

# Human cellular model systems of $\beta$ -thalassemia enable in-depth analysis of disease phenotype

Received: 27 August 2022

Accepted: 26 September 2023

Published online: 06 October 2023

 Check for updates

Deborah E. Daniels<sup>1</sup>, Ivan Ferrer-Vicens<sup>1</sup>, Joseph Hawksworth<sup>1</sup>, Tatyana N. Andrienko<sup>1</sup>, Elizabeth M. Finnie<sup>1</sup>, Natalie S. Bretherton<sup>1</sup>, Daniel C. J. Ferguson<sup>1</sup>, A. Sofia. F. Oliveira<sup>1</sup>, Jenn-Yeu A. Szeto<sup>1</sup>, Marieangela C. Wilson<sup>1</sup>, John N. Brewin<sup>2,3</sup> & Jan Frayne<sup>1</sup>✉

$\beta$ -thalassemia is a prevalent genetic disorder causing severe anemia due to defective erythropoiesis, with few treatment options. Studying the underlying molecular defects is impeded by paucity of suitable patient material. In this study we create human disease cellular model systems for  $\beta$ -thalassemia by gene editing the erythroid line BEL-A, which accurately recapitulate the phenotype of patient erythroid cells. We also develop a high throughput compatible fluorometric-based assay for evaluating severity of disease phenotype and utilize the assay to demonstrate that the lines respond appropriately to verified reagents. We next use the lines to perform extensive analysis of the altered molecular mechanisms in  $\beta$ -thalassemia erythroid cells, revealing upregulation of a wide range of biological pathways and processes along with potential novel targets for therapeutic investigation. Overall, the lines provide a sustainable supply of disease cells as research tools for identifying therapeutic targets and as screening platforms for new drugs and reagents.

$\beta$ -thalassemia syndromes are a heterogeneous range of anemias and a major source of morbidity, mortality, and substantial financial burden globally. The disease is characterized by variable reduced ( $\beta^+$ ) or absent ( $\beta^0$ )  $\beta$ -globin chain synthesis, with more than 300 different mutations in and around the  $\beta$  globin gene identified<sup>1</sup>. Mutations resulting in  $\beta^0$ -thalassemia cause a severe chronic anemia, whereas that resulting from  $\beta^+$  varies from clinically benign to severe.

Reduction in  $\beta$ -globin results in excess  $\alpha$ -globin, forming insoluble aggregates which undergo auto-oxidation leading to increased reactive oxygen species (ROS) and a range of intracellular downstream events that result in ineffective erythropoiesis (IE), the hallmark and primary cause of  $\beta$ -thalassemia disease pathophysiology. IE is manifest as increased expansion of erythroid progenitors due to increased EPO production in vivo, but with a maturation block at the polychromatic stage of differentiation<sup>2,3</sup> with increased apoptosis (reviewed by Ribeil et al.<sup>4</sup>), and thus impaired erythropoiesis.

However, the underlying molecular mechanisms that cause IE are not fully understood.

In recent years there has been much renewed interest in studying  $\beta$ -thalassemia, however there are still few treatment options. The mainstay is blood transfusion, but this has associated complications of alloimmunisation and iron-mediated potentially fatal organ damage, requiring life-long iron-chelation therapy.

Non-curative approaches have aimed to correct globin imbalance by promoting  $\gamma$ -globin synthesis, such as with hydroxyurea<sup>5</sup>, but responses are inconsistent<sup>6–8</sup>. More recently, luspatercept has been shown to increase hemoglobin values, but modestly by 1–2 g/dl in non-transfusion dependent thalassemia patients (NTDT). However, only 20% of transfusion-dependent (TDT) patients showed a reduction of blood transfusion by more than a third<sup>9</sup>. The only definitive cure is bone marrow transplant, best applicable for some young patients with a matched sibling lacking TDT<sup>5</sup>. Gene therapy, either using gene

<sup>1</sup>School of Biochemistry, University of Bristol, Bristol BS8 1TD, UK. <sup>2</sup>Haematology Department, King's college Hospital NHS Foundation, London SE5 9RS, UK.

<sup>3</sup>Red Cell Biology Group, Kings College London, London SE5 9NU, UK. ✉e-mail: [Jan.Frayne@Bristol.ac.uk](mailto:Jan.Frayne@Bristol.ac.uk)

replacement or editing with associated bone marrow transplantation holds much promise for those lacking a matched donor<sup>10</sup>. However, efficacy and safety issues are still concerns and given the complexity and high cost of these approaches, they are unlikely to scale to the massive global unmet need for  $\beta$ -thalassaemia therapeutics, particularly in low to middle income countries<sup>11</sup>. Thus, new cost-effective treatment strategies and drug targets are desperately required to deliver optimal therapies to the greatest numbers of people.

However, studying the molecular defects underlying the  $\beta$ -thalassaemia phenotype, and identifying drug targets, is severely impeded by paucity of suitable material from patients, and lack of suitable cell lines. Although erythroid cells can be generated in vitro from peripheral blood stem cells, the approach is severely limited by the restricted expansion potential of the cells<sup>12</sup> and thus number of erythroid cells generated, with repeat collections required, a particularly unsuitable approach for anemic patients. Thus, currently much of our understanding is based on mouse models which are also used to investigate and evaluate the effect of drugs. However, fundamental differences between mouse and human erythropoiesis are becoming increasingly apparent<sup>13,14</sup>. Therefore, new cellular models for  $\beta$ -thalassaemia are essential to enable in-depth investigation of underlying molecular mechanisms, to aid identification of new therapeutic targets and as screening platforms for the effect and efficacy of potential new drugs and reagents.

In this study we created and validated human disease erythroid cell lines for  $\beta$ -thalassaemia. Homozygous ( $HBB^{-/-}$ ) and heterozygous ( $HBB^{+/-}$ )  $HBB$  gene knockout mutations, and the prevalent  $\beta$ -thalassaemia mutations CD41/42 -TTCT and IVS-1-1 G  $\rightarrow$  T, were introduced into the erythroid cell line BEL-A, which undergoes normal erythropoiesis and expresses a normal adult globin profile<sup>15</sup>, via CRISPR genome editing. All disease lines accurately recapitulate the disease phenotype of  $\beta^0$ -thalassaemia patient erythroid cells, whilst the  $HBB^{+/-}$  lines recapitulate cells of individuals with  $\beta$ -thalassaemia trait. These lines provide a sustainable and consistent supply of disease cells as unique research tools, and as drug screening platforms. For the latter we also developed a high throughput compatible flow cytometry assay for evaluation of severity of IE. We utilized the assay to demonstrate that the  $HBB^{-/-}$  line responds appropriately to drugs known to increase  $\gamma$ -globin (hydroxyurea, pomalidomide and decitabine) and to  $BCL11A$  + 58 enhancer editing, with a correlating improvement to IE, validating the line for screening applications.

We also performed TMT-based comparative proteomic analysis. This confirmed the same profile of proteins previously reported dysregulated in  $\beta$ -thalassaemia erythroid cells, providing further validation, whilst also providing more detailed insights on the altered molecular mechanisms in  $\beta$ -thalassaemia erythroid cells. This included increased levels of antioxidant response enzymes, heat shock proteins and chaperones, and many proteins of the ubiquitin-proteasome and autophagy pathways along with proteins associated with lysosome biogenesis, together with up-regulation of respective pathways. The data also revealed upregulated pathways and increased levels of proteins for heme biosynthesis, along with associated upregulated endosome biogenesis and trafficking pathways, which may contribute to the disease phenotype. In addition, interrogation of the dataset revealed potential new therapeutic targeting opportunities.

Finally, we also describe a novel  $\beta$ -globin splice variant arising from the IVS-1-1 G  $\rightarrow$  T mutation which is unable to form stable tetramers with  $\alpha$ -globin, the variant instead precipitating and potentially contributing to the IE phenotype in IVS-1-1 erythroid cells.

## Results

The erythroid cell line BEL-A recapitulates normal adult erythroid cell differentiation and has a normal adult globin expression profile<sup>15</sup>. It is therefore an excellent founder line to introduce mutations to create

cellular model systems of red blood cell diseases, and so was used in this study to create  $\beta$ -thalassaemia disease lines.

### Generating $HBB^{+/-}$ and $HBB^{-/-}$ erythroid progenitor cell lines

As  $\beta$ -thalassaemia is caused by a wide range of mutations that inhibit or reduce the production of  $\beta$ -globin, in the first instance homozygous ( $HBB^{-/-}$ ) and control heterozygous ( $HBB^{+/-}$ )  $\beta$ -globin knockout lines were created using CRISPR-Cas9 mediated genome editing of BEL-A. This resulted in biallelic 1 base pair and monoallelic 2 base pair deletions for the  $HBB^{-/-}$  and  $HBB^{+/-}$  clonal lines respectively (Fig. 1a), leading to frameshift and premature stop codons in exon 1 in both cases.

The resultant  $HBB^{+/-}$  and  $HBB^{-/-}$  lines had a predominantly proerythroblast morphology with some basophilic erythroblasts, in line with the WT BEL-A line (Fig. 1b). Doubling times and viability were invariant across the lines (Fig. 1c and d).

Analysis of globin levels by RP-HPLC showed that  $HBB^{+/-}$  cells produced similar levels of all globin subunits to WT cells whereas, as expected,  $\beta$ -globin was eliminated in the  $HBB^{-/-}$  line (Fig. 1e; see also Supplementary Fig. 1 for clarification of residual peak overlapping position of  $\beta$ -globin).  $\alpha$ -globin was also reduced in the  $HBB^{-/-}$  cells as, in absence of  $\beta$ - (or  $\gamma$ -) globin it forms insoluble aggregates (Fig. 1f), a feature of  $\beta$ -thalassaemia patient erythroid cells.

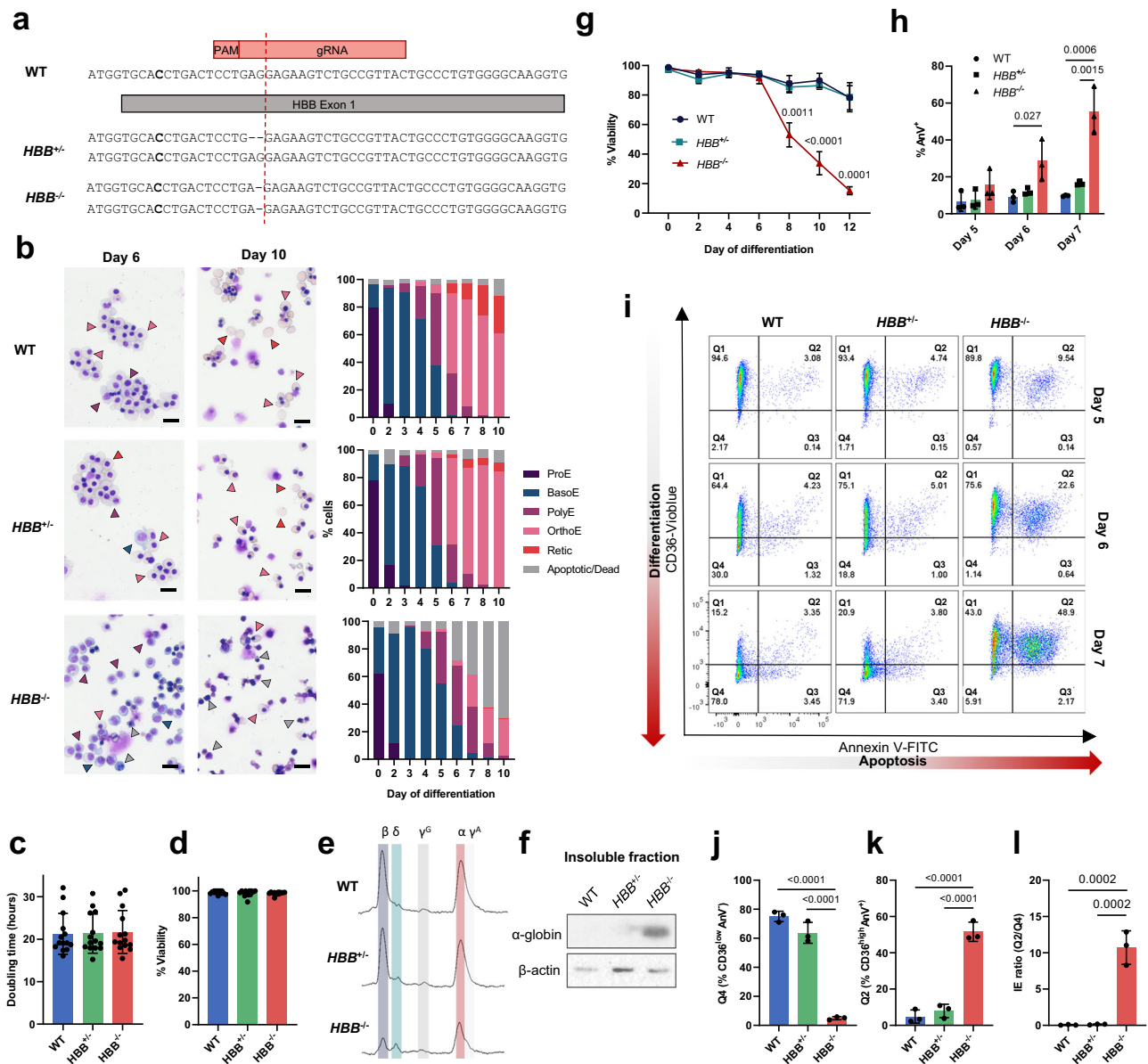
When differentiated, viability of the WT and  $HBB^{+/-}$  line remained high throughout the culture duration (Fig. 1g). In contrast the  $HBB^{-/-}$  line recapitulated the phenotype of  $\beta^0$ -thalassaemia erythroid cells<sup>2</sup> with clear ineffective erythropoiesis, as demonstrated by a dramatic and significant decline in viability from day 6 of differentiation (Fig. 1g), corresponding temporally with differentiation to polychromatic erythroblasts (Fig. 1b), and with a significant increase in Annexin V (AnV) cell surface abundance (Fig. 1h), indicative of apoptosis (Fig. 1b). In line, growth curves during differentiation showed unchanged expansion in  $HBB^{+/-}$  and  $HBB^{-/-}$  lines compared to WT up until day 6, but with greater reduction in cell numbers post-day 6 in  $HBB^{-/-}$  line due to cell death (Supplementary Fig. 2). Similarly, kinetics of the  $HBB^{+/-}$  cultures were invariant to that of WT cells during differentiation (Fig. 1b). In contrast, although the majority of  $HBB^{-/-}$  cells differentiated to polychromatic erythroblasts, few differentiated further to orthochromatic erythroblasts (Fig. 1b), as found for cultured patient erythroid cells<sup>2</sup>.

