scientific reports

Check for updates

OPEN

A long-range enhancer at -52Kb drives expression of the COUP-TFII transcription factor in erythroid cells

Valentina Pastori[1,10], Agata Labedz[1,7,8,10], Maria A. Simanovich[2,10], Martina Fabiano[1,10], Carlotta Frigo[1], Thijs Verheul[2], Ludovica Proietti[3,9], Florian Grebien[3,4,5], Elisabetta Citterio[1,6], Sjaak Philipsen[2]✉ & Antonella Ellena Ronchi[1]✉

COUP-TFII, encoded by the *NR2F2* gene, is an orphan nuclear receptor highly expressed during embryonic development in several tissues. In erythropoiesis, COUP-TFII is active in yolk sac-derived cells prior to the switch to adult globin expression. Its broad expression pattern suggests a complex transcriptional regulation involving multiple, yet poorly defined, regulatory elements. Using integrative *in silico* and chromatin accessibility analyses, here we identified an erythroid-specific enhancer located 52Kb upstream of the *NR2F2* transcription start site. This element shows epigenetic features of an active enhancer in K562 erythroid cells. Notably, in subclones derived from adult HUDEP2 erythroid progenitor cells that spontaneously re-express fetal gamma globin, *NR2F2* is reactivated, concomitantly with the opening of the -52Kb enhancer. We also identify the transcription factor ZBTB7A as a repressor of *NR2F2*, as knock-out of *ZBTB7A* in HUDEP2 cells restores *NR2F2* expression and active chromatin marks at the -52Kb region. Our findings uncover a novel distal enhancer controlling *NR2F2* expression in erythroid cells.

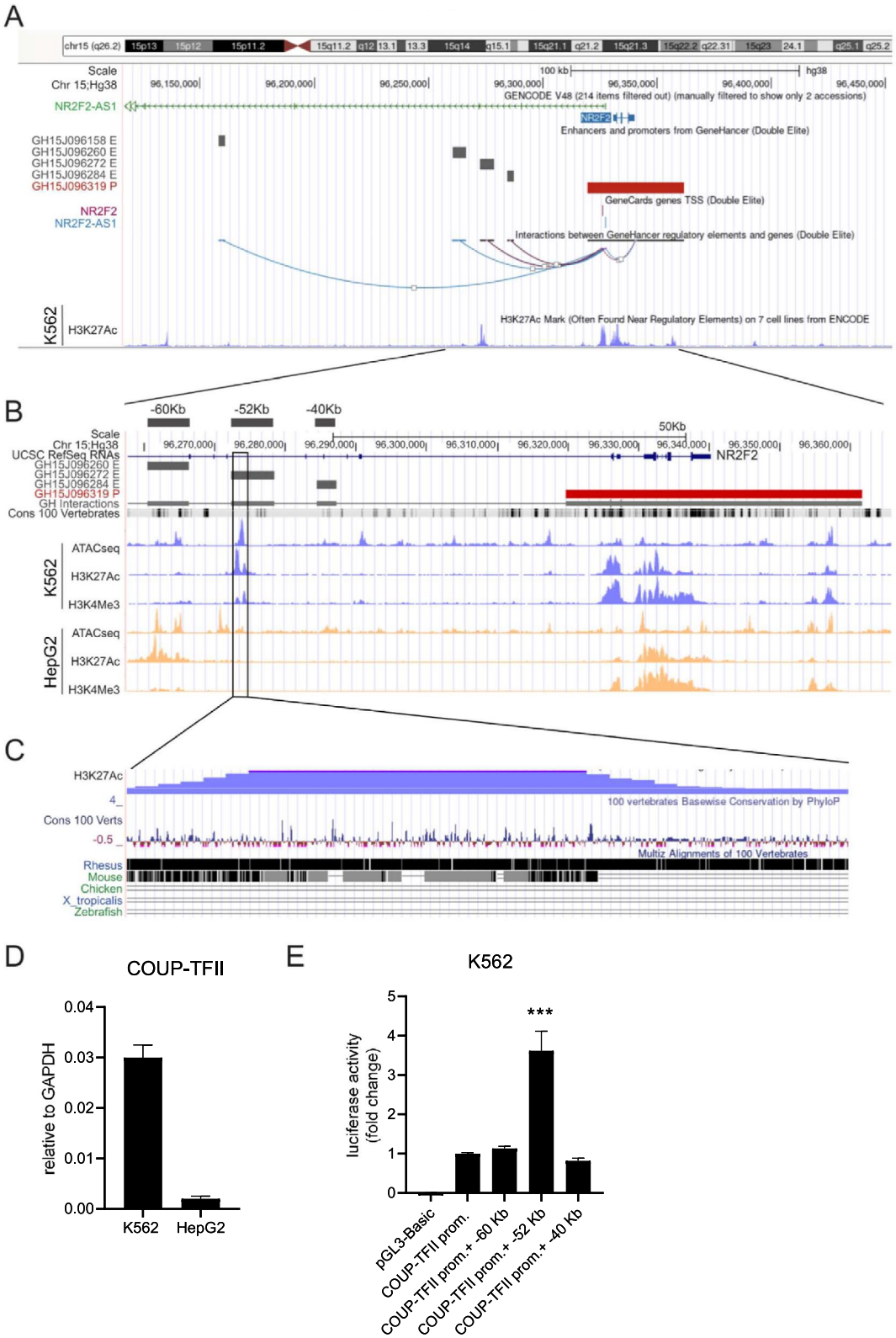
COUP-TFII (encoded by *NR2F2*) is an orphan nuclear receptor highly expressed during organogenesis and development, particularly in the mesenchymal components of various tissues. Knock-out studies in mice reveal that COUP-TFII deletion leads to early embryonic lethality (~E10.5) accompanied by extensive brain and heart hemorrhages[1]. Heterozygous mutant mice exhibit growth restriction and reduced postnatal viability. In contrast, COUP-TFII expression in adult tissues is minimal or absent[2]. Conditional knock-out experiments and functional studies have further demonstrated that COUP-TFII is crucial for venous identity, cardiovascular and brain development[1,3,4], as well as for male and female sexual differentiation[5].

In humans, rare heterozygous mutations in COUP-TFII are associated with a broad spectrum of malformations that significantly overlap with the phenotypes observed in mutant mice. These include heart defects, diaphragmatic hernias, sexual developmental abnormalities, brain malformations, and growth restriction, underscoring the conserved and vital roles of COUP-TFII across species[6].

In line with its early developmental functions, in murine and human erythropoiesis COUP-TFII is expressed in cells of yolk sac origin, before the definitive switch to adult globin expression[7,8]. When ectopically expressed in adult erythroid cells COUP-TFII specifically activates embryonic/fetal globins, effectively overcoming the repressive chromatin environment of adult cells. This characteristic highlights its potential as a therapeutic target for reactivating fetal globin in hemoglobinopathies[7].

The broad developmentally regulated expression profile of COUP-TFII suggests a complex transcriptional regulation of the *NR2F2* gene, likely involving multiple and distinct regulatory regions that remain poorly characterized. Previous studies have shown that the *NR2F2* promoter responds to different signaling pathways

[1]Dipartimento di Biotecnologie e Bioscienze, Università degli Studi di Milano-Bicocca, Milan, Italy. [2]Department of Developmental Biology, Erasmus MC, P.O. Box 2040, 3000CA Rotterdam, The Netherlands. [3]Centre of Biological Sciences, University of Veterinary Medicine Vienna, Vienna, Austria. [4]St. Anna Children's Cancer Research Institute (CCRI), Vienna, Austria. [5]CeMM Research Center for Molecular Medicine of the Austrian Academy of Sciences, Vienna, Austria. [6]Department of Life Science, Health, and Health Professions, Link Campus University, Rome, Italy. [7]International Centre for Translational Eye Research, Warsaw, Poland. [8]Institute of Physical Chemistry, Polish Academy of Sciences, Warsaw, Poland. [9]Present address: Computational Biology Research Centre, Human Technopole Milan, Milan, Italy. [10]Valentina Pastori, Agata Labedz, Maria A. Simanovich and Martina Fabiano contributed equally to this work. ✉email: j.philipsen@erasmusmc.nl; antonella.ronchi@unimib.it



depending on the cellular context: retinoic acid in P19 pluripotent stem cells[9], SHH in neuronal cells[10], and ETS1 during angiogenesis(5). To date, the only tissue-specific enhancer described is located about 60Kb upstream of the transcriptional start site (TSS) and drives *NR2F2* expression in hepatic HepG2 cells[11].

