HEATR3 variants impair nuclear import of uL18 (RPL5) and drive Diamond-Blackfan anemia

O'Donohue et al., Blood (2022)

Supplementary Data

Supplementary Materials and Methods
Table S1
Figures S1-S6

SUPPLEMENTARY MATERIALS AND METHODS

Screening EuroDBA network registries and exome sequencing. After routine diagnostic referral, exomes of affected children and parents from Family A were enriched using the SureSelectXT Clinical Research Exome V2 (Agilent, elid S30409818, genome build GRCh37) and sequenced on an Illumina Novaseq 6000. The sequencing data was processed with an in-house developed pipeline DxNextflowWES v1.0.1, based on the Genome Analysis Toolkit (GATK v3.8-1-0-gf15c1c3ef) best practices guidelines.[1,2] The read pairs were mapped with BWA-MEM v0.7.17,[3] marking duplicates and merging lanes using Sambamba v0.7.0 and realigning indels using GATK IndelRealigner.[4] GATK Haplotypecaller was used to call single nucleotide polymorphisms and indels, creating variant call formatted (vcf) files.

Analysis was performed using Alissa Interpret software (Agilent). In the first step, an intellectual disability gene panel comprising 816 genes (available on request) was analyzed. Variant were filtered using a population allele frequency of 0.05% or 0.5% (gnomAD database[5]) for the dominant or recessive inheritance model, respectively.

In a second step, often performed when the first step did not result in a diagnosis, the whole exome was analyzed, where variant filtering for a de novo or recessive inheritance model was done. All de novo variants were analyzed, but variants in putative recessive genes were only analyzed if both alleles showed a potential deleterious effect on the coding regions. Larger deletions/duplications, missense, synonymous, and intronic variants affecting protein function of other genes cannot be excluded. Variants that matched predefined and validated quality criteria were not validated by Sanger sequencing. Based on putative function, inheritance and segregation only one homozygous variant in the HEATR3 gene remained as candidate variant.

For the Turkish samples (Family B), the DBA registry in Hacettepe University Medical Genetics Exome Facility was used to search for a putative mutation in *HEATR3*. In this registry, WES data from a total of 22 individuals with unsolved DBA were screened. Briefly, DNA libraries from peripheral blood leukocytes for WES was prepared by Ion AmpliSeq Exome RDY kit (ThermoFisher Scientific) and subsequently sequenced by Ion Proton semiconductor sequencer (ThermoFisher Scientific). The IonReporter software was used for alignment, variant calling and annotation. *HEATR3* variant observed in P3 was verified and available family members were genotyped by Sanger sequencing. Further peripheral leukocyte DNA obtained from 20 individuals were sequenced by Sanger's method for all 15

coding exons of *HEATR3*. Identified variants were screened in gnomAD and in-house exome databases for non-DBA individuals (n=131).

For French samples (Family C), out of 412 alive DBA affected patients, we were able to identify one patient from the DBA patients with no characterized genotype after an extensive targeted NGS and CGH screening. P4, the French DBA patient and relatives have been identified by WES and confirmed by Sanger technique. Genomic DNA was extracted from leucocytes. Exome capture was performed with the SureSelect Human All Exon kit (Agilent Technologies). Agilent Sure Select Human All Exon (51Mb V5 or 54Mb Clinical Research Exome) libraries were prepared from 3 µg of genomic DNA sheared with an Ultrasonicator (Covaris) as recommended by the manufacturer. Barcoded exome libraries were pooled and sequenced with a HiSeq2500 system (Illumina) generating paired-end reads. After demultiplexing, sequences were mapped on the human genome reference (NCBI build 37, hg19 version) with BWA. After demultiplexing, Variant calling was carried out with the Genome Analysis Toolkit (GATK), SAMtools, and Picard tools. The mean depth of coverage of the exome libraries was greater than ~100-120X with >98 to 99% of the targeted exonic bases covered at least by 15 independent reads and >93 to 98% at least 30 independent sequencing reads (98-99% at 15X and 93 to 98% at 30X). Single-nucleotide variants were called with GATK Unified Genotyper, whereas indel calls were made with the GATK IndelGenotyper_v2. All variants with a read coverage ≤ 2x and a Phred-scaled quality of ≤ 20 were filtered out. All the variants were annotated and filtered with PolyWeb, an in-house-developed annotation software.

For Family D, blood sample of P5 was analyzed by the whole exome plus test service (version 2, Feb 9, 2018) from Blueprint Genomics. Individual P6 from Family D was diagnosed by Sanger sequencing verification of the variant identified by exome sequencing in P5 and this variant was verified by Sanger sequencing in the parents of P5 and P6.

Identification of homozygous stretches in consanguineous families. Unannotated vcf files obtained from WES data from individuals P1 (family A), P2 (family B), P3 (family B) and P5 (family D) with consanguineous parents were analyzed for long contiguous stretches of homozygosity (LCSH) >5Mbp[6] using HomSI[7]. LCSHs were identified visually for each individual and common regions were compared for all autosomes. Variants between [GRCh37]chr16:15,000,000 and [GRCh37]chr16:65,000,000, which surround the only common region ([GRCh37]chr16:25,263,278-53,191,470) was compared and illustrated through MS excel spread sheets. Illustrations were colored according to allele fractions (AF) for individual variants. Following assumptions were mad efor determining zygosity of varaiants: AF<0.20, homozygous for reference (grey); 0.20≤AF<0.80, heterozygous (yellow);

AF≥0.80 homozygous for alternative. Individual variants with no AF were considered either homozygous for reference or uncovered (grey).