Analysis of a second  $HBB^{-/-}$  clonal line ( $HBB^{-/-}$ C2) displayed the same phenotype, with similar decrease in viability from day 6 of culture and associated significant increase in AnV levels and apoptosis at the polychromatic stage (Supplementary Fig. 3A–C).

Overall, the data confirm the  $HBB^{+/-}$  lines as robust cellular models of  $\beta^0$ -thalassaemia, providing a sustainable and consistent supply of disease cells.

### Assay for assessing severity of ineffective erythropoiesis

To evaluate drugs and reagents that reduce severity of IE, thus improving disease phenotype, the WT,  $HBB^{+/-}$  and  $HBB^{-/-}$  lines were used to establish a flow cytometric assay for IE utilizing fluorophore tagged antibodies to membrane markers. For this, the abundance of membrane protein CD36, which decreases during erythroid cell differentiation<sup>16</sup>, was first evaluated. CD36 levels decreased from day 6 of differentiation in WT and  $HBB^{+/-}$  cultures, but was retained in  $HBB^{-/-}$  cultures, correlating with arrested differentiation of  $HBB^{-/-}$  cells at the polychromatic stage (Supplementary Fig. 4), and providing effective resolution of transition from polychromatic to orthochromatic erythroblasts. Dual staining of CD36 with Annexin V then provided a robust and high throughput compatible assay for monitoring IE, as illustrated in Fig. 1i. To clarify, as differentiation proceeds WT and  $HBB^{+/-}$  cells show progressively decreasing levels of CD36 and consistently low levels of AnV, with the majority of cells by day 7 in quadrant 4, which delineates viable orthochromatic erythroblasts (CD36<sup>low</sup>AnV<sup>-</sup>; Fig. 1i and j). In contrast very few  $HBB^{-/-}$  cells proceed to quadrant 4, instead shifting from quadrant 1 to 2 (CD36<sup>high</sup>AnV<sup>+</sup>;



**Fig. 1 | Generation and characterization of  $HBB^{-/-}$  and  $HBB^{+/-}$  erythroid progenitor cell lines. a** Schematic of DNA sequence for  $HBB^{+/-}$  and  $HBB^{-/-}$  compared to WT BEL-A cells. **b** Representative cytopsin of WT,  $HBB^{+/-}$  and  $HBB^{-/-}$  BEL-A with morphology quantification (ProE, proerythroblast; BasoE, basophilic erythroblast; PolyE, polychromatic erythroblast; OrthoE, orthochromatic erythroblast; Retic, reticulocyte). Arrowheads indicate the following cell types: blue, BasoE; purple, PolyE; pink, OrthoE; red, Retic; gray, dead/apoptotic. Scale bars 20  $\mu$ m. Doubling time (c) and % viability (d) of expanding cell lines as determined by trypan blue exclusion assay,  $n = 14$  independent time points. **e** RP-HPLC traces for WT,  $HBB^{+/-}$  and  $HBB^{-/-}$  BEL-A at day 6 of differentiation. Peaks are identified for  $\beta$ -globin ( $\beta$ ),  $\delta$ -globin ( $\delta$ ),  $\gamma$ -globin ( $\gamma$ ),  $\alpha$ -globin ( $\alpha$ ) and  $\gamma^H$ -globin ( $\gamma^H$ ). **f** Representative western

blot of aggregate proteins from WT,  $HBB^{+/-}$  and  $HBB^{-/-}$  BEL-A harvested at day 6 of differentiation, incubated with  $\alpha$ -globin antibody.  $\beta$ -actin was used as a protein loading control. **g** Percentage viability of WT,  $HBB^{+/-}$  and  $HBB^{-/-}$  BEL-A during differentiation by trypan blue exclusion assay. **h** Percentage of Annexin V positive cells (AnV<sup>+</sup>) in WT,  $HBB^{+/-}$  and  $HBB^{-/-}$  BEL-A during differentiation as determined by flow cytometry. **i** Representative IE flow cytometry plots at day 5, 6 and 7 of differentiation. Quantification of Q4 CD36<sup>low</sup> AnV<sup>-</sup> cells (j), Q2 CD36<sup>high</sup> AnV<sup>+</sup> (k) and IE ratio (Q2/Q4) (l) at day 7 of differentiation. Results show mean  $\pm$  SD,  $n = 3$  independent experiments unless otherwise stated. Statistical significance was determined using one-way ANOVA with Tukey multiple comparison testing. Source data are provided as a Source Data file.

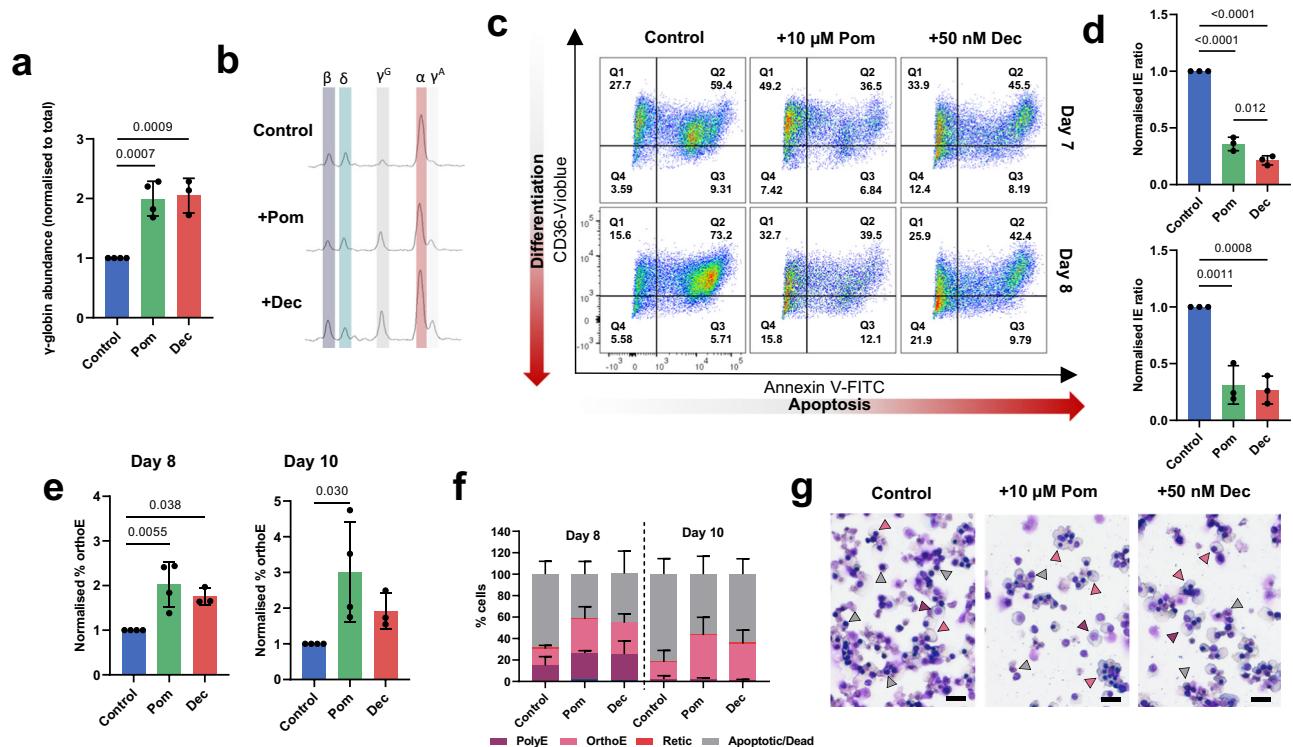
Fig. 1i and k) as levels of apoptosis increase with cells reaching the polychromatic stage. The number of cells in Q2 (CD36<sup>high</sup> AnV<sup>+</sup>):Q4 (CD36<sup>low</sup> AnV<sup>-</sup>) provides an IE ratio and numerical value for the severity of IE (Fig. 1l). The second  $HBB^{-/-}$  clone showed similar profile and IE ratio in the assay (Supplementary Fig. 3D–F).

### Validating $\beta^0$ -thalassaemia line as a drug and reagent screening platform

To test the potential of the  $HBB^{-/-}$  line as a screening platform for drugs and therapeutic approaches for reducing the disease severity of

$\beta$ -thalassaemia, we employed two established approaches that increase the level of  $\gamma$ -globin, and thus fetal hemoglobin (HbF) in erythroid cells. Upregulation of  $\gamma$ -globin was selected as our gauge, as it is the most promising approach for the treatment of  $\beta$ -thalassaemia, and also sickle cell disease (SCD), compensating for loss of, or mutated,  $\beta$ -globin.

Pomalidomide and Decitabine have both been reported to increase  $\gamma$ -globin expression in human primary erythroid cells<sup>17,18</sup>. The  $HBB^{-/-}$  cells responded accordingly to both drugs, with significant increase in  $\gamma$ -globin levels (by RP-HPLC, Fig. 2a and b). To equate with



**Fig. 2 | *HBB*<sup>-/-</sup> cells respond to pomalidomide and decitabine treatment with increased  $\gamma$ -globin expression and reduced ineffective erythropoiesis.** Differentiating *HBB*<sup>-/-</sup> cells were treated with 10  $\mu$ M pomalidomide (Pom), 50 nM decitabine (Dec) or vehicle only control (0.1% DMSO) and harvested for RP-HPLC analysis and IE analysis by flow cytometry. **a** Quantification of  $\gamma$ -globin levels ( $^c\gamma$ -globin +  $^A\gamma$ -globin) at day 6 measured by RP-HPLC, normalized to total globin levels and shown relative to control samples. **b** Representative RP-HPLC traces. Peaks are identified for  $\beta$ -globin ( $\beta$ ),  $\delta$ -globin ( $\delta$ ),  $^c\gamma$ -globin ( $^c\gamma$ ),  $\alpha$ -globin ( $\alpha$ ) and  $^A\gamma$ -globin ( $^A\gamma$ ). **c** Representative IE flow cytometry plots at day 7 and 8 of differentiation. **d** Quantification of IE ratio (%CD36<sup>high</sup>AnV<sup>+</sup>/CD36<sup>low</sup>AnV<sup>-</sup>). **e** Percentage of

orthochromatic erythroblasts in Pom and Dec treated cultures normalized to Control cultures at day 8 and 10 of differentiation. **f** Morphology analysis of Pom and Dec treated *HBB*<sup>-/-</sup> cells at day 8 and 10 of differentiation. **g** Representative cytospin images of Pom and Dec treated *HBB*<sup>-/-</sup> cells at day 10 of differentiation. Arrowheads indicate the following cell types: purple, PolyE; pink, OrthoE; red, Retic; gray, dead/apoptotic. Scale bars 20  $\mu$ m. Results show mean  $\pm$  SD,  $n = 4$  (control and Pom)  $n = 3$  (Dec) independent experiments. Statistical significance was determined using one-way ANOVA with Tukey multiple comparison testing. Source data are provided as a Source Data file.

effect on disease severity, the *HBB*<sup>-/-</sup> cells were evaluated by our IE assay, which revealed a concurrent decrease in IE ratio (Fig. 2c and d), and increased number of cells terminally differentiating (Fig. 2e–g). Similar results were seen with hydroxyurea (HU) treatment, a drug approved for the treatment of non-transfusion dependent  $\beta$ -thalassaemia and SCD, with increase in  $\gamma$ -globin levels and decrease in IE ratio paralleling increasing concentrations of HU (Supplementary Fig. 5).

