HHS Public Access

Author manuscript

Nat Genet. Author manuscript; available in PMC 2011 March 01.

Published in final edited form as:

Nat Genet. 2010 September ; 42(9): 801–805. doi:10.1038/ng.630.

Haploinsufficiency for the erythroid transcription factor KLF1 causes Hereditary Persistence of Fetal Hemoglobin

Joseph Borg[1,2,*], **Petros Papadopoulos**[3,*], **Marianthi Georgitsi**[4,*], **Laura Gutiérrez**[3], **Godfrey Grech**[1,2], **Pavlos Fanis**[3], **Marios Phylactides**[5], **Annemieke J.M.H. Verkerk**[6], **Peter J. van der Spek**[6], **Christian A. Scerri**[1,2], **Wilhelmina Cassar**[1,2], **Ruth Galdies**[1,2], **Wilfred van IJcken**[7], **Zeliha Özgür**[7], **Nynke Gillemans**[3], **Jun Hou**[3,8], **Marisa Bugeja**[1,2], **Frank G. Grosveld**[3,8,9], **Marieke von Lindern**[10], **Alex E. Felice**[1,2,#], **George P. Patrinos**[4,#], and **Sjaak Philipsen**[3,8,#]

[1] Laboratory of Molecular Genetics, Department of Physiology & Biochemistry, University of Malta, Msida, MSD 2080, Malta [2] Thalassaemia Clinic, Section of Pathology, Mater Dei Hospital, Msida, MSD 2080, Malta [3] Erasmus MC, Department of Cell Biology, P.O. Box 2040, 3000 CA Rotterdam, The Netherlands [4] University of Patras, Department of Pharmacy, University Campus, Rion GR-265 04, Patras, Greece [5] Cyprus Institute of Neurology and Genetics, P.O. Box 23462, 1683 Nicosia, Cyprus [6] Erasmus MC, Department of Bioinformatics, P.O. Box 2040, 3000 CA Rotterdam, The Netherlands [7] Erasmus MC, Center for Biomics, P.O. Box 2040, 3000 CA Rotterdam, The Netherlands [8] Netherlands Consortium for Systems Biology, P.O. Box 2040, 3000 CA Rotterdam, The Netherlands [9] Center for Biomedical Genetics, P.O. Box 2040, 3000 CA Rotterdam, The Netherlands [10] Erasmus MC, Department of Hematology, P.O. Box 2040, 3000 CA Rotterdam, The Netherlands

Abstract

Hereditary Persistence of Fetal Hemoglobin (HPFH) is characterized by persistent high levels of fetal hemoglobin (HbF) in adults. Several contributory factors, both genetic and environmental, have been identified [1], but others remain elusive. Ten of twenty-seven members from a Maltese family presented with HPFH. A genome-wide SNP scan followed by linkage analysis revealed a candidate region on chromosome 19p13.12–13. Sequencing identified a nonsense mutation in the *KLF1* gene, p.K288X, ablating the DNA binding domain of this key erythroid transcriptional regulator [2]. Only HPFH family members were heterozygote carriers of this mutation. Expression

[#]Corresponding authors: alex.felice@um.edu.mt (Tel. 356-23402774), gpatrinos@upatras.gr (Tel. +30-2610-969834), j.philipsen@erasmusmc.nl (Tel. +31-10-7044282).

[*]These authors contributed equally to this work

AUTHOR CONTRIBUTIONS

FGG, AEF, GPP and SP designed experiments; JB, PP, MG, LG, GG, PF, MP, CAS, WC, RG, ZÖ, NG, and ML performed experiments; JB, PP, MG, LG and GG analyzed experiments; PJS FGG, AEF, GPP and SP supervised data analysis, PJS, WIJ and MB, provided expertise, analysis tools and infrastructure; AJMHV, JH and MB analyzed data; JB, PP, MG, FGG, ML, AEF, GPP and SP wrote the paper.

Declaration of competing financial interests

The authors declare no competing financial interests.

Database accession number

GSE22109

profiling on primary erythroid progenitors revealed down-regulation of KLF1 target genes in HPFH samples. Functional assays demonstrated that, in addition to its established role in adult globin expression, KLF1 is a critical activator of the *BCL11A* gene, encoding a suppressor of HbF expression [3]. These observations provide a rationale for the effects of *KLF1* haploinsufficiency on HbF levels.

Hemoglobin (Hb) is composed of two α-like and two β-like globin chains, encoded by genes in the *HBA* and *HBB* clusters, respectively. Developmental regulation of globin genes results in expression of stage-specific Hb variants (Supplementary Fig. 1). HbF ameliorates the symptoms of β-thalassemia and sickle cell disease, and reactivation of the *HBG1*/*HBG2* genes in adults is therefore of significant interest for the clinical management of β-type hemoglobinopathies. After birth, HbF is gradually replaced by adult hemoglobin (HbA) [4]. Residual amounts of HbF continue to be synthesized throughout adult life. In the majority of adults, HbF contributes <2% to total Hb, but there is considerable variation [5]. Genetic studies have revealed three loci that control HbF levels in adults: *HBB* (11p15.4) [6–7], *HBS1L-MYB* (6q23.3) [6,8–9] and *BCL11A* (2p16.1) [10–11]. Together, these loci account for <50% of the variation in HbF, indicating that additional loci are involved [5]. Genetic analysis of HPFH families is a particularly powerful approach to identify novel modifiers of HbF levels [8]. Here, we describe a Maltese pedigree with HPFH. The proband (II-5, Fig. 1a) was referred to the clinic because of microcytosis. She presented with high HbF levels (19.5%). Additional family members were recruited and ten of twenty nine tested were identified with HPFH (Fig. 1a and Supplementary Table 1), suggesting an autosomal dominant inheritance of the trait. Linkage to the *HBB* locus (Supplementary methods) was excluded, indicating involvement of a *trans*-acting factor. We performed a genome-wide linkage analysis on twenty seven family members to identify candidate loci for the HPFH modifier. Whole genome multi-point parametric linkage analysis was conducted using the Merlin programme [12] with two software packages, easyLINKAGE [13] and dChip [14]. The analyses resulted in one significant linkage peak with LOD scores of 2.7 and 4.2, respectively, on chromosome 19p13.12–13 (Fig. 1b and Supplementary Fig. 2). These analyses were performed using an autosomal dominant model, assuming a penetrance of 90% and 1% phenocopy rate. No evidence of significant linkage was observed to the previously reported *trans*-acting HPFH loci of chromosomes 2p16.1 [10–11] and 6q23.3 [6,8–9]. This was further investigated by genotyping of the five individual SNPs linked to increased HbF levels. This ruled out involvement of the *HBS1L-MYB* locus and revealed that heterozygosity at SNP rs766432 in the *BCL11A* locus may have contributed to the increased HbF levels, but it was not the major determinant (Supplementary Table 1). HPFH individuals had a consistent haplotype at 19p13.12–13, and the inferred haplotypes revealed that all HPFH individuals shared one copy of an identical chromosome segment, presumably containing the putative HPFH locus (Supplementary Fig. 2). Recombination events delineating the linkage region are indicated with arrows. The distal boundary is determined by a recombination event in individuals IV-3 and IV-5 (Supplementary Fig. 2, white arrow). The proximal boundary is determined by individuals III-12, III-18, IV-6 and IV-7 (Supplementary Fig. 2, black arrow). These narrowed the region down to a 663 kb interval between rs7247513 and rs12462609. The *KLF1* gene, encoding a key erythroid transcriptional regulator [2], resides in this area. Mutations in *KLF1* have been reported to form the molecular basis of the rare blood group