By combining in silico and chromatin accessibility approaches, we identified an *NR2F2* erythroid-specific enhancer located approximately 52Kb upstream of its TSS. This enhancer exhibits hallmark features of an active erythroid regulatory element in K562 cells. Adult erythroid HUDEP2 progenitor cells[12] do not express *NR2F2* and the locus exhibits a closed chromatin conformation. From these cells, we isolated clones that spontaneously express high gamma globin (HbF) and concomitant *NR2F2* reactivation along with −52Kb element opening. In

◀**Fig. 1**. Identification of the −52Kb *NR2F2* erythroid enhancer. (**A**) long-range interactions within the *NR2F2* locus mapped by the GeneHancer database to the reference human genome GRCh38/hg38 in UCSC Genome Browser (https://genome.ucsc.edu/). NR2F2-AS1 is a long non-coding RNA (lncRNA) with no defined role in erythropoiesis. The four candidate regulatory regions detected within −200Kb from the NR2F2 TSS (−164Kb, −60Kb, −52Kb and −40Kb) are shown as grey rectangles and are linked to the promoter (red rectangle) by continuous lines indicating their long-range interactions mapped by GeneHancer. The −60Kb region corresponds to the *NR2F2* candidate liver enhancer identified in Baroukh et al.(11). (**B**) ATAC-seq, H3K27Ac and H3K4Me3 profiles in human erythroid K562 (violet) and hepatic HepG2 (yellow) cells. (**C**) Evolutionary conservation under the −52Kb H3K27Ac peak. (**D**) RT-qPCR: expression levels of COUP-TFII relative to GAPDH in K562 and HepG2 cells (n= 3, ***$p < 0.001$). (**E**) Luciferase reporter experiments in K562 cells show that only the −52Kb region (sequence in Fig. 2A) activates transcription when linked to the *NR2F2* minimal promoter (n≥ 3; ***$p < 0.001$). Dataset numbers are in Data Availability section. The sequences of the primers used to amplify the different regions from K562 genomic DNA are in Table 2.

parental HUDEP2 cells, the knock-out of ZBTB7A, a known gamma-globin repressor, restores active chromatin marks at both the −52Kb enhancer and *NR2F2* promoter.

Results

Identification of a −52Kb enhancer driving *NR2F2* expression in erythroid cells

To identify potential regulatory regions responsible for the expression of *NR2F2* in erythroid cells, we utilized the GeneHancer prediction tool[13]. This tool integrates data from ENCODE, Ensembl, FANTOM, and VISTA to infer long-range interactions between predicted enhancers and target genes. GeneHancer identified four high-confidence putative long-range enhancers located approximately around −164Kb, −60Kb, −52Kb, and −40Kb upstream of the *NR2F2* TSS, which are predicted to interact with the *NR2F2* promoter. Among them, only the −52Kb site exhibited marks of open chromatin, including H3K27ac and H3K4me3 in the erythroleukemia K562 cell line, which are absent in non-erythroid cells (Fig. 1A-C and Suppl. Fig. 1). The −60Kb region likely corresponds to the evolutionarily conserved regions described by Baroukh and colleagues[11] containing a hepatic enhancer. In line with this report, ChIP-seq and ATAC-seq data from ENCODE in the hepatic HepG2 cell line confirmed that chromatin in the −60Kb region is acetylated (H3K27ac) and accessible in this cell line (Fig. 1B)[11]. Nevertheless, we find that COUP-TFII expression levels in HepG2 cells are much lower than those observed in K562 cells as measured by RT-qPCR (Fig. 1 D). Marks of active chromatin and genome accessibility were absent in the −40Kb region in either HepG2 or K562 cells. The −52Kb region showed significant conservation within primates (94.6% identity with Rhesus macaque) and is overall poorly evolutionarily conserved between humans and mouse (Fig. 1C). To test the enhancer activity of the −52Kb element in erythroid cells, we cloned a fragment surrounding the ATAC-seq peak upstream of a minimal 321bp *NR2F2* promoter[9,14] and tested its activity in Luciferase reporter assays in K562 cells. We found that the −52Kb region robustly activated luciferase expression in K562 cells (Fig. 1E), whereas the −60Kb and −40Kb regions failed to do so.

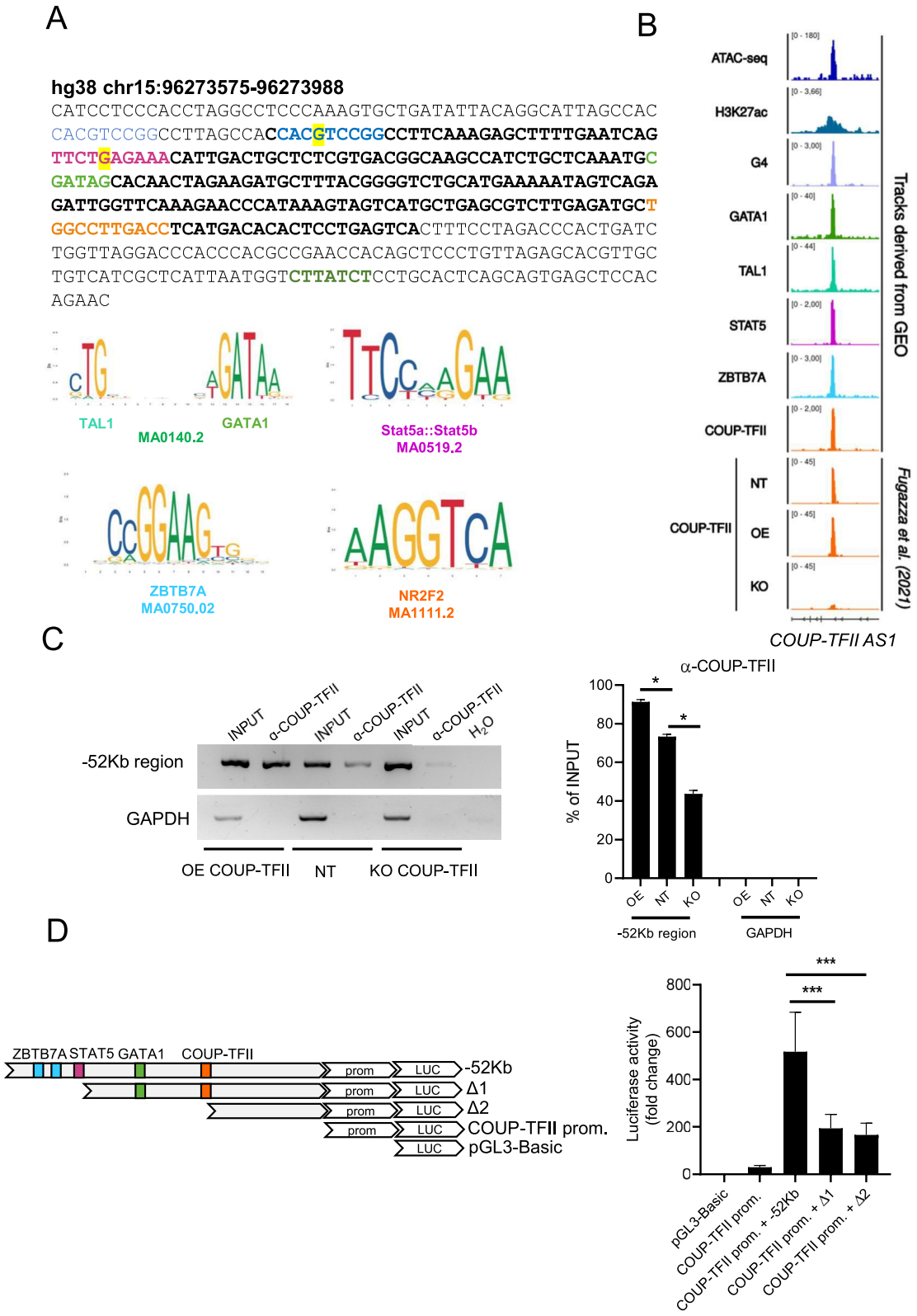
Overall, these features establish the −52Kb region as a candidate enhancer element of the *NR2F2* gene that is specifically active in human erythroid cells.