We also used CRISPR-Cas9 gene editing to disrupt the +58 enhancer of the  $\gamma$ -globin repressor *BCL11A* in *HBB*<sup>-/-</sup> cells, using a previously validated sgRNA<sup>19</sup>, an approach granted PRIME designation for TDT and SCD, with planned regulatory submission this year. A clonal population of +58 enhancer edited *HBB*<sup>-/-</sup> cells containing a compound heterozygous deletion (Supplementary Fig. 6A) showed an ~25% increase in  $\gamma$ -globin levels (Supplementary Fig. 6B and C) in line with previous findings<sup>20</sup>, and a decrease in the severity of IE, as indicated by an ~50% decrease in the IE ratio at day 7 of differentiation and an increase in cell viability at day 8 and 10 of culture (Supplementary Fig. 6D–F). Improvement in erythroid differentiation was also observed, with a significant increase in the percentage of cells progressing to the orthochromatic stage at day 10 (Supplementary Fig. 6G–I).

These findings confirm that the *HBB*<sup>-/-</sup> disease line responds effectively to both pharmacological and genome editing approaches to increase  $\gamma$ -globin expression, whilst also directly demonstrating that the treatments reduce IE and thus the severity of  $\beta^0$ -thalassaemia

disease phenotype, collectively providing proof-of-principle for utilizing the line for screening applications.

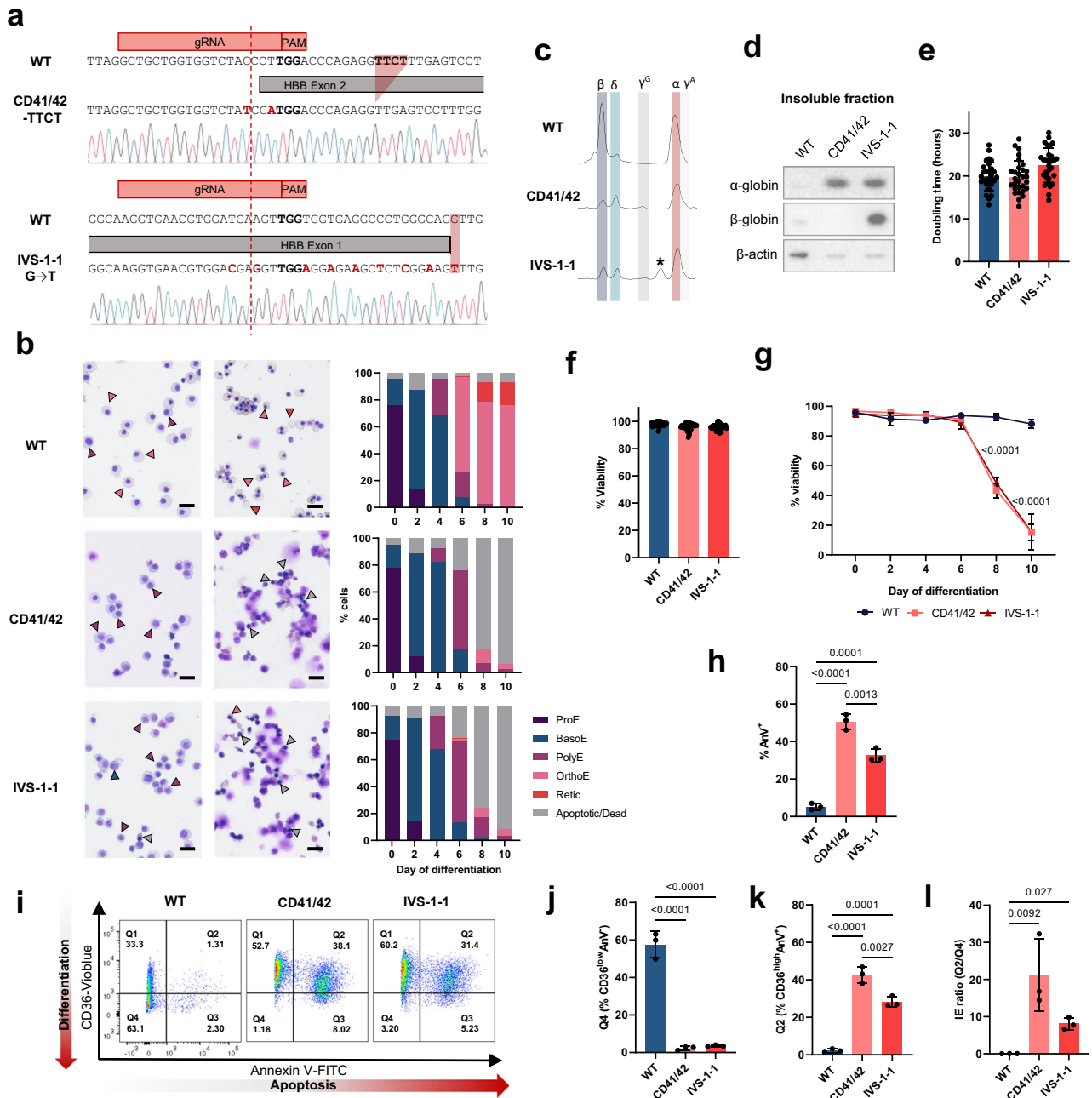
### Generating $\beta^0$ -thalassaemia cell lines with specific prevalent mutations

While the *HBB*<sup>-/-</sup> lines do successfully recapitulate the expected phenotype of  $\beta^0$ -thalassaemia erythroid cells, the deletion mutation is not one of over 300 different mutations that cause reduced or absent  $\beta$ -globin production in  $\beta$ -thalassaemia patients<sup>21</sup>. Therefore, we next selected two prevalent  $\beta^0$ -thalassaemia mutations to introduce into BEL-A, to more accurately replicate patient genotypes and corroborate data from *HBB*<sup>-/-</sup> lines; CD41/42 -TTCT deletion which results in a frameshift and introduction of a premature stop codon and IVS-1-1 G  $\rightarrow$  T substitution which prevents correct splicing of *HBB* mRNA<sup>22</sup>.

Both homozygous mutations were successfully introduced into BEL-A using RNP-based CRISPR-Cas9 genome editing with a single-strand oligodeoxynucleotide (ssODN) donor template (Fig. 3a). In the case of IVS-1-1, due to the relatively large 24 base pair distance from the cleavage site of the closest possible *HBB*-specific sgRNA and the intended edit, the ssODN was recoded with silent mutations, which has been shown to enhance incorporation efficiency of distal edits<sup>23</sup>.

As for the WT line, both disease lines had a predominantly proerythroblast morphology with some basophilic erythroblasts (Fig. 3b).

Similar to the *HBB*<sup>-/-</sup> line,  $\beta$ -globin protein was eliminated in both the CD41/42 and IVS-1-1 clonal lines (Fig. 3c), and  $\alpha$ -globin reduced due to insoluble aggregation (Fig. 3c and d), with doubling times and



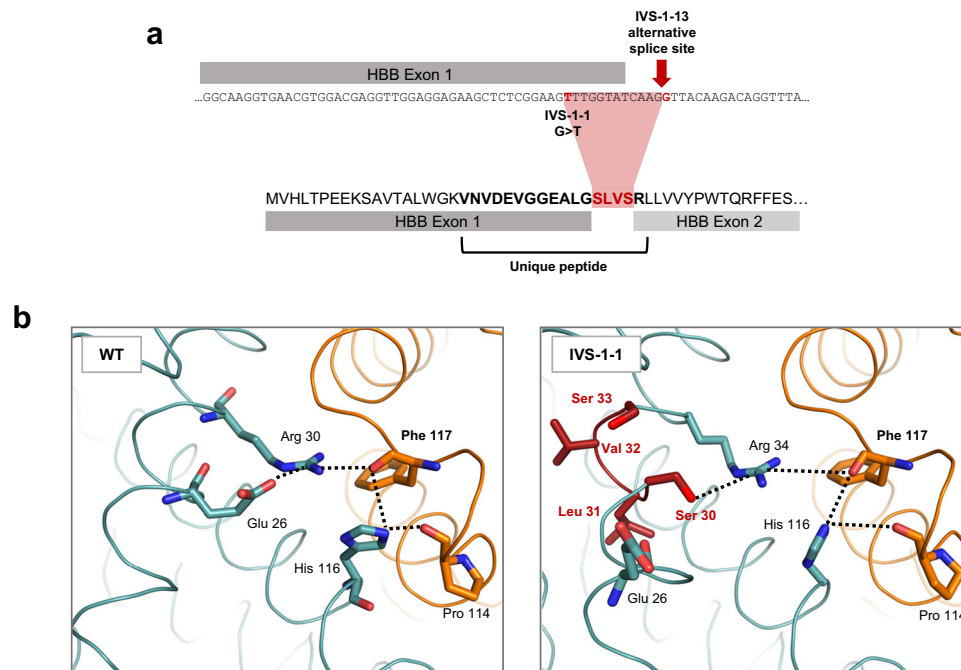
**Fig. 3 | Generation and characterization  $\beta^0$ -thalassemia cell lines with specific prevalent mutations.** **a** Schematic of DNA sequence for CD41/42-TTCT and IVS-1-1 G → T compared to WT BEL-A cells. **b** Representative cytopins of WT, CD41/42, and IVS-1-1 BEL-A with morphology quantification (ProE, proerythroblast; BasoE, basophilic erythroblast; PolyE, polychromatic erythroblast; OrthoE, orthochromatic erythroblast; Retic, reticulocyte). Arrowheads indicate the following cell types: blue, BasoE; purple, PolyE; pink, OrthoE; red, Retic; gray, dead/apoptotic. Scale bars 20  $\mu$ m. **c** RP-HPLC traces for WT, CD41/42, and IVS-1-1 BEL-A at day 6 of differentiation. Peaks are identified for  $\beta$ -globin ( $\beta$ ),  $\delta$ -globin ( $\delta$ ),  $\gamma$ -globin ( $\gamma$ ),  $\alpha$ -globin ( $\alpha$ ) and  $\gamma^*$ -globin ( $\gamma^*$ ). \* Denotes splice  $\beta$ -globin variant (see Supplementary Fig. 7). **d** Representative western blot of aggregate proteins from WT, CD41/42, and IVS-1-1 BEL-A harvested at day 6 of differentiation, incubated with  $\alpha$ -globin or  $\beta$ -globin

antibody.  $\beta$ -actin was used as a protein loading control. Doubling time (**e**) and % viability (**f**) of expanding cell lines as determined by trypan blue exclusion assay,  $n = 32$  independent time points (**g**) Percentage viability of WT, CD41/42, and IVS-1-1 BEL-A during differentiation by trypan blue exclusion assay. **h** Percentage of Annexin V positive (AnV<sup>+</sup>) cells in WT, CD41/42, and IVS-1-1 BEL-A at day 7 of differentiation as determined by flow cytometry. **i** Representative IE flow cytometry plots at day 7 of differentiation. Quantification Q4 CD36<sup>low</sup>AnV<sup>+</sup> cells (**j**), Q2 CD36<sup>high</sup>AnV<sup>+</sup> (**k**) and IE ratio (Q2/Q4) (**l**) at day 7 of differentiation. Results show mean  $\pm$  SD,  $n = 3$  independent experiments unless otherwise stated. Statistical significance was determined using one-way ANOVA with Tukey multiple comparison testing. Source data are provided as a Source Data file.

viability during the expansion phase not significantly different (Fig. 3e and f).

As for the *HBB*<sup>-/-</sup> lines, both the CD41/42 and IVS-1-1 lines demonstrated IE, with significantly reduced cell viability at the polychromatic cell stage (from day 6; Fig. 3b and g), due to significantly

increased apoptosis compared to control cells (Fig. 3h), with very few cells surviving and differentiating to orthochromatic erythroblasts (Fig. 3g). Analysis using the flow cytometry IE assay showed similar profiles to the *HBB*<sup>-/-</sup> line, with very few CD41/42 and IVS-1-1 cells shifted to quadrant 4 by day 7 (CD36<sup>low</sup>AnV<sup>+</sup>; Fig. 3i and j), instead



**Fig. 4 | Identification of IVS-1-1  $\beta$ -globin variant.** **a** Schematic of the IVS-1-1 variant amino acid sequence resulting from use of the IVS-1-13 alternative splice site. The resulting peptide sequence contains a 4 amino acid insertion (red) from 12 base pairs of translated *HBB* intron. **b** Models of the  $\alpha$ - $\beta$ -globin interface ( $\alpha$ -globin in

orange and  $\beta$ -globin in teal) showing increased distances between residues contributing to interactions with Phe I17 (dashed lines) on the  $\alpha$ -globin chain in IVS1-1 compared to WT dimers. Inserted amino acids in the IVS-1-1 variant are shown in dark red.

having shifted to quadrant 2 (CD36<sup>high</sup>AnV<sup>+</sup>; Fig. 3i and k), due to levels of apoptosis increasing when cells reached the polychromatic stage, giving significantly increased IE ratios compared to WT cells (Fig. 3l).