In(Lu) phenotype [15], but a connection with HPFH has not been made. DNA sequencing revealed two linked mutations in *KLF1* present exclusively in all HPFH individuals (Fig. 1c). The first mutation, p.M39L, is most likely a neutral substitution since mouse Klf1 contains a leucine at this position [16]. The second mutation, p.K288X, ablates the complete zinc finger domain and therefore abrogates DNA binding of the mutant protein [17]. The *KLF1* p.K288X variant was absent in a random sample from the general Maltese population (n=400). To identify differentially expressed genes, RNA was isolated from erythroid progenitors (HEPs) cultured from peripheral blood [18] of four HPFH and four non-HPFH family members and used for genome-wide expression analysis. Comparison to the reported gene expression profiles of mouse Klf1 *null* erythroid progenitors [19] identified a set of common differentially regulated genes (Supplementary Table 2). Cluster analysis with this set of genes separated the non-HPFH samples from the HPFH samples (Fig. 2a), consistent with the notion that KLF1 activity is compromised in the HPFH individuals. Deregulation of these KLF1 target genes could explain the mild hypochromic microcytic indices displayed by the HPFH individuals (Supplementary Table 1). Of note, the embryonic *Hbb-y* and *HBE1* genes were highly upregulated (Supplementary Table 2) while expression of the fetal globin repressor BCL11A [3] was downregulated (Supplementary Table 2 and Supplementary Fig. 3). Expression of fetal/adult globins can not be measured quantitatively on the microarrays owing to saturation effects. Quantitative RT-PCR (qPCR) confirmed downregulation of BCL11A, and showed increased expression of *HBG1*/*HBG2* genes in the HPFH samples (Fig. 2b). Next, we investigated the effects of KLF1 knockdown in HEPs derived from healthy donors. Efficient knockdown of KLF1 was observed with two out of five lentiviral shRNA constructs [20] tested (Fig. 3a). Quantitative S1 nuclease protection assays [21] demonstrated significantly increased HBG1/HBG2 expression upon KLF1 knockdown (Fig. 3b–d), which was confirmed by qPCR (Fig. 3e). In addition, we found that BCL11A expression was diminished after KLF1 knockdown, both at the protein- (Fig. 3a) and at the mRNA level (Fig. 3e). Thus, the effects of KLF1 insufficiency on HBG1/HBG2 and BCL11A expression in HEPs from healthy donors were similar to those observed in *KLF1* p.K288X heterozygotes, supporting the causative role of this mutation in the HPFH phenotype. To further investigate this, we transduced HEPs with lentiviral vectors expressing the KLF1 p.K288X truncation mutant or full length KLF1. The transgenic proteins were expressed at physiological levels in control HEPs (Supplementary Fig. 4a). This did not affect HBG1/HBG2 expression levels (Supplementary Fig. 4b, c), indicating that the truncated form of KLF1 does not act as a dominant-negative factor. In HPFH HEPs, lentivirus-mediated expression of full-length KLF1 resulted in considerable downregulation of HBG1/HBG2 mRNA levels, while expression of truncated KLF1 had no effect (Fig. 4). BCL11A protein levels were increased after transduction with full-length KLF1 lentivirus, while no such changes were observed upon transduction with either GFP- or truncated KLF1 lentiviral vectors (Fig. 4a). We noted that the endogenous truncated KLF1 protein was not or at best barely detectable in HPFH HEPs. This suggested that RNA transcribed from the *KLF1* p.K288X allele was subject to nonsense-mediated decay [22], further emphasizing that it was dysfunctional. Consistent with this notion, we found that KLF1 mRNA expression was reduced in HPFH HEPs (Supplementary Fig. 3). It is well established that KLF1 preferentially activates the *HBB* gene, at the expense of *HBG1/HBG2* gene expression, through direct interactions with regulatory elements in the *HBB* promoter [23–25]. The

molecular analysis of the Maltese HPFH family is consistent with this function of KLF1. In addition, it revealed a novel mechanism by which KLF1 tips the balance from *HBG1/HBG2* to *HBB* expression: through activation of the gene encoding the *HBG1/HBG2* repressor BCL11A [3]. The promoter area of the *BCL11A* gene contains several putative KLF1 binding sites (CACC boxes; Fig. 5a). We performed chromatin immunoprecipitation (ChIP) assays to investigate whether KLF1 was bound to the *BCL11A* promoter *in vivo*. We used human fetal liver erythroid progenitors, which express high levels of HBG1/HBG2, and HEPs from adult peripheral blood in which the *HBG1/HBG2* genes are suppressed. In adult HEPs, we observed strong binding of KLF1 to the *BCL11A* promoter (Fig. 5b). This was similar to the binding of KLF1 to the *HBB* promoter which served as a positive control [26]. Neither promoter appeared to be bound by KLF1 in fetal liver-derived erythroid progenitors. ChIP reactions with the unrelated CD71 antibody were negative in all cases. We conclude that in adult HEPs KLF1 is bound to the *BCL11A* promoter *in vivo*. Diminished KLF1 activity, either through mutation of one *KLF1* allele as occurs in the Maltese HPFH individuals or experimentally through shRNA-mediated knockdown in HEPs from normal donors, decreased *BCL11A* expression. Conversely, BCL11A levels increased upon restoration of KLF1 activity in Maltese HPFH HEPs. This identifies KLF1 as a double-barreled regulator of fetal-to-adult globin switching in humans (Supplementary Fig. 5). Firstly, it acts on the *HBB* locus as a preferential activator of the *HBB* gene [27]. Secondly, it activates expression of *BCL11A*, which in turn represses the *HBG1/HBG2* genes. This “double whammy” ensures that in most adults HbF levels are <2% of total Hb. In conclusion, we have identified haploinsufficiency for KLF1 as a cause of HPFH. We suggest that attenuation of KLF1 activity may be a fruitful approach to raise HbF levels in patients with β-type hemoglobinopathies.