Sanger sequencing of blood samples and LCLs. mRNA from LCLs was isolated with Trizol (ThermoFisher) for generation of cDNA using the QuantiTect Reverse Transcription Kit 400 (Qiagen) per the manufacturers' instructions. Primers used for Sanger sequencing of the *HEATR3* variant in LCLs were for P2 FW: 5'-TGCAAAGCAGGCATTCATAG-3' and REV: 5'-GAGGCTCTTTCGGCTTCTTT-3', and for P3 FW: 5'-TCTGCCTCTCCCATGAAGTT-3' and REV:5'- ACATCCAGGAGGGACACAAG-3'. For P4, genomic DNA was extracted from a whole blood sample and analyzed with primers for exon 3 FW: 5'-GGAAATCTCTGGATTGGTTCTG-3' and exon 3 REV: 5'-AGCAGATTGGTGTTTCCAGG-3'; exon 6 FW: 5'- ATTTGCGCATATCACCCTG-3' and exon 6 REV: 5'-CAACTGCTTGAAGTAAGGGTCTC-3'. For P5 and P6 and their parents, genomic DNA extracted from whole blood samples was analyzed using *HEATR3* primers FW: 5'-GTTTGTAAGGATTCGACTGGCTTG-3' and REV: 5'-TAGCAGCCTGGGTACTGCTAT-3'.

Yeast strains. To produce yeast strains expressing wild-type or mutated versions of HA-tagged Syo1, a *syo1Δ* and its BY4741 haploid isogenic strains obtained from Euroscarf were transformed with suitable plasmids (see Table A). To produce the plasmids, the following steps were followed: first, a plasmid expressing Syo1-Gly522Ala (pDL1008) was generated by homologous recombination in yeast cells by co-transforming a synthetic DNA fragment (gDL0019) harboring the Syo1-Gly522Ala gene and a linearized recipient low copy (ARS-CEN)/HIS3 plasmid (pFL36 digested with *Xho* I and *Bam* HI)[8] into the haploid strain BY4741. To produce plasmids expressing wild-type Syo1 and Syo1-Gly522Glu, point mutations were introduced directly into pDL1008 using a site directed mutagenesis kit (Quick change, Agilent, #200523). To produce plasmids expressing HA-tagged versions of Syo1, DNA sequences carrying wild-type or mutated versions of Syo1 were recovered by PCR amplification with oligonucleotides LD4496 and LD4497 using plasmids pDL1020 (Syo1), pDL1008 (Syo1-Gly522A), or pDL1077 (Syo1-Gly522Glu) as template, and co-transformed with a recipient linearized plasmid (pTL6/pDL923 digested with *Nco* I and *Sac* I)[9] into strain BY4741. All yeast clones were selected on synthetic medium lacking leucine. All plasmids were diagnosed by restriction digest and DNA sequencing. The plasmids and the oligonucleotides used for site-directed mutagenesis and DNA sequencing are listed below in Tables A and B.

Name	Information
pFL36	ARS-CEN LEU2
pDL923	pTL6
pDL1020	Syo1-WT
pDL1008	Syo1-G522A
pDL1077	Syo1-G522E
pDL1071	HA-Syo1
pDL1072	HA-Syo1-G522A
pDL1073	HA-Syo1-G522E
pDL1034	HA-HEATR3
pDL1035	HA-HEATR3-C446Y

Table A. Plasmids used for studies in yeast

Protein or mutant	Name	Sequence	Purpose
Syo1	LD4471	GCAAATAAGGAGATTGGGCAGTTCTTCATACAAAC	Site directed mutagenesis
Syo1	LD4472	GTTTGTATGAAGAACTGCCCAATCTCCTTATTTGC	Site directed mutagenesis
Syo1-G522A	gDL0019	GTCACGACGTTGTAAAACGACGGCCAGTGAATTCGAGCTCGGTACCCGGG…[on chromosome IV SYO1 DNA sequence from 338272 bp to 340134 bp mutated at position 339957 (G in C)] …AGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCC	Plasmid construction
Syo1-G522E	LD4473	GCAAATAAGGAGATTGAGCAGTTCTTCATAC	Site directed mutagenesis
	LD4474	GTATGAAGAACTGCTCAATCTCCTTATTTGC	Site directed mutagenesis
Syo1-HA	LD4496	CGTCGTACCTTTTATATCGATACAAATAAAAATAAAATTATACCATGTACCCCT ACGACGTGCCCGACTACGCCGGCAGATCAAAGAAAAGATCAAGAGC	Plasmid construction
	LD4497	CTTAGGTAAATAGTATACAAGCACTTACATAATTGATAAGAGAGCCTATTGGTT AGTAGAATTTTCATTCATTTTGTAGTG	Plasmid construction
Syo1	LD4477	CTGAACAATTCAGGATTCGC	Sequencing
Syo1	LD4486	AGAAATCGACTTTTTAGGGC	Sequencing
Syo1	LD4487	TTACTGTCGTTAGTGGTGGC	Sequencing
Syo1	LD4488	GCAATCATTGAGATAGTTGC	Sequencing

Table B. Oligonucleotides used for studies in yeast

Polysome gradient analysis of yeast. Yeast cells were grown overnight in synthetic medium lacking leucine to OD_{600} of ~0.8. Cycloheximide (250 µg/ml final) was added to the cultures which were incubated on ice for 30 min. Cell extracts and sucrose gradients (10–50%) were prepared exactly as described previously.[10] Fractions were collected on a density gradient fractionation system (Teledyne ISCO).

Yeast drop test assay. For serial dilution growth assays, yeast cells were grown overnight in synthetic medium lacking leucine. Cultures were harvested at OD_{600} of ~0.3 and serially diluted 10x. 2.5 µl of diluted cultures was spotted on agar plates containing synthetic medium lacking leucine and incubated at 16°C for 6 days.

Western blot analysis of yeast. Yeast strains expressing wild-type and mutants HA-tagged versions of Syo1 were grown to mid-log phase in synthetic medium lacking leucine. Growth was monitored by OD_{600} with an Ultrospec 10 (Amersham Biosciences). Total protein fractions were separated on 10% SDS-polyacrylamide gels and transferred to a PVDF membrane. Western blotting signals were captured with a Chemidoc (Biorad). The following antibodies were used: anti-GAPDH (Sigma, 1:5000) and anti-HA antibodies (Sigma, 1:5000).