The −52Kb element possesses an erythroid enhancer signature

The −52Kb element is composed of a core of 217 bp surrounded by repetitive elements (Fig. 2A). According to ENCODE data, chromatin in this region is open (ATAC-seq) and acetylated (H3K27ac) in K562 cells (Fig. 2B). In addition, two independent G-quadruplex (G4) prediction tools identified sequences capable of forming G4 structures within this region (Suppl. Fig. 2). G4 structures are often found at active enhancers and have been proposed to favor the interaction with promoters[15,16]. In support of this hypothesis, CUT&Tag data showed that G4 structures are detected both at the promoter and at the −52Kb element (Suppl. Fig. 2) only in K562 cells, where *NR2F2* is actively transcribed. The −52Kb sequence contains DNA consensus sequences for the major erythroid transcription factors GATA1/TAL1, STAT5A and ZBTB7A, and these factors were bound to this element *in vivo* (ENCODE ChIP-seq data; Fig. 2A, B). Interestingly, the sequence also contains a COUP-TFII consensus site, suggesting auto-regulation (Fig. 2A, B). COUP-TFII binding to this site was confirmed by ChIP-seq data from ENCODE and re-analysis of our data from K562 cells expressing different levels of COUP-TFII (Fig. 2B)[7], as well as by ChIP-qPCR experiments with anti-COUP-TFII antibodies in COUP-TFII overexpressing (OE), untransduced (NT) or knock-out (KO) K562 cells (Fig. 2C). In K562 cells, the activity of a Luciferase reporter driven by the −52Kb element was reduced upon progressive deletions of the fragment, indicating that it contains binding sites that are essential for activator function (Fig. 2D). Overall, this suggests that the activity of the enhancer likely depends on multiple transcription factor binding sites and that the 5' region is crucial for the enhancer activity of the −52Kb element in erythroid cells.

Genomic deletion of the −52Kb element greatly reduces *NR2F2* expression

To evaluate the *in vivo* contribution of the −52Kb element to *NR2F2* transcription in a genomic context, we deleted it by CRISPR/Cas9-mediated genome editing in K562 cells. Using a K562 clone that stably expresses the Cas9 protein[17], we deleted the −52Kb element using viral vectors carrying different combinations of sgRNAs targeting its 5' and the 3' ends. The vectors also contained distinct fluorescent reporter genes (iRFP670 for 5' sgRNA and GFP for 3' sgRNA) for tracking and isolation of targeted cells (Fig. 3A). To ensure independent validation, we employed two distinct 3' sgRNAs (sgRNA2 and sgRNA3) in combination with the same 5' sgRNA (sgRNA1) (Fig. 3B). Following viral infection, double-positive cells were isolated by FACS sorting and COUP-



TFII expression was assessed via RT-qPCR and Western blot analysis in the edited bulk population (Fig. 3C). Deletion of the −52Kb enhancer using either pair of sgRNAs (sg1+sg2 and sg1+sg3) resulted in ≈80% reduction in COUP-TFII expression at both the RNA and protein level (Fig. 3C) as well as a significant reduction of H3K27 acetylation levels at the *NR2F2* promoter, as measured by ChIP (Fig. 3D). Targeted editing of the COUP-TFII binding site (sgMut.COUP-TFII) led to a ~40% decrease in *NR2F2* expression, supporting the relevance of autoregulation through COUP-TFII binding to this site (Fig. 3B, C). Finally, the deletion of the 5'region of the −52Kb element (sg1+sg4) drastically reduced COUP-TFII expression in the edited cells, similarly to longer deletions (Fig. 3B, C).

◀**Fig. 2**. (**A**) Sequence of the human −52Kb *NR2F2* enhancer region. The 217bp core region (in bold) is surrounded by a flanking repetitive region. The putative DNA consensus sequences for relevant transcription factors and their Jaspar matrix (https://jaspar.elixir.no/) are indicated. Mismatches with respect to the canonical site are highlighted in yellow. (**B**) ATAC peaks, G4 structures and binding peaks for various erythroid transcription factors in K562 cells, including COUP-TFII itself (ATAC-seq and ChIP-seq datasets are listed in Data Availability section). The lower panels show COUP-TFII ChIP-seq peaks from cells expressing different levels of COUP-TFII (K562 cells untransduced, NT; overexpressing, OE; knock-out, KO) generated previously(7). (**C**) COUP-TFII ChIP on the −52Kb element in K562 overexpressing (OE), untransduced (NT) or knock-out (KO) for COUP-TFII: representative gel (left) and statistical analysis (right): three biological replicates, each of them in three technical replicates, as analyzed by PCR. The GAPDH locus served as a negative control. Uncropped gel in Suppl. Fig.7C. (**D**) Deletional analysis in K562 cells: the pGL3 plasmid constructs were transiently transfected into K562 cells and luciferase activity was measured 48-72h after transfection ($n \geq 3$; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$). The primers used to generate the different constructs are listed in Table 2.

HUDEP2 clones spontaneously expressing high γ-globin have activated COUP-TFII expression and display an open −52Kb enhancer

HUDEP2 cells are a well-established model of adult human erythroid progenitors, as they exclusively express adult beta globin[12].

During the generation of a HUDEP2 reporter cell line in which the HiBiT tag was inserted in-frame at the C-terminus of the *HBG1* gene (encoding Aγ-globin)[18], we isolated three subclones (clone #10, #25, #28) that spontaneously reactivated HBG (as well as embryonic ε- and ζ-globin genes, Suppl. Fig. 3) at very high levels.

In control HUDEP2 cells and in other clones (such as clone #1) the expression of the embryonic/fetal globin genes was virtually undetectable (Fig. 4A). In all three clones exhibiting embryonic/fetal globins activation (#10, #25, #28), we observed a parallel re-expression of *NR2F2* (RNA-seq tracks) and gain in accessibility of the −52Kb enhancer (ATAC-seq tracks). RT-qPCR and Western blot, using HUDEP1 as γ-globin expressing control cells, confirmed that in these clones the activation of *HBG* expression coincides with COUP-TFII spontaneous re-expression (Fig. 4C). These results are in line with the findings described in Fig. 2 and 3 and further support a role of the −52Kb enhancer in the control of *NR2F2* expression in erythroid cells.

ZBTB7A is a NR2F2 repressor in adult HUDEP2 cells

The −52Kb element contains two DNA consensus sites for ZBTB7A in its 5' region and ZBTB7A binding was observed in ChIP-seq data from K562 cells (Fig. 2A-B and Fig. 5). ZBTB7A is a known repressor of the γ-globin genes. In fetal liver and adult erythroid cells, it shows a mutually exclusive expression pattern with COUP-TFII, consistent with a repressive role in the regulation of embryonic/fetal globin genes (Suppl. Fig. 4 and 5).

