Molecular Cell, Volume *81*

Supplemental Information

ZNF410 Uniquely Activates the NuRD Component CHD4 to Silence Fetal Hemoglobin Expression

Xianjiang Lan, Ren Ren, Ruopeng Feng, Lana C. Ly, Yemin Lan, Zhe Zhang, Nicholas Aboreden, Kunhua Qin, John R. Horton, Jeremy D. Grevet, Thiyagaraj Mayuranathan, Osheiza Abdulmalik, Cheryl A. Keller, Belinda Giardine, Ross C. Hardison, Merlin Crossley, Mitchell J. Weiss, Xiaodong Cheng, Junwei Shi, and Gerd A. Blobel

Figure S1

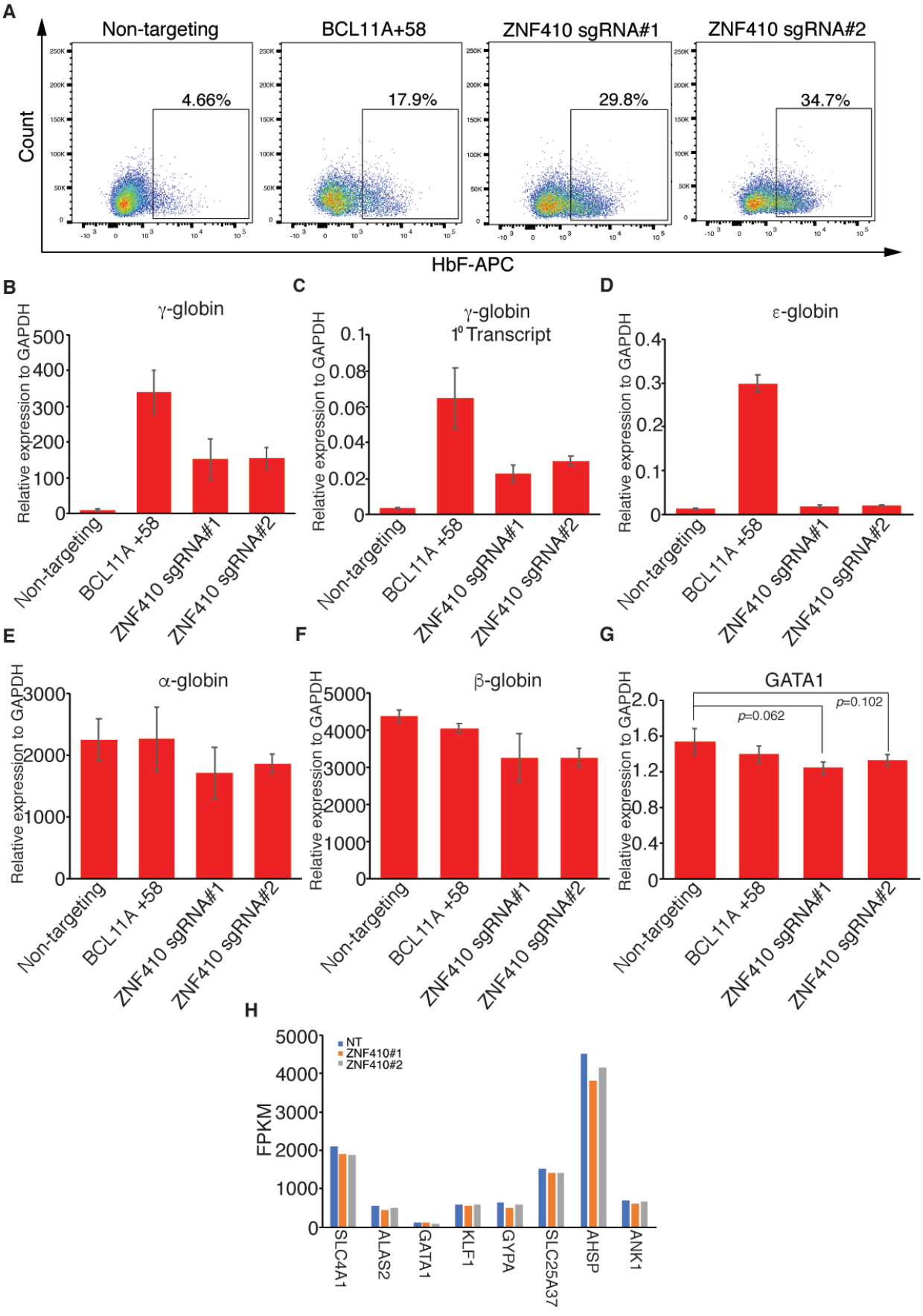
A
Non-targeting
BCL11A+58
ZNF410 sgRNA#1
ZNF410 sgRNA#2
4.66%
17.9%
29.8%
34.7%
Count
HbF-APC
B
γ-globin
Relative expression to GAPDH
C
γ-globin
1⁰ Transcript
D
ε-globin
E
α-globin
F
β-globin
G
GATA1
p=0.062
p=0.102
Non-targeting
BCL11A +58
ZNF410 sgRNA#1
ZNF410 sgRNA#2
H
NT
ZNF410#1
ZNF410#2
FPKM
SLC4A1
ALAS2
GATA1
KLF1
GYPA
SLC25A37
AHSP
ANK1

