genome Browser<sup>27</sup> using autoscale settings.

observed in the *HMIP*-2 and -3 regions and in coincidence with GATA-1 signal.

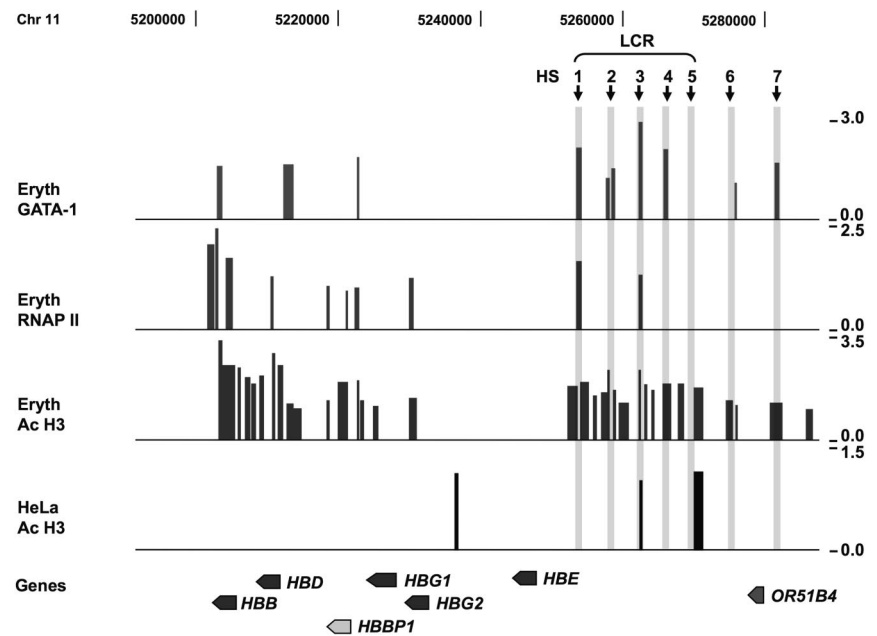
In contrast to the strong signals observed in erythroid precursors, histone H3 acetylation was absent in the intergenic region in

HeLa cells with the exception of a restricted but significant peak in the *HMIP*-2 region, coincident with HMHS3 in K562 cells and a strong GATA-1 signal in erythroid precursors. This peak is in the vicinity of the putative promoter region of *HBS1L* exon 1a.



**Figure 3.** ChIP-chip data for the *HBS1L*-*MYB* intergenic region. Results from ChIP-chip experiments for GATA-1, RNAP II, AcH4, and AcH3 in erythroid precursors and AcH3 in HeLa cells. The figure covers a 210-kb region (position chr6: 135 373 000-135 582 000) and includes a part of the *HBS1L* gene, the intergenic region, and the *MYB* gene. The *HMIP* block 1, 2, and 3 are shown above indicated by the horizontal brackets. The figure also includes conserved GATA-1 motifs (●) and DNase I hypersensitive sites (↓) as identified in hemin-induced K562 cells.

**Figure 4. ChIP-chip data for the  $\beta$  globin locus in erythroid precursor and HeLa cells.** Results from ChIP-chip experiments for GATA-1, RNAP II, and AcH3 in erythroid precursors and AcH3 in HeLa cells analyzed on the custom designed array. The figure covers a 120-kb region (position chr11: 5 180 000-5 302 000) that includes the entire  $\beta$  globin locus. The globin genes are indicated below and include  $\beta$  globin (*HBB*),  $\delta$  globin (*HBD*),  $\gamma$  globin (*HBG1* and *HBG2*) and  $\epsilon$  globin (*HBE*), and the  $\beta\psi$  pseudogene (*HBBP1*). The region also includes the olfactory gene (*OR51B4*). Hypersensitive sites in the  $\beta$ -globin LCR (HS1-HS5) and the upstream HS6 and HS7 are marked with  $\downarrow$ .



