bjh short report

ASH1L (a histone methyltransferase protein) is a novel candidate globin gene regulator revealed by genetic study of an English family with beta-thalassaemia unlinked to the beta-globin locus

Amandine Breton,[1] Andria Theodorou,[1] Suleyman Aktuna,[1] Laura Sonzogni,[1] David Darling,[2] Lucas Chan,[2] Stephan Menzel,[1] Peter J. van der Spek,[3] Sigrid M.A. Swagemakers,[3] Frank Grosveld,[4,5,6] Sjaak Philipsen[4,5] and Swee Lay Thein[1,2,*]

[1]Molecular Haematology, Division of Cancer Studies, King's College London Faculty of Life Sciences and Medicine, London SE5 9NU, UK, [2]Department of Haematological Medicine, King's College Hospital NHS Foundation Trust, London SE5 9RS, UK, [3]Department of Bioinformatics, [4]Department of Cell Biology, [5]Netherlands Consortium for Systems Biology and [6]Centre for Biomedical Genetics, Erasmus MC, Rotterdam, The Netherlands

Received 5 February 2016; accepted for publication 6 June 2016

Correspondence: Swee Lay Thein, Sickle Cell Branch, National Heart, Lung and Blood Institute, The National Institutes of Health, Building 10-CRC, Room 5E-5142, 10 Center Drive, Bethesda, MD 20892, USA
E-mail: sl.thein@nih.gov

*Present Address: Sickle Cell Branch, National Heart, Lung and Blood Institute/NIH, Bethesda, MD, USA

In 1993, we described an English family with beta-thalassaemia that was not linked to the beta-globin locus. Whole genome sequence analyses revealed potential causative mutations in 15 different genes, of which 4 were consistently and uniquely associated with the phenotype in all 7 affected family members, also confirmed by genetic linkage analysis. Of the 4 genes, which are present in a centromeric region of chromosome 1, *ASH1L* was proposed as causative through functional mRNA knock-down and chromatin-immunoprecipitation studies in human erythroid progenitor cells. Our data suggest a putative role for ASH1L (Trithorax protein) in the regulation of globin genes.

Keywords: *ASH1L*, Trithorax protein, beta-thalassaemia.

Beta-thalassaemia is among the most common monogenic disorders worldwide. Over 200 mutations down-regulating the β-globin gene (*HBB*) causing β-globin deficiency have been characterized; the majority of these mutations are *cis*-acting and reside within *HBB* itself, or within its immediate flanking sequences (Giardine *et al*, 2011; Thein, 2013). Several rare *trans*-acting mutations have also been described, including mutations affecting *GATA1* on the X chromosome, *KLF1* on chromosome 19 (Perseu *et al*, 2011) and those affecting XPD protein, or implicated from family studies (Thein, 2013).

In 1993, we described one such family of English descent with β -thalassaemia trait phenotype that was not linked to the *HBB* locus (Thein *et al*, 1993). Further genetic analyses revealed 4 candidate genes (*LRIG2*, *ANKRD35*, *NUP210L* and *ASH1L*), which were inherited as a block spanning the centromere on chromosome 1. Of the 4 genes, 2 were expressed in erythroid progenitors, but only *ASH1L* correlated with differentiation of human erythroid progenitor cells (hEPCs) and *HBB* expression. Our data suggest a role for ASH1L, a methyl transferase protein and member of the trithorax

 doi: 10.1111/bjh.14256

(Trx) family, in regulation of the *HBB* gene and as a potential modifier of beta-thalassaemia severity.

Material and methods

Patient samples were sent for work-up of unexplained anaemia.

Genetic analyses

We performed whole genome sequence (WGS) analysis in 2 affected and 2 unaffected members of the family (Data S1, and Figure S1). Filtering approaches included: exclusion of variants that are not shared by affected members, and absent in unaffected members; exclusion of non-coding variants and non-rare coding variants (frequency > 0·5%) (Ott *et al*, 2015). Following the dominant mode of inheritance suggested by the familial segregation, we found mutations in 15 genes [four on chromosome (chr)1, four on chr3, one on chr6, one on chr8, one on chr10, one on chr19, and three on chr20] (Table SI). Using Sanger sequencing we screened 27 family members spanning three generations, including the four individuals that were subjected to WGS, for the presence of these 15 genetic variants. This revealed firstly, that all 15 variants were present in the two affected, and absent in the two unaffected WGS individuals, confirming that these mutations were not artefacts of the WGS assemblies; and secondly, that four genetic variants were consistently and uniquely present in all seven affected family members (Table SII).