The overall karyotypes of the three  $\beta$ -thalassaemia lines (*HBB*<sup>-/-</sup>, CD41/42 and IVS-1-1) are similar to that of the original BEL-A line<sup>16</sup>, with trisomy of chromosome 6 and 19 and a small number of partial chromosome losses/abnormalities.

Interestingly, an additional peak was observed in the IVS-1-1 line RP-HPLC traces (indicated by \* in Fig. 3c). Mass spectrometry sequencing revealed this to be a splice variant of  $\beta$ -globin, created by splicing machinery using a previously identified IVS-1-13 alternative splice site<sup>24</sup>, resulting in a 4 amino acid insertion in  $\beta$ -globin between exon 1 and 2 (G29\_R30insSLVS; Fig. 4a and Supplementary Fig. 7). Molecular modelling revealed disruption of interactions between the  $\beta$ -globin variant and  $\alpha$ -globin, including with  $\alpha$ -globin Phe 117, critical for  $\alpha\beta$ 1 dimerization<sup>25</sup>, due to increased distances between the network of key residues (Fig. 4b). Of note, the variant  $\beta$ -globin was detected in the insoluble fraction of IVS-1-1 cells (Fig. 3d), further supporting the hypothesis that it is unable to form a stable hemoglobin tetramer with  $\alpha$ -globin and thus precipitating, potentially contributing to the IE phenotype of  $\beta$ -thalassaemia erythroid cells with the IVS-1-1 mutation.

### Culture of $\beta$ -thalassaemia patient cells further validates $\beta$ -thalassaemia cell lines

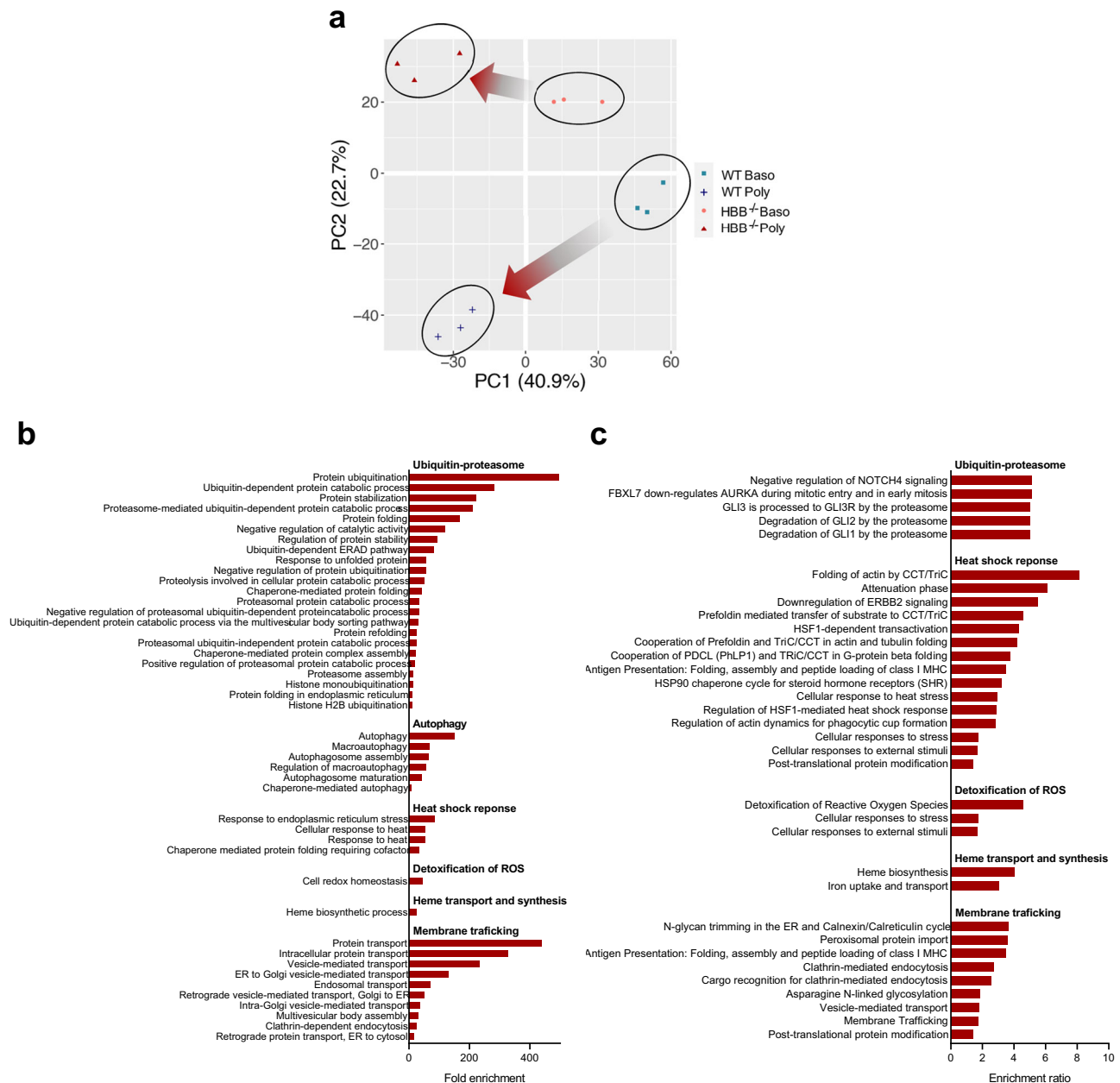
Due to the limited availability of suitable patient material for in vitro studies, there are few reports analyzing cultures of  $\beta$ -thalassaemia patient erythroid cells<sup>23</sup>. To provide further analysis, and to validate our model  $\beta$ -thalassaemia lines under the same conditions, peripheral blood mononuclear cells (PBMCs) from two  $\beta$ -thalassaemia patients (IVS-1-110 G  $\rightarrow$  A [ $\beta^+$ ] and IVS-1-1 G  $\rightarrow$  T [ $\beta^0$ ]) were isolated and differentiated using our erythroid culture system. As seen with the  $\beta$ -thalassaemia lines, and as reported previously for in vitro cultured  $\beta$ -thalassaemia ( $\beta^0$ ) erythroid cells<sup>2</sup>, there was drastically increased cell death and onset of apoptosis correlating with differentiation to

polychromatic erythroblasts. This is exemplified using the IE assay, with increased IE ratio clearly apparent from day 11 of differentiation onwards (Supplementary Fig. 8A–D). The IE assay distinguished degree of severity between the  $\beta^0$  (IVS-1-1) and  $\beta^+$  (IVS 1-110) patient erythroid cells, with the IE ratio corresponding with genotype (Supplementary Fig. 8B). In addition,  $\alpha$ -globin was found at high abundance in the insoluble fraction of whole cell lysates from differentiated cells (Supplementary Fig. 8E), due to aggregation in the absence of sufficient  $\beta$ -globin. The mutant  $\beta$ -globin identified in the IVS-1-1 cell line, was also detected in the IVS-1-1 ( $\beta^0$ ) patient erythroid cells (Supplementary Fig. 8E) and established as the same splice variant by mass spectrometry, confirming that this mutated  $\beta$ -globin is indeed produced in patients.

### Quantitative proteomic comparison of WT and *HBB*<sup>-/-</sup> BEL-A

To obtain a global assessment of differences between WT and *HBB*<sup>-/-</sup> erythroid cell proteome, verify alterations in proteins previously identified, and reveal novel proteins and processes for potential therapeutic targeting, TMT LC-MS/MS comparative proteomic analysis was performed. For this, cells were evaluated during early differentiation (basophilic erythroblasts), prior to manifestation of IE phenotype, and in mid-differentiation (polychromatic erythroblasts), when the  $\alpha$ - $\beta$ -globin chain imbalance approaches maximal levels and apoptosis is induced. Due to differences in kinetics of the WT and *HBB*<sup>-/-</sup> cultures, basophilic and polychromatic erythroblasts were isolated by FACS, using inverse levels of GPA and CD36 (representative cytospin images of these populations are shown in Supplementary Fig. 9).

Of the 3364 unique proteins quantified, 1048 (31%) were significantly different in *HBB*<sup>-/-</sup> compared to WT basophilic erythroblasts, and 1001 (30%) in *HBB*<sup>-/-</sup> compared to WT polychromatic erythroblasts (Supplementary Data 1). Despite the similar proportion of significantly different proteins in both cell types, principal component analysis (PCA; Fig. 5a) illustrates the increased magnitude of proteome changes between WT and *HBB*<sup>-/-</sup> polychromatic erythroblasts compared to



**Fig. 5 | Comparative proteomic analysis of *HBB*<sup>-/-</sup> and WT BEL-A.** **a** Principal component analysis (PCA) plot of multiplex TMT-based comparative proteomics of FACS-isolated basophilic and polychromatic erythroblasts from *HBB*<sup>-/-</sup> and WT BEL-A. **b** and **c** Overrepresented pathways in *HBB*<sup>-/-</sup> vs WT BEL-A polychromatic cells

from TMT-based comparative proteomic data. Pathways shown were significantly enriched by FDR (<0.05) from **(b)** DAVID **(c)** Webgestalt proteomics pathway analysis tools. Source data are provided as a Source Data file.

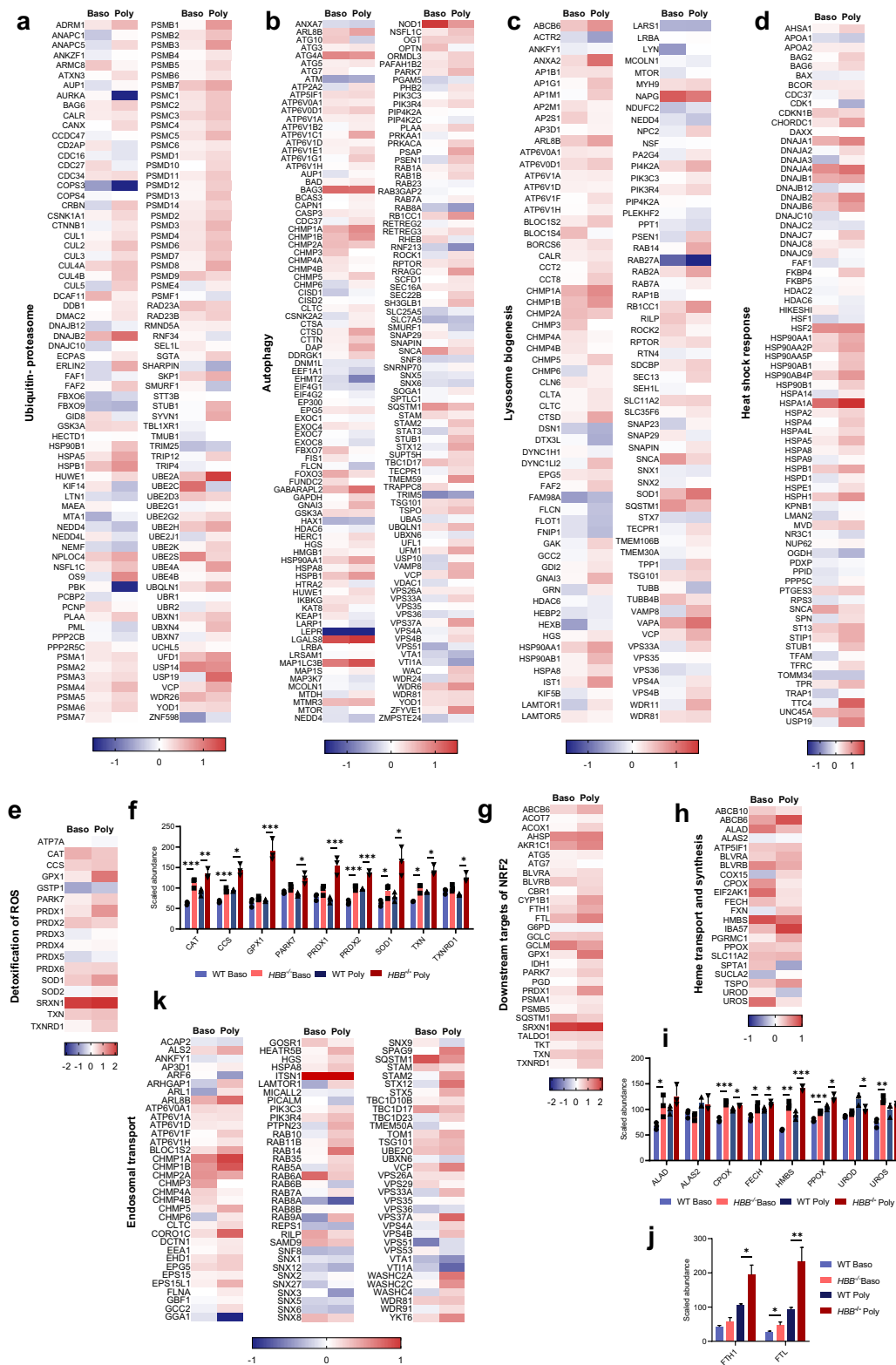
basophilic erythroblasts, reflecting the more severe phenotype of the former.