Figure S1. HbF flow cytometric analysis and RT-qPCR in HUDEP-2 cells, related to Figure 1

(A) Representative flow cytometric analysis of differentiated HUDEP-2 cells stained with anti-HbF antibody. Positive control: sgRNA against BCL11A +58: Negative control: Non-targeting sgRNA.

(B-G) mRNA levels of γ-globin, γ-globin primary transcripts, ε-globin, α-globin, β-globin and GATA1 by RT-qPCR. Results are shown as mean ± SD (n=3). GAPDH was used for normalization. 1^0 transcript: primary transcript. P values were calculated by Prism (GraphPad) with unpaired student's *t*-test.

(H) Expression levels of SLC4A1(Band3), ALAS2, GATA1, KLF1, GYPA, SLC25A37, AHSP and ANK1 in differentiated HUDEP-2 cells transduced with indicated sgRNAs by RNA-seq analysis. NT: non-targeting.

Figure S2

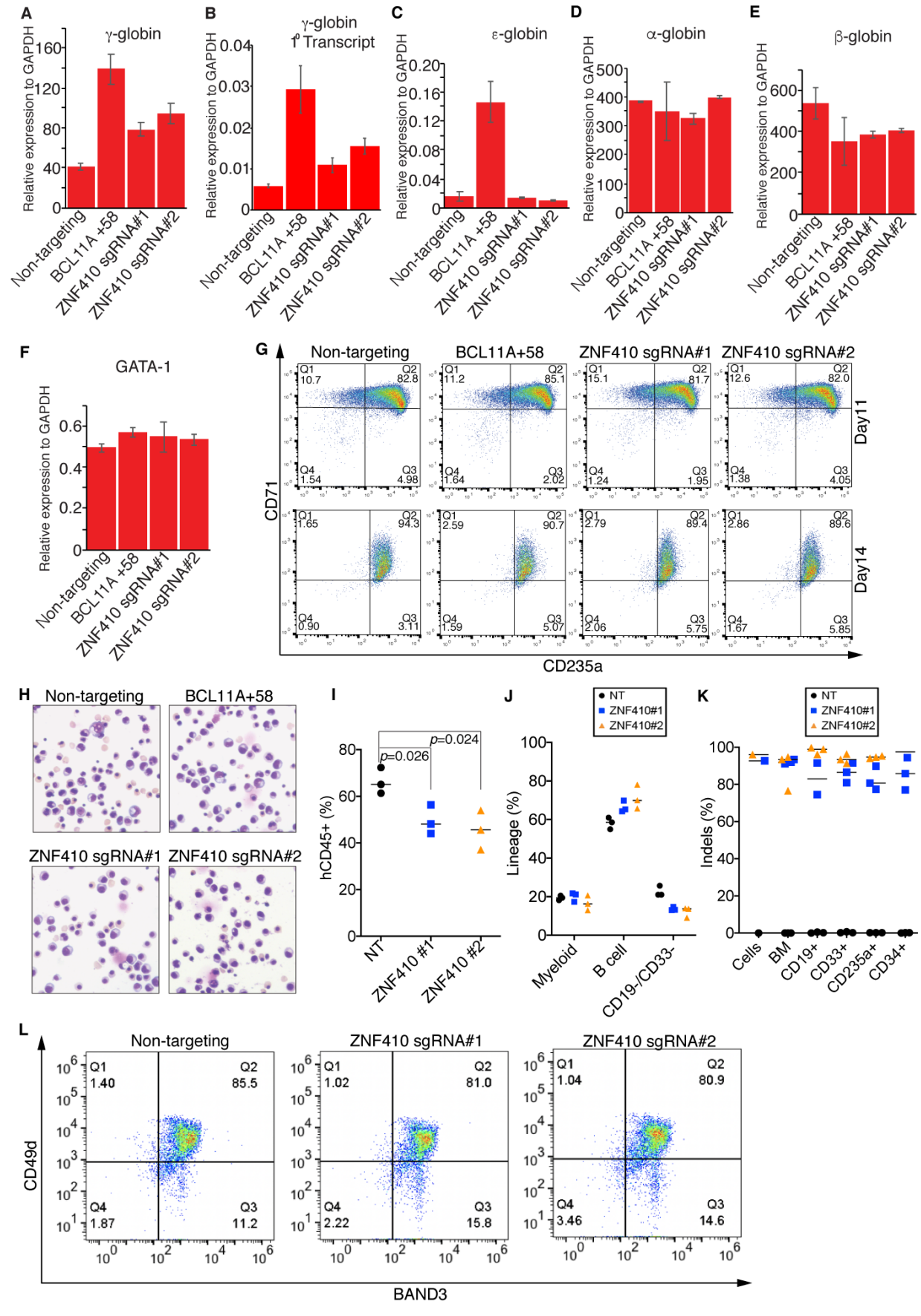
A
γ-globin
Relative expression to GAPDH
Non-targeting
BCL11A +58
ZNF410 sgRNA#1
ZNF410 sgRNA#2
B
γ-globin
1° Transcript
C
ε-globin
D
α-globin
E
β-globin
F
GATA-1
G
Non-targeting
BCL11A+58
ZNF410 sgRNA#1
ZNF410 sgRNA#2
Day11
Day14
CD71
CD235a
H
Non-targeting
BCL11A+58
ZNF410 sgRNA#1
ZNF410 sgRNA#2
I
hCD45+ (%)
p=0.024
p=0.026
NT
ZNF410 #1
ZNF410 #2
J
NT
ZNF410#1
ZNF410#2
Lineage (%)
Myeloid
B cell
CD19-/CD33-
K
Indels (%)
Cells
BM
CD19+
CD33+
CD235a+
CD34+
L
Non-targeting
ZNF410 sgRNA#1
ZNF410 sgRNA#2
CD49d
BAND3