Online Methods

Molecular genetic analysis

The proband (II-5; Fig. 1a) was referred to the clinic because of microcytosis. She presented with 19.5% HbF, and therefore additional family members were approached to participate. Blood samples were obtained with informed consent and standard hematological indices were determined (Supplementary Table 1). Genomic DNA was extracted from ~1×10^6 cells from whole blood using a modified salting out procedure [29]. Control DNA samples isolated from 400 random Maltese individuals were available from the Laboratory of Molecular Genetics, Biomedical Sciences Building, University of Malta. The family members were genotyped in the *HBB, HBD* genes and the *HBG1/HBG2* gene promoters to detect point mutations and small insertions/deletions leading to β-, δ-thalassemia or HPFH, respectively, following routine procedures [30]. Gap PCR was carried out to detect possible genomic rearrangements leading to deletional HPFH or δ β-thalassemia [31]. This excluded linkage of the HPFH phenotype to the *HBB* locus. Occurrence of common α-thalassemic mutations (SEA, 3.7 and 4.2 deletions) was also excluded. The *Nsp*I mapping 250K set (Affymetrix, Santa Clara, CA, USA) was used to analyze twenty seven DNA samples from the HPFH family, starting with 250 ng of genomic DNA per array. Individual SNPs in the *HBS1L-MYB* (rs28384513, rs9399137, rs4895441) and *BCL11A* (rs766432, rs11886868) loci [32] were genotyped manually.

DNA Linkage Analysis

Multipoint parametric linkage analyses was performed using the Merlin v1.0.1 [12] program with two software packages; EasyLinkage v5.05 Beta [13] and dChip [14], in order to calculate parametric LOD scores. Parametric analysis was carried out using an autosomal dominant mode of inheritance. Penetrances used for the dominant model were 0.01, 0.90 and 0.90 for the wild-type homozygote, mutant heterozygote and mutant homozygote, respectively. We assumed a disease allele prevalence frequency of 0.0001, and a phenocopy rate of 1%. A co-dominant allele frequency algorithm was used for the analysis. These analyses were carried out using the sex-averaged 500K Marshfield genetic map provided with the easyLINKAGE software package [13]. Mendelian inheritance check was performed for all family members, using the program PedCheck [33] and incompatibilities were omitted from the analysis. This increased the power and accuracy. The analysis was performed by taking HbF as a quantitative hematological value and classifying family members as 'affected' with HbF >2%, and 'non-affected' with HbF <2%. Replicates of the linkage analysis and inferred haplotypes were constructed and visualized using dChip [14]. DNA from 400 random Maltese individuals was used to check for the presence of the KLF1 p.K288X mutation in the population.

Cell culture

Human erythroid progenitor cells (HEPs) were cultured as described [18] in the presence of recombinant human Epo (1 unit/ml, kind gift of Ortho-Biotech, Tilburg, The Netherlands), recombinant human SCF (50 ng/ml, kind gift of Amgen, Breda, The Netherlands) and dexamethasone (5×10^{-7} M; Sigma, St. Louis, MO, USA). Cells were counted with an electronic cell counter (CASY-1, Schärfe System, Reutlingen, Germany).

Transcription profiling

A minimum of 1.5×10^6 HEPs were harvested at day12 of culture and RNA was extracted with Trizol reagent (Sigma) and purified using the RNeasy Mini Kit (Qiagen, Crawley, UK), including an on-column DNaseI digestion, according to the manufacturer's instructions. RNA yield was determined using the 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). 8–10μg of total RNA was analysed by microarrays using cells from day 12 of culture. Quality of the total RNA samples and the resulting cRNA was assessed on the Bioanalyzer. Fragmented biotinylated cRNA was prepared and 15μg hybridized to HG-U133 plus 2 GeneChips, according to the manufacturer's protocols (Affymetrix). The data files have been deposited in MIAME-compliant format in the NCBI GEO database (GSE22109). Single Array Expression Analysis was performed using the Affymetrix GeneChip Operating Software (GCOS). A global scaling strategy was used to give an average target intensity of 500 for each array. Data from all eight arrays were filtered to exclude probe sets called either absent or marginal in all arrays. Control probe sets with the prefix AFFX were also removed prior to subsequent data analysis. Filtered data were transformed to a log2 scale and analysed to determine differentially expressed genes. A 1.5-fold change threshold and test statistic of $p<0.05$ were used as cut-off. A list of genes differentially expressed in mouse *Klf1* null erythroid progenitors ($p<0.05$) 19 was downloaded from http://data.genome.duke.edu/EKLFDef.

Quantitative S1 nuclease protection assays

To measure globin mRNA levels directly, we used quantitative S1 nuclease protection assays [21]. The probe fragment for detection of HBG1/HBG2 mRNAs was amplified by PCR using the primers S1-HBG-S and S1-HBG-A (Supplementary Table 3). Sizes of probes/protected fragments are: HBA1/HBA2: 700 nt/218 nt; HBG1/HBG2: 350 nt/165 nt; HBB: 525 nt/155 nt [21]. Quantitation was performed using a Typhoon Trio Phosphorimager (GE Healthcare, Chalfont St Giles, UK) and corrected for specific activity of the probes.

qPCR analysis

Total RNA (1 μg) isolated from HEPs was converted to cDNA using SuperScript II reverse transcriptase according to the manufacturer's instructions (Invitrogen, Carlsbad, CA, USA). Expression levels of mRNAs were analysed by quantitative real-time PCR (qPCR). Amplification reactions were performed with primers designed with Primer Express software v2.0 (Applied Biosystems, Foster City, CA, USA). All amplifications used SYBR Green PCR Master Mix (Applied Biosystems). qPCR was performed with an Optical IQ Thermal Cycler (Bio-Rad Laboratories, Hercules, CA, USA) with the following conditions: 50°C for 2 minutes and 95°C for 10 minutes, followed by 45 cycles of 95°C for 15 seconds and 62°C for 45 seconds. All reactions were performed in triplicate. Gene expression levels were calculated with the 2 (−DeltaDeltaC(T)) method [34]. Target gene expression was normalized to GAPDH expression, unless indicated otherwise. Primers used are listed in Supplementary Table 3.