Co-translational capturing assay. HeLa cells were transfected with plasmids expressing HA-HEATR3 (pDL1034), or HA-HEATR3-C446Y (pDL1035), or with an empty plasmid (control), using lipofectamine 3000 (Thermofischer). Total extract of transfected HeLa cells were used in ChIP assays as described.[11] Ribosomes were stalled by brief cycloheximide treatment. The material co-precipitated following affinity purification with an anti-HA antibody was amplified by RTqPCR. The qScript cDNA SuperMix kit (QuantaBio) for the reverse transcription and the Perfecta SYBR Green SuperMix (QuantaBio) for the RTqPCR, were used. Amplicons specific to the uL18/RPL5 mRNA or, as control to the uL5/RPL10 mRNA were amplified and quantitated. cDNAs were amplified with the following oligonucleotides: RPL5-fwd 5'-TTGTCAGATTGCTTATGCCC-3', RPL5-rev 5'- GCCAAACCTATTGAGAAGCC-3', RPL10-fwd 5'-ACTTCCTGTCCACCTATGTC-3', RPL10-rev 5'-AGTGTTTTTCAACCTAC TGCC-3'. The experiment was performed on a Roche Applied Science LightCycler 480. The data shown are a representative example of a triplicate. Calculation of the relative enrichment of RP encoding mRNAs was performed as in Pausch et *al*., 2015. The threshold cycle (C_t) was determined for each qPCR. Then, the ΔC_t between the average of the triplicate 'total RNA' qPCRs and each of the triplicate eluate qPCRs were calculated. These values were then expressed as fold difference in template abundance between eluate and total extract (ratio eluate/total) for *RPL5* and *RPL10* mRNAs. These values represent the relative enrichment of the RPL5 and RPL10 encoding mRNAs.

uL18 subcellular distribution by widefield and confocal microscopy analysis. Cells were fixed with 4% paraformaldehyde for 10 min. Cells were incubated overnight at 4°C with anti-uL18/RPL5 antibodies (Abcam, Ref. ab86863) at 1:200, and then one hour at room

temperature with chicken anti-rabbit IgG (H+L)/Alexa Fluor 488 secondary antibody (Invitrogen, Ref. A-21441) at 1:1000. Imaging was performed on a Zeiss Axio Observer.Z1 microscope with a motorized stage, driven by MetaMorph (MDS Analytical Technologies, Canada). Images were captured on an HQ2 camera in widefield mode (with CoolLED illumination), or in confocal mode using a Yokogawa spin-disk head and a laser bench from Roper (405 nm 100 mW Vortran, 491 nm 50 mW Cobolt Calypso and 561 nm 50 mW Cobolt Jive) and a 20x, or 40x objective (Plan NeoFluar, Zeiss).

Analysis of total RNA extracts by northern blot. Total RNAs were extracted with Trizol from cell pellets containing ~20 x 10^6 cells. The aqueous phase was extracted with phenol-chloroform-isoamylic alcohol (25:24:1; Sigma), then with chloroform. Total RNAs were recovered after precipitation with 2-propanol. For northern blot analyses, total RNAs were separated on a 1.2% agarose gel containing 1.2% formaldehyde and Tri/Tri buffer (30 mM triethanolamine, 30 mM tricine, pH 7.9) (3 µg/lane), prior to be transferred to a Hybond N^+ nylon membrane (GE Healthcare). Sequential hybridizations with the following 5'-radiolabeled oligonucleotide probes were performed as previously described: 5'ITS1 (5'-CCTCGCCCTCCGGGCTCCGTTAATGATC-3'), ITS1-5.8S (5'-CTAAGAGTCGTACGAGG TCG-3'), ITS2 (5'-CTGCGAGGGAACCCCCAGCCGCGCA-3' and 5'-GCGCGACGGC GGACGACACCGCGGCGTC-3'), 18S (5'-TTTACTTCCTCTAGATAGTCAAGTTCGACC-3'), 28S (5'-CCCGTTCCCTTGGCTGTGGTTTCGCTAGATA-3'). After exposure to an intensifying screen, signals were acquired with a Typhoon Trio PhosphoImager (GE Healthcare) and quantified using the MultiGauge software.

Polysome profiling analysis of LCLs. 600 µg of total protein from freshly lysed LCLs was loaded onto 10-50% sucrose gradients as previously described.[12] The tubes were centrifuged at 4°C and at 36,000 rpm for 2 hours in a SW41 rotor (Optima L100XP ultracentrifuge; Beckman Coulter). The gradient fractions were measured at OD254 using a syringe pump and UV detector (Brandel) and collected with a Foxy Jr. gradient collector (Teledyne Isco).

Western blot analysis of LCLs. LCLs were plated at 100,000 per well of a 6 well dish overnight. For p53 stabilization assays, camptothecin (100 nM in DMSO) (Sigma), MG-132 (1 µM in DMSO) (Sigma) or an equal volume of DMSO as a vehicle control was added to wells for 4-6 hours. For all assays cells were lysed in RIPA buffer supplemented with phosphatase and protease inhibitor cocktails (Sigma). 4-12% Bis-Tris polyacrylamide gels (Invitrogen) were run for detection of HEATR3 (Abcam # ab221085), p53 (Thermo Fisher #MA5-12557), uL18 (RPL5) (Abcam, #ab86863), uL5 (RPL11) (Santa Cruz #sc-25931), and actin (Sigma #A4700).