Linkage analysis

Two-point linkage analysis was performed with MLINK (Lathrop *et al*, 1984). Marker genotypes were obtained from Sanger sequencing data (seven affected and 20 unaffected individuals), thalassaemia phenotypes were as previously described (Thein *et al*, 1993). The model assumed autosomal-dominant inheritance of the β thalassaemia trait with complete penetrance and phenocopy rate 0·000001. The *ASH1L* variant (T>C, dbSNP *rs151028549*) identified in the English family is extremely rare with a population frequency of 0·001, present only in British and Spanish populations (http://www.ncbi.nlm.nih.gov/variation/tools/1000genomes/?chr=NC_000001.10&from=155447318&to=155448318&mk=155447818:155447818|rs151028549&gts=rs151028549). The minor allele, consisting of a block spanning the centromere on chromosome 1 containing *LRIG2*, *ANKRD35*, *NUP210L* and *ASH1L* genes, was assumed to be extremely rare, and the population frequency was set to 0·00001.

Logarithm of odds (LOD) scores were calculated for recombination rates (Θ) of 0·5 (no linkage), 0 (marker is the causative variant), 0·01, 0·02, 0·03, 0·04 and 0·05 between marker and causative variant. The block on chromosome 1 displays a LOD score of 3·01 for a Θ of 0, giving evidence for linkage between the rare sequence variants and the causative variant acting in this family (Table SIII).

Gene expression and chromatin immunoprecipitation

hEPCs were cultured from buffy coats from healthy adults using a two-phase culture system (van den Akker *et al*, 2010). K562 cells were maintained in Dulbecco's Modified Eagle medium and RPMI medium, respectively, supplemented with 10% fetal calf serum and penicillin/streptomycin.