Figure S2. ZNF410 depletion induces HbF with minimal impact on erythroid maturation in vitro and in vivo, related to Figure 2

(A-F) mRNA levels of γ-globin, γ-globin primary transcripts, ε-globin, α-globin, β-globin and GATA1 by RT-qPCR in cultured primary human erythroblasts on day 12 of differentiation. Positive control: sgRNA against BCL11A +58: Negative control: Non-targeting sgRNA. Results are shown as mean ± SD (n=3). GAPDH was used for normalization.

(G) Representative flow cytometric analysis of erythroid maturation markers CD71 and CD235a in cultured primary human erythroblasts on day 11 and day 14 of differentiation.

(H) Wright-Giemsa staining in cultured primary human erythroblasts at day 16 of differentiation.

(I) Normalized human chimerism in bone marrow from NBSGW mice at 16 weeks after transplantation, shown as percentage of human (h) CD45+ cells. P values were calculated by Prism (GraphPad) with unpaired student's *t*-test.

(J) Human myeloid (CD33+), B cells (CD19+) and other cell types (CD19-/CD33-) shown as percentages of the human CD45+ population in bone marrow (BM) from NBSGW mice at 16 weeks after transplantation.

(K) Indels measured by next generation sequencing (NGS) in input cells, bone marrows (BM) and specific hematopoietic lineages including B cells (CD19+), myeloid (CD33+), erythroblasts (CD235a+) and hematopoietic stem cells (CD34+).

(L) Representative flow cytometric analysis of erythroid maturation markers CD49d/Band3 in human CD235a+ erythroblasts from recipient mouse bone marrow. For (I-K) panels, each dot in graphs represents a separate mouse. n=3 mice per sgRNA. NT: non-targeting.