Statistical analysis

Statistical analysis of gene expression data obtained from quantitative S1 nuclease protection assays and qPCRs was performed with Mann Whitney tests using STATA data analysis and statistical software (StataCorp LP, College Station, TX, USA).

KLF1 expression constructs

Human KLF1 cDNA clone (BC040000, Imagenes, Berlin, Germany) was amplified by PCR with an *att*-specific set of primers (Invitrogen) in order to fuse the cDNA with a V5 tag at the C-terminus of the protein. Primers used were KLF1-F and KLF1-R1 (Supplementary Table 3). In parallel, part of the clone was amplified, truncating the protein at amino acid 288, with *att*-specific primers using a different reverse primer KLF1-R2. The PCR products were introduced into the lentiviral expression vector pRRLsin.sPPT.CMV.Wpre [35] modified for Gateway cloning (Invitrogen). The final clones were verified by sequencing.

Lentiviral transduction of human erythroid progenitors

Lentivirus was produced by transient transfection of 293T cells according to standard protocols [36]. Two days after transfection, the supernatant was collected, filtered and concentrated by centrifugation at 20 krpm for 2h at 4°C. HEPs cultured for one week were transduced in 24 well plates. We used 0.5×10^6 cells per well and sufficient amounts of virus to transduce ~80% of the cells. When appropriate, puromycin (1μg/ml final concentration) was added to the cells after 2 days, and selection was performed for 2–3 days. At day 5–7 after transduction cells were harvested and nuclear extracts were prepared [37]. RNA was

extracted with the Trizol reagent. For knockdown experiments, clones from The RNAi Consortium (TRC [20]; Sigma) were used. The non-target SHC002 vector was used as a control. (SHC002: 5'-CAACAAGATGAAGAGCACCAA-3'). Five shRNA clones targeting KLF1 were tested: TRCN0000016273, TRCN0000016274, TRCN0000016275, TRCN0000016276 and TRCN0000016277. Efficient knockdown of KLF1 expression was observed with TRCN0000016276 (sh1) and TRCN0000016277 (sh2). Sequences are listed in Supplementary Table 3.

Western blotting

Nuclear extracts were separated on denaturing polyacrylamide gels followed by semi-dry blotting to PVDF or nitrocellulose membranes. The membranes were probed with the following primary antibodies: BCL11A (sc-56013, Santa Cruz Biotechnology, Santa Cruz, CA, USA), NPM1 (ab10530, Abcam, Cambridge, UK), KLF1 [26], and anti-V5-HRP (R961-25, Invitrogen). For detection, the appropriate secondary antibodies were used. The enhanced chemoluminescence kit (GE Healthcare) or the Odyssey Infrared Imaging System (Li-Cor Biosciences, Lincoln, NE, USA) was used to develop the membranes.

Chromatin immunoprecipitations

Fetal liver and adult HEPs were cultured [18] and used for ChIP reactions were performed as described [38] with the KLF1 antibody, and a CD71 antibody (347510, BD Biosciences, San Jose, CA, USA) as a negative control. qPCR was performed on the input and immuno-precipitated samples using primers for the *RASSF1A*, *HBB* and *BCL11A* genes. The relative fold enrichment was calculated as $2^{-[(CT\ x\ ChIP\ y - CT\ input\ y)-(CT\ KLF1\text{-}ChIP\ HEP\ RASSF1A - CT\ input\ HEP\ RASSF1A]}$ (where 'x' is the antibody and 'y' the sample), i.e. setting the relative fold-enrichment of the *RASSF1A* amplicon by the KLF1 antibody in HEPs to 1. Primers used are listed in Supplementary Table 3.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We are grateful to the familymembers for their co-operation. We thank Mr. Teus van Gent for help with statistical analysis, Dr. Thamar van Dijk for sharing his expertise in lentiviral technology, Ms. Natalie Papazian for preparation of human fetal liver samples and Ms. Monica Pizzuto for administrative support. The population samples from Malta were obtained from the Malta BioBank (EuroBioBank). This work was supported by institutional funds of the University of Malta and Mater Dei Hospital to AEF; a Malta Government Scholarship to JB; European Molecular Biology Organization (EMBO) short-term fellowships to JB and PP; the Netherlands Scientific Organization (NWO) to LG (VENI 863.09.012); the Netherlands Genomics Initiative (NGI), Erasmus MC (MRace; 296088), the Landsteiner Foundation for Blood Transfusion Research (LSBR; 0615), and NWO (DN 82-294 and 912-07-019) to SP; the Centre for Biomedical Genetics (CBG), the European Commission FP6 EuTRACC consortium (037445) and the NIH (R01-HL073455) to FG; and the Research Promotion Foundation of Cyprus (πΔ E046_02) and the European Commission FP7 GEN2PHEN (200754) projects to GPP.