By including the well-characterized  $\alpha$  and  $\beta$  globin loci on our custom-designed array, we introduced positive controls for histone acetylation and GATA-1 binding in erythroid cells to evaluate the quality of our ChIP material and data analysis. In addition, the custom array allowed us to compare patterns of histone acetylation and transcription factor binding between the distal regulatory regions of the globin loci and the *HBSIL-MYB* intergenic interval that would facilitate the evaluation of the region upstream of *HBSIL* and *MYB* as a potential distal regulatory element. Interestingly, strong similarities in patterns of histone acetylation as well as GATA-1 and RNAP II interactions were observed between the *HMIP* block 2 and 3 regions and the  $\alpha$  and  $\beta$  globin control regions.

In the  $\beta$  globin locus, 2 isolated domains of activity (AcH3 signals) were clearly discernable and these were concentrated around the  $\beta$  LCR and a region covering the  $\beta$  and  $\delta$  globin genes (*HBB* and *HBD*) and the area around the  $\beta\psi$  pseudogene (*HBBP1*; Figure 4). These 2 domains of activity were separated by an inactive domain comprising the  $\epsilon$  globin (*HBE*) gene and the region up to the  $\gamma$  globin genes (*HBG1* and 2). This pattern was similar to previously published data of H3 acetylation in human erythroid precursors.<sup>29</sup> HeLa cells showed little signal for H3 acetylation in the  $\beta$  globin locus. Some signal was detected in the LCR at HS3 and HS5, and upstream of the  $\gamma$ -globin genes. Acetylation at HS5 in HeLa cells is consistent with this site being ubiquitously active.

Strong GATA-1 signal was detected in the  $\beta$  globin LCR at HS1, HS2, HS3, and HS4, and at the upstream hypersensitive sites 6 and 7. In addition, GATA-1 signals were observed upstream of the *HBD* and *HBBP1* as well as in *HBB*. A closer view revealed that the GATA-1 signal in *HBB* coincided with exon 3 (Figure 4), which has previously been shown to contain an enhancer element involved in the control of  $\beta$  globin expression.<sup>30</sup> RNAP II binding in the  $\beta$  globin locus coincided with GATA-1 signals at HS1 and HS3 and was also associated with *HBG1* and the adult globin genes. The strongest RNAP II signal was observed around *HBB*, consistent with high expression of  $\beta$  globin in erythroid precursors at the time of harvest.

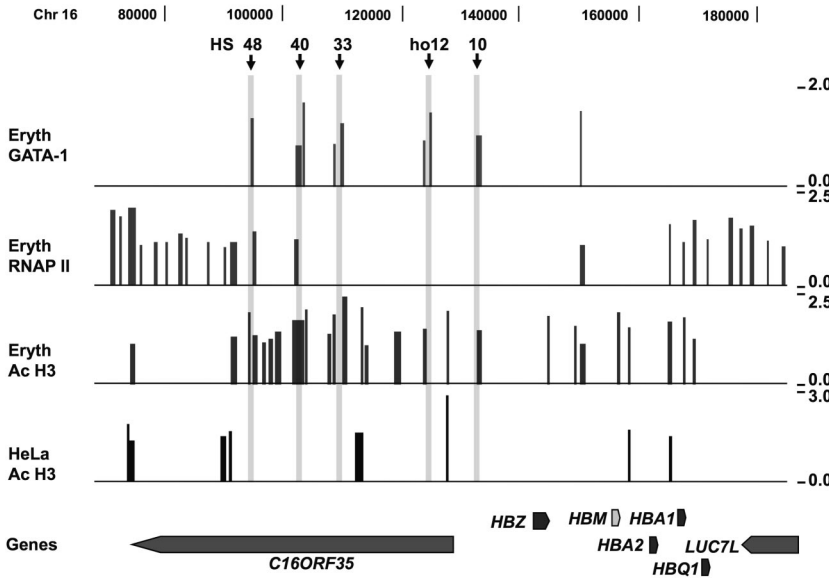
The  $\alpha$  globin locus showed strong H3 acetylation around the  $\alpha$  globin genes and the upstream regulatory domain in erythroid precursors but not in HeLa cells (Figure 5). Consistent with

previous observations,<sup>26</sup> we observed GATA-1 signals at HS48, HS40, HS33, and HS10 upstream of the  $\alpha$  globin genes. GATA-1 signal was also found at the human orthologous region corresponding to the mouse HS12 (hoHS12) and at a site in-between the  $\zeta$  globin (*HBZ*) and the  $\psi\alpha 2$  pseudogene (*HBM*). RNAP II signals were observed at the  $\alpha$  globin promoters and the upstream elements HS-48 and HS-40, again consistent with previous observations.