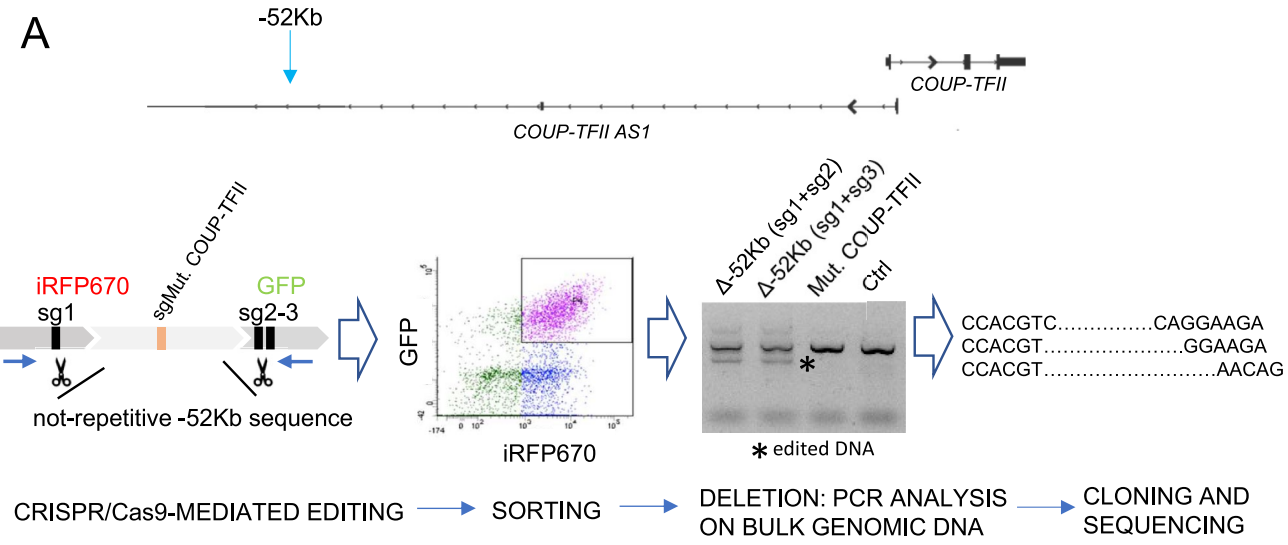
To investigate the potential role of ZBTB7A in regulating the activity of the −52Kb element within its genomic context, we analyzed chromatin accessibility (ATAC-seq) and H3K27Ac profiles from wild-type and ZBTB7A-knock-out HUDEP2 cells[19,20]. In wild-type HUDEP2 cells, which do not express COUP-TFII, the *NR2F2* locus was neither accessible nor acetylated. In ZBTB7A knock-out cells, chromatin in the −52Kb region was accessible, and the promoter acquired H3K27Ac marks, indicating active transcription of *NR2F2* in the absence of ZBTB7A. Moreover, in high-HbF HUDEP2 clones where NR2F2 expression was reactivated (#10, #25, #28, Fig. 4), RNA-seq and RT-qPCR showed reduced ZBTB7A transcription, further supporting a repressive role of ZBTB7A on NR2F2 expression (Fig. 5B-C). A Luciferase reporter construct driven by the −52Kb element, when transfected in K562 cells stably overexpressing ZBTB7A approximately threefold (Supplementary Fig. 6), showed a modest but consistent decrease in activity, indicating that the −52Kb element is a target of ZBTB7A-mediated repression.

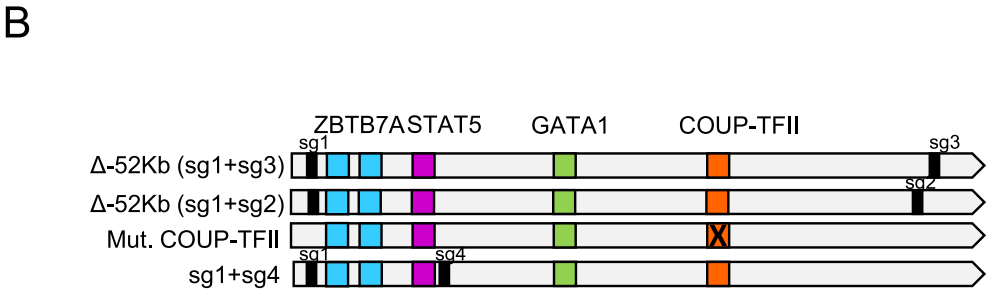
Discussion

Here, we identified an erythroid-specific enhancer located at −52Kb from the *NR2F2* TSS by using independent approaches in two human erythroid cellular model systems: K562 cells, expressing embryonic/fetal globins, and HUDEP2-derived clones spontaneously expressing high levels of HbF. The −52Kb sequence is predicted to establish long-range interaction with the *NR2F2* promoter, is accessible (ATAC-seq) and displays active histone marks (H3K27Ac) in K562 cells, where COUP-TFII is actively transcribed. This element is bound *in vivo* by core erythroid transcription factors and by COUP-TFII itself, pointing to an autoregulatory loop. The −52Kb region also contains a sequence capable of forming G-quadruplexes, non-canonical secondary DNA structures implicated in chromatin looping and transcriptional regulation[21]. When present at both enhancers and promoters, G-quadruplexes were proposed to drive elevated expression of associated genes by providing docking sites for epigenetic modifiers and transcription factors[22]. Notably, CUT&Tag data revealed the simultaneous presence of G4 structures at both the *NR2F2* promoter and the −52Kb region exclusively in K562 cells, supporting the functional relevance of the −52Kb element in regulating COUP-TFII promoter activity. Together, these features confer an erythroid-specific regulatory signature to the −52Kb region, consistent with enhancer activity in erythroid cells which we confirmed functionally by its CRISPR/Cas9-mediated deletion in K562 cells.

The −52Kb region contains two consensus binding sites for ZBTB7A, a protein involved in chromatin organization and remodeling[23,24] essential for erythroid development[25] and fetal globin repression[20,26].

In K562 cells, stable ectopic expression of ZBTB7A reduced the activity of the −52KKb reporter construct. This appears in contrast with the results obtained in parental K562 cells, where the deletion of the ZBTB7A binding sites (Δ1) partially reduced luciferase activity, suggesting an activating role for these sequences. This