Focussing on the polychromatic erythroblasts where changes in proteome were more pronounced, we interrogated the set of proteins significantly up-regulated and significantly down-regulated in *HBB*<sup>-/-</sup> cells, also using Gene Ontology and Pathway Enrichment Analysis toolkits WebGestalt<sup>26</sup> and DAVID<sup>27</sup> to identify overrepresented pathways. No down-regulated pathways reached significance, but multiple significantly upregulated pathways were identified, as detailed below.

$\beta$ -thalassemia has been described as a protein-aggregation disorder<sup>28</sup> with increased activity of the ubiquitin-proteasome pathway and autophagy<sup>29–31</sup> as cells attempt to clear the  $\alpha$ -globin aggregates. Here we show that the ubiquitin-proteasome and autophagy pathways are significantly upregulated in *HBB*<sup>-/-</sup> polychromatic

erythroblasts (Fig. 5b and c), with the majority of proteins identified in our dataset involved in these pathways increased in level (Fig. 6a and b). In addition, the abundance of the majority of proteins in the lysosome biogenesis pathway, associated with autophagy, was increased (Fig. 6c). We also found enrichment of pathways associated with heat shock response (Fig. 5b), with a striking increase in the majority of identified proteins including many chaperones (Fig. 6d). Changes in abundance of proteins within these pathways were also detected in basophilic erythroblasts, although in general were less apparent than in polychromatic erythroblasts (Fig. 6), in line with increasing  $\alpha$ -globin levels.

Of note, the most significantly over-expressed chaperone protein in both *HBB*<sup>-/-</sup> polychromatic and basophilic erythroblasts was HSP70 (HSPA1A; Supplementary Fig. 10A). HSP70 has previously been



**Fig. 6 | Analysis of alterations in abundance of proteins from significantly enriched pathways of *HBB*<sup>-/-</sup> vs WT BEL-A.** Heatmaps show log<sub>2</sub> fold-changes (FC) of all WT vs *HBB*<sup>-/-</sup> normalized protein abundances in both basophilic (Baso) and polychromatic (Poly) erythroblasts from significantly upregulated pathways identified by pathway analysis, using the following search terms on <http://geneontology.org/>: (a) ubiquitin-proteasome, (b) autophagy, (c) lysosome biogenesis, (d) heat shock response, (e) detoxification of reactive oxygen species

(ROS), (h) heme transport and synthesis, and (k) endosomal transport. **g** Heat map of all downstream target proteins of NRF2 identified in the dataset. Bar charts show abundance values normalized to total protein and scaled relative to 100 shown as mean ± SD, *n* = 3 independent experiments, for antioxidant enzymes (f), Heme processing enzymes (i), and ferritin heavy and light chain (j). \**P* < 0.05, \*\**P* < 0.01, \*\*\**P* < 0.001. *P* values represent results from two-tailed heteroscedastic *t*-tests performed on log<sub>2</sub> normalized data. Source data are provided as a Source Data file.



reported to interact with aggregated  $\alpha$ -globin in  $\beta$ -thalassemia erythroblasts sequestering it to the cytosol, thus preventing its translocation to the nucleus to protect GATA1 from degradation<sup>3</sup>. Consistent with this finding, GATA1 levels were significantly decreased in  $HBB^{-/-}$  basophilic and polychromatic erythroblasts (Supplementary Fig. 10A) and also in the nuclear fraction of these cells (Supplementary Fig. 10B and C). In addition, HSP70 was found to colocalise with  $\alpha$ -globin in the insoluble fraction of  $HBB^{-/-}$  cell lysate (Supplementary Fig. 10D) and in the insoluble fraction of  $\beta$ -thalassemia patient primary erythroid cell lysates (Supplementary Fig. 8E).

Few pathways involved in oxidative stress have been studied in detail in  $\beta$ -thalassemia erythroid cells<sup>32,33</sup> although auto oxidation of  $\alpha$ -globin aggregates is hypothesized to lead to increased ROS<sup>34</sup>. Consistent with this, pathways associated with detoxification of ROS were significantly overrepresented (Fig. 5b and c), with the majority of proteins involved in detoxification of ROS found at increased levels (Fig. 6e), including eight antioxidant enzymes significantly increased in polychromatic erythroblasts, with four also significantly increased in basophilic erythroblasts, prior to onset of apoptosis (Fig. 6f). In addition, target genes of NRF2<sup>35-37</sup>, an activator of the antioxidant response programme<sup>38</sup> previously shown to be upregulated in  $\beta^{IVS-2-654}$  thalassaemic mice<sup>37</sup>, were increased (Fig. 6g).

Interestingly, pathways involved in heme transport and synthesis were overrepresented in our dataset (Fig. 5b and c), with proteins involved in heme biosynthesis found at overall higher levels in both  $HBB^{-/-}$  basophilic and polychromatic erythroblasts (Fig. 6h). This included seven of the eight core heme synthesis enzymes, with just one in basophilic and two in polychromatic erythroblasts failing to reach significance (Fig. 6i). Levels of both ferritin light chain and ferritin heavy chain were also significantly increased (Fig. 6j). In relation, pathways associated with endosomal transport and membrane trafficking were significantly enriched in polychromatic erythroblasts (Fig. 5b and c) and the majority of proteins involved in endosomal transport found at increased levels (Fig. 6k), indicating that more general transport mechanisms may be upregulated in order to elevate heme transport capacity. To further investigate the increased heme biosynthesis observed in  $\beta$ -thalassemia erythroblasts, WT and  $HBB^{-/-}$  cells were treated with 50  $\mu$ M succinylacetone during differentiation to inhibit the second step of heme synthesis. Succinylacetone significantly reduced the viability of WT cells, but conversely did not affect viability of  $HBB^{-/-}$  cells. Instead there was a small but significant decrease in CD36 levels indicating increased maturation of the  $HBB^{-/-}$  cells. (Supplementary Fig. 11). These data indicate that heme biosynthesis is upregulated above functionally required levels in  $\beta$ -thalassemia erythroid cells, and that normalizing levels may have a therapeutic effect. Overall, the findings suggest that increased heme biosynthesis in  $\beta$ -thalassemia erythroblasts, likely in response to decreased hemoglobin levels, could exacerbate IE as well as erythroid intrinsic iron overload, contributing to the disease phenotype.

## Discussion

In this study we created, characterized, and validated human disease cellular model systems for  $\beta$ -thalassemia, which recapitulate the phenotype of patient erythroid cells. The lines provide a sustainable and consistent supply of disease cells as unique research tools for studying the underlying molecular defects, for identifying new therapeutic targets, and as screening platforms for drug and therapeutic reagents. For the latter we also developed a high throughput compatible fluorometric-based assay for evaluation of severity of IE, and thus of disease phenotype. We then utilized the assay to demonstrate our  $\beta$ -thalassemia lines respond appropriately to drug and gene therapy approaches previously shown to increase  $\gamma$ -globin, with resultant improvement to IE, providing validation for such applications.

Proteomic studies have previously been utilized to identify potential biomarkers of  $\beta$ -thalassemia disease severity, analyzing

serum<sup>39</sup>, plasma<sup>40-43</sup>, or peripheral blood microvesicles<sup>44,45</sup>, for improved clinical management. Such analyses, although important, address downstream sequelae of the root cause of disease pathophysiology, IE of the erythroid cells. Our  $\beta$ -thalassemia disease lines overcome the hurdle of obtaining adequate numbers of bone marrow erythroid cells or haematopoietic stem cells from patients, enabling in depth and accurate analysis of molecular changes underlying the defective differentiation. In addition, an issue with comparing control and  $\beta$ -thalassemia erythroid cells arises from differences in kinetics of the cultures<sup>2</sup> (Figs. 1b and 3b), whereby differences due to stage of differentiation can confound identification of those due to disease phenotype. However, as we show, our lines enable sufficient size cultures to isolate specific, stage-matched cells.

Using TMT-based comparative proteomics we reveal the changes in biological pathways and processes in  $\beta$ -thalassemia erythroid cells, including upregulated pathways associated with, and increased levels of, antioxidant response enzymes. In addition, we identified increased heat shock proteins and chaperones, and increased levels of many proteins of the ubiquitin-proteasome and autophagy pathways along with proteins of associated lysosome biogenesis, together with respective upregulated pathways. The changes in proteome are manifest in earlier erythroid cells, but with the magnitude of change increased by the polychromatic erythroid cell stage, in line with increased levels of  $\alpha$ -globin. However, this plethora of cellular responses are unable to adequately remove the aggregated  $\alpha$ -globin or resolve the downstream sequelae, resulting in manifestation of the disease phenotype.

Previously, only few proteins associated with some of these processes have been reported upregulated in  $\beta$ -thalassemia erythroid cells<sup>29,31,46</sup>, although increased abundance of transcripts for most proteasome subunits has been reported in the  $\beta$ -globin<sup>Th3/+</sup> and  $\beta$ -globin<sup>Th3/Th3</sup> mouse<sup>30</sup>, in line with increased levels of these proteins in human cells in the present study. In contrast, our data reveal extensive upregulation of proteins involved in wide range of biological response pathways, in a biologically relevant human system.

We also reveal upregulated pathways and increased levels of proteins for heme biosynthesis, including all eight core heme synthesis enzymes, along with associated upregulated endosome biogenesis and trafficking pathways. These findings support the hypothesis that erythroid-intrinsic iron overload may be contributing to oxidative stress in  $\beta$ -thalassemia, as heme molecules bind insoluble  $\alpha$ -globin chains exacerbating the toxicity of these aggregates<sup>47</sup>. Heme also inhibits HRI, increasing globin synthesis, which will cause further aggregate formation<sup>48</sup>. Succinylacetone is an inhibitor of the second step of heme synthesis and has shown promise in alleviating disease phenotype in Diamond Blackfan Anemia (DBA) erythroid cells *in vitro*<sup>49</sup>. Our data show that succinylacetone treatment also has a positive, albeit small, effect on  $\beta$ -thalassemia erythroid cells. The detrimental effect of the drug on viability of WT, but not  $\beta$ -thalassemia erythroid cells, supports the pathological upregulation of heme biosynthesis observed in the  $HBB^{-/-}$  erythroid proteomic data. More targeted strategies to reduce heme synthesis may provide improved effects and thus potential novel avenues for therapeutic intervention in  $\beta$ -thalassemia. For example, catalytic activity of erythroid specific ALAS2, the first enzyme in the heme synthesis pathway, can be modulated by small molecules binding a hotspot around its auto-inhibitory Ct-extension<sup>50</sup>.

Amongst the 10 most aberrantly increased proteins in  $\beta$ -thalassemia erythroid cells ( $HBB^{-/-}$ ; Supplementary Table 1) those with a possible causative role in disease phenotype, rather than resultant change, and thus potential application as novel therapeutic targets for follow up, include: OTUD5 a deubiquitinase, increased levels of which could exacerbate production of  $\alpha$ -globin aggregates directly and thus ineffective erythropoiesis, and also via mTORC1 pathway, OTUD5 being a positive regulator of mTORC, with inhibition

of mTORC1 increasing autophagy and reducing  $\alpha$ -globin aggregation and disease phenotype in a mouse model of  $\beta$ -thalassemia;<sup>31</sup> Cathepsin B (CTSB), a negative regulator of autophagy, also over expressed in Sickle Cell Disease erythroid cells where it is being investigated as a therapeutic target;<sup>51</sup> Exportin 7 (XPO7) a broad-spectrum exportin and a nuclear import receptor implicated in erythroid differentiation, with increased levels potentially perturbing critical intracellular processes<sup>52</sup>. Proteins that are reduced in level are more challenging to target therapeutically, but of interest ZNF648 was ~6-fold lower ( $p < 0.0001$ ) in  $\beta$ -thalassemia polychromatic erythroid cells, with reduced levels of ZNF648 in wild type erythroid cells shown to impede erythroid cell survival and differentiation<sup>53</sup>.

As aberrant activation/inactivation of proteins can also perturb intracellular processes we also compared the phosphorylation status of proteins in the proteomic datasets. Proteins with greatest alteration in phosphorylation status (twofold or more in  $HBB^{-/-}$  compared to WT polychromatic erythroid cells) are shown in Supplementary Table 2, along with peptide and residues phosphorylated. Although the function of many of these proteins in erythropoiesis is presently unknown, such data has potential to provide information on changes that lead to ineffective erythropoiesis, follow up analyses also distinguishing causative from consequential effects on phosphorylation status. For example, the GTPase RAB8A had significantly increased phosphorylation at S111 in  $\beta$ -thalassemia erythroid cells, a modification which inactivates this GTPase<sup>54</sup>, disrupting associated signalling cascades; OSBP2 has been implicated as potential driver of erythroid cell differentiation<sup>55</sup>, with phosphorylation modulating its function; Slc29a1 promotes erythroblast survival and erythroid differentiation<sup>56</sup>, and is known to be phosphorylated at multiple sites by both PKA and PKC<sup>57</sup>, however, how phosphorylation regulates the function of this adenosine transporter is not yet known. Proteins with significantly reduced levels of phosphorylation were found to be predominantly associated with cellular stress and apoptosis pathways, likely consequential changes, as well as with the cell cycle.