Given the unusually high and concentrated levels of H3 acetylation and evidence of RNAPII binding in the intergenic region, we decided to investigate the intergenic region for evidence of transcription using a high-resolution tiling array (supplemental methods). Primary human erythroid precursor cells (phase II, days 3 and 5 liquid culture) from 2 individuals were analyzed. All samples gave essentially identical results. As expected, there is strong transcriptional activity at the exons of *MYB* and *HBSIL*, with relatively little in the introns. However, very strong and well-defined transcriptional activity was identified in the intergenic locus spanning *HMIP*-2 and -3. In several areas, the signal intensity is even greater than from the *MYB* exons (Figure 6).

## Discussion

Here, we provide a first characterization of the intergenic sequences upstream of the *HBSIL* and *MYB* genes, strongly supporting the hypothesis of a regulatory region being located in this interval. The conclusion is supported by parallel analysis of histone acetylation, GATA-1, and RNAP II interaction patterns across the erythroid-specific  $\alpha$  and  $\beta$  globin loci. Chromatin immunoprecipitation showed significant histone acetylation in the intergenic region in a restricted interval that encompasses *HMIP*-2 and -3 linkage disequilibrium blocks as defined from genetic analysis. The H3 acetylation was particularly well defined and concentrated in *HMIP*-2. Several GATA-1 binding sites were also identified in the *HBSIL-MYB* intergenic interval; within *HMIP*-2, all the GATA-1 signals coincided with the 3 DNase I hypersensitive sites identified in induced K562 cells, providing strong support for the sites being active regulatory elements in erythroid cells. Interestingly some of the GATA-1 sites also coincided with weak RNAP II binding. Weak