Figure S3

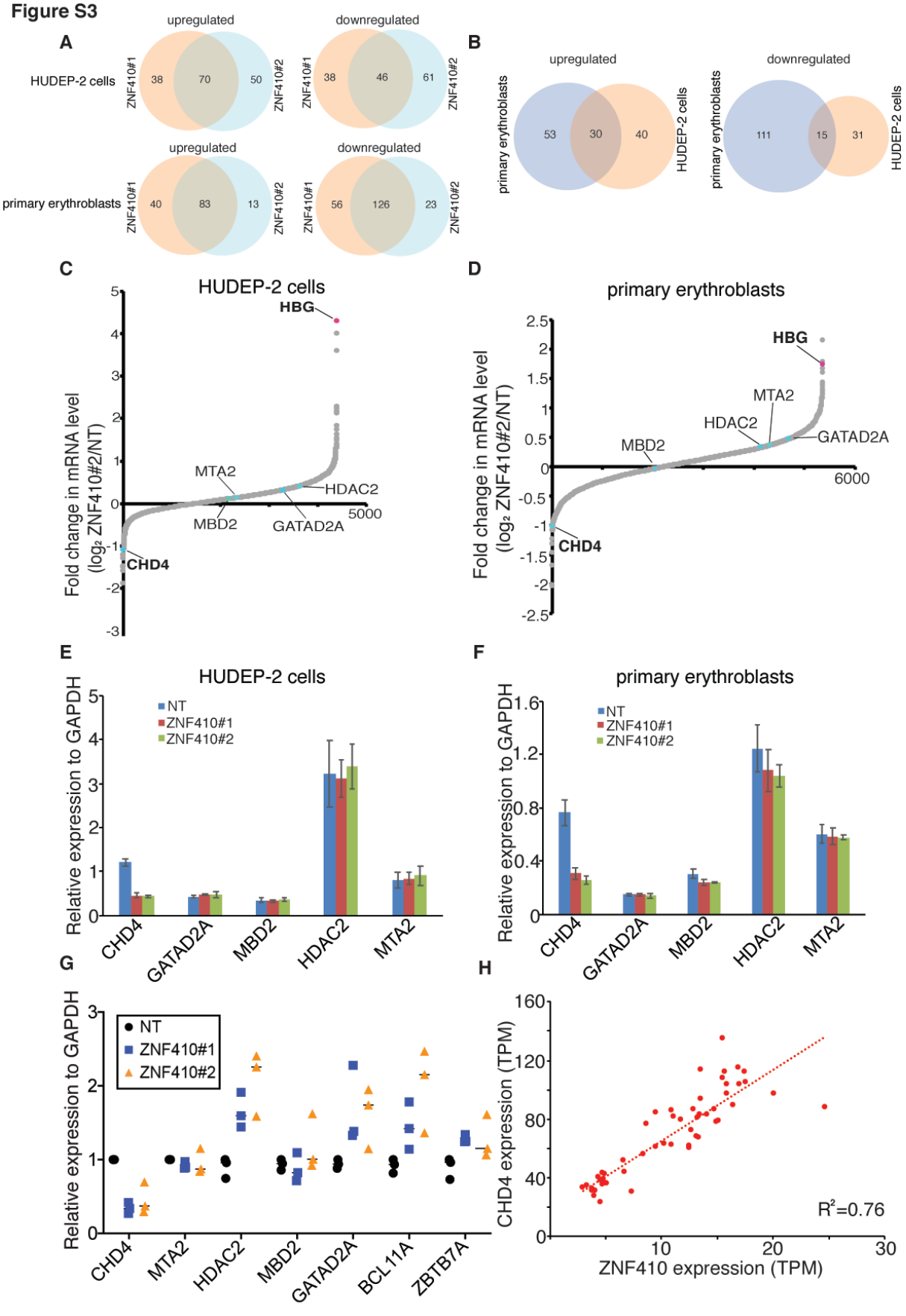
A
upregulated
downregulated
HUDEP-2 cells
primary erythroblasts
ZNF410#1
ZNF410#2
38
70
50
38
46
61
40
83
13
56
126
23
B
upregulated
downregulated
primary erythroblasts
HUDEP-2 cells
53
30
40
111
15
31
C
HUDEP-2 cells
HBG
MTA2
HDAC2
MBD2
GATAD2A
CHD4
5000
Fold change in mRNA level
(log2 ZNF410#2/NT)
D
primary erythroblasts
HBG
MTA2
HDAC2
MBD2
GATAD2A
CHD4
6000
Fold change in mRNA level
(log2 ZNF410#2/NT)
E
HUDEP-2 cells
NT
ZNF410#1
ZNF410#2
Relative expression to GAPDH
CHD4
GATAD2A
MBD2
HDAC2
MTA2
F
primary erythroblasts
NT
ZNF410#1
ZNF410#2
Relative expression to GAPDH
CHD4
GATAD2A
MBD2
HDAC2
MTA2
G
NT
ZNF410#1
ZNF410#2
Relative expression to GAPDH
CHD4
MTA2
HDAC2
MBD2
GATAD2A
BCL11A
ZBTB7A
H
CHD4 expression (TPM)
ZNF410 expression (TPM)
R2=0.76