CRISPR editing to KO the  $HBB$  gene in the HUDEP-2 erythroid cell line has been shown to result in a significant increase in  $\gamma$ -globin or HbF due to a stress erythropoiesis response<sup>33,58</sup>. However, increased  $\gamma$ -globin expression is not an inherent feature of  $\beta$ -thalassemia. Indeed,  $\beta$ -thalassemia patients who co-inherit other genetic variants that cause increased  $\gamma$ -globin expression and HbF have reduced or ameliorated disease severity (reviewed by Vinjamur et al.<sup>59</sup>), with methods to increase HbF the most promising therapeutic approaches for the disease. A  $\beta$ -thalassemia phenotype with  $\alpha$ -globin aggregates and IE was not reported for these HUDEP-2 KO lines. In contrast, Zhang et al.<sup>60</sup> did not get increased  $\gamma$ -globin following differentiation, when instead used prime editing to introduce  $\beta$ -thalassemia mutations into the  $HBB$  gene of HUDEP-2 cells, with most lines having lower expression of  $\gamma$ -globin than the unedited control cells. However, introduction of even silent mutations, that do not result in a  $\beta$ -thalassemia disease phenotype, into the  $HBB$  gene of HUDEP-2 by prime editing resulted in significant apoptosis and cell death of the resultant lines when differentiated, as much or even more so than lines with mutations that cause  $\beta^0$  thalassaemia introduced. The cell death therefore appears to be associated with the prime editing process itself, with the lines not therefore representing a  $\beta$ -thalassemia phenotype.

Intriguingly, we found a novel mutant  $\beta$ -globin splice variant arising from the IVS-1-1 G  $\rightarrow$  T mutation, predicted to reduce the stability of dimer and tetramer and thus still resulting in a  $\beta^0$ -thalassaemia phenotype. A transcript for this variant has previously been reported from overexpressing IVS-1-1 G  $\rightarrow$  A  $HBB$  coding region construct in HeLa cells<sup>24</sup>. That the variant does not have a beneficial role correlates with the lack of evidence for a milder clinical phenotype in IVS-1-1 G  $\rightarrow$  T patients compared to those with frameshift mutations<sup>61</sup>. In addition, there was no evidence of reduced insoluble  $\alpha$ -globin in our IVS-1-1 compared to CD41/42 line, further supporting the hypothesis that the

IVS-1-1  $\beta$ -globin variant cannot form stable interactions with  $\alpha$ -globin. Instead, as the variant  $\beta$ -globin precipitates it could exacerbate the disease phenotype. As there are many splicing mutations that cause  $\beta$ -thalassaemia, it will be interesting to determine if other in-frame mutant  $\beta$ -globin splice variants are produced and if so, how they impact disease phenotype.

Finally, the findings presented here also provide proof of principle for CRISPR-Cas9 genome editing of BEL-A to create sustainable model cellular systems of other red blood cell disorders, expanding the toolkit available to erythropoiesis researchers.

## Methods

### Cell culture

**Primary cells.** Peripheral blood mononuclear cells (PBMCs) were isolated from LRS cones for control cells and whole blood for  $\beta$ -thalassaemia cells. LRS cones were obtained with informed consent from all donors and used in accordance with the Declaration of Helsinki and approved by the National Health Service National Research Ethics Committee (reference number 08/HO102/26) and the Bristol Research Ethics Committee (reference 12/SW/0199). Whole blood from  $\beta$ -thalassaemia patients was obtained under ethics board North of Scotland REC (18/NS/0005). PBMCs were thawed into EDM supplemented with 1 ng.ml<sup>-1</sup> IL-3 and 10 ng.ml<sup>-1</sup> SCF and a lineage depletion was performed on day 5 according to manufacturer's instructions (Miltenyi Biotec; 130-092-211). On day 9, cells were transferred to EDM supplemented with 10 ng.ml<sup>-1</sup> SCF only, and on day 13 cells were transferred to EDM supplemented with holo-transferrin to a final concentration of 500  $\mu$ g.ml<sup>-1</sup>.

**BEL-A.** BEL-A cells were cultured as previously described<sup>62</sup>. In brief, for expansion phase cells were cultured in StemSpan™ SFEM [Stemcell Technologies] containing 50 ng.ml<sup>-1</sup> SCF, 3 U.ml<sup>-1</sup> EPO, 1  $\mu$ M dexamethasone and 1  $\mu$ g.ml<sup>-1</sup> doxycycline. To induce differentiation, expanding cells were transferred to erythroid differentiation medium (Iscove's medium with stable glutamine [Merck] containing 3% (v/v) AB serum [Merck], 2% (v/v) fetal bovine serum [Hyclone], 10  $\mu$ g.ml<sup>-1</sup> Insulin [Merck], Heparin 3 U.ml<sup>-1</sup> [Merck], 200  $\mu$ g.ml<sup>-1</sup> holo-transferrin [Sanquin, NED] and 3 U.ml<sup>-1</sup> EPO [Roche]—EDM) supplemented with 1 ng.ml<sup>-1</sup> IL-3 and 10 ng.ml<sup>-1</sup> SCF (both R&D Systems) and 1  $\mu$ g.ml<sup>-1</sup> doxycycline for 4 days, and for a further 4 days without doxycycline. Cells were then transferred to EDM supplemented with holo-transferrin to a final concentration of 500  $\mu$ g.ml<sup>-1</sup>. For hydroxyurea treatment, hydroxyurea (H8627; Sigma) was added to cells at 50  $\mu$ M every 2 days from 4 days prior to the start of differentiation until harvesting of cells for analysis. Pomalidomide (10  $\mu$ M; P0018; Sigma-Aldrich) was added every 2 days and Decitabine (50 nM; A3656; Sigma-Aldrich) was added to cells daily during differentiation.

### CRISPR-Cas9 genome editing of BEL-A

**Plasmid-based editing.** To generate  $HBB^{-/-}$  and  $HBB^{+/-}$  lines, Guide RNA oligos (GTAACGGCAGACTTCTCCTC) were ligated into the pSpCas9(BB)-2A-GFP (px458) vector (48138; Addgene). BEL-A cells ( $1.0 \times 10^6$ ) were resuspended in CD34+ nucleofection kit buffer (Lonza Biosciences) containing 2.5  $\mu$ g px458 (with gRNA) and 2.5  $\mu$ M ssODN donor template, and electroporated using program U-008 of AMAXA Nucleofactor™ 2b. After 48 h, dead and dying cells were removed using a dead cell removal kit (130-090-101; Miltenyi). After a further 24 h, single-cell cloning of GFP + /DRAQ5- cells into 96-well plates was performed using a BDInflux Cell Sorter.

**Ribonucleoprotein (RNP)-based editing.** A sub-cloned population of BEL-A cells ( $1.0 \times 10^5$ ) were resuspended in CD34+ nucleofection kit buffer (Lonza Biosciences) containing 18 pmol Cas9, 45 pmol gRNA. To generate the  $HBB^{-/-}$   $HBD^{-/-}$  line, three gRNAs previously validated by Boontanrart et al.<sup>58</sup> were used simultaneously, two of which were

specific to *HBB* (e87 (g10) CUUGCCCCACAGGGCAGUAA; e383 GCUC AUGGCAAGAAAGUGCU), and one of which targeted both *HBB* and *HBD* (e297 UGGUCUACCCUUGGACCCAG). For CD41/42 -TCTT IVS1-1 G → T *HBB* edits, gRNAs ((CD41/42 *HBB*—GGCUGCUGGUGGUCUA CCCU; IVS1-1 *HBB*—AAGGUGAACGUGGAUGAAGU, Synthego) were used along with 100 pmol ssODN template containing the mutation of interest. ssODN templates were antisense for both the CD41/42 and IVS1-1 edits with homology arms 36 nucleotides away from the edit and 91 nucleotides toward the edit, in accordance with the design principles described previously<sup>23</sup>. A recoded template was chosen for IVS1-1 G → T, where silent mutations are incorporated into the ssODN between the site of the DSB and the intended edit to increase editing efficiency<sup>23</sup>, due to the relatively large distance from the edit to DSB.

*BCL11A* + 58 enhancer editing, was performed as above with *HBB*<sup>-/-</sup> BEL-A as a founder population using a previously validated gRNA targeting the *BCL11A* + 58 enhancer (CUAACAGUUGCUUUUAUCAC)<sup>2,13</sup>.

### Clone screening

Genomic DNA was isolated and the gene region of interest was amplified using GoTaq DNA polymerase (Promega) (*HBB* Fw: 5'- TG GTATGGGGCCAAGAGATA-3', *HBB* Rv: 5'-GAGCCAGGCCATCACTA AAG-3'; *BCL11A* + 58 Fw: GGCAGCTAGACAGGACTTGG, *BCL11A* + 58 Rv: GGAGGCAAGTCAGTTGGAA). PCR products were cleaned up (QIAquick PCR purification Kit, Qiagen) and sequenced by EuroFins Scientific. Sequencing data were analyzed by TIDE<sup>63</sup> and TIDER<sup>64</sup> web tools (<https://tide.nki.nl/>).

### Cytospin preparation, imaging, and morphology analysis

Aliquots of 2–8 × 10<sup>4</sup> cells in 200 μl of appropriate cell culture medium were used to prepare cytospin preparations on coated slides, by centrifugation at 400 g for 5 min with a Thermo Scientific Shandon 4 Cytospin. The slides were stained in Leishman's Eosin-Methylene blue solution (VWR). The stained slides were imaged using an Olympus CX43 microscope mounted with Olympus SCS0 camera. Morphology analyses were determined by counting at least 200 cells per slide.

### Flow cytometry

**Ineffective erythropoiesis (IE) assay.** Aliquots of 2–3 × 10<sup>5</sup> cells were incubated with CD36-Vioblu conjugated antibody (Clone AC106, 130-095-482; Miltenyi; dilution 1:11) in PBS containing 1% (w/v) BSA (Park Scientific Ltd) and 2 mg·ml<sup>-1</sup> glucose (PBS-AG), followed by incubation with Annexin V-FITC conjugated antibody (130-093-060; Miltenyi; dilution 1:11) in Annexin V binding buffer (10 mM HEPES, pH7.4; 140 mM NaCl; 2.5 mM CaCl<sub>2</sub>). Cells were analyzed on a BD LSR Fortessa or BD LSR II flow cytometer and data analyzed using FlowJo v10.6.1 (FlowJo LLC). Gating strategy is depicted in Supplementary Fig. 12.

### RP-HPLC

RP-HPLC was performed as described by Loucari et al.<sup>65</sup> with the exception of the injection size (600,000 cells in 30 μl dH<sub>2</sub>O) and the LC column used (Jupiter 5 μm C18 300 A, size 250 × 4.4 mm, protected with a Security Guard analytical guard system [KJ0-4282, Phenomenex]). Purified HbA2 and HbA0 controls (H0266 and H0267; Sigma Aldrich) were reconstituted in dH<sub>2</sub>O at 200 ng/μl, with 30 μl used per run.