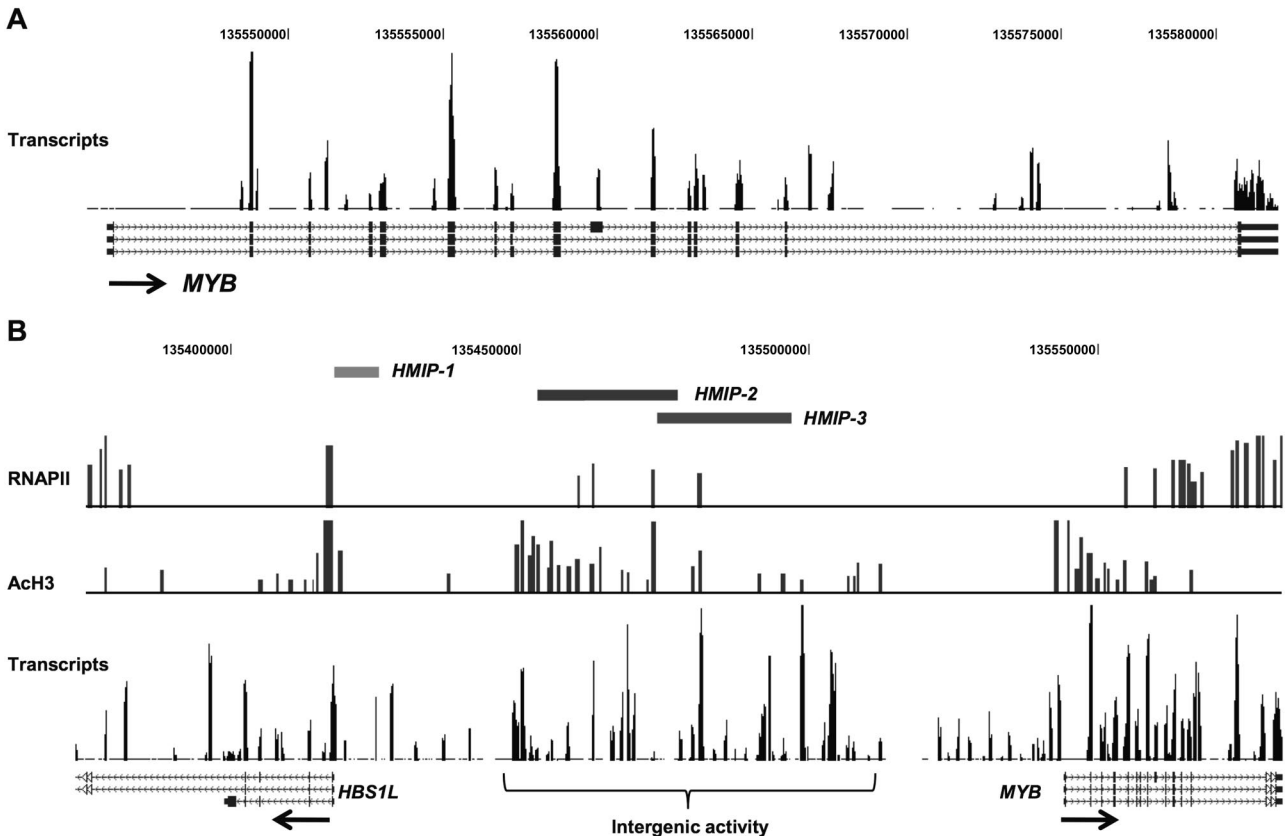


**Figure 5. CHIP-chip data for the  $\alpha$  globin locus in erythroid precursor and HeLa cells.** Results from ChIP-chip experiments for GATA-1, RNAP II, and AcH3 in erythroid precursors and AcH3 in HeLa cells analyzed on the custom-designed array. The figure covers a 120-kb region (position chr16: 70 000-187 000) that includes the  $\alpha$ -globin locus. The globin genes are indicated below and include  $\zeta$  globin (*HBZ*),  $\alpha$  globins (*HBA1* and *HBA2*), and  $\theta$  globin (*HBQ1*), and the  $\psi\alpha 2$  pseudogene (*HBM*). Other genes include *C16ORF35* and *LUC7L*. Hypersensitive sites in the human  $\alpha$  globin locus (HS48, HS40, HS33, and HS10) and the human orthologous mouse HS12 (ho12) are marked with  $\downarrow$ .

RNAP-II signals have previously been observed from ChIP-chip experiments at the site of enhancers and could be a marker of physical interactions with active promoters.<sup>31</sup> Such enhancer elements could affect distal transcriptional control via long-range physical interaction as supported by the observations of GATA-1 involvement in looping formation within the  $\beta$  globin locus.

GATA-1, together with FOG-1, functions as an anchor in the formation of chromatin looping, and is required for physical interactions between the  $\beta$  LCR and  $\beta$  globin promoter.<sup>32</sup>

We show that the pattern of H3 acetylation in the *HBS1L-MYB* region differs between erythroid precursors and HeLa cells. HeLa cells that do not express hematopoietic transcription factors,



**Figure 6. Transcript tiling array data from erythroid precursor cells.** Results from 1 representative experiment are presented. (A) Transcript signal from the *MYB* genomic region shows very clear activity in exons and demonstrates that the array design and methodology are robust, given that *MYB* is highly expressed in the cells used. Occasional strong intronic signals were not coincident with spliced or unspliced ESTs and may therefore represent uncharacterized control elements or ncRNA. (B) Transcript data including *HBS1L*, *MYB*, and the intergenic region. The HbF-associated *HMIP* blocks 1 and 2 cover a region of high transcriptional activity. Again, there is no evidence of either spliced or unspliced ESTs. Given their magnitude in relation to the *MYB* signals and their complete consistency between different cultures tested, the intergenic signals are likely to represent important regulatory sequences.

including *MYB*, showed substantially less H3 acetylation in the intergenic region, suggesting a link between activity in the region and *MYB* expression. We identified a restricted peak of H3 acetylation in the intergenic region in HeLa cells that coincided with GATA-1 signal in erythroid precursors and HMHS3 in induced K562 cells. It is possible that this peak represents a basal level of chromatin acetylation and could conceivably be the crucial activation core for the locus in erythroid cells. Binding of specific transcription factors including GATA-1 to this site could lead to the activation of the regulatory region with consequent induction of expression of *MYB* and other genes in the domain.