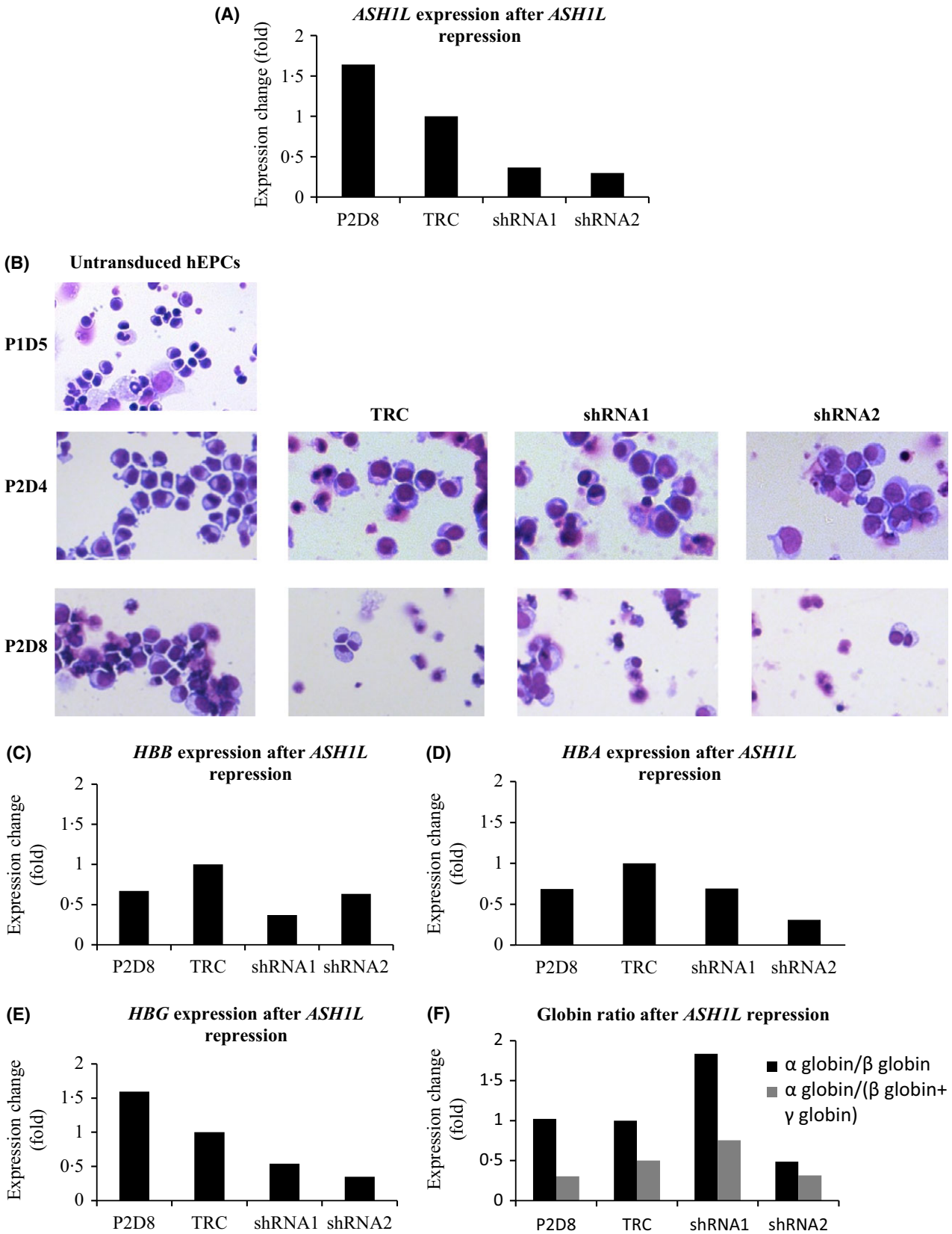
shRNA constructs against *ASH1L* were obtained from The RNAi Consortium (TRC) shRNA library (Sigma, St Louis, MO, USA). Four different shRNA lentiviruses were tested in K562 cells, of which 2 (referred to as shRNA1 and shRNA2) were retained for further experiments. A non-targeting TRC shRNA was used as control. hEPCs were transduced on Phase 2 Day 2 (P2D2) culture, selected in medium supplemented with puromycin, and processed for gene analysis on Phase 2 Day 8 (P2D8).

Binding of ASH1L and histone H3 trimethylated-lysine 4 (H3K4me3) at *HBB* and *HBA* promoter regions in K562 and hEPCs was evaluated following previously described procedures. hEPCs at P1D5, P2D4 and P2D8 were selected for chromatin immunoprecipitation-quantitative polymerase chain reaction (ChIP-qPCR) analysis. Cells were crosslinked using 1% formaldehyde and the chromatin sonicated using a Covaris (Woburn, MA, USA) E2000 machine. Quantitation of immunoprecipitated samples was performed by real-time PCR using SYBR green dye on a Prism 7900H instrument (ABI, Foster City, CA, USA).

Results and discussion

Linkage analysis in conjunction with WGS (Ott *et al*, 2015) has allowed us to narrow potential trans-acting candidates for the beta-thalassaemia phenotype to 4 genes. Candidacy of

Fig 1. ASH1L and globin expression in hEPCs after silencing of ASH1L. (A) - Relative expression of ASH1L expressed as fold-change compared to The RNAi Consortium (TRC) control in Phase 2 Day 8 (P2D8) human erythroid progenitor cells (hEPCs), after transduction with shRNA1 and shRNA2; (B) - Cytospins of untransduced hEPCs, TRC-transduced, shRNA1 and shRNA2 transduced hEPCs at P1D5, P2D4 and P2D8 of differentiation. Briefly the cells were cytospun onto a slide and then stained with eosin and methylene blue before observation and acquisition of images. Relative expression of *HBB* (C); *HBA* (D), *HBG* (E) expressed as fold-change compared to TRC control in Phase 2 Day 8 (P2D8) hEPCs, after transduction with shRNA1 and shRNA2; (F) globin ratios in P2D8 hEPCS and P2D8 hEPCS transduced with TRC, shRNA1 and shRNA2, α/β (black bars) and α/(β+γ) (grey bars). The results represent an average of the repeats from 3 independent cultures. *GAPDH* expression was used for normalization purposes. Results are presented as fold-change over TRC samples.