Quantitative PCR analysis. LCLs were plated six times at 100,000 cells per well of a six-well dish overnight and lysed the following day with Trizol for total RNA extraction. cDNA was generated from 1 µg total RNA with the QuantiTect® Reverse Transcription Kit (Qiagen) according to the manufacturer's specifications for a final volume of 14 µL (diluted 1:20 for qPCR reaction). Six cDNA samples were generated, one from each well of the six-well dish, and analyzed in tandem each with a technical duplicate. Per qPCR reaction, a 4 µL LC480 SYBR Green® I Master mix (Roche #4707516001) was added with 1 µL forward and reverse primer mix (final primer concentration 10µM of each), 2 µL diluted cDNA, and 1 µL H_20. Reactions were run on a Lightcycler 480 (Roche) and according to the manufacturer's instructions for SYBR Green® use (Roche #4707516001). Using a relative quantification method using the quantification cycle (Cq), healthy control LCL (NFT), the mean Reference Cq of three control genes, *ACTB*, *36B4*, and *GAPDH* was calculated for each of the six cDNA samples, a ratio was calculated compared to the mean Cq of *HEATR3* from the same cDNA samples, and normalized to one. Similar ratios were then calculated for a second healthy control LCL (NhnM) as well as both LCLs carrying the *HEATR3* variants and compared to the normalized ratio. Primers used for HEATR3 p.(Gly584Glu) were FW: 5'-TGCAAAGCAGGCATTCATAG-3' and REV: 5'- GAGGCTCTTTCGGCTTCTTT-3', for HEATR p.(Cys446Tyr) FW: 5'- TCTGCCTCTCCCATGAAGTT-3' and REV: 5'-ACATCCAGGAGGGACACAAG-3'. Pre-designed SYBR Green® primers for control genes were from Sigma.

Lentiviral constructs. Plasmid (pDONR223-HEATR3) containing HEATR3 open reading frame (ORF) from the human ORFEOME library[13] was generously provided by the Broad Institute Genetic Perturbation Platform (Boston, USA). The HEATR3 ORF was subcloned in the pTRIP lentiviral vector[14] downstream a CMV promoter. An internal ribosomal entry site followed by the ORF encoding eGFP (retrieved from pIRES2-EGFP, Clonetech-Takara, Japan) was subcloned in 3' of the HEATR3 ORF in order to construct a bicistronic gene. This pTRIP-HEATR3-IRES-eGFP plasmid as well as a pTRIP-GFP plasmid[15] were used together with plasmids pCMV-dR8.91 and pCMV-VSVG to produce lentiviral vector stocks by transient tri-transfection of HEK 293T cells.[16] The lentiviral titer was estimated by ddPCR to be ~ 4.10^6 TU/ml for both constructs. Patient and control LCLs were transduced with vector particles at a MOI 10. Briefly, $4x10^5$ cells cultured in OPTI-MEM (Gibco-Invitrogen) were transduced with 1mL aliquots of lentiparticles in the presence of 10µg/mL of protamine sulfate (Merck) overnight. Medium was changed the next day and cells were cultured until flow cytometry. Sorting of transduced cells was performed by FACS based on GFP fluorescence on a BD Influx cell sorter with the CellQuest Pro software (BD Biosciences FACS Software V1.2).

HEATR3 synthesis was confirmed by western blot using an anti-HEATR3 antibody from Abcam (Ref. ab221085).

REFERENCES

1. Ernst, R.F. & Elferink, M. UMCUGenetics/DxNextflowWES: v1.0.1 (v1.0.1). Zenodo (2020) https://doi.org/10.5281/zenodo.4552029 (2020).
2. McKenna, A. et al. The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 20, 1297–1303 (2010).
3. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv:1303.3997v2 [q-bio.GN] (2013).
4. Tarasov, A., Vilella, A. J., Cuppen, E., Nijman, I. J. & Prins, P. Sambamba: Fast processing of NGS alignment formats. Bioinformatics 31, 2032–2034 (2015).
5. Karczewski, K. J. et al. The mutational constraint spectrum quantified from variation in 141,456 humans. Nature 581, 434–443 (2020).
6. McQuillan, R. et al. Runs of homozygosity in European populations. Am J Hum Genet. 83, 359-372 (2008).
7. Görmez, Z. et al. HomSI: a homozygous stretch identifier from next-generation sequencing data. Bioinformatics 30, 445-447 (2014).
8. Bonneaud, N. et al. A family of low and high copy replicative, integrative and single-stranded S. cerevisiae/E. coli shuttle vectors. Yeast 7, 609–615 (1991).
9. Lafontaine, D. L. J., Preiss, T. & Tollervey, D. Yeast 18S rRNA Dimethylase Dim1p: a Quality Control Mechanism in Ribosome Synthesis? Mol. Cell. Biol. 18, 2360–2370 (1998).
10. Kressler, D. et al. Synchronizing nuclear import of ribosomal proteins with ribosome assembly. Science 338, 666–671 (2012).
11. Pausch, P. et al. Co-translational capturing of nascent ribosomal proteins by their dedicated chaperones. Nat. Commun. 6:7494 (2015).
12. Lezzerini, M. et al. Ribosomal protein gene RPL9 variants can differentially impair ribosome function and cellular metabolism. Nucleic Acids Res. 48, 770–787 (2020).
13. Yang, X. et al. A public genome-scale lentiviral expression library of human ORFs. Nat. Methods 8, 659–661 (2011).
14. Zennou, V. et al. The HIV-1 DNA flap stimulates HIV vector-mediated cell transduction in the brain. Nat. Biotechnol. 19, 446–450 (2001).
15. Pencalet, P. et al. Integration of genetically modified adult astrocytes into the lesioned rat spinal cord. J. Neurosci. Res. 83, 61–67 (2006).
16. Hantelys, F. et al. Vasohibin1, a new mouse cardiomyocyte IRES trans-acting factor that regulates translation in early hypoxia. Elife 8:e50094 (2019).

Family	**A**		**B**	**C**	**D**	
Individual	P1	P2	P3	P4	P5	P6
PHKG2	c.227G>A (p.Arg76Gln) homozygous	c.227G>A (p.Arg76Gln) homozygous	None	None	None	N/A
ZNF267	c.1072G>A (p.Val358Ile) homozygous	c.1072G>A (p.Val358Ile) homozygous	None	None	None	N/A
ABCC12	c.3224G>C (p.Arg1075Pro) homozygous	c.3224G>C (p.Arg1075Pro) homozygous	None	None	None	N/A
HEATR3	c.1751G>A (p.Gly584Glu) homozygous	c.1751G>A (p.Gly584Glu) homozygous	c.1337G>A (p.Cys446Tyr) homozygous	c.399+1G>T (p.?); c.719C>T (p.Pro249Leu) compound hetero.	c.400T>C (p.Cys134Arg) homozygous	c.400T>C (p.Cys134Arg) homozygous
PARP4	c.3667-1G>A (p.?); c.3295T>C (p.Cys1099Arg) compound hetero.	c.3667-1G>A (p.?); c.3295T>C (p.Cys1099Arg) compound hetero.	None	None	None	N/A

Table S1: Candidate variants in families A-D.