Our study has also provided, for the first time, a large-scale analysis of GATA-1 occupancy in human erythroid cells. The abundant GATA-1 signals over the entire 70-Mb interval of chromosome 6q suggest that GATA-1 has an important regulatory role in erythroid cells. The powerful influence of GATA-1 on erythroid commitment and development was recently suggested in overexpression studies of GATA-1 that resulted in the transformation of HeLa cells to a more erythroid phenotype, including formation of the  $\beta$  globin LCR and expression of globin mRNA.<sup>33</sup> Our findings are supportive of GATA-1 having a role as a general transcription factor in erythroid cells in regulating the transcription of ubiquitous as well as erythroid-specific genes.

The identification of high levels of intergenic transcription provides further evidence that the *HBSIL-MYB* region contains a distal regulatory locus. Except for *HBSIL*-exon 1a, there is no evidence of ESTs, spliced or unspliced, or GeneScan gene predictions in the region (NCBI database; <http://www.ncbi.nlm.nih.gov><sup>34</sup>), suggesting there are no undiscovered genes in the interval. Several studies on locus control regions have suggested that intergenic transcription may be involved in chromatin decondensation and looping, which is fundamental to gene activation. Alternatively, it may represent a “tracking” mechanism that enables a transcription complex to move along the locus until a transcriptionally competent promoter is encountered.<sup>29,35,36</sup> The patterns of histone acetylation, RNAP II binding, and GATA-1 interactions, in coincidence with the multiple DNase I hypersensitive sites and the intergenic transcripts in this defined interval, are highly similar to previously well-characterized erythroid control regions.

The regulatory potential of the region upstream of the *MYB* region and its influence on *MYB* expression has previously been emphasized. In murine models, *Myb* has been shown to be a key target for transcriptional activation by long-range upstream and downstream retroviral insertion.<sup>37</sup> Integration of proviruses in a region 25- to 90-kb upstream of *Myb* in mice is associated with tumorigenesis, suggesting a functional importance of these sequences.<sup>38</sup> Further, it has been observed that increased expression of the flanking genes occurred only in the presence of *Myb* overexpression. The observations suggest the possibility that regulation of *Myb* may affect a wider chromatin domain surrounding the gene. Alternatively, there may be common transcription factors or a common *cis*-regulatory element(s) that controls the expression of *Myb* and another gene(s) in its vicinity.<sup>37</sup> Further evidence supporting the regulatory potential of this region comes from a serendipitous insertion of a transgene in this intergenic region, 77-kb upstream of the mouse *Myb* gene, that resulted in reduced *Myb* expression and markedly decreased megakaryocyte/erythrocyte lineage-restricted progenitors of the homozygous mutant mice.<sup>39</sup>

Previously, we showed that *MYB* is a quantitative trait gene, with variable expression in healthy adults.<sup>11</sup> Our previous studies also showed that human erythroid precursor cells from individuals with higher HbF and higher F cell levels have lower *MYB* expression that was also associated with lower erythrocyte count but higher erythrocyte

volume, and higher platelet count.<sup>11</sup> Further, mouse models in which *Myb* activity was reduced, due to either mutation or integration of a transgene near the *Myb* locus, displayed anemia and thrombocytosis.<sup>39-44</sup> It is clear that *MYB*, a transcription factor that is also involved in oncogenesis, has multiple essential roles throughout the different stages of erythropoiesis.