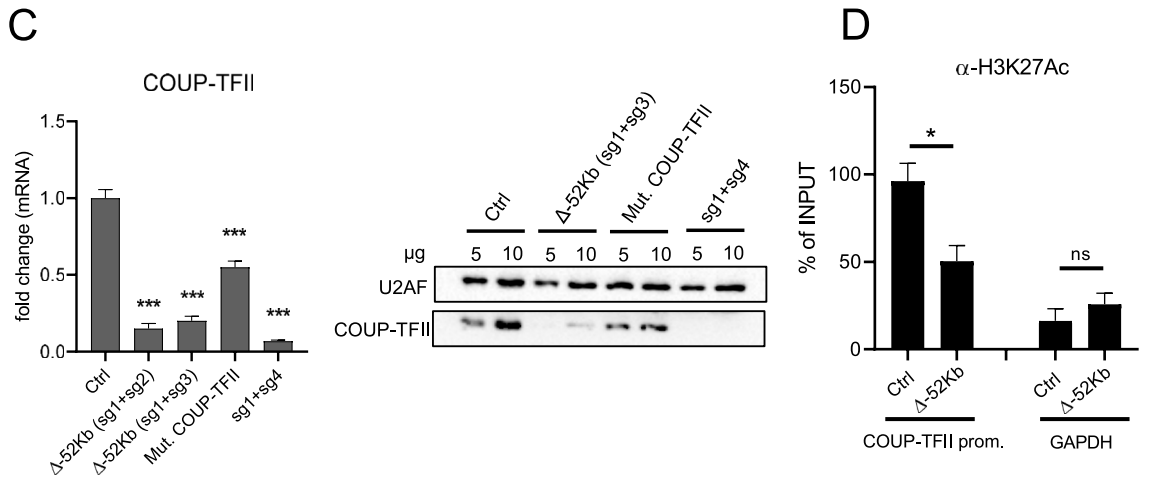


Fig. 3. (**A**) CRISPR editing strategy of the −52Kb region in K562 cells. Deletions and COUP-TFII binding site editing were confirmed by sequencing (details provided in Materials and Methods). Control experiments using a non-targeting AAVS1 sgRNA (Ctrl) produced no significant reduction in COUP-TFII expression, supporting the specificity of the −52Kb element in regulating *NR2F2* transcription. Arrows indicate the position of PCR primers. (**B**) Schematic representation showing the position of the different sgRNAs used (black rectangles). Their sequence is in Table 3. (**C**) COUP-TFII expression levels upon editing with the indicated sgRNAs. RT-qPCR (left): mRNA analysis normalized to non-targeting sgRNA AAVS1 (mean ±SEM***$p < 0.001$); Western blot (right): protein levels (representative gel, uncropped gel in Suppl. Fig. 7A). (**D**) Chromatin IP: H3K27Ac profile at the *NR2F2* promoter in K562 cells wt (ctrl) or deleted for the −52Kb enhancer (sg1+sg3, Δ−52Kb) ($n \geq 3$, *$p < 0.05$; **$p < 0.01$).

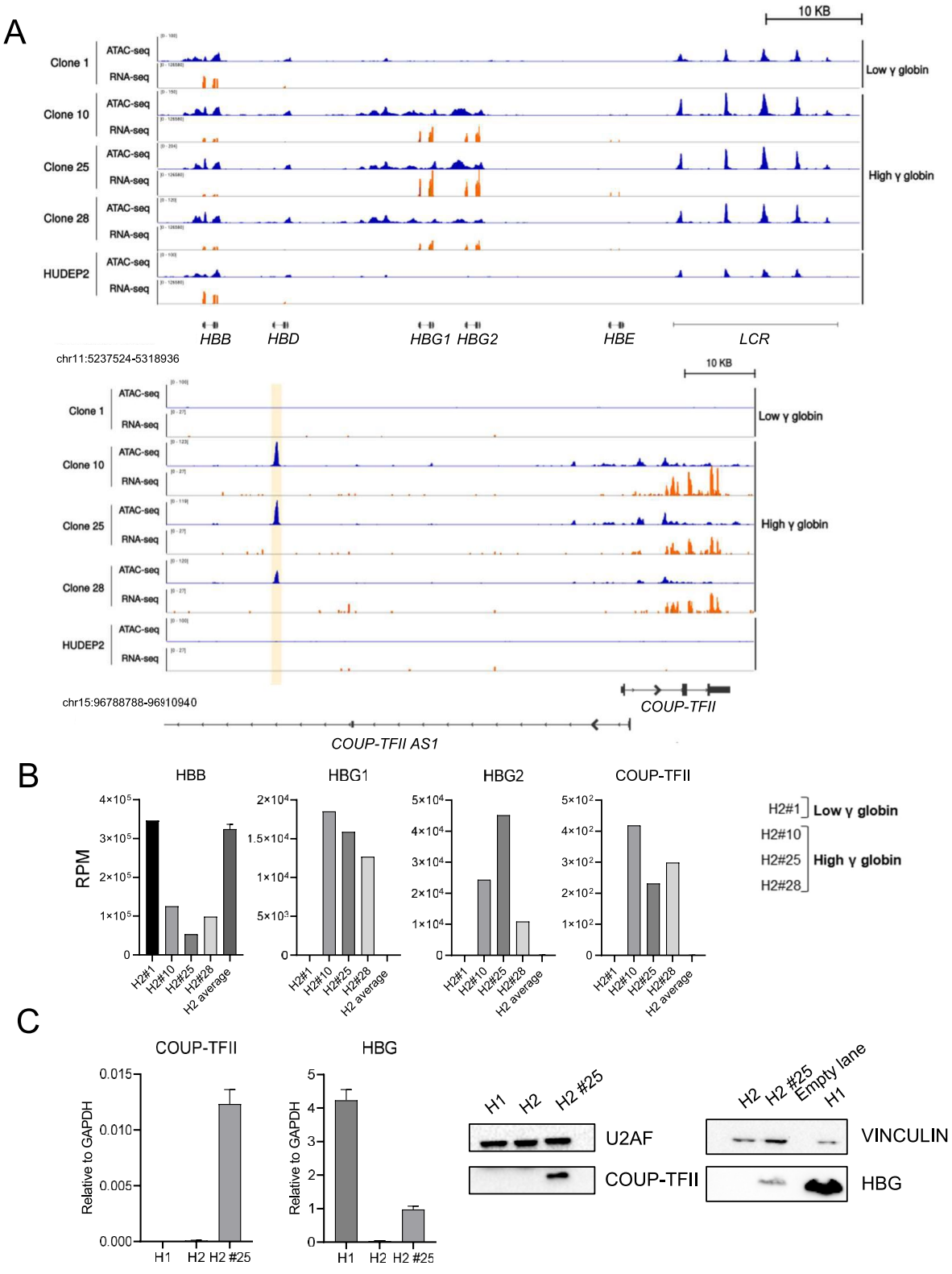


Fig. 4. (**A**) ATAC-seq (blue tracks) and RNA-seq (orange tracks) profiles at the HBB locus (upper panel) and at the *NR2F2* locus (lower panel). HUDEP2 clones #10, #25 and #28: high-HbF clones. Clone #1 and parental HUDEP2 cells: non-HbF expressing control cells. Tracks were loaded on IGV. (**B**) Reads per million (RPM) for HBB, HBG1, HBG2 and NR2F2 in the same single cell-derived clones as in A. For HUDEP2 average levels from 3 independent experiments are shown. (**C**) RT-qPCR (n≥ 3 *p <0.05) and Western blots on COUP-TFII and γ-globin in HUDEP1 (H1), HUDEP2 (H2) and HUDEP2 clone #25 cells. U2AF and Vinculin: loading controls, uncropped gels in Suppl. Fig. 7B.