References

1. Serjeant, G. Geographic heterogeneity of sickle cell disease. In: Steinberg, MH.; Forget, BG.; Higgs, DR.; Nagel, RL., editors. Disorders of Hemoglobin. Cambridge University Press; Cambrige: 2001. p. 895-906.
2. Bieker JJ. Probing the onset and regulation of erythroid cell-specific gene expression. Mt Sinai J Med. 2005; 72:333–8. [PubMed: 16184297]
3. Sankaran VG, et al. Human fetal hemoglobin expression is regulated by the developmental stage-specific repressor BCL11A. Science. 2008; 322:1839–42. [PubMed: 19056937]
4. Stamatoyannopoulos, G.; Grosveld, F. Hemoglobin switching. In: Stamatoyannopoulos, G.; Majerus, PW.; Perlmutter, RM.; Varmus, H., editors. The Molecular Basis of Blood Diseases. WB Saunders Company; Philadelphia PA: 2001. p. 135-182.
5. Thein SL, Menzel S, Lathrop M, Garner C. Control of fetal hemoglobin: new insights emerging from genomics and clinical implications. Hum Mol Genet. 2009; 18:R216–23. [PubMed: 19808799]
6. Craig JE, et al. Dissecting the loci controlling fetal haemoglobin production on chromosomes 11p and 6q by the regressive approach. Nat Genet. 1996; 12:58–64. [PubMed: 8528252]
7. Gilman JG, Huisman TH. DNA sequence variation associated with elevated fetal G gamma globin production. Blood. 1985; 66:783–7. [PubMed: 2412616]
8. Close J, et al. Genome annotation of a 1.5 Mb region of human chromosome 6q23 encompassing a quantitative trait locus for fetal hemoglobin expression in adults. BMC Genomics. 2004; 5:33. [PubMed: 15169551]
9. Garner C, et al. Haplotype mapping of a major quantitative-trait locus for fetal hemoglobin production, on chromosome 6q23. Am J Hum Genet. 1998; 62:1468–74. [PubMed: 9585587]
10. Lettre G, 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; 105:11869–74. [PubMed: 18667698]
11. Menzel S, et al. A QTL influencing F cell production maps to a gene encoding a zinc-finger protein on chromosome 2p15. Nat Genet. 2007; 39:1197–9. [PubMed: 17767159]
12. Abecasis GR, Cherny SS, Cookson WO, Cardon LR. Merlin--rapid analysis of dense genetic maps using sparse gene flow trees. Nat Genet. 2002; 30:97–101. [PubMed: 11731797]
13. Hoffmann K, Lindner TH. easyLINKAGE-Plus--automated linkage analyses using large-scale SNP data. Bioinformatics. 2005; 21:3565–7. [PubMed: 16014370]
14. Leykin I, et al. Comparative linkage analysis and visualization of high-density oligonucleotide SNP array data. BMC Genet. 2005; 6:7. [PubMed: 15713228]
15. Singleton BK, Burton NM, Green C, Brady RL, Anstee DJ. Mutations in EKLF/KLF1 form the molecular basis of the rare blood group In(Lu) phenotype. Blood. 2008; 112:2081–8. [PubMed: 18487511]
16. Miller IJ, Bieker JJ. A novel, erythroid cell-specific murine transcription factor that binds to the CACCC element and is related to the Kruppel family of nuclear proteins. Mol Cell Biol. 1993; 13:2776–86. [PubMed: 7682653]
17. Feng WC, Southwood CM, Bieker JJ. Analyses of beta-thalassemia mutant DNA interactions with erythroid Kruppel-like factor (EKLF), an erythroid cell-specific transcription factor. J Biol Chem. 1994; 269:1493–500. [PubMed: 8288615]
18. Leberbauer C, et al. Different steroids co-regulate long-term expansion versus terminal differentiation in primary human erythroid progenitors. Blood. 2005; 105:85–94. [PubMed: 15358620]
19. Pilon AM, et al. Failure of terminal erythroid differentiation in EKLF-deficient mice is associated with cell cycle perturbation and reduced expression of E2F2. Mol Cell Biol. 2008; 28:7394–401. [PubMed: 18852285]
20. Moffat J, et al. A lentiviral RNAi library for human and mouse genes applied to an arrayed viral high-content screen. Cell. 2006; 124:1283–98. [PubMed: 16564017]
21. Antoniou M. Induction of erythroid-specific expression in murine erythroleukemia (MEL) cell lines. Methods in Molecular Biology. 1991; 7:421–434. [PubMed: 21416373]

22. Muhlemann O, Eberle AB, Stalder L, Zamudio Orozco R. Recognition and elimination of nonsense mRNA. Biochim Biophys Acta. 2008; 1779:538–49. [PubMed: 18657639]
23. Donze D, Townes TM, Bieker JJ. Role of erythroid Kruppel-like factor in human gamma- to beta-globin gene switching. J Biol Chem. 1995; 270:1955–9. [PubMed: 7829533]
24. Drissen R, et al. The active spatial organization of the beta-globin locus requires the transcription factor EKLF. Genes Dev. 2004; 18:2485–90. [PubMed: 15489291]
25. Zhou D, Pawlik KM, Ren J, Sun CW, Townes TM. Differential binding of erythroid Krupple-like factor to embryonic/fetal globin gene promoters during development. J Biol Chem. 2006; 281:16052–7. [PubMed: 16606611]
26. Bottardi S, Ross J, Pierre-Charles N, Blank V, Milot E. Lineage-specific activators affect beta-globin locus chromatin in multipotent hematopoietic progenitors. EMBO J. 2006; 25:3586–95. [PubMed: 16858401]
27. Wijgerde M, et al. The role of EKLF in human beta-globin gene competition. Genes Dev. 1996; 10:2894–902. [PubMed: 8918890]
28. Drissen R, et al. The erythroid phenotype of EKLF-null mice: defects in hemoglobin metabolism and membrane stability. Mol Cell Biol. 2005; 25:5205–14. [PubMed: 15923635]
29. Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res. 1988; 16:1215. [PubMed: 3344216]
30. Trifillis P, Ioannou P, Schwartz E, Surrey S. Identification of four novel delta-globin gene mutations in Greek Cypriots using polymerase chain reaction and automated fluorescence-based DNA sequence analysis. Blood. 1991; 78:3298–305. [PubMed: 1742490]
31. Craig JE, Barnetson RA, Prior J, Raven JL, Thein SL. Rapid detection of deletions causing delta beta thalassemia and hereditary persistence of fetal hemoglobin by enzymatic amplification. Blood. 1994; 83:1673–82. [PubMed: 7510147]
32. Higgs DR, Wood WG. Genetic complexity in sickle cell disease. Proc Natl Acad Sci U S A. 2008; 105:11595–6. [PubMed: 18695233]
33. O'Connell JR, Weeks DE. PedCheck: a program for identification of genotype incompatibilities in linkage analysis. Am J Hum Genet. 1998; 63:259–66. [PubMed: 9634505]
34. Schmittgen TD, Livak KJ. Analyzing real-time PCR data by the comparative C(T) method. Nat Protoc. 2008; 3:1101–8. [PubMed: 18546601]
35. Follenzi A, Sabatino G, Lombardo A, Boccaccio C, Naldini L. Efficient gene delivery and targeted expression to hepatocytes in vivo by improved lentiviral vectors. Hum Gene Ther. 2002; 13:243–60. [PubMed: 11812281]
36. Zufferey R, Nagy D, Mandel RJ, Naldini L, Trono D. Multiply attenuated lentiviral vector achieves efficient gene delivery in vivo. Nat Biotechnol. 1997; 15:871–5. [PubMed: 9306402]
37. Andrews NC, Faller DV. A rapid micropreparation technique for extraction of DNA-binding proteins from limiting numbers of mammalian cells. Nucleic Acids Res. 1991; 19:2499. [PubMed: 2041787]
38. Follows GA, et al. Epigenetic consequences of AML1-ETO action at the human c-FMS locus. EMBO J. 2003; 22:2798–809. [PubMed: 12773394]