Figure S3. ZNF410 depletion diminishes CHD4 transcription, related to Figure 3

(A) Venn diagrams of the commonly upregulated and downregulated genes in both ZNF410 sgRNAs in HUDEP-2 cells or primary erythroblasts.

(B) Venn diagrams of the commonly upregulated and downregulated genes in both HUDEP-2 and primary erythroblasts.

(C-D) RNA-seq analysis of HUDEP-2 cells transduced with ZNF410 sgRNA#2 and primary erythroblasts with ZNF410 sgRNA#2. Plotted is the average fold-change in mRNA levels of two biological replicates. FPKM value was used to calculate fold change for each gene. NuRD complex subunits and γ-globin genes are indicated. NT: non-targeting. X axis indicates the number of genes.

(E-F) mRNA levels of the NuRD complex subunits including CHD4, HDAC2, GATAD2A, MBD2 and MTA2 by RT-qPCR in HUDEP-2 cells and primary erythroblasts. Results are shown as mean ± SD (n=2). GAPDH was used for normalization. NT: non-targeting.

(G) mRNA levels of the NuRD complex subunits and BCL11A and ZBTB7A by RT-qPCR in human CD235a+ erythroblasts from recipient NBSGW mouse bone marrow. Each dot in graph represents a separate mouse. n=3 mice per sgRNA. GAPDH was used for normalization. NT: non-targeting.

(H) Correlation of ZNF410 and CHD4 mRNA levels across 53 human tissues. TPM: transcripts per kilobase million. Each dot indicates one tissue. Expression data were obtained from the Genotype-Tissue Expression database.