### SDS-PAGE and western blot

Proteins (either from whole cell lysates or insoluble fraction) were resolved by SDS-PAGE and transferred to PVDF membrane (Merck) by western blotting. The insoluble fraction was obtained by centrifugation (17,000 g) of whole cell lysate homogenized in RIPA buffer, before direct resuspension in SDS-PAGE sample buffer, with the equivalent of 250,000 cells loaded per lane. To obtain nuclear lysates, 1 × 10<sup>7</sup> cells were resuspended with 800 μl of 4 °C HLBN buffer (10 mM Tris-HCl [pH 7.5], 10 mM NaCl, 2.5 mM MgCl<sub>2</sub>, 0.5% (v/v) nonyl

phenoxyethoxyethanol-40 [NP-40]) and incubated for 5 min on ice. 200 μl HLBN buffer + 10% sucrose was then pipetted to under-ly the HLBN suspension. Following centrifugation at 400 g for 5 min, the supernatant was removed, and the remaining nuclear pellet was lysed in RIPA buffer. For each lane, 3 μg of whole cell lysate and 20 μg nuclear lysate was loaded. Membranes were blocked with 5% (w/v) milk powder before incubation in primary and HRP-secondary antibodies. Primary antibodies used were: Monoclonal mouse anti-hemoglobin α (Clone D-4, Cat. No. sc-514378, Lot No. I0916, Santa Cruz; dilution 1:3000), monoclonal mouse anti-hemoglobin β (Clone 37-8, Cat. No. sc-21757, Lot No. G2418, Santa Cruz; dilution 1:3000), monoclonal mouse anti-hemoglobin γ-globin (Clone 51-7, Cat. No. sc-21756, Lot No. G0318, Santa Cruz; dilution 1:1500), monoclonal rat anti-GATA1 (Clone N1, Cat. No. sc-266, Lot No. E1713, Santa Cruz; dilution 1:1500), polyclonal rabbit anti-Lamin A/C (Clone H-110, Cat. No. sc-20681, Lot No. B0111, Santa Cruz; dilution 1:1000), monoclonal mouse anti-β-actin (AC-15, Cat. No. A1978, Batch No. 0000086303, Merck; dilution 1:3000), monoclonal mouse anti-Ctip1/BCL-11A (Clone 14B5, Cat. No. ab19487, Lot No. GR3383959-1, Abcam; dilution 1:2000), mouse anti-CD233/Band3 (Clone BRIC 170, IBGRL; dilution 1:500), polyclonal rabbit anti-HSP70 (Cat. No. ADI-SPA-812-D, Lot No. 09011646, Enzo Life Sciences; dilution 1:1000). Secondary antibodies used were: Polyclonal rabbit anti-mouse immunoglobulins-HRP conjugated (Cat. No. P0260, Lot No. 41424309, DAKO; dilution 1:3000), polyclonal swine anti-rabbit immunoglobulins-HRP conjugated (Cat. No. P0399, Lot No. 41270083, DAKO; dilution 1:3000), polyclonal goat anti-rat immunoglobulins-HRP conjugated (Cat. No. 112-035-003, Jackson ImmunoResearch; dilution 1:10000). Bands were visualized using enhanced chemiluminescence (G.E. Healthcare) with a G:BOX Chemiluminescence imager (Syngene).

### Cell stage matching for comparative proteomics

Mouse anti-Glycophorin A (GPA; Clone BRIC 256; IBGRL; dilution 1:1), vs CD36 cell surface marker analysis by flow cytometry was used to isolate stage-matched basophilic and polychromatic erythroblasts for WT and *HBB*<sup>-/-</sup> BEL-A cells. Cells were obtained from 3 independent cultures.

### TMT labelling, mass spectrometry and data analysis

Aliquots of 100 μg of protein digested with trypsin (2.5 μg trypsin; 37 °C, overnight) and labelled with Tandem Mass Tag (TMT) Pro tag reagents according to the manufacturer's protocol (Thermo Fisher Scientific, Loughborough, LE11 5RG, UK), and the labelled samples pooled. For the Total proteome analysis, an aliquot of 50 μg of the pooled sample was desalted using a SepPak cartridges according to the manufacturer's instructions (Waters, Milford, MA, USA). Eluate from the SepPak cartridge was evaporated to dryness and resuspended in buffer A (20 mM ammonium hydroxide, pH 10) prior to fractionation by high pH reversed-phase chromatography using an Ultimate 3000 liquid chromatography system (Thermo Fisher Scientific). In brief, the sample was loaded onto an XBridge BEH C18 Column (130 Å, 3.5 μm, 2.1 × 150 mm, Waters, UK) in buffer A and peptides eluted with an increasing gradient of buffer B (20 mM ammonium hydroxide in acetonitrile, pH 10) from 0 to 95% over 60 min. The resulting fractions were evaporated to dryness and resuspended in 1% formic acid prior to analysis by nano-LC MSMS using an Orbitrap Fusion Tribrid mass spectrometer (Thermo Scientific).

For the Phospho-proteome analysis, the remainder of the TMT-labelled pooled sample was also desalted using a SepPak cartridge (Waters, Milford, Massachusetts, USA). Eluate from the SepPak cartridge was evaporated to dryness and subjected to TiO<sub>2</sub>-based phosphopeptide enrichment according to the manufacturer's instructions (Thermo Fisher Scientific). The flow-through and washes from the TiO<sub>2</sub>-based enrichment were then subjected to FeNTA-based phosphopeptide enrichment according to the manufacturer's instructions

(Thermo Fisher Scientific). The phospho-enriched samples were evaporated to dryness and resuspended in 1% formic acid prior to analysis by nano-LC MSMS using an Orbitrap Fusion Tribrid mass spectrometer (Thermo Scientific).

High pH RP fractions (Total proteome analysis) or the phospho-enriched fractions (Phospho-proteome analysis) were further fractionated using an Ultimate 3000 nano-LC system in line with an Orbitrap Fusion Tribrid mass spectrometer (Thermo Scientific). In brief, peptides in 1% (vol/vol) formic acid were injected onto an Acclaim PepMap C18 nano-trap column (Thermo Scientific). After washing with 0.5% (vol/vol) acetonitrile 0.1% (vol/vol) formic acid, peptides were resolved on a 250 mm × 75 μm Acclaim PepMap C18 reverse-phase analytical column (Thermo Scientific) over a 150-min organic gradient, using seven gradient segments (1–6% solvent B over 1 min, 6–15% B over 58 min, 15–32% B over 58 min, 32–40% B over 5 min, 40–90% B over 1 min, held at 90% B for 6 min and then reduced to 1% B over 1 min) with a flow rate of 300 nl/min. Solvent A was 0.1% formic acid, and Solvent B was aqueous 80% acetonitrile in 0.1% formic acid. Peptides were ionized by nano-electrospray ionization at 2.0 kV using a stainless-steel emitter with an internal diameter of 30 μm (Thermo Scientific) and a capillary temperature of 275 °C.

All spectra were acquired using an Orbitrap Fusion Tribrid mass spectrometer controlled by Xcalibur 2.1 software (Thermo Scientific) and operated in data-dependent acquisition mode using an SPS-MS3 workflow. FTMS1 spectra were collected at a resolution of 120,000, with an automatic gain control (AGC) target of 200,000 and a max injection time of 50 ms. Precursors were filtered with an intensity threshold of 5000, according to charge state (to include charge states 2–7) and with monoisotopic precursor selection. Previously interrogated precursors were excluded using a dynamic window (60 s ± 10 ppm). The MS2 precursors were isolated with a quadrupole mass filter set to a width of 1.2 m/z. ITMS2 spectra were collected with an AGC target of 10,000, max injection time of 70 ms and CID collision energy of 35%.

For FTMS3 analysis, the Orbitrap was operated at 50,000 resolution with an AGC target of 50,000 and a max injection time of 105 ms. Precursors were fragmented by high energy collision dissociation (HCD) at a normalized collision energy of 60% to ensure maximal TMT reporter ion yield. Synchronous precursor selection (SPS) was enabled to include up to ten MS2 fragment ions in the FTMS3 scan.

The raw data files were processed and quantified using Proteome Discoverer software v2.4 (Thermo Scientific) and searched against the UniProt Human database (downloaded January 2022: 178,486 entries) using the SEQUEST algorithm. Peptide precursor mass tolerance was set at 10 ppm, and MS/MS tolerance was set at 0.6 Da. Search criteria included oxidation of methionine (+ 15.995 Da), acetylation of the protein N-terminus (+ 42.011 Da) and methionine loss plus acetylation of the protein N-terminus (−89.03 Da) as variable modifications and carbamidomethylation of cysteine (+ 57.0214) and the addition of the TMTpro mass tag (+ 304.207) to peptide N-termini and lysine as fixed modifications. For the Phospho-proteome analysis, phosphorylation of serine, threonine and tyrosine (+ 79.966) was also included as a variable modification. Searches were performed with full tryptic digestion and a maximum of 2 missed cleavages were allowed. The reverse database search option was enabled and all data was filtered to satisfy a false discovery rate (FDR) of 5%.

Data analyses were performed on log<sub>2</sub> normalized abundances using an un-paired, two-tailed, heteroscedastic Student's *t* test in Microsoft Excel. A 5% false discovery rate (FDR) threshold for each comparison was calculated using R Studio. For data visualization, principal component analysis (PCA) plots and volcano plots were generated using R Studio. Heatmaps were generated using GraphPad Prism 8. For overrepresentation analysis, significantly differentially expressed proteins (*P* < 0.05) were functionally categorized based on

Reactome (<https://reactome.org>) via gene ontology (GO) classification using WebGestalt<sup>26</sup> (<http://www.webgestalt.org/>). Gene ontology biological process was used as a functional database. Functional annotation of biological pathways was also performed using DAVID<sup>27</sup> (<https://david.ncifcrf.gov>). WebGestalt and DAVID use the Benjamini-Hochberg Procedure to decrease the false discovery rate (FDR). Significantly enriched pathways are presented (FDR ≤ 0.05).

### Molecular modelling

The model for the IVS-1-1 mutant was built using AlphaFold<sup>66</sup>. The IVS-1-1 mutant model was then aligned with the β-globin proteins in the high-resolution X-ray structure of human hemoglobin (pdb code: 1A00<sup>67</sup>). The tetramer was relaxed by energy minimization with the Gromacs software<sup>68</sup> using the Gromos 54A7 force-field<sup>69</sup>. The system was energy minimized with the steepest-descent method in order to remove excessive strain by performing 100 steps of minimization with harmonic restraints applied to all non-hydrogen atoms, followed by further 100 steps restraining the Cα atoms only, ending with 100 steps with no restraints.

### Statistical analysis

For statistical analysis of >2 groups, ANOVA with Tukey multiple comparison testing was performed. For analysis of 2 groups, a two-tailed Welch's *t*-test was performed. In both cases, statistical tests were performed in GraphPad Prism 8. In all cases, measurements were taken from distinct samples.

### Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

### Data availability

The proteomic data set is provided as a Supplementary Data 1, with raw files available through PRIDE (accession PXD044730). Source data are provided with this paper.

### References

1. Thein, S. L. Molecular basis of β thalassemia and potential therapeutic targets. *Blood Cells Mol. Dis.* **70**, 54–65 (2018).
2. Mathias, L. A. et al. Ineffective erythropoiesis in beta-thalassemia major is due to apoptosis at the polychromatophilic normoblast stage. *Exp. Hematol.* **28**, 1343–1353 (2000).
3. Arlet, J.-B. et al. HSP70 sequestration by free alpha-globin promotes ineffective erythropoiesis in beta-thalassaemia. *Nature* **514**, 242–246 (2014).
4. Ribeil, J.-A. et al. Ineffective erythropoiesis in β -thalassemia. *ScientificWorldJournal* **2013**, 394295 (2013).
5. Porter, J. Beyond transfusion therapy: new therapies in thalassemia including drugs, alternate donor transplant, and gene therapy. *Hematol. Am. Soc. Hematol. Educ. Progr.* **2018**, 361–370 (2018).
6. Foong, W. C., Ho, J. J., Loh, C. K. & Viprakasit, V. Hydroxyurea for reducing blood transfusion in non-transfusion dependent beta thalassaemias. *Coch. Database Syst. Rev.* **10**, CD011579 (2016).
7. Algiragiri, A. H., Wright, N. A. M., Paolucci, E. O. & Kassam, A. Hydroxyurea for lifelong transfusion-dependent β-thalassaemia: A meta-analysis. *Pediatr. Hematol. Oncol.* **34**, 435–448 (2017).
8. Ansari, S. H., Lassi, Z. S., Khowaja, S. M., Adil, S. O. & Shamsi, T. S. Hydroxyurea (hydroxycarbamide) for transfusion-dependent β-thalassaemia. *Coch. Database Syst. Rev.* **3**, CD012064 (2019).
9. Cappellini, M. D. & Taher, A. T. The use of luspatercept for thalassaemia in adults. *Blood Adv.* **5**, 326–333 (2021).
10. Christakopoulos, G. E., Telange, R., Yen, J. & Weiss, M. J. Gene Therapy and Gene Editing for β-Thalassaemia. *Hematol. Oncol. Clin. North Am.* **37**, 433–447 (2023).