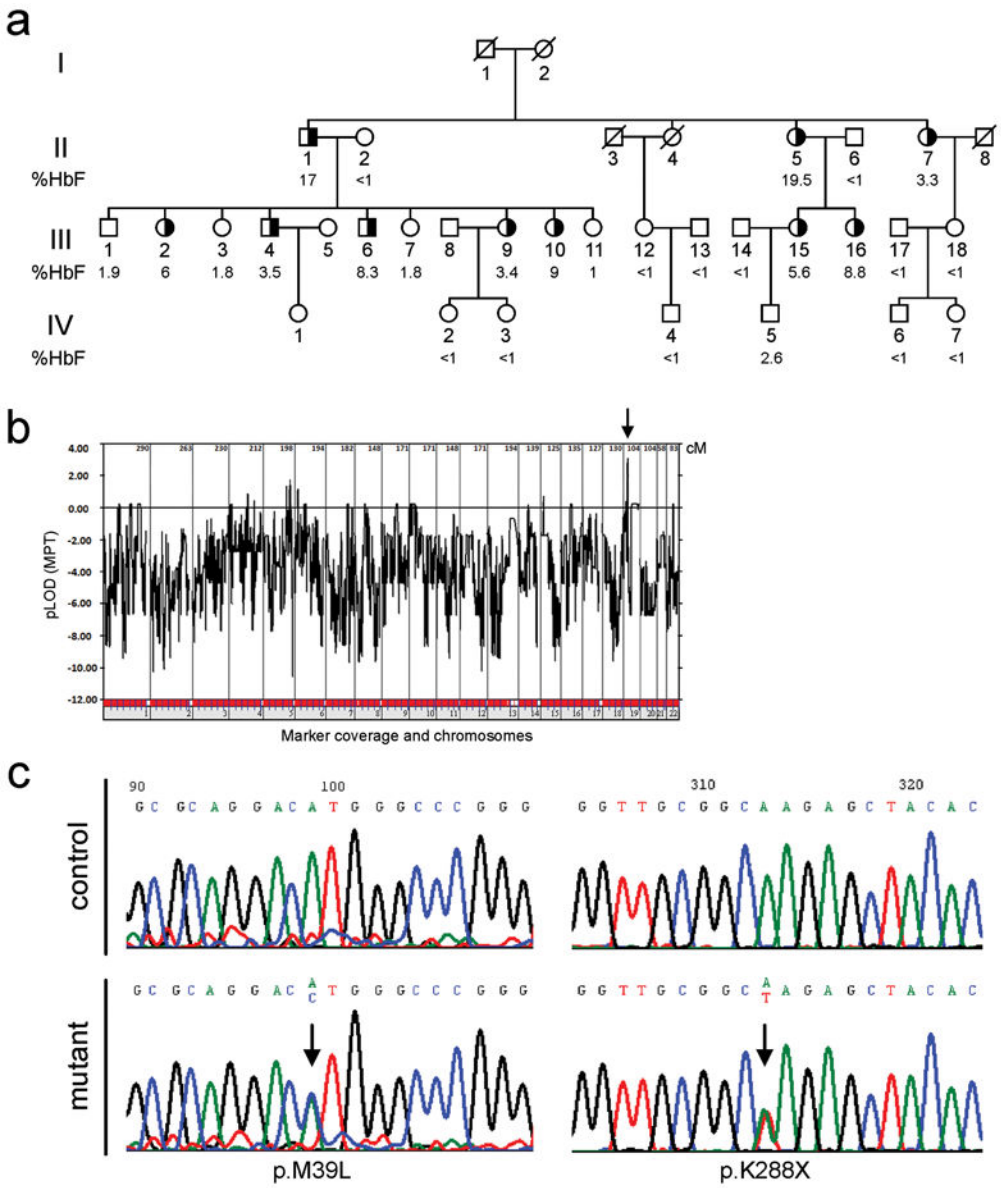

Figure 1. Chromosome 19 locus linked to HPFH in a Maltese family

a) The Maltese HPFH pedigree. HbF levels are indicated as percentage of total Hb (%HbF). HPFH individuals are shown as half-filled symbols.

b) LOD scores derived from genome-wide linkage analysis. The putative HPFH locus on chr. 19 is indicated by an arrow. pLOD = parametric LOD score; MPT = multi point test; cM = centiMorgan.

c) Sequence analysis of *KLF1*. HPFH individuals were heterozygous for two mutations (arrows; Supplementary Table 1). The predicted effects of the mutations on KLF1 protein are shown below.

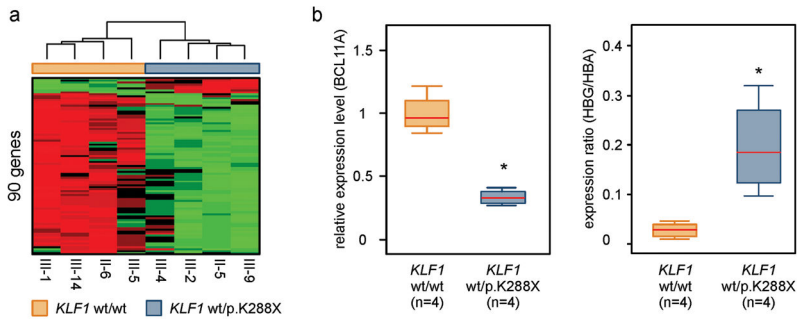

Figure 2. KLF1 target genes are downregulated in wt/KLF1 p.K288X HEPs

a) RNA isolated from HEPs derived from normal (wt/wt) and HPFH (wt/KLF1 p.K288X) family members was used for genome-wide expression analysis. Deregulated genes common between wt/wt and wt/KLF1 p.K288X and mouse wt/wt versus *Klf1* null mutant erythroid progenitors [19] (Supplementary Table 2) were used for cluster analysis.

b) Validation of key target genes by qPCR. Expression levels of BCL11A were normalized using GAPDH as a reference. Expression levels of HBG1/HBG2 (HBG) were calculated as ratio to HBA1/HBA2 (HBA) expression. Medians are indicated by red lines in the box plots. Asterisk: p=0.0209.