Figure S4

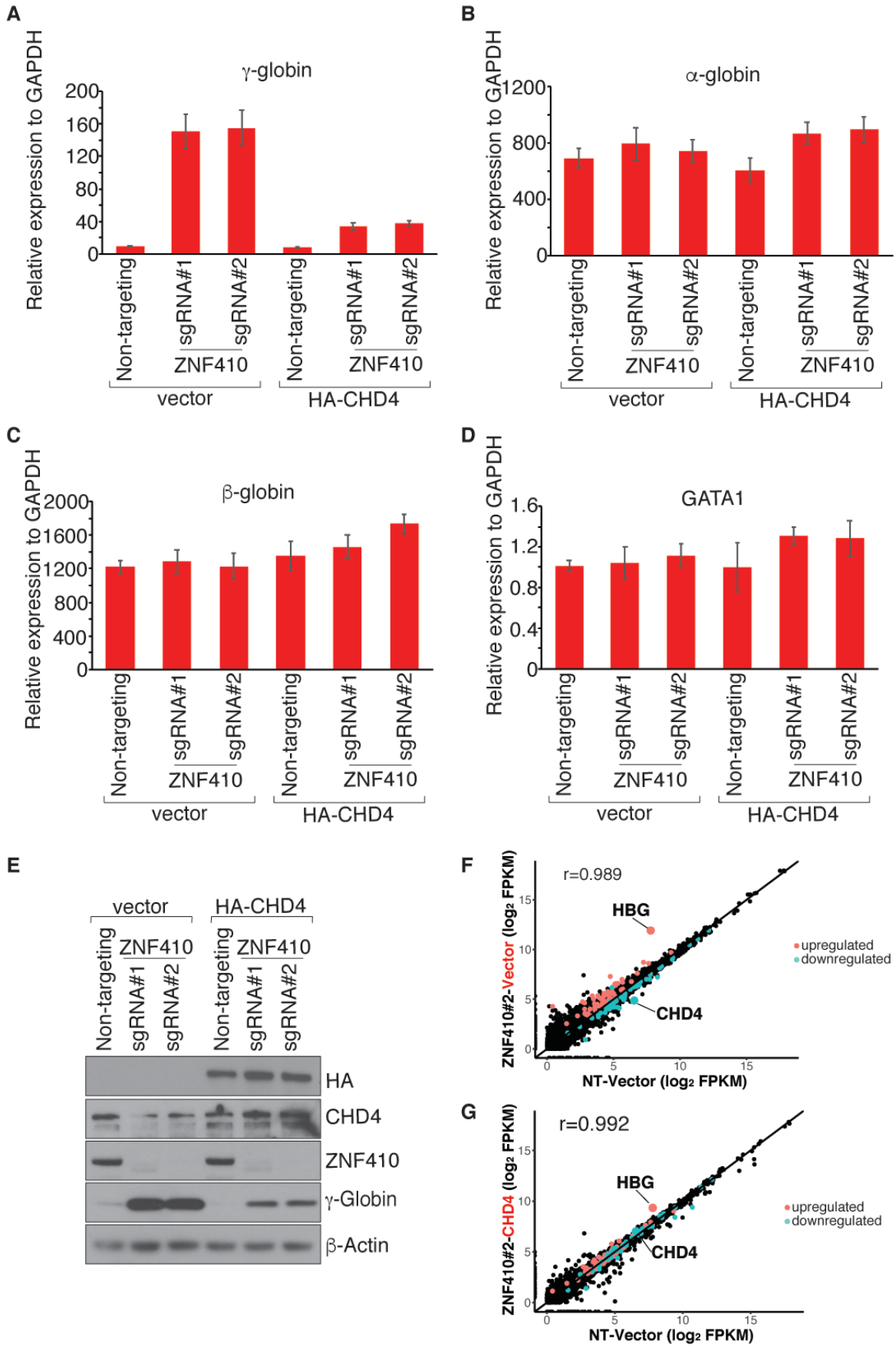
A
γ-globin
Relative expression to GAPDH
200
160
120
80
40
0
Non-targeting
sgRNA#1
sgRNA#2
ZNF410
vector
HA-CHD4
B
α-globin
1200
800
400
C
β-globin
2000
1600
1200
800
400
D
GATA1
1.6
1.2
0.8
0.4
E
HA
CHD4
ZNF410
γ-Globin
β-Actin
F
r=0.989
HBG
CHD4
upregulated
downregulated
ZNF410#2-Vector (log2 FPKM)
NT-Vector (log2 FPKM)
G
r=0.992
ZNF410#2-CHD4 (log2 FPKM)

Figure S4. Re-introduction of CHD4 cDNA restores the silencing of HbF and transcriptome in ZNF410 deficient HUDEP-2 cells, related to Figure 3

(A-D) mRNA levels of γ-globin, α-globin, β-globin and GATA1 by RT-qPCR in ZNF410 deficient HUDEP-2 cells transduced with lentiviral vector containing CHD4 cDNA or empty vector. Results are shown as mean ± SD (n=2). GAPDH was used for normalization.

(E) Western blot using whole-cell lysates from ZNF410 deficient HUDEP-2 cells with empty vector or forced CHD4 expression. HA-CHD4: N-terminal HA tagged CHD4.

(F) Scatter plot of RNA-seq analysis in ZNF410 deficient HUDEP-2 cells (by ZNF410 sgRNA#2) with empty vector. Each gene is depicted according to averaged FPKM value from 2 biological replicates. r: Pearson's correlation coefficient. NT: non-targeting.

(G) Scatter plot of RNA-seq analysis in ZNF410 deficient HUDEP-2 cells (by ZNF410 sgRNA#2) with re-introduction of CHD4 cDNA. Each gene is depicted according to averaged FPKM value from 2 biological replicates. r: Pearson's correlation coefficient. NT: non-targeting.