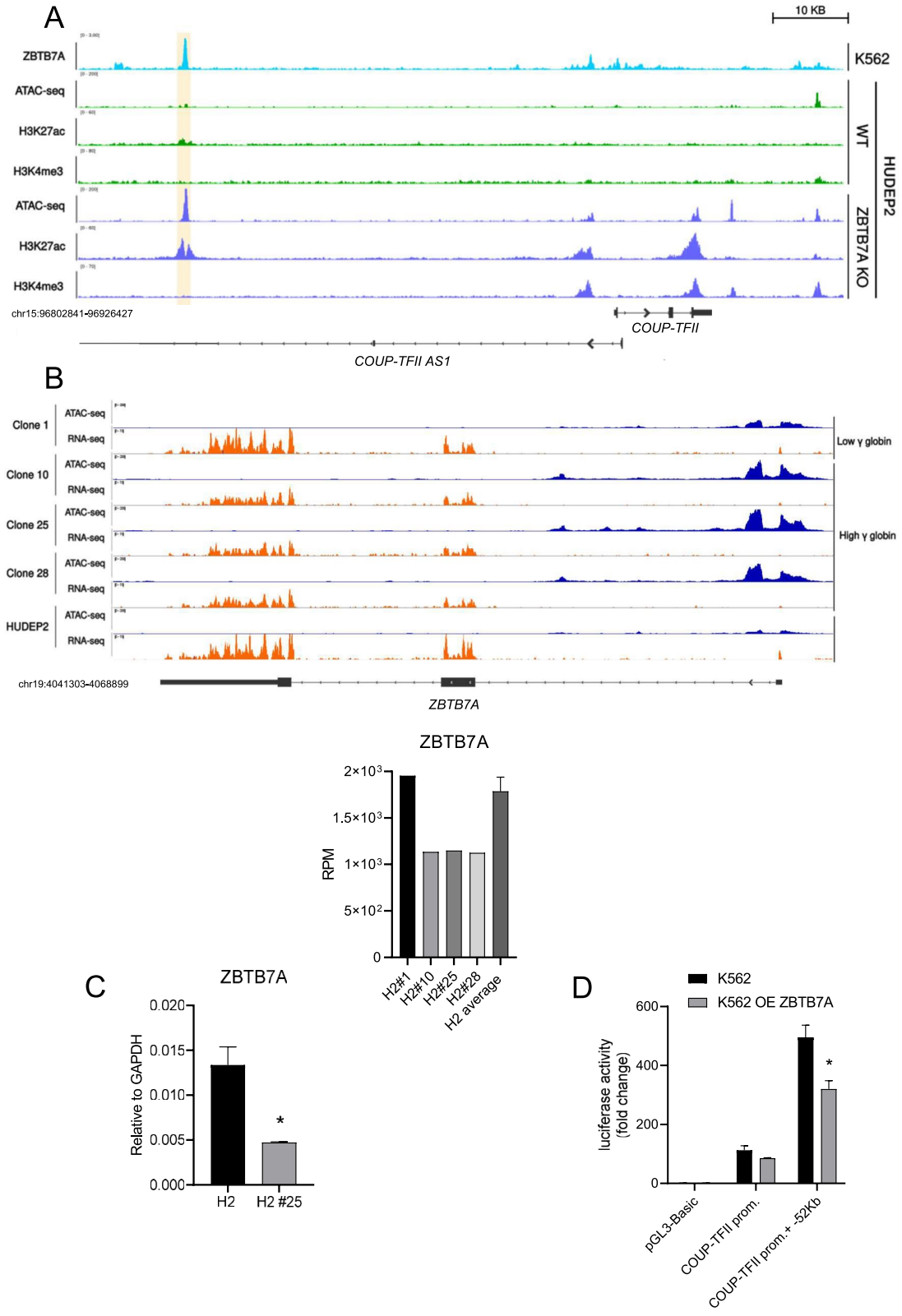


Fig. 5. Chromatin accessibility and occupancy profiles at the *NR2F2* locus in wt cells (K562 and HUDEP2) and in HUDEP2 cells knock-out for ZBTB7A. (**A**) Upper track (light blue): ZBTB7A ChIP-seq data from K562; lower lines: accessibility and chromatin modification in wt (green tracks) and knock-out for ZBTB7A (violet tracks) HUDEP2 cells (**B**) ATAC-seq and RNA-seq profiles at the ZBTB7A locus in high-HbF clones and in control HUDEP2 cells and clone #1. The corresponding RPM are shown below tracks. (**C**) RT-qPCR showing ZBTB7A expression relative to GAPDH in HUDEP2 and in clone #25 (n≥ 3 *p <0.05). (**D**) Luciferase reporter assay in K562 cells, either wild-type or overexpressing ZBTB7A at ≈ three times the endogenous level (Suppl. Fig. 6). The reporter constructs are described in Fig. 2D. n≥ 3, *p < 0.05.

discrepancy likely reflects the inability of luciferase assays to recapitulate endogenous transcriptional regulation in these two distinct experimental cellular models. An additional explanation is that deletions, by perturbing the architectural structure of the −52Kb region, could have effects not directly superimposable to those of the ZBTB7A knockout.

HUDEP2 cells express adult globin genes and do not express *NR2F2*. In clonal HUDEP2 lines with spontaneous high-HbF expression[18], we found reactivation of *NR2F2* expression. This was accompanied by the opening of chromatin at the −52Kb enhancer. These observations strongly support the essential role of the −52Kb enhancer in the regulation of COUP-TFII expression in erythroid cells. ZBTB7A expression level is reduced in high-HbF HUDEP2 clones, where NR2F2 expression is reactivated. Consistently, we found that in parental HUDEP2 cells knock-out for ZBTB7A, the −52Kb *NR2F2* enhancer became accessible and acetylated and its promoter acquired strong H3K4Me3 enrichment, indicating a transcriptionally active state.

Overall, our work identifies an erythroid-specific enhancer that controls *NR2F2* transcriptional activation in erythroid cells. Given the ability of COUP-TFII to reactivate γ-globin expression in adult cells, the identification of the molecular mechanisms governing *NR2F2* expression in erythroid cells is a prerequisite for attempts to force COUP-TFII-mediated reactivation of γ-globin expression in adult erythroid cells.

Materials & methods

Plasmid construction

The *NR2F2* promoter and the upstream regulatory sequences were amplified from K562 cells genomic DNA by using Q5 High-Fidelity DNA Polymerase. Oligonucleotides used for the amplification were designed by using the Bisearch tool (http://bisearch.enzim.hu/?run). Flanking restriction sites were added to primers for further cloning into the pGL3-Basic vector (Promega). Clones with deleted regions of the −52Kb element were obtained by NEBuilder HiFi DNA Assembly procedure. The sequence of all primers used is listed in Table 2.

Cells

K562 cells[27] are a human immortalized erythroleukemia embryonic-like cell line (ATCC-CCL-243). They were grown in RPMI+5%FBS. HUDEP1 and HUDEP2 are human immortalized erythroid progenitor cells[12]. They were grown in StemSpan-SFEM +75ng/ml SCF, 1μM Dexamethasone, 2U/ml Epo, 1μg/ml Doxycyclin or in-house prepared Cellquin medium[18]. HUDEP2 clones #1, #10,#25, #28 were obtained while generating the HUDEP2 reporter cell line in which the HiBiT tag was inserted in-frame at the C-terminus of the *HBG1*gene, as detailed in ref[18].

RT-qPCR

Total RNA from ≥10^5 cells was purified with TRIzol Reagent and Direct-zol RNA MiniPrep kit (Euroclone) and retrotranscribed (High Capacity cDNA Reverse Transcription Kit, Applied Biosystems). Real time analysis was performed on a StepOne instrument (Applied Biosystems) by using SsoAdvanced Universal SYBR Green Supermix (Bio-Rad) fluorescent dye. Primers are listed in Table 2.

Western blots

Total and nuclear extracts were prepared according to standard protocols. Protein extracts (10–15 μg/lane) were resolved by SDS/PAGE in a 10–15% acrylamide gel and blotted onto Hybond-ECL Nitrocellulose membrane (GE healthcare). Membranes were blocked, incubated with the appropriate antibodies and after washing, ECL reagent (Millipore) was used for detection. Antibodies are listed in Table 1.

Transfection experiments and Luciferase assays

3x10^5 cells were seeded in a 24 -well plate in 0.5mL of Opti-MEM medium. Cells were transfected by using 0.8μg of DNA of reporter plasmids and 2μL of Lipofectamine 2000 in 100μL of Opti-MEM per well. 4 h after transfection, 0.6mL of RPMI+20% FBS per well was added. Cells were collected 48-72h after transfection, pellets were re-suspended in 80μL of Reporter Lysis Buffer 1X (Promega kit #E4030), followed by two freeze-thaw cycles. Luciferase activity was measured by using the Promega kit in a GloMax 20/20 Luminometer. Transfection experiments were done in biological (n≥ 3) and technical (n= 3) replicates, using at least two independent DNA preparations. To generate K562 cells stably overexpressing ZBTB7A, cells were transfected with a pcDNA3.1-ZBTB7A expressing vector (gift from Drs. Saccani and Van Essen(24)). 48 h after transfections neomycin was added and cells were selected for 15 days (Supplemental Fig. 6).