the 4 genes (*LRIG2*, *ANKRD35*, *NUP210L* and *ASH1L*) was analysed using Encode and Ensembl databases accessed through the UCSC Genome Browser (http://genome.ucsc.edu/) – *ANKRD35* and *NUP210L* did not appear to be involved in erythropoiesis. *LRIG2* is expressed at early stages of erythroid differentiation and in monocytes

at low levels (Wang *et al*, 2012). *ASH1L* encodes a protein from the Trithorax family and acts as a histone methyl transferase. It has been found to be recruited at the transcribed portion of the mouse beta-major globin gene of G1E cells under GATA-1-mediated erythroid differentiation and is absent in undifferentiated cells (Gregory *et al*, 2007).

Both *LRIG2* and *ASH1L* are expressed during differentiation of hEPCs; *LRIG2* being the less expressed of the two genes (Figure S2A). While expression of *LRIG2* remained at low levels throughout erythroid differentiation, *ASH1L* expression clearly increased, reaching a peak at day 6 of erythroid differentiation, just before the rise in globin gene expression (Figure S2A,B). Based on the *ASH1L* expression profile during erythroid differentiation and its known functions, it was selected for further studies.

shRNA1 and shRNA2 induced a decrease in 60% and 80%, respectively, of *ASH1L* mRNA expression in transduced hEPCs (Fig 1A) and the differentiation of the hEPCs does not appear to be affected (Fig 1B). Overall, a decrease in *HBB*, *HBA* and *HBG* expression was observed on suppression of *ASH1L* with shRNA1 and shRNA2 compared to the TRC hEPCs (Fig 1 C–E), but there was minimal change in the *HBA*/non-*HBA* ratio on ASH1L repression (Fig 1F).

We performed ASH1L ChIP on K562 cells and on hEPCs at different stages of erythroid differentiation. K562 cells appear as a good control for evaluating ASH1L enrichment in the *HBB* promoter as they do not express β-globin and express *HBA* in a low quantity. In K562 cells, there was no enrichment of ASH1L to the *HBB* promoter but a slight enrichment was observed for the *HBA* promoter (Fig 2A).

hEPCs at P1D5, P2D4 and P2D8 were selected for ChIP-qPCR analysis as both *HBA* and *HBB* are expressed but at different levels in these cells. In P1D5 hEPCs, there is no expression of *HBB* and no evidence of ASH1L binding to the *HBB*