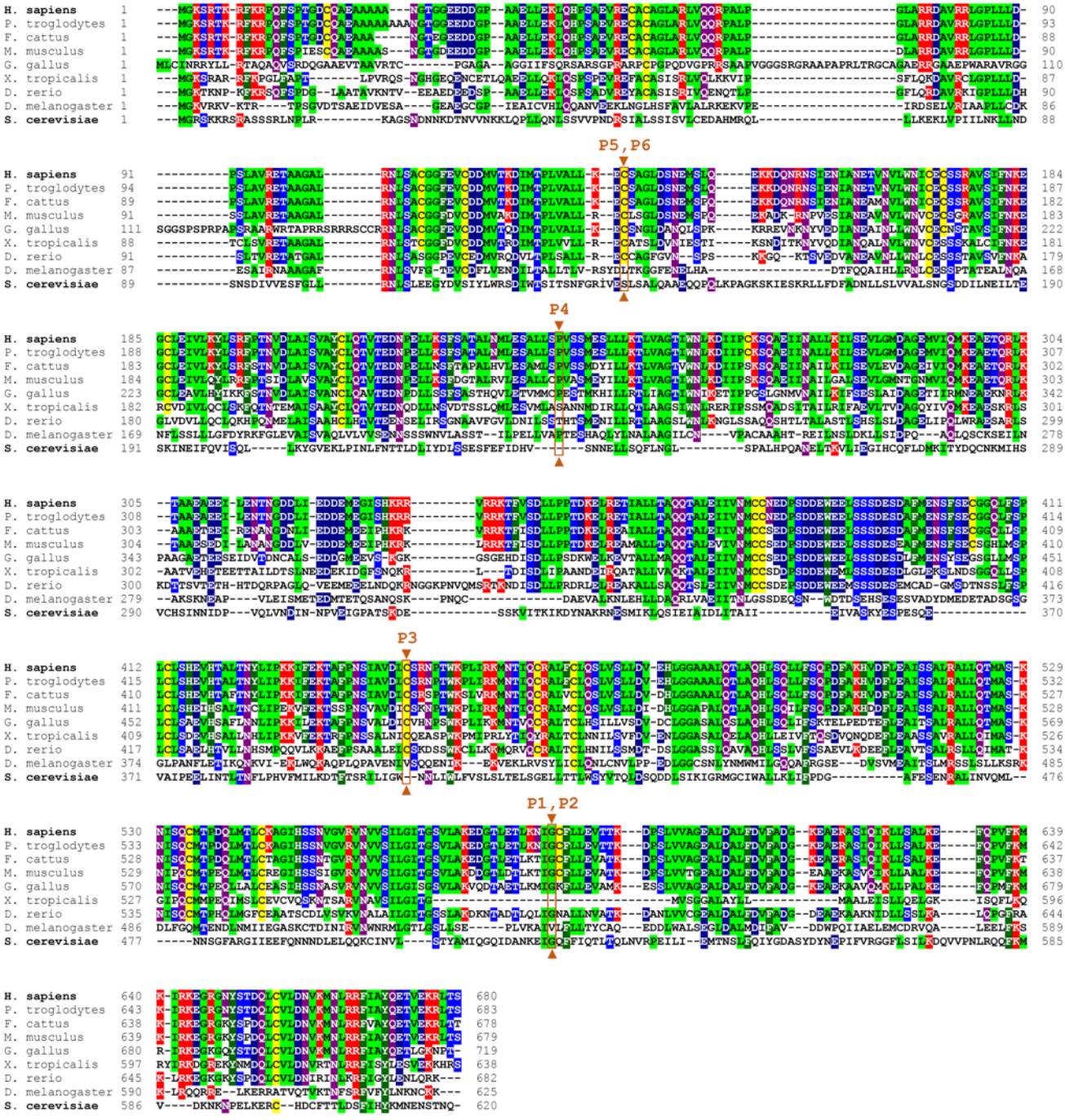

Figure S1. Multiple sequence alignment for human HEATR3 and orthologs

HEATR3 orthologs for *H. sapiens* (ENSP00000299192), *P. troglodytes* (ENSPTRP00000013827), *F. cattus* (ENSFCAP00000006688), *M. musculus* (ENSMUSP00000034079), *G. gallus* (ENSGALP00000066[45]6), *X. tropicalis* (ENSXETP00000015435), *D. rerio* (ENSDARP00000052367), *D. melanogaster* (FBpp0081214) and *S. cerevisiae* (YDL063C) were used to perform multiple sequence alignment using Constraint-based Multiple Alignment Tool (COBALT). The position of the modified amino acid in each affected individual (P1-P6) is boxed in red and labeled. Note that only the modifed Gly584 in P1 and 2 (Gly584Glu) has a homologous residue in the yeast Syo1 sequence, Gly522.