ChIP and ChIP-seq assay and analysis

The detailed protocol is described in ref. (7). Briefly, K562 expressing physiological levels (NT), overexpressing (OE), knock–out (KO) for *NR2F2* or Δ−52Kb K562 were fixed with 1% formaldehyde for 10 minutes at room temperature. Chromatin was sonicated to a size of about 500bp. Immunoprecipitation (10^6cells/sample) was performed after overnight incubation with anti-COUP-TFII anti-H3K27Ac or isotypic IgG as a control.

Immunoprecipitated DNA was prepared, analyzed by PCR or sequenced on an Illumina platform and data were uploaded to the Galaxy web platform (https://usegalaxy.org/).

CRISPR/Cas9-mediated targeting

Cloning of the guide RNAs, lentiviral production and transduction were performed as described(17). Briefly, single guide RNAs (Table 2,3) were designed using CRISPR Targets track at the UCSC Genome Browser (https://genome.ucsc.edu/) and cloned into a lentiviral expression vector containing GFP (LentiGuide-Puro-P 2A-EGFP, Addgene, 137729) or into a lentiviral expression vector containing iRFP670 (pLenti hU6-sgRNA-IT-PGK-iRFP670). For lentiviral production, Lenti-X cells were transfected with indicated constructs together

Antibodies and reagents	Cat n°	Manufacturer
Anti-COUP-TFII	ab41859	Abcam
Anti U2AF	U4758	Sigma-aldrich
Anti HBG (γ-globin)	ab283313	Abcam
Anti β actin	5125	Cell signaling
rIgG	PP64	Millipore
Anti H3K27Ac	ab4729	Abcam
Opti-MEM	31985-047	Thermofisher
Lipofectamine 2000	11668019	Invitrogen
Luciferase assay kit	3040	Promega
RPMI 1640	ECB9006L	Euroclone
Phosphate-buffered saline (PBS)	ECB4004L	Euroclone
Fetal bovine serum (FBS)	F7524	Sigma-aldrich
L-glutamine	ECB3000D	Euroclone
Penicillin-streptomycin	ECB3001D	Euroclone
High capacity cDNA RT Kit	4368814	Applied biosystems
SsoAdvanced universal SYBR green	1725274	Bio-Rad
StemSpan SFEM	09650	Voden medical instruments
rh SCF	11343325	ImmunoTools
rh EPO	11344795	ImmunoTools
dexamethasone	D4902	Sigma-Aldrich
doxycycline	D9891	Sigma-Aldrich
Cellquin medium	P04-20251K	PAN biotech
NEBuilder HiFi DNA As sembly Cloning Kit	E5520S	New England Biolabs
Q5 high-fidelity DNA polymerase	M0491S	New England Biolabs

Table 1. List of antibodies and reagents.

with psPAX2 and pMD2.G packaging plasmids (Addgene 12260 and 12259) using polyethyleneimine (PEI, Polysciences Europe, 24765-1). After 24 h, Lenti-X cell medium was substituted with the target cell medium. The supernatant containing lentiviral pseudo-particles was collected 48 h after transfection. Next, approximately $3x10^6$ K562 cells stably expressing Cas9 (K562-Cas9) were seeded per well of a 6-well plate and transduced by spinfection (1000 rcf, 90 min) with lentiviral supernatants (diluted 1:2) supplemented with Polybrene Transfection Reagent (Merck, TR-1003-G). Cells were collected 6 days after transduction and sorted for GFP and/or iRFP670 expressing cells, followed by RNA and genomic DNA extraction. PCR on genomic DNA with primers flanking the −52Kb region (Table 2) was used to verify the presence of the −52Kb deletion at the different steps as outlined in Fig. 3A. Editing of the COUP-TFII site was obtained designing a primer (sgMut.COUP-TFII, Table 3) that uses the two CC within the COUP-TFII consensus as PAM sequence.

Primers used for RT-qPCR		
GENE	F/R	Sequence (5'–3')
GAPDH	F	ACGGATTTGGTCGTATTGGG
	R	TGATTTTGGAGGGATCTCGC
HBG (γ-globin)	F	CTTCAAGCTCCTGGGAAATGT
	R	GCAGAATAAAGCCTATCTTGAAAG
COUP-TFII	F	TTGACTCAGCCGAGTACAGC
	R	AAAGCTTTCCGAATCTCGTC
Primers used for amplification from genomic K562 DNA and for cloning into pGL3 luciferase reporter vector		
	F/R	Sequence (5'–3')
Region –60Kb	F	ACGTGCTAGCGGCACAAATCGCTCATTAGC
	R	ACGTGTCGACTGCTTAACTCAAGCCTCAGAAAGC
Region –40Kb	F	ACGTGCTAGCAACACACAGCATCGTGCATG
	R	ACGTGTCGACTATGCATACACTGTTTCCTCATG
Region –52Kb	F	ACGTGCTAGCGTCACTGCAGCCTTTACCCC
	R	ACGTGTCGACGGAGGCATGGATTATACTCG
NR2F2 promoter	F	ACGTAAGCTTCCGCTCGCGCTAGGACC
	R	ACGTAAGCTTGAAAAGAACAGAGAATCAAAGTGATC
Δ1	F	GAGCTCTTACGCGTGCTAGCCATTGACTGCTCTCGTG
	R	CTTACTTAGATCGCAGATCTCGAGGGAGGCATGGATTATAC
Δ2	F	GAGCTCTTACGCGTGCTAGCTCATGACACACTCCTG
	R	CTTACTTAGATCGCAGATCTCGAGGGAGGCATGGATTATAC
Primers used to amplify the indicated target regions in chromatin immunoprecipitation experiments		
Region –52Kb ChIP	F	CGTCCGGCCTTCAAAGAG
	R	GAGCGATGACAGCAACG
GAPDH ChIP	F	CGGAGTCAACGGATTTGGTCGTAT
	R	AGCCTTCTCCATGGTGGTGAAGAC
NR2F2 prom. ChIP	F	GTTCGTGGCACCTTCCCACC
	R	GCCTCTCGCAACTTGGCG

Table 2. List of primers.

sgRNA	Sequence (5'–3')
sg1	AGCTCTTTGAAGGCCGGACG
sg2	TGGAGCTCACTGCTGAGTGC
sg3	TGCATGCTCTTCCTGTTCTG
sg4	AACATTGACTGCTCTCGTGA
sgMut.COUP-TFII	TGACTCAGGAGTGTGTCATG

Table 3. List of sgRNAs used for the editing of the –52Kb enhancer in K562 cells (Fig. 3).

Data availability

The datasets analysed during the current study are available in: K562: from ENCODE REGULATION,/gbdb/hg38/bbi/wgEncodeReg/wgEncodeRegMarkH3k27ac/wgEncodeBroadHistoneK562H3k27acStdSig.bigWig,/gbdb/hg38/bbi/wgEncodeReg/wgEncodeRegMarkH3k4me3/wgEncodeBroadHistoneK562H3k4me3StdSig.bigWig, from ENCODE PROJECT, ENCFF252GZO, HepG2: from GEO Gene Expression Omnibus, GSE170012, GSM5597548, GSM5597550 (Data in Fig. 1) K562: from GEO Gene Expression Omnibus, GSM7974396, GSM8292067, GSM5501186, GSM4818700, GSM6463414, GSM1010877, GSM2771534, GSM1010782 (Data in Fig. 2). HUDEP2: from European Nucleotide Archive (ENA), PRJEB67342, PRJEB31728, from EMBL-EBI ArrayExpress, E-MTAB-15779, E-MTAB-15772 (Data in Fig. 4) K562: from GEO Gene Expression Omnibus, GSM2771534, HUDEP2: from GEO Gene Expression Omnibus, GSM5267234, GSM5267250, GSM5267246, GSM5267230, GSM5267249, GSM5267245 (Data in Fig. 5) K562: from GEO Gene Expression Omnibus, GSM7974396, GSM8292067, GSM5501186, U2OS: from GEO Gene Expression Omnibus, GSM5501189, MCF7: from GEO Gene Expression Omnibus, GSM5501223 (Data in Suppl. Fig. 2).