What is not clear, however, are the regulatory sequences controlling *MYB* expression. Recent studies show that *MYB* is a major target of the microRNA 150 (miR-150), and that one pathway of *MYB* regulation is through the 2 conserved miR-150 binding sites in the 3' UTR of *MYB* mRNA.<sup>45</sup> miR-150 repression of *MYB* in CD34<sup>+</sup> human bone marrow cells not only supported *MYB*'s key role in erythroid and megakaryocytic differentiation, but also suggested that modulations of its level are critical to its role.<sup>45</sup> We propose that the *HBSIL-MYB* region upstream of *MYB* contains distal regulatory elements that form a key part of the overall control of *MYB* expression. The intergenic variants may account for some of the *cis*-control of the intrinsic quantitative variation in *MYB* expression. Genetic variants in the *HBSIL-MYB* interval on chromosome 6q have been shown to be highly associated, not only with HbF levels, but also with the control of other hematologic parameters. Taken together, our data have provided a functional basis for this association and strongly support the hypothesis of a regulatory locus upstream of the *HBSIL* and *MYB* genes, located within *HMIP*-2 and -3 blocks as identified in genetic association studies.

Delineation of the key variants in this *HBSIL-MYB* control region may lead to an improved understanding of *MYB* control and dysregulation<sup>46</sup> that underlies many of the leukemias and cancers, and may also provide targets for therapeutic activation of HbF<sup>47</sup> in the treatment of sickle cell disease and  $\beta$  thalassemia.

## Acknowledgments

We thank Claire Steward for help in preparation of the manuscript; Drs Marco de Gobbi, Jim Hughes and David Garrick for their help with ChIP-chip experiments, and Dr Mike Antoniou, Dr Stephan Menzel, Professors Doug Higgs and Bill Wood for helpful discussions.

This work was supported by a grant from the Medical Research Council, United Kingdom (MRC G0000111 and ID51640) to S.L.T. and an MRC training studentship to K.W. The research at the Center for Genomic Medicine, Kyoto University, is partly supported by the Ministry of Education, Culture, Sports, Science and Technology of Japan. We also thank the London University Central Research Fund (CRF) and British Society for Hematology for support (S.B.).

## Authorship

Contribution: K.W. performed research, analyzed data, and wrote the paper; J.J., H.R., K.J., F.M., and M.Y. performed research; M.L. contributed to data analysis and writing of the paper; S.L.T. codirected research and wrote the paper; and S.B. codirected research, analyzed data, and wrote the paper.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

Correspondence: Swee Lay Thein, King's College London School of Medicine, James Black Centre, 125 Coldharbour Lane, London SE5 9NU, United Kingdom; e-mail: sl.thein@kcl.ac.uk.