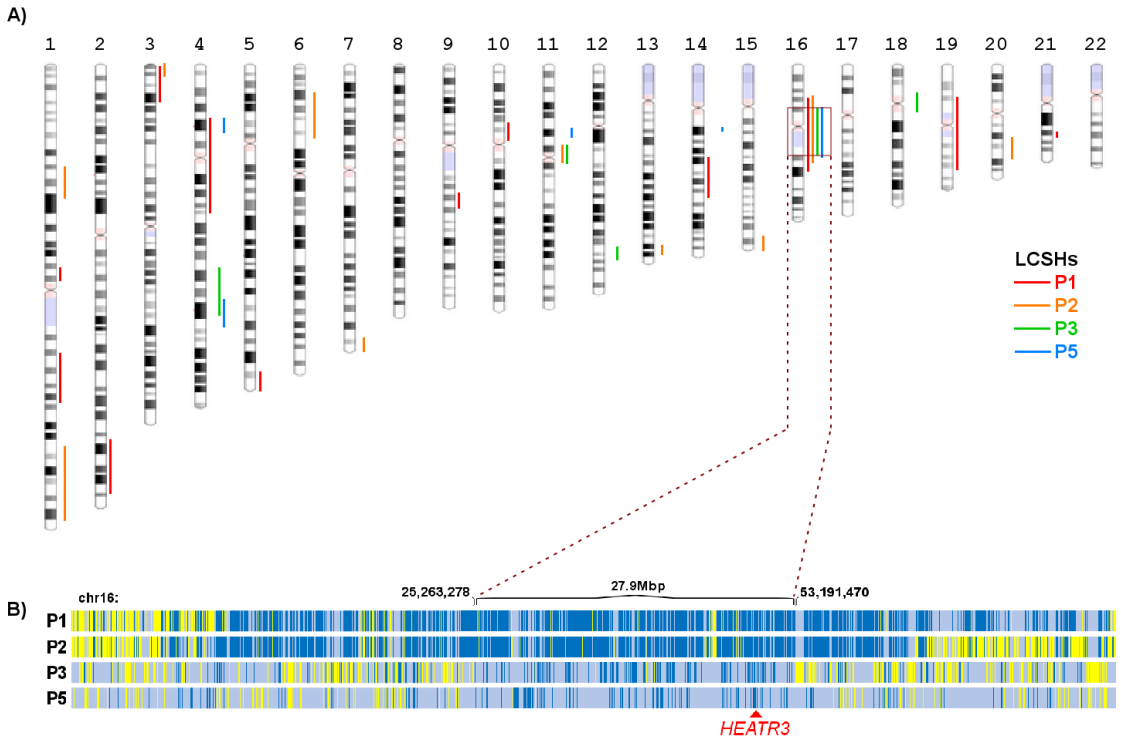

Figure S2. Identification of LCSHs in consanguineous families

A) Distribution of all LCSHs within all autosomes for individuals P1, P2, P3 and P5 are shown in the right side of each chromosomal ideogram. The colors associated with each individual are indicated. The brown box shows the single common LCSH region.
B) The only common LCSH region in all individuals is shown in detail. Each horizontal lane represents WES data from the indicated individual and colors of horizontal variants represent the zygosity of the variant: blue for homozygous alternative allele, yellow for heterozygous, and grey for either homozygous reference allele or uncovered regions. The red variants are the pathogenic homozygous HEATR3 variants. Note that [GRCh37]chr16:25,263,278-53,191,470 is a common LCSH region made up of exceedingly homozygous variants (blue and grey) for all individuals.

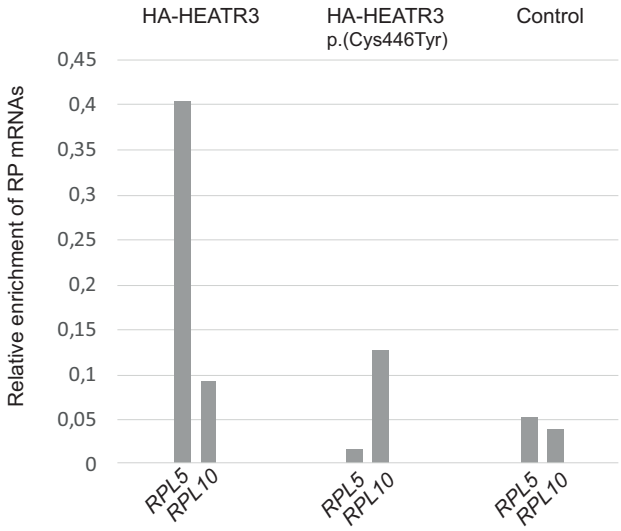

Figure S3. The co-translational capturing of RPL5 is evolutionarily conserved and impaired by a HEATR3 variant

Total extract of HeLa cells transfected with an HA-HEATR3, or an HA-HEATR3-p.(Cys446Tyr), or an empty plasmid (control), were used in ChIP assay. Ribosomes were stalled by brief cycloheximide treatment. The material coprecipitated following affinity purification with an anti-HA antibody was amplified by RT-qPCR with dedicated amplicons specific to the *RPL5* mRNA or, as control for specificity, to the *RPL10* mRNA. The data shown are a representative example of a triplicate.