Received: 24 October 2025; Accepted: 25 March 2026
Published online: 01 April 2026

References

1. Pereira, F. A., Qiu, Y., Zhou, G., Tsai, M. J. & Tsai, S. Y. The orphan nuclear receptor COUP-TFII is required for angiogenesis and heart development. *Genes Dev.* **13**(8), 1037–1049 (1999).
2. Pereira, F. A., Qiu, Y., Tsai, M. J. & Tsai, S. Y. Chicken ovalbumin upstream promoter transcription factor (COUP-TF): Expression during mouse embryogenesis. *J. Steroid Biochem. Mol. Biol.* **53**(1–6), 503–508 (1995).
3. You, L. R. et al. Suppression of Notch signalling by the COUP-TFII transcription factor regulates vein identity. *Nature* **435**(7038), 98–104 (2005).
4. You, L. R. et al. Mouse lacking COUP-TFII as an animal model of Bochdalek-type congenital diaphragmatic hernia. *Proc. Natl. Acad. Sci. U.S.A.* **102**(45), 16351–16356 (2005).
5. Petit, F. G., Salas, R., Tsai, M. J. & Tsai, S. Y. The regulation of COUP-TFII gene expression by Ets-1 is enhanced by the steroid receptor co-activators. *Mech. Ageing Dev.* **125**(10–11), 719–732 (2004).
6. Al Turki, S. et al. Rare variants in NR2F2 cause congenital heart defects in humans. *Am. J. Hum. Genet.* **98**(3), 592 (2016).
7. Fugazza, C., Barbarani, G., Elangovan, S., Marini, M.G., Giolitto, S., Font-Monclus, I. et al. The Coup-TFII orphan nuclear receptor is an activator of the gamma-globin gene. *Haematologica* (2021).
8. Frigo, C., Pastori, V., Zambanini, G., Fabiano, M., Ahmed, S., Citterio, E. et al. COUP-TFII regulates hemoglobin switching by activating the BCL11A-XL repressor Lin28B and directly binding delta and beta globin promoters in fetal versus adult erythroid cells. *Haematologica* (2025).
9. Qiu, Y., Krishnan, V., Pereira, F. A., Tsai, S. Y. & Tsai, M. J. Chicken ovalbumin upstream promoter-transcription factors and their regulation. *J. Steroid. Biochem. Mol. Biol.* **56**(1-6 Spec No), 81–85 (1996).
10. Krishnan, V., Elberg, G., Tsai, M. J. & Tsai, S. Y. Identification of a novel sonic hedgehog response element in the chicken ovalbumin upstream promoter-transcription factor II promoter. *Mol. Endocrinol.* **11**(10), 1458–1466 (1997).
11. Baroukh, N. et al. Comparative genomic analysis reveals a distant liver enhancer upstream of the COUP-TFII gene. *Mamm Genome.* **16**(2), 91–95 (2005).
12. Kurita, R. et al. Establishment of immortalized human erythroid progenitor cell lines able to produce enucleated red blood cells. *PLoS ONE* **8**(3), e59890 (2013).
13. Fishilevich, S., Nudel, R., Rappaport, N., Hadar, R., Plaschkes, I., Iny Stein, T. et al. GeneHancer: genome-wide integration of enhancers and target genes in GeneCards. *Database (Oxford)* (2017).
14. Jonk, L. J., de Jonge, M. E., Vervaart, J. M., Wissink, S. & Kruijer, W. Isolation and developmental expression of retinoic-acid-induced genes. *Dev. Biol.* **161**(2), 604–614 (1994).
15. Varshney, D., Spiegel, J., Zyner, K., Tannahill, D. & Balasubramanian, S. The regulation and functions of DNA and RNA G-quadruplexes. *Nat. Rev. Mol. Cell Biol.* **21**(8), 459–474 (2020).
16. DeMeis, J. D. et al. Long G4-rich enhancers target promoters via a G4 DNA-based mechanism. *Nucl. Acid. Res.* https://doi.org/10.1093/nar/gkae1180 (2025).
17. Proietti, L., Manhart, G., Heyes, E., Troester, S., Grebien, F. Arrayed CRISPR/Cas9 screening for the functional validation of cancer genetic dependencies. *Bio. Protoc.* **12**(24), (2022).
18. Verheul, T. C. J. et al. A cellular reporter system to evaluate endogenous fetal hemoglobin induction and screen for therapeutic compounds. *Hemasphere.* **8**(8), e139 (2024).
19. King, A. J. et al. Reactivation of a developmentally silenced embryonic globin gene. *Nat. Commun.* **12**(1), 4439 (2021).
20. Masuda, T. et al. Transcription factors LRF and BCL11A independently repress expression of fetal hemoglobin. *Science* **351**(6270), 285–289 (2016).
21. Roy, S. S. et al. Artificially inserted strong promoter containing multiple G-quadruplexes induces long-range chromatin modification. *Elife* https://doi.org/10.7554/eLife.96216 (2024).
22. Antariksa, N. F. & Di Antonio, M. The emerging roles of multimolecular G-Quadruplexes in transcriptional regulation and chromatin organization. *Acc. Chem. Res.* **57**(23), 3397–3406 (2024).
23. Constantinou, C. et al. The multi-faceted functioning portrait of LRF/ZBTB7A. *Hum. Genom.* **13**(1), 66 (2019).
24. Ramos Pittol, J. M., Oruba, A., Mittler, G., Saccani, S. & van Essen, D. Zbtb7a is a transducer for the control of promoter accessibility by NF-kappa B and multiple other transcription factors. *PLoS Biol.* **16**(5), e2004526 (2018).
25. Maeda, T. et al. LRF is an essential downstream target of GATA1 in erythroid development and regulates BIM-dependent apoptosis. *Dev. Cell.* **17**(4), 527–540 (2009).
26. Maeda, T. Regulation of hematopoietic development by ZBTB transcription factors. *Int. J. Hematol.* **104**(3), 310–323 (2016).
27. Lozzio, B. B. & Lozzio, C. B. Properties and usefulness of the original K-562 human myelogenous leukemia cell line. *Leuk Res.* **3**(6), 363–370 (1979).

Acknowledgments

We thank Rocco Piazza for helping in the preliminary in silico analysis and Sajeel Ahmed for sharing of results.

Author contributions

V.P., A.L., M.A.S., M.F., C.F., T.V., L.P. performed experiments, data analysis and data interpretation. E.C., S.P., F.G. conceived the experiments and interpreted data. E.C., S.P. and A.E.R. conceived the experiments, interpreted data and wrote the manuscript.

Funding

This work was supported by the European Union's Horizon 2020 Research and Innovation Program under the Marie Skłodowska Curie grant agreement No. 813091 (A.E.R. and F.G.), by MUR - Ministero dell'Università e della Ricerca (PRIN: 2021-NAZ-0253) (A.E.R.) and by the European Union - NextGenerationEU through the Italian Ministry of University and Research under PNRR - M4C2-I1.3 Project PE_00000019 "HEAL ITALIA" to A.E.R. (H43C22000830006. The views and opinions expressed are those of the authors only and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the European Commission can be held responsible for them. Work in the laboratory of SP was supported by the ZonMw PSIDER programme (TRACER consortium, grant 10250022110001, M.A.S. and S.P.), and NWO Applied and Engineering Sciences Open Technology Programme (grant 18947, T.V. and S.P.).

Declarations

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/s41598-026-46308-0.

Correspondence and requests for materials should be addressed to S.P. or A.E.R.

Reprints and permissions information is available at www.nature.com/reprints.

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