Figure S5

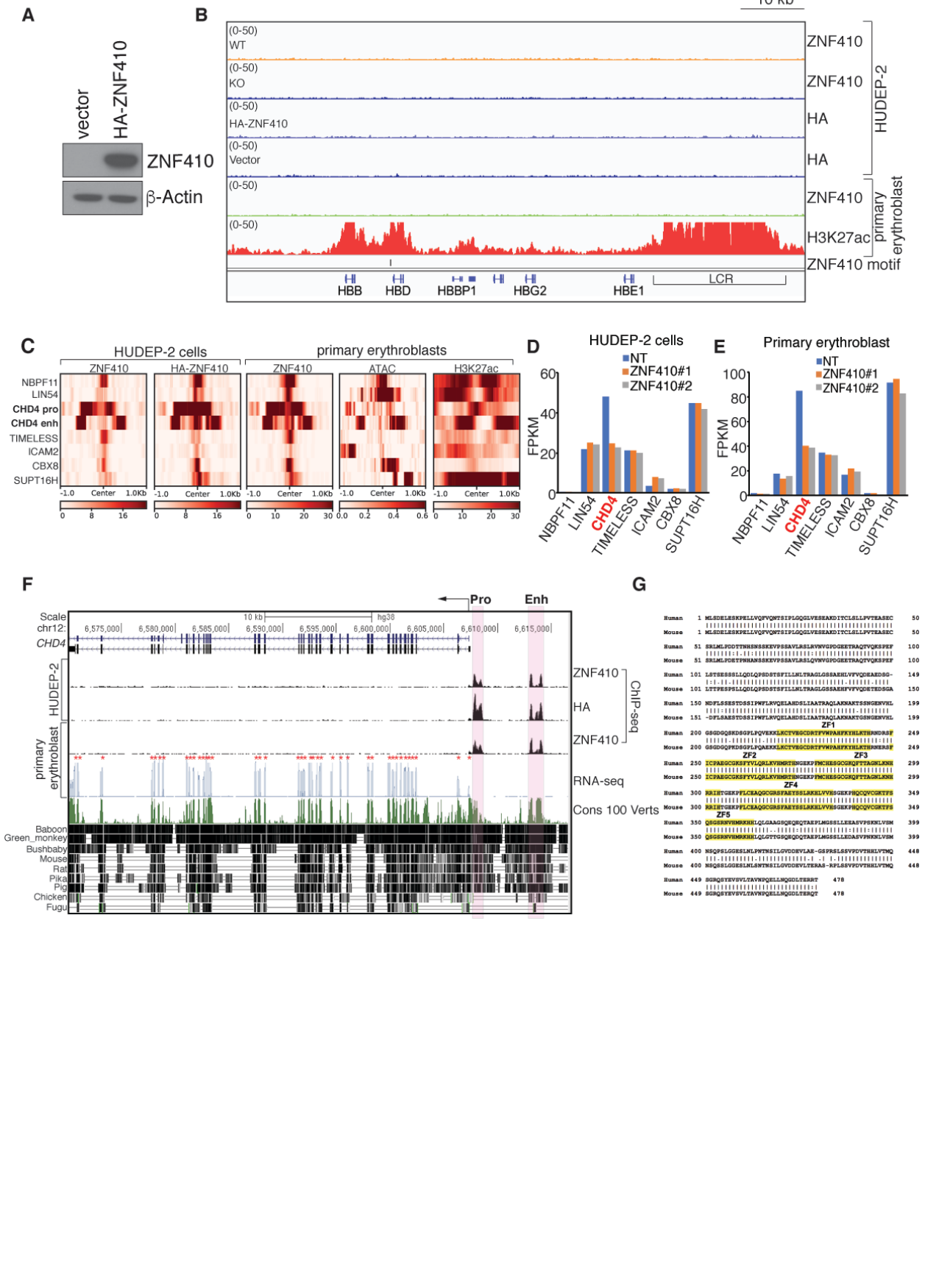
A
vector
HA-ZNF410
ZNF410
β-Actin
B
10 kb
(0-50)
WT
ZNF410
(0-50)
KO
ZNF410
(0-50)
HA-ZNF410
HA
(0-50)
Vector
HA
HUDEP-2
(0-50)
ZNF410
(0-50)
H3K27ac
primary erythroblast
ZNF410 motif
HBB
HBD
HBBP1
HBG2
HBE1
LCR
C
HUDEP-2 cells
primary erythroblasts
ZNF410
HA-ZNF410
ZNF410
ATAC
H3K27ac
NBPF11
LIN54
CHD4 pro
CHD4 enh
TIMELESS
ICAM2
CBX8
SUPT16H
-1.0
Center
1.0Kb
D
HUDEP-2 cells
NT
ZNF410#1
ZNF410#2
FPKM
NBPF11
LIN54
CHD4
TIMELESS
ICAM2
CBX8
SUPT16H