11. Thuret, I., Ruggeri, A., Angelucci, E. & Chabannon, C. Hurdles to the Adoption of Gene Therapy as a Curative Option for Transfusion-Dependent Thalassemia. *Stem Cells Transl. Med.* **11**, 407–414 (2022).
12. Anstee, D. J., Gampel, A. & Toye, A. M. Ex-vivo generation of human red cells for transfusion. *Curr. Opin. Hematol.* **19**, 163–169 (2012).
13. An, X. et al. Global transcriptome analyses of human and murine terminal erythroid differentiation. *Blood* **123**, 3466–3477 (2014).
14. McColl, B. & Vadolas, J. Animal models of  $\beta$ -hemoglobinopathies: utility and limitations. *J. Blood Med.* **7**, 263–274 (2016).
15. Trakarnsanga, K. et al. An immortalized adult human erythroid line facilitates sustainable and scalable generation of functional red cells. *Nat. Commun.* **8**, 14750 (2017).
16. Daniels, D. E. et al. Comparing the two leading erythroid lines BEL-A and HUDEP-2. *Haematologica* **105**, e389–e394 (2020).
17. Moutouh-de Parseval, L. A. et al. Pomalidomide and lenalidomide regulate erythropoiesis and fetal hemoglobin production in human CD34+ cells. *J. Clin. Investig.* **118**, 248–258 (2008).
18. Theodorou, A. et al. Proteomic Studies for the Investigation of  $\gamma$ -Globin Induction by Decitabine in Human Primary Erythroid Progenitor Cultures. *J. Clin. Med.* **9**, 134 (2020).
19. Canver, M. C. et al. BCL11A enhancer dissection by Cas9-mediated in situ saturating mutagenesis. *Nature* **527**, 192–197 (2015).
20. Psatha, N. et al. Disruption of the BCL11A Erythroid Enhancer Reactivates Fetal Hemoglobin in Erythroid Cells of Patients with  $\beta$ -Thalassemia Major. *Mol. Ther. Methods Clin. Dev.* **10**, 313–326 (2018).
21. Kountouris, P. et al. IthaGenes: an interactive database for haemoglobin variations and epidemiology. *PLoS ONE* **9**, e103020 (2014).
22. Kazazian, H. H. J. et al. Molecular characterization of seven beta-thalassemia mutations in Asian Indians. *EMBO J.* **3**, 593–596 (1984).
23. Paix, A. et al. Precision genome editing using synthesis-dependent repair of Cas9-induced DNA breaks. *Proc. Natl Acad. Sci. USA.* **114**, E10745–E10754 (2017).
24. Treisman, R., Orkin, S. H. & Maniatis, T. Specific transcription and RNA splicing defects in five cloned beta-thalassaemia genes. *Nature* **302**, 591–596 (1983).
25. Thom, C. S., Dickson, C. F., Gell, D. A. & Weiss, M. J. Hemoglobin variants: biochemical properties and clinical correlates. *Cold Spring Harb. Perspect. Med.* **3**, a011858 (2013).
26. Zhang, B., Kirov, S. & Snoddy, J. WebGestalt: an integrated system for exploring gene sets in various biological contexts. *Nucleic Acids Res.* **33**, W741–W748 (2005).
27. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57 (2009).
28. Khandros, E. & Weiss, M. J. Protein quality control during erythropoiesis and hemoglobin synthesis. *Hematol. Oncol. Clin. North Am.* **24**, 1071–1088 (2010).
29. Lithanatudom, P. et al. Enhanced activation of autophagy in  $\beta$ -thalassemia/Hb E erythroblasts during erythropoiesis. *Ann. Hematol.* **90**, 747–758 (2011).
30. Khandros, E., Thom, C. S., D'Souza, J. & Weiss, M. J. Integrated protein quality-control pathways regulate free  $\alpha$ -globin in murine  $\beta$ -thalassemia. *Blood* **119**, 5265–5275 (2012).
31. Lechauve, C. et al. The autophagy-activating kinase ULK1 mediates clearance of free  $\alpha$ -globin in  $\beta$ -thalassemia. *Sci. Transl. Med.* **11**, eaav4881 (2019).
32. Maia de Oliveira da Silva, J. P. et al. Global gene expression reveals an increase of HMGB1 and APEX1 proteins and their involvement in oxidative stress, apoptosis and inflammation pathways among beta-thalassaemia intermedia and major phenotypes. *Br. J. Haematol.* **186**, 608–619 (2019).
33. Zhou, G. et al. Multi-Omics Analysis in  $\beta$ -Thalassemia Using an HBB Gene-Knockout Human Erythroid Progenitor Cell Model. *Int. J. Mol. Sci.* **23**, 2807 (2022).
34. Scott, M. D. et al. Effect of excess alpha-hemoglobin chains on cellular and membrane oxidation in model beta-thalassemic erythrocytes. *J. Clin. Investig.* **91**, 1706–1712 (1993).
35. Ma, Q. Role of nrf2 in oxidative stress and toxicity. *Annu. Rev. Pharmacol. Toxicol.* **53**, 401–426 (2013).
36. Song, M.-Y., Lee, D.-Y., Chun, K.-S. & Kim, E.-H. The Role of NRF2/KEAP1 Signaling Pathway in Cancer Metabolism. *Int. J. Mol. Sci.* **22**, 4376 (2021).
37. Han, G. et al. Nrf2 expands the intracellular pool of the chaperone AHSP in a cellular model of  $\beta$ -thalassemia. *Redox Biol.* **50**, 102239 (2022).
38. Lee, J.-M. et al. Nrf2, a multi-organ protector? *FASEB J. Publ. Fed. Am. Soc. Exp. Biol.* **19**, 1061–1066 (2005).
39. Khan, F., Ali, A., Iqbal, A. & Musharraf, S. G. Profiling of hydroxyurea-treated  $\beta$ -thalassemia/serum proteome through nano-LC-ESI-MS/MS in combination with microsol-isoelectric focusing. *Biomed. Chromatogr.* **34**, e4753 (2020).
40. Hatairaktham, S. et al. Differential plasma proteome profiles of mild versus severe  $\beta$ -thalassemia/Hb E. *Ann. Hematol.* **92**, 365–377 (2013).
41. Ponnikorn, S. et al. Comparative Proteome-Wide Analysis of Bone Marrow Microenvironment of  $\beta$ -Thalassemia/Hemoglobin E. *Proteomes* **7**, 8 (2019).
42. Panachan, J. et al. Differentially expressed plasma proteins of  $\beta$ -thalassemia/hemoglobin E patients in response to curcuminoids/vitamin E antioxidant cocktails. *Hematology* **24**, 300–307 (2019).
43. Li, N. et al. Plasma proteome profiling combined with clinical and genetic features reveals the pathophysiological characteristics of  $\beta$ -thalassemia. *iScience* **25**, 104091 (2022).
44. Chaichompoo, P. et al. Characterizations and proteome analysis of platelet-free plasma-derived microparticles in  $\beta$ -thalassemia/hemoglobin E patients. *J. Proteom.* **76 Spec No**, 239–250 (2012).
45. Kittivorapart, J. et al. Quantitative proteomics of plasma vesicles identify novel biomarkers for hemoglobin E/ $\beta$ -thalassemic patients. *Blood Adv.* **2**, 95–104 (2018).
46. Leecharoenkiat, A. et al. Increased oxidative metabolism is associated with erythroid precursor expansion in  $\beta$ 0-thalassaemia/Hb E disease. *Blood Cells Mol. Dis.* **47**, 143–157 (2011).
47. Rivella, S. Iron metabolism under conditions of ineffective erythropoiesis in  $\beta$ -thalassemia. *Blood* **133**, 51–58 (2019).
48. Chen, J.-J. Regulation of protein synthesis by the heme-regulated eIF2alpha kinase: relevance to anemias. *Blood* **109**, 2693–2699 (2007).
49. Yang, Z. et al. Delayed globin synthesis leads to excess heme and the macrocytic anemia of Diamond Blackfan anemia and del(5q) myelodysplastic syndrome. *Sci. Transl. Med.* **8**, 338ra67 (2016).
50. Bailey, H. J. et al. Human aminolevulinate synthase structure reveals a eukaryotic-specific autoinhibitory loop regulating substrate binding and product release. *Nat. Commun.* **11**, 2813 (2020).
51. Ramasamy, J. et al. Cathepsin B, a Negative Regulator of Autophagy, Identified As a Novel Therapeutic Drug Target in Sickle Cell Disease. *Blood* **136**, 32 (2020).
52. Hattangadi, S. M. et al. Histones to the cytosol: exportin 7 is essential for normal terminal erythroid nuclear maturation. *Blood* **124**, 1931–1940 (2014).
53. Ferguson, D. C. J. et al. Characterization and evolutionary origin of novel C(2)H(2) zinc finger protein (ZNF648) required for both erythroid and megakaryocyte differentiation in humans. *Haematologica* **106**, 2859–2873 (2021).
54. Lai, Y.-C. et al. Phosphoproteomic screening identifies Rab GTPases as novel downstream targets of PINK1. *EMBO J.* **34**, 2840–2861 (2015).

55. Zhu, F. et al. Modeling dynamic functional relationship networks and application to ex vivo human erythroid differentiation. *Bioinformatics* **30**, 3325–3333 (2014).
56. Zwifelhofer, N. M. et al. GATA factor-regulated solute carrier ensemble reveals a nucleoside transporter-dependent differentiation mechanism. *PLoS Genet.* **16**, e1009286 (2020).
57. Reyes, G. et al. The Equilibrative Nucleoside Transporter (ENT1) can be phosphorylated at multiple sites by PKC and PKA. *Mol. Membr. Biol.* **28**, 412–426 (2011).
58. Boontanrart, M. Y. et al. ATF4 Regulates MYB to Increase  $\gamma$ -Globin in Response to Loss of  $\beta$ -Globin. *Cell Rep.* **32**, 107993 (2020).
59. Vinjamur, D. S., Bauer, D. E. & Orkin, S. H. Recent progress in understanding and manipulating haemoglobin switching for the haemoglobinopathies. *Br. J. Haematol.* **180**, 630–643 (2018).
60. Zhang, H., Zhou, Q., Chen, H. & Lu, D. Prime Editor 3 Mediated Beta-Thalassemia Mutations of the HBB Gene in Human Erythroid Progenitor Cells. *Int. J. Mol. Sci.* **23**, 5002 (2022).
61. Traivaree, C., Monsereenusorn, C., Rujkijyanont, P., Prasertsin, W. & Boonyawat, B. Genotype-phenotype correlation among beta-thalassemia and beta-thalassemia/HbE disease in Thai children: predictable clinical spectrum using genotypic analysis. *J. Blood Med.* **9**, 35–41 (2018).
62. Hawksworth, J. et al. Enhancement of red blood cell transfusion compatibility using CRISPR-mediated erythroblast gene editing. *EMBO Mol. Med.* **10**, e8454 (2018).
63. Brinkman, E. K., Chen, T., Amendola, M. & van Steensel, B. Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Res.* **42**, e168 (2014).
64. Brinkman, E. K. et al. Easy quantification of template-directed CRISPR/Cas9 editing. *Nucleic Acids Res.* **46**, e58 (2018).
65. Loucari, C. C. et al. Rapid and Sensitive Assessment of Globin Chains for Gene and Cell Therapy of Hemoglobinopathies. *Hum. Gene Ther. Methods* **29**, 60–74 (2018).
66. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
67. Kavanaugh, J. S., Weydert, J. A., Rogers, P. H. & Arnone, A. High-resolution crystal structures of human hemoglobin with mutations at tryptophan 37beta: structural basis for a high-affinity T-state. *Biochemistry* **37**, 4358–4373 (1998).
68. Abraham, M. J. et al. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1–2**, 19–25 (2015).
69. Schmid, N. et al. Definition and testing of the GROMOS force-field versions 54A7 and 54B7. *Eur. Biophys. J.* **40**, 843–856 (2011).
- Research Fund via the University of Bristol's GCRF and Newton Fund Consolidation Account (GNCA) and BBSRC Pathfinder award BB/X511195/1 via the University of Bristol. The work of J.A.S. was supported by BBSRC-funded SWBio DTP. A.S.F.O. acknowledges the BBSRC grant BB/X009831/1 and ERC (Advanced Grant PREDACTED <https://cordis.europa.eu/project/id/101021207>) for support. ASFO also thanks Oracle for Research for the Research Fellowship.

### Author contributions

J.F. conceived and supervised the study. Experiments were conceived and designed by J.F. and D.E.D. with contribution from I.F.-V. and J.H. *HBB*<sup>-/-</sup> and *HBB*<sup>+/-</sup> cell lines were generated by I.F.-V. and D.C.J.F. *HBB*<sup>-/-</sup>*HBD*<sup>-/-</sup> cell line was generated by J.H. *HBB*<sup>-/-</sup> *BCL11A* +58 edited cell line was generated by E.M.F. CD41/42 and IVS-1-1 cell lines were generated by D.E.D. D.E.D. conducted the majority of experiments, analyzed data and prepared figures. T.N.A., E.M.F., and N.S.B. conducted experiments. M.C.W. carried out the HPLC experiments and contributed to proteomics analysis. A.S.F.O. carried out molecular modelling with contribution from J.A.S. J.N.B. provided whole blood from  $\beta$ -thalassaemia patients. J.F. and D.E.D. wrote the paper.

### Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-023-41961-9>.

**Correspondence** and requests for materials should be addressed to Jan Frayne.

**Peer review information** *Nature Communications* thanks Syed Musharraf and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023

### Acknowledgements

The authors would like to thank Dr. Kate Heesom, Director of the University of Bristol Proteomics Facility, for performing mass spectrometry and analysis of raw data, Dr. Phil Lewis, School of Cellular and Molecular Medicine, University of Bristol, Bristol, UK for mass spectrometry data analysis, the University of Bristol Flow Cytometry Facility for use of equipment, and Dr. Andrew Herman, director of, and Helen Rice of the Flow Cytometry Facility for performing cell sorting. The gRNAs used to generate the *HBB*<sup>-/-</sup>*HBD*<sup>-/-</sup> cell line were a gift from Professor Jacob Corn, ETH Zurich, Switzerland. We also thank the patients for their valuable participation in this study. The study was supported by MRC grant MR/S021140/1, Wellcome Trust grant 209739/Z/17/Z, the Global Challenges