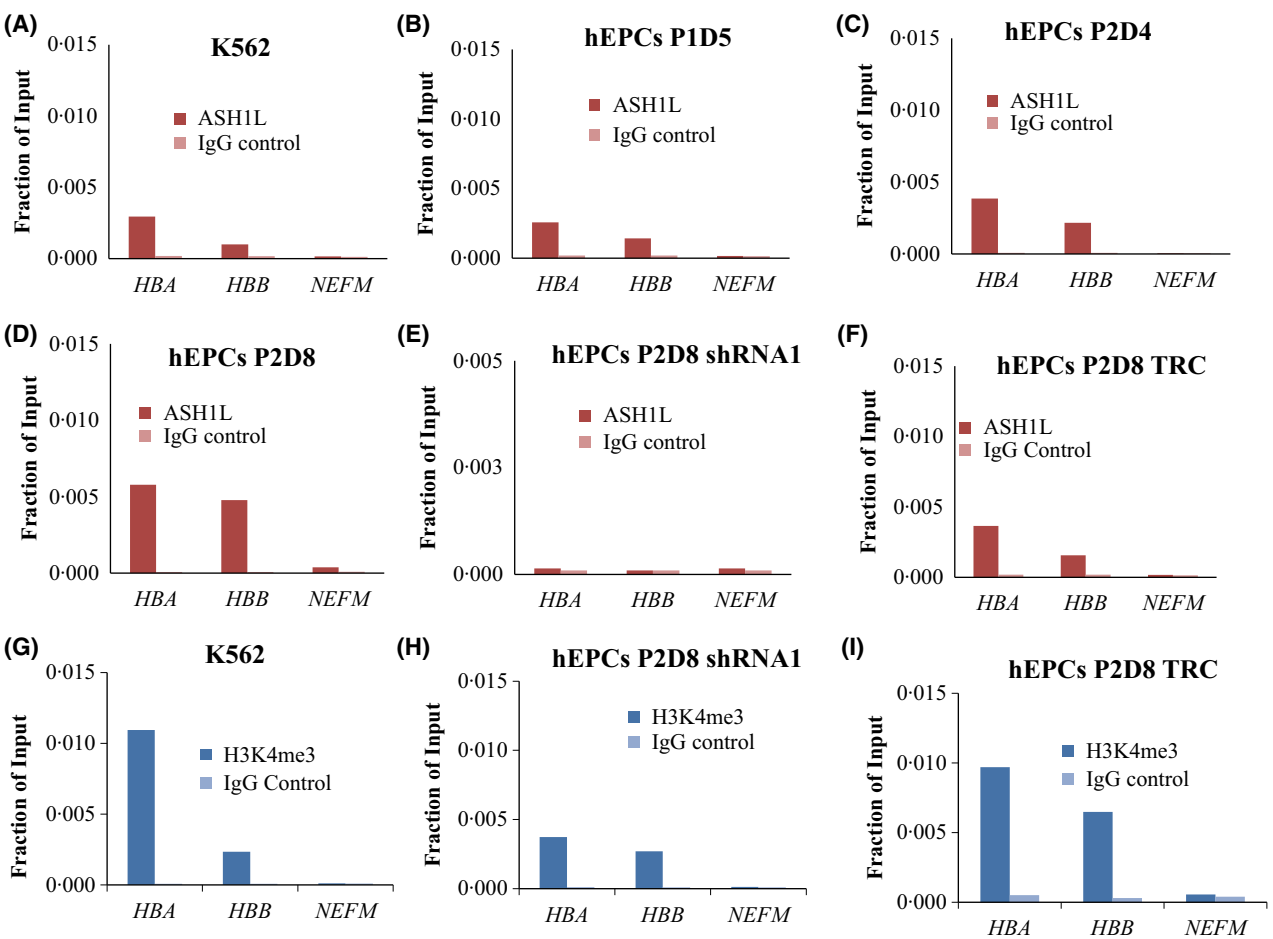


Fig 2. ASH1L protein and H3K4me3 association with *HBA* and *HBB* promoters in K562 and hEPCs at different stages of culture assessed by ChIP-qPCR. ASH1L enrichment (in red) and mouse IgG control (pink) at the *HBA*, *HBB* and *NEFM* promoters in (A) K562 cells, (B) human erythroid progenitor cells (hEPCs) in Phase 1 Day 5 (P1D5), (C) hEPCs P2D4, (D) hEPCs P2D8, (E) hEPCs P2D8 shRNA1, (F) hEPCs P2D8 TRC, and H3K4me3 enrichment (in blue) and rabbit IgG control (light blue) to the *HBA*, *HBB* and *NEFM* promoters in (G) K562 cells, (H) hEPCs P2D8 shRNA1 and (I) hEPCs P2D8 TRC. All reactions were performed in triplicate and results represent the average of two or more independent experiments. Enrichment at a specific primer target (amplicon) in ChIP samples was expressed as a fraction of chromatin input. The enrichment of non-target gene (*NEFM*) is presented. Details of the primer sequences and antibodies used in qRT-PCR and ChIP-qPCR experiments are listed in Tables SIV and SV.

promoter (Fig 2B). With differentiation, there is increased binding of ASH1L to the *HBB* and *HBA* promoters in both P2D4 and P2D8 cells (Fig 2C and D). On repression of *ASH1L* with shRNA1, ASH1L binding to the *HBA* and *HBB* promoters is lost (Fig 2E); it is still present in the TRC transduced hEPCs (Fig 2F). To further investigate the impact of *ASH1L* repression, we assessed H3K4me3 enrichment at the *HBA* and *HBB* promoters in K562 and P2D8 cells transduced with TRC or *ASH1L* shRNA1. In K562, H3K4me3 enrichment is observed at the *HBA* promoter with modest enrichment at the *HBB* promoter (Fig 2G). Upon repression of *ASH1L* with shRNA1, H3K4me3 enrichment at the *HBA* and *HBB* promoters was reduced by a third compared to the TRC P2D8 hEPCs (Fig 2H and I). Our data support the genetic studies that suggest *ASH1L* as a putative candidate underlying the beta-thalassaemia in this English family. ASH1L is associated with the active *HBB* promoter in differentiating human erythroid precursor cells, just as in the mouse. However, the imbalanced globin ratio observed in the family was not replicated in our experiments. Compared to the missense mutation observed in the English family, we propose that 60–80% silencing of *ASH1L* expression has a larger impact on the entire genome as ASH1L protein is associated with a large number of active genes in the cells (Gregory *et al*, 2007).