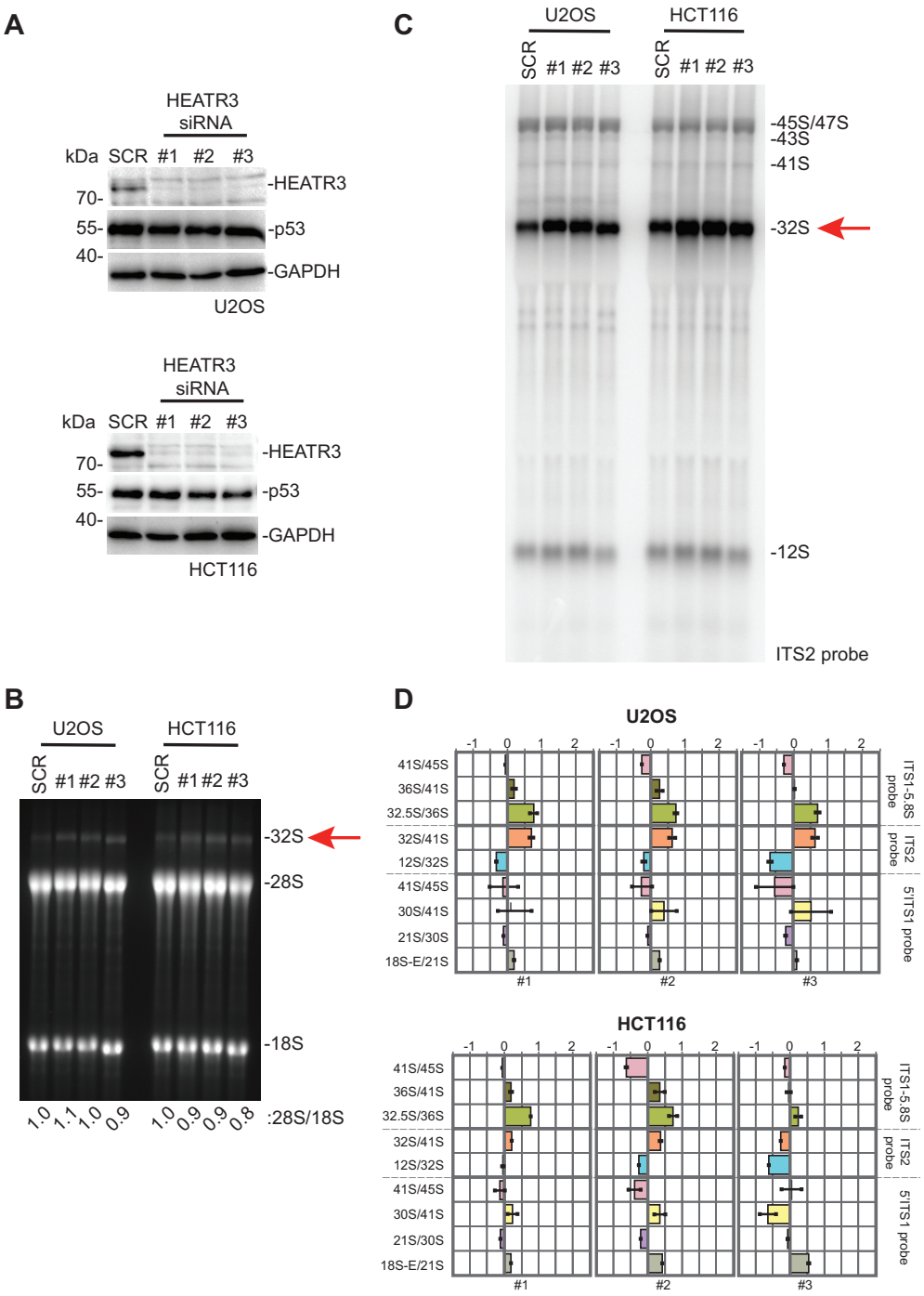

Figure S4. Knockdown of HEATR3 in cell lines impairs pre-rRNA processing without stabilizing p53

A) Western blots of U2OS and HCT116 cells transfected with three different siRNAs targeting *HEATR3* or a scrambled control siRNA (SCR) probed with antibodies against HEATR3, p53 and GAPDH.
B) Steady-state accumulation of mature rRNAs as revealed by ethidium bromide staining. The 28S/18S rRNA ratio was computed from Bioanalyzer electropherograms. Note that the 32S pre-rRNA accumulation is so important upon HEATR3 depletion that it is already visible by ethidium staining (read arrow).
C) Northern blots of RNAs from U2OS and HCT116 cells. Radiolabeled probes targeting ITS2 used to detect pre-rRNA precursors. Abnormal 32S accumulation is indicated by the red arrow.
D) Quantification of pre-rRNA processing by RAMP analysis (see Figure 4 for details). Mean values ± s.e.m. from three independent experiments.