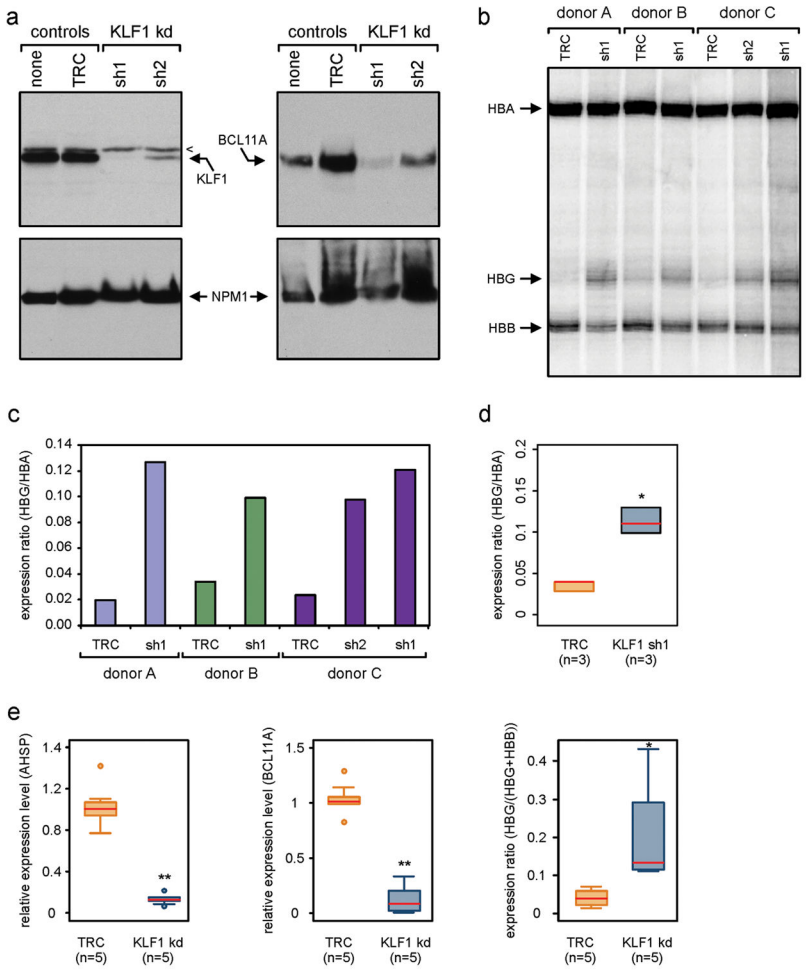

Figure 3. Increased HBG1/HBG2 expression after knockdown of KLF1 in normal HEPs

a) HEPs derived from normal donors were transduced with shRNA-expressing lentiviruses. Cells were harvested five days after transduction, and nuclear extracts prepared. Top panels: KLF1 protein expression assessed by Western blot analysis. Bottom panels: BCL11A protein levels were reduced upon KLF1 knockdown. NPM1 served as a loading control. None = mock transduction; TRC = control non-specific shRNA; sh1 and sh2 = two independent shRNAs targeting KLF1. A non-specific band is indicated (<).

b) RNA was isolated from HEPs five days after transduction with the indicated lentiviruses, and used in quantitative S1 nuclease protection assays to measure globin expression. Arrows indicate protected fragments diagnostic for HBA1/HBA2 (HBA), HBG1/HBG2 (HBG) and HBB mRNAs.

c) Quantitation of data shown in b) by Phosphorimager analysis.

d) Box plots of HBG/HBA ratios after sh1-mediated KLF1 knockdown in HEPs derived from three independent donors. Medians are indicated by red lines. Asterisk: p=0.0463.

e) Box plots of qPCR analysis of AHSP, BCL11A and HBG expression after sh1/sh2-mediated KLF1 knockdown. *AHSP* is a known KLF1 target gene [28] and serves as a positive control. Expression levels of AHSP and BCL11A were normalized using GAPDH as a reference. Expression levels of HBG were calculated as ratio to total β-like globin

expression (HBG + HBB) expression. Medians are indicated by red lines. Circles: points outside the range of the error bars. Asterisk: p=0.020; double asterisks: p<0.003.