The *ASH1L* variant within exon 3, replaces Glycine for Arginine, introducing a larger and charged amino acid at location 1615, in a serine rich domain before the SET domain. The human *ASH1L* gene, like its *Drosophila* counterpart, encodes a SET domain–containing trithorax protein with a histone methyl-transferase activity that can associate with actively transcribed loci, including at several *HOX* genes (Gregory *et al*, 2007; Tanaka *et al*, 2008, 2011). The ASH1L SET domain was reported to have intrinsic H3K36 dimethyltransferase activity using *in-vitro* biochemical assays (Tanaka *et al*, 2007; An *et al*, 2011; Yuan *et al*, 2011) and methylates H3K4 *in vitro* (Gregory *et al*, 2007). Here, we established that enrichment of H3K4me3 at *HBB* and *HBA* promoters is reduced upon repression of *ASH1L*. The shRNA experiments may have also differential effects on the *HBA* and *HBB* promoters because they are fundamentally different promoters, α being part of a methylation-free island with many CG sequences (Quante & Bird, 2016). Further studies on the role of ASH1L and its role in beta globin regulation will clarify the contribution of this gene to globin gene regulation.

Acknowledgements

We thank members of the Grosveld laboratory for training and advice on the ChIP-qPCR procedure and members of the Farzaneh laboratory for technical support in lentiviral production for RNA interference experiments. This work was supported by the EU-FP7 THALAMOSS consortium (SP and SLT) and an EMBO travel fellowship to visit the Grosveld laboratory (AB). FG and SP were supported by the Landsteiner Foundation for Blood Transfusion Research (LSBR 1040); SP by the Netherlands Organization for Health Research and Development (NWO/ZonMw TOP 40-00812-98-12128); SLT, by the Medical Research Council, UK (MRC Grant No. G0000111, ID51640). Bioinformatics activities were supported from the Dutch CTMM Biochip grant.

Author contributions

AB, AT, SA, and LS performed experiments; AB, SP, FG and SLT analysed the data; AB, FG, SP, and SLT designed the research study; SS and PS processed and analysed the WGS data; SM performed genetic analysis; AB, SLT, SP and FG wrote the paper. All authors contributed to the final version of the paper.

Conflict of interest

The authors declare that they have no competing interests.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Data S1. Methods.

Fig S1. Pedigree of the English family showing nonlinkage of the β-thalassaemia phenotype with the HBA and HBB gene clusters.

Fig S2. Relative expression of ASH1L, LRIG2 and globin genes during phase2 hEPCs culture.

Table SI. List of the 15 candidate genes for the bor the f the 15 candidate genes EPCs reltand their chromosomal location in GRCh38/hg38.

Table SII. Analyses of the SNPs in the 15 genes in the English family.

Table SIII. LOD scores for linkage between the fragment containing the four mutated genes at block spanning the centromeric region of chr 1 and beta-thalassaemia trait at defined recombination fractions.

Table SIV. List of primers used for qRT-PCR and ChIP-qPCR experiments.

Table SV. List of antibodies used for the ChIP-qPCR experiments.

References

van den Akker, E., Satchwell, T.J., Pellegrin, S., Daniels, G. & Toye, A.M. (2010) The majority of the *in vitro* erythroid expansion potential resides in CD34(-) cells, outweighing the contribution of CD34(+) cells and significantly increasing the erythroblast yield from peripheral blood samples. *Haematologica*, **95**, 1594–1598.