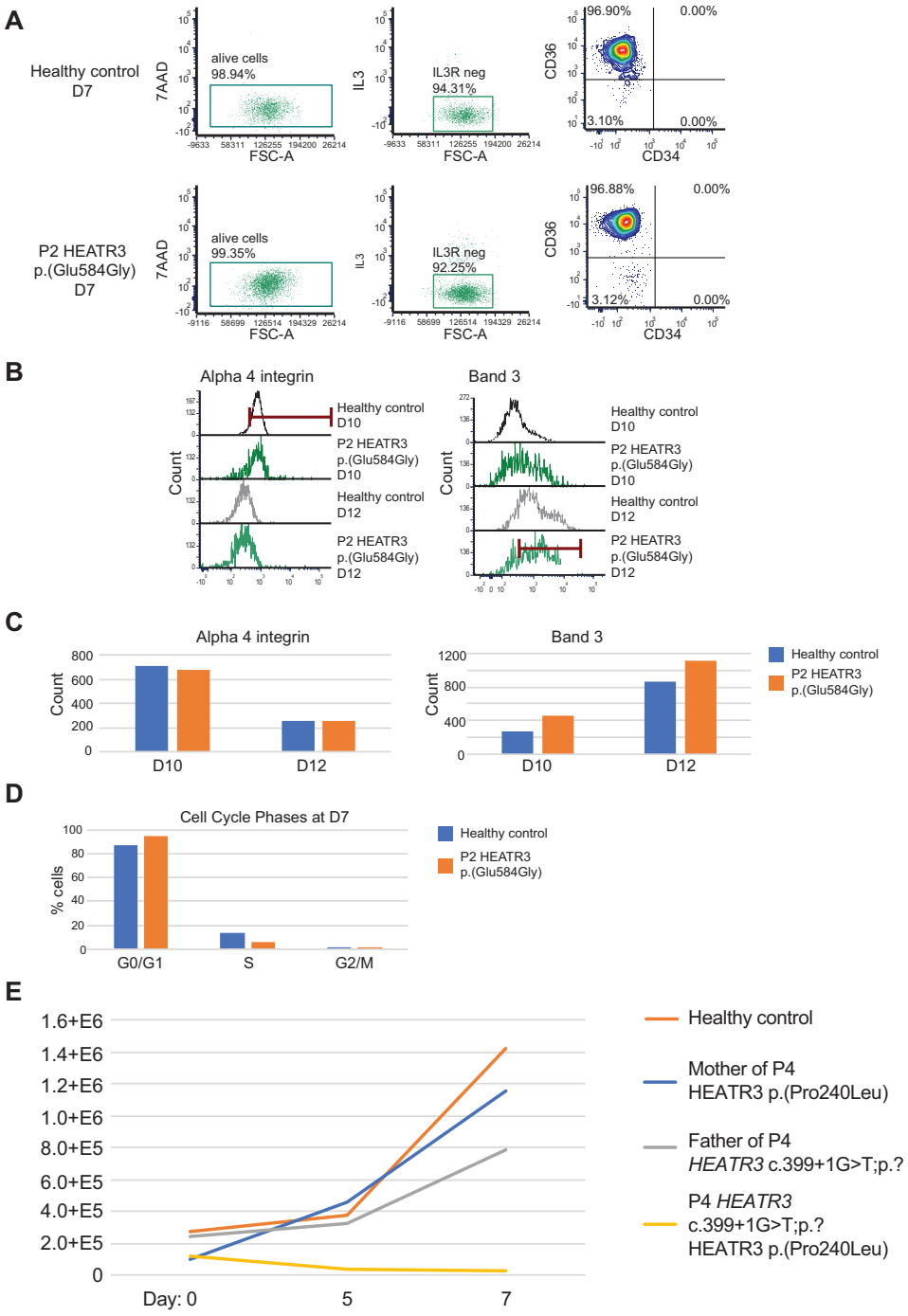

Figure S5. Erythroid cell culture assay initiated from peripheral blood-purified CD34+ cells

A) Flow cytometry at day 7 (D7) with erythroid cells from P2 [HEATR3 p.(Glu584Gly)] or from a healthy control in the erythroid cell culture assay using antibodies against IL3R, CD34, CD36, or stained with 7-AAD.
B) Histogram display of flow cytometry analysis of the same erythroid cell culture assay as in (A) at days 10 and 12 (D10, D12) with antibodies against alpha 4 integrin and Band 3.
C) Quantification of the results shown in (B).
D) Cell cycle analysis showing percentage of cells at day 7 (D7) in the G0/G1, S, or G2/M phases.
E) Cell counts of purified $CD34^{+}$ cells from the peripheral blood of P4 (yellow line), the mother of P4 (blue line), the father of P4 (grey line) and a healthy control (orange line) subjected to the erythroid culture assay. Shown are cell counts on days 0, 5, and 7.

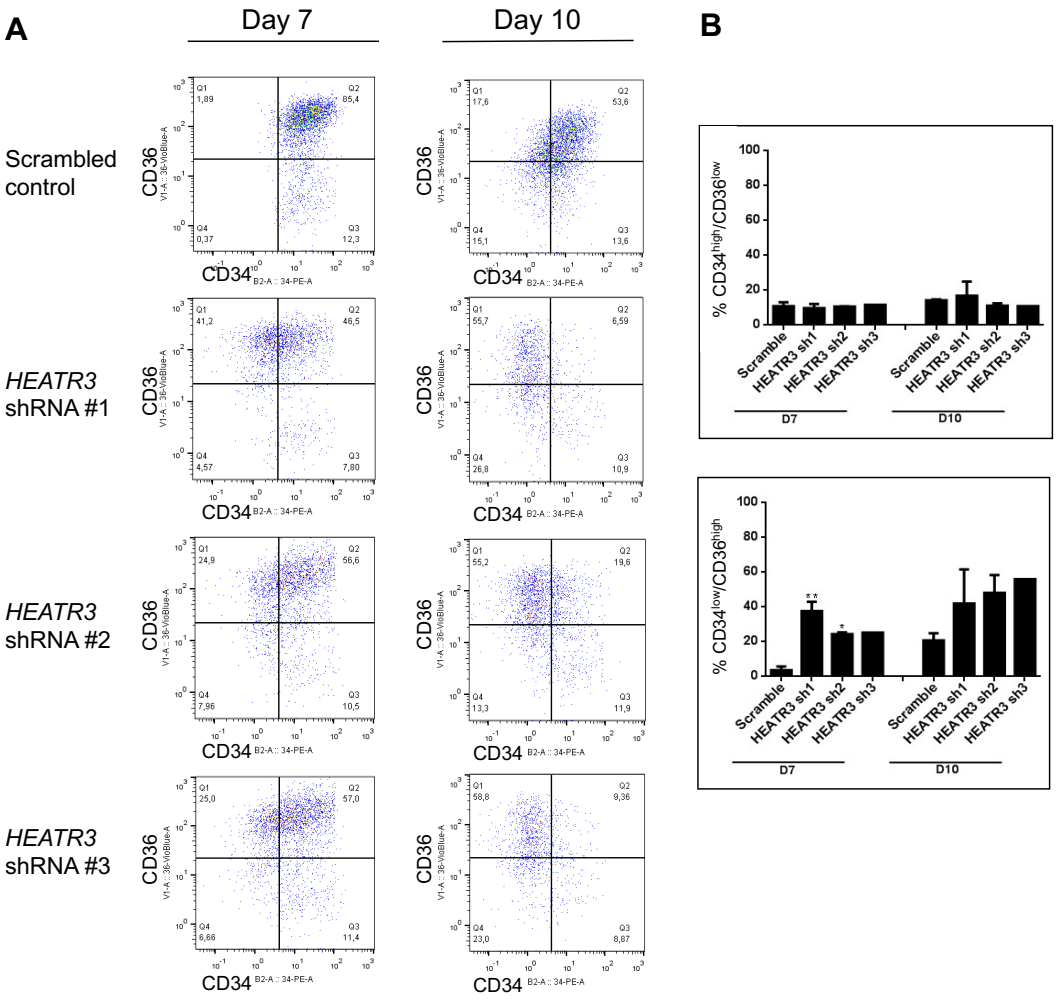

Figure S6. Knockdown of *HEAT3* in CD34⁺ cells impairs erythroid differentiation

A) FACS plots of CD34⁺ cord blood cells transduced with lentiviruses expressing either a scrambled shRNA or three independent shRNAs targeting *HEATR3* (*HEATR3* sh1-3) and subjected to the erythroid cell culture assay. Cells collected at days 7 and 10 were stained with antibodies against CD34 or CD36.
B) Quantification of the results shown in (A). Three independent experiments were performed with each siRNA, but only two of them were taken into account with shRNA3 as too few cells were available for analysis in the third experiment. * p<0.05, ** p<0.01.