## References

- Platt OS, Brambilla DJ, Rosse WF, et al. Mortality in sickle cell disease: life expectancy and risk factors for early death. *N Engl J Med*. 1994;330(23):1639-1644.
- Ho PJ, Hall GW, Luo LY, Weatherall DJ, Thein SL. Beta thalassaemia intermedia: is it possible to consistently predict phenotype from genotype? *Br J Haematol*. 1998;100(1):70-78.
- Weatherall DJ, Clegg JB. *The Thalassemia Syndromes*. 4th Ed. Oxford, United Kingdom: Blackwell Science; 2001.
- Garner C, Tatu T, Reittie JE, et al. Genetic influences on F cells and other hematologic variables: a twin heritability study. *Blood*. 2000;95(1):342-346.
- Menzel S, Garner C, Gut I, et al. A QTL influencing F cell production maps to a gene encoding a zinc-finger protein on chromosome 2p15. *Nat Genet*. 2007;39(10):1197-1199.
- Lettre G, Sankaran VG, Bezerra MA, et al. DNA polymorphisms at the BCL11A, HBS1L-MYB, and beta-globin loci associate with fetal hemoglobin levels and pain crises in sickle cell disease. *Proc Natl Acad Sci U S A*. 2008;105(33):11869-11874.
- Sedgewick AE, Timofeev N, Sebastiani P, et al. BCL11A is a major HbF quantitative trait locus in three different populations with beta-hemoglobinopathies. *Blood Cells Mol Dis*. 2008;41(3):255-258.
- Thein SL, Menzel S, Peng X, et al. Intergenic variants of HBS1L-MYB are responsible for a major quantitative trait locus on chromosome 6q23 influencing fetal hemoglobin levels in adults. *Proc Natl Acad Sci U S A*. 2007;104(23):11346-11351.
- Uda M, Galanello R, Sanna S, et al. Genome-wide association study shows BCL11A associated with persistent fetal hemoglobin and amelioration of the phenotype of beta-thalassaemia. *Proc Natl Acad Sci U S A*. 2008;105(3):1620-1625.
- So CC, Song YQ, Tsang ST, et al. The HBS1L-MYB intergenic region on chromosome 6q23 is a quantitative trait locus controlling fetal haemoglobin level in carriers of beta-thalassaemia. *J Med Genet*. 2008;45(11):745-751.
- Jiang J, Best S, Menzel S, et al. cMYB is involved in the regulation of fetal hemoglobin production in adults. *Blood*. 2006;108(3):1077-1083.
- Menzel S, Jiang J, Silver N, et al. The HBS1L-MYB intergenic region on chromosome 6q23.3 influences erythrocyte, platelet, and monocyte counts in humans. *Blood*. 2007;110(10):3624-3626.
- Wallrapp C, Verrier S-B, Zhouravleva G, et al. The product of the mammalian orthologue of the *Saccharomyces cerevisiae* HBS1 gene is phylogenetically related to eukaryotic release factor 3 (eRF3) but does not carry eRF3-like activity. *FEBS Lett*. 1998;440(3):387-392.
- Oh IH, Reddy EP. The myb gene family in cell growth, differentiation and apoptosis. *Oncogene*. 1999;18(19):3017-3033.
- Cantor AB, Orkin SH. Transcriptional regulation of erythropoiesis: an affair involving multiple partners. *Oncogene*. 2002;21(21):3368-3376.
- Vegiopoulos A, Garcia P, Emambokus N, Frampton J. Coordination of erythropoiesis by the transcription factor c-Myb. *Blood*. 2006;107(12):4703-4710.
- Fibach E, Manor D, Oppenheim A, Rachmilewitz EA. Proliferation and maturation of human erythroid progenitors in liquid culture. *Blood*. 1989;73(1):100-103.
- McArthur M, Gerum S, Stamatoyannopoulos G. Quantification of DNaseI-sensitivity by real-time PCR: quantitative analysis of DNaseI-hypersensitivity of the mouse beta-globin LCR. *J Mol Biol*. 2001;313(1):27-34.
- Dorschner MO, Hawrylycz M, Humbert R, et al. High-throughput localization of functional elements by quantitative chromatin profiling. *Nat Methods*. 2004;1(3):219-225.
- National Center for Biotechnology Information. Gene Expression Omnibus (GEO). <http://www.ncbi.nlm.nih.gov/geo>. Accessed June 2, 2009.
- O'Geen H, Nicolet CM, Blahnik K, Green R, Farnham PJ. Comparison of sample preparation methods for ChIP-chip assays. *Biotechniques*. 2006;41(5):577-580.
- Kapranov P, Drenkow J, Cheng J, et al. Examples of the complex architecture of the human transcriptome revealed by RACE and high-density tiling arrays. *Genome Res*. 2005;15(17):987-997.
- National Center for Biotechnology Information. Build 36. [http://www.ncbi.nlm.nih.gov/mapview/map\\_search.cgi?taxid=10090](http://www.ncbi.nlm.nih.gov/mapview/map_search.cgi?taxid=10090). Accessed March 14, 2007.