An, S., Yeo, K.J., Jeon, Y.H. & Song, J.J. (2011) Crystal structure of the human histone methyltransferase ASH1L catalytic domain and its implications for the regulatory mechanism.

Journal of Biological Chemistry, **286**, 8369–8374.

Giardine, B., Borg, J., Higgs, D.R., Peterson, K.R., Philipsen, S., Maglott, D., Singleton, B.K., Anstee, D.J., Basak, A.N., Clark, B., Costa, F.C., Faustino, P., Fedosyuk, H., Felice, A.E., Francina, A., Galanello, R., Gallivan, M.V., Georgitsi, M., Gibbons, R.J., Giordano, P.C., Harteveld, C.L., Hoyer, J.D., Jarvis, M., Joly, P., Kanavakis, E., Kollia, P., Menzel, S., Miller, W., Moradkhani, K., Old, J., Papachatzopoulou, A., Papadakis, M.N., Papadopoulos, P., Pavlovic, S., Perseu, L., Radmilovic, M., Riemer, C., Satta, S., Schrijver, I., Stojiljkovic, M., Thein, S.L., Traeger-Synodinos, J., Tully, R., Wada, T., Waye, J.S., Wiemann, C., Zukic, B., Chui, D.H., Wajcman, H., Hardison, R.C. & Patrinos, G.P. (2011) Systematic documentation and analysis of human genetic variation in hemoglobinopathies using the microattribution approach. *Nature Genetics*, **43**, 295–301.

Gregory, G.D., Vakoc, C.R., Rozovskaia, T., Zheng, X., Patel, S., Nakamura, T., Canaani, E. & Blobel, G.A. (2007) Mammalian ASH1L is a histone methyltransferase that occupies the transcribed region of active genes. *Molecular and Cellular Biology*, **27**, 8466–8479.

Lathrop, G.M., Lalouel, J.M., Julier, C. & Ott, J. (1984) Strategies for multilocus linkage analysis in humans. *Proceedings of the National Academy of Sciences of the United States of America*, **81**, 3443–3446.

Ott, J., Wang, J. & Leal, S.M. (2015) Genetic linkage analysis in the age of whole-genome sequencing. *Nature Reviews Genetics*, **16**, 275–284.

Perseu, L., Satta, S., Moi, P., Demartis, F.R., Manunza, L., Sollaino, M.C., Barella, S., Cao, A. & Galanello, R. (2011) KLF1 gene mutations cause borderline HbA2. *Blood*, **118**, 4454–4458.

Quante, T. & Bird, A. (2016) Do short, frequent DNA sequence motifs mould the epigenome? *Nature Reviews Molecular Cell Biology*, **17**, 257–262.

Tanaka, Y., Katagiri, Z., Kawahashi, K., Kioussis, D. & Kitajima, S. (2007) Trithorax-group protein ASH1 methylates histone H3 lysine 36. *Gene*, **397**, 161–168.

Tanaka, Y., Nakayama, Y., Taniguchi, M. & Kioussis, D. (2008) Regulation of early T cell development by the PHD finger of histone lysine methyltransferase ASH1. *Biochemical and Biophysical Research Communications*, **365**, 589–594.

Tanaka, Y., Kawahashi, K., Katagiri, Z., Nakayama, Y., Mahajan, M. & Kioussis, D. (2011) Dual function of histone H3 lysine 36 methyltransferase ASH1 in regulation of Hox gene expression. *PLoS ONE*, **6**, e28171.

Thein, S.L. (2013) The Molecular Basis of beta-Thalassemia. *Cold Spring Harbor Perspectives in Medicine*, **3**, 159–182.

Thein, S.L., Wood, W.G., Wickramasinghe, S.N. & Galvin, M.C. (1993) Beta-thalassemia unlinked to the beta-globin gene in an English family. *Blood*, **82**, 961–967.

Wang, D., Rendon, A., Ouwehand, W. & Wernisch, L. (2012) Transcription factor co-localization patterns affect human cell type-specific gene expression. *BMC Genomics*, **13**, 263.

Yuan, W., Xu, M., Huang, C., Liu, N., Chen, S. & Zhu, B. (2011) H3K36 methylation antagonizes PRC2-mediated H3K27 methylation. *Journal of Biological Chemistry*, **286**, 7983–7989.