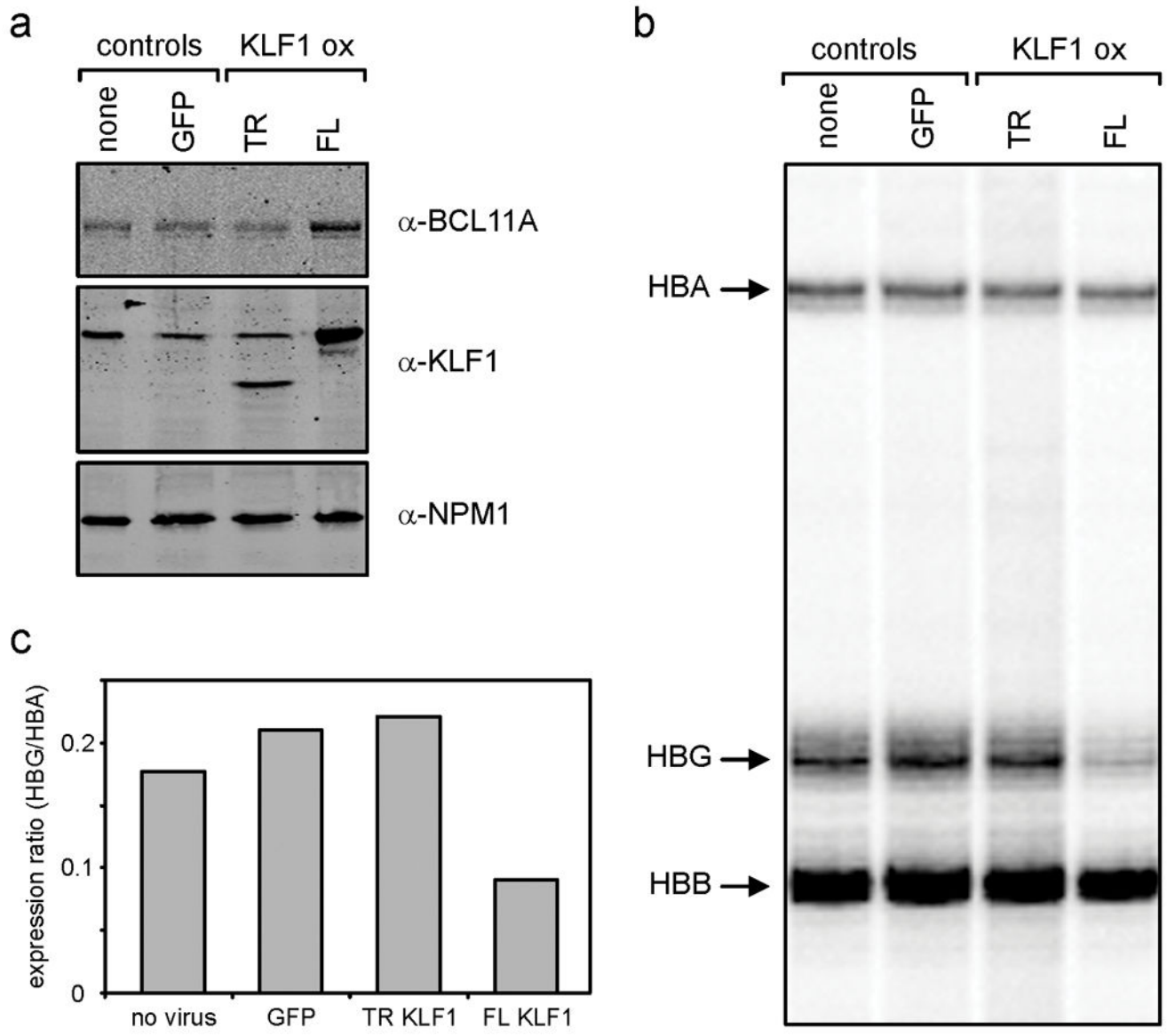

Figure 4. Expression of exogenous KLF1 in HPFH HEPs

a) HEPs derived from individual II-5 were transduced with lentiviral constructs expressing GFP, KLF1 truncated at amino acid 288 (TR) or full-length KLF1 (FL). Seven days after transduction, nuclear extracts were prepared and expression of BCL11A and KLF1 was assessed by Western blot. NPM1 served as a loading control.

b) RNA was isolated from II-5 HEPs seven days after transduction with the indicated lentiviruses, and was used for quantitative S1 nuclease protection assays to measure globin mRNA expression. Arrows indicate protected fragments diagnostic for HBA1/HBA2 (HBA), HBG1/HBG2 (HBG) and HBB mRNAs.

c) Quantitation of data shown in b) by Phosphorimager analysis.

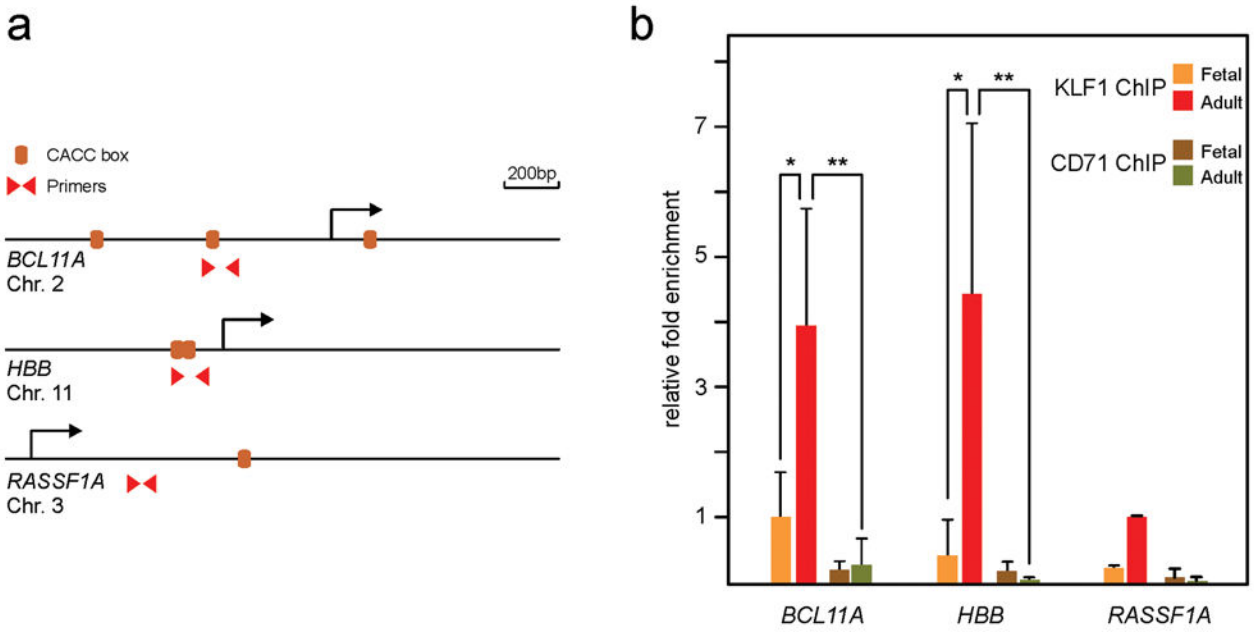

Figure 5. KLF1 binds to the promoter of the BCL11A gene in vivo

a) Schematic drawings of the promoter areas of the *BCL11A*, *HBB* and *RASSF1A* genes. Positions of potential KLF1 binding sites (CACC boxes) and PCR primers used are indicated. Arrows indicate transcription start sites.

b) ChIP analysis of KLF1 binding to the *BCL11A* promoter in human fetal liver cells and adult HEPs. The *HBB* promoter served as a positive control [26]. *RASSF1A* was used as a negative control, and the unrelated CD71 antibody served as a control for the specificity of the KLF1 antibody. Asterisk: $p<0.05$; double asterisks: $p<0.01$. Error bars: standard deviation.