- Eisbruch A, Blick M, Evinger-Hodges MJ, et al. Effect of differentiation-inducing agents on oncogene expression in a chronic myelogenous leukemia cell line. *Cancer*. 1988;62(6):1171-1178.
- Villeval JL, Pelicci PG, Tabilio A, et al. Erythroid properties of K562 cells: effect of hemin, butyrate and TPA induction. *Exp Cell Res*. 1983;146(2):428-435.
- De Gobbi M, Anguita E, Hughes J, et al. Tissue-specific histone modification and transcription factor binding in {alpha} globin gene expression. *Blood*. 2007;110(13):4503-4510.
- University of California Santa Cruz. UCSC Genome Browser. <http://genome.ucsc.edu>. Accessed May 15, 2008.
- Lian Z, Karpikov A, Lian J, et al. A genomic analysis of RNA polymerase II modification and chromatin architecture related to 3' end RNA polyadenylation. *Genome Res*. 2008;18(8):1224-1237.
- Miles J, Mitchell JA, Chakalova L, et al. Intergenic transcription, cell-cycle and the developmentally regulated epigenetic profile of the human beta-globin locus. *PLoS ONE*. (<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?>) 2007;2(7):e630.
- Behringer RR, Hammer RE, Brinster RL, Palmiter RD, Townes TM. Two 3' sequences direct adult erythroid-specific expression of human beta-globin genes in transgenic mice. *Proc Natl Acad Sci U S A*. 1987;84(20):7056-7060.
- Heintzman ND, Stuart RK, Hon G, et al. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat Genet*. 2007;39(3):311-318.
- Vakoc CR, Letting DL, Gheldof N, et al. Proximity among distant regulatory elements at the beta-globin locus requires GATA-1 and FOG-1. *Mol Cell*. 2005;17(3):453-462.
- Layon ME, Ackley CJ, West RJ, Lowrey CH. Expression of GATA-1 in a non-hematopoietic cell line induces beta-globin locus control region chromatin structure remodeling and an erythroid pattern of gene expression. *J Mol Biol*. 2007;366(3):737-744.
- National Center for Biotechnology Information. <http://www.ncbi.nlm.nih.gov>. Accessed February 20, 2009.
- Ashe HL, Monks J, Wijgerde M, Fraser P, Proudfoot NJ. Intergenic transcription and transinduction of the human beta-globin locus. *Genes Dev*. 1997;11(19):2494-2509.
- Gribnau J, Diderich K, Pruzina S, Calzolari R, Fraser P. Intergenic transcription and developmental remodeling of chromatin subdomains in the human beta-globin locus. *Mol Cell*. 2000;5(2):377-386.
- Hanlon L, Barr NI, Blyth K, et al. Long-range effects of retroviral insertion on c-myb: overexpression may be obscured by silencing during tumor growth in vitro. *J Virol*. 2003;77(2):1059-1068.
- Haviernik P, Festin SM, Opavsky R, et al. Linkage on chromosome 10 of several murine retroviral integration loci associated with leukaemia. *J Gen Virol*. 2002;83(4):819-827.
- Mukai HY, Motohashi H, Ohneda O, Suzuki N, Nagano M, Yamamoto M. Transgene insertion in proximity to the c-myb gene disrupts erythroid-megakaryocytic lineage bifurcation. *Mol Cell Biol*. 2006;26(21):7953-7965.
- Mucenski ML, McLain K, Kier AB, et al. A functional c-myb gene is required for normal murine fetal hepatic hematopoiesis. *Cell*. 1991;65(4):677-689.
- Emambokus N, Vegiopoulos A, Harman B, Jenkinson E, Anderson G, Frampton J. Progression through key stages of haemopoiesis is dependent on distinct threshold levels of c-Myb. *EMBO J*. 2003;22(17):4478-4488.
- Carpinelli MR, Hilton DJ, Metcalf D, et al. Suppressor screen in Mpl<sup>-/-</sup> mice: c-Myb mutation causes supraphysiological production of platelets in the absence of thrombopoietin signaling. *Proc Natl Acad Sci U S A*. 2004;101(17):6553-6558.
- Sandberg ML, Sutton SE, Pletcher MT, et al. c-Myb and p300 regulate hematopoietic stem cell proliferation and differentiation. *Dev Cell*. 2005;8(2):153-166.
- Kasper LH, Boussouar F, Ney PA, et al. A transcription-factor-binding surface of coactivator p300 is required for haematopoiesis. *Nature*. 2002;419(6908):738-743.
- Lu J, Guo S, Ebert BL, et al. MicroRNA-mediated control of cell fate in megakaryocyte-erythrocyte progenitors. *Dev Cell*. 2008;14(6):843-853.
- Garcia P, Frampton J. Hematopoietic lineage commitment: miRNAs add specificity to a widely expressed transcription factor. *Dev Cell*. 2008;14(6):815-816.
- Perrine SP. Hemoglobin F: new targets, new path. *Blood*. 2006;108(3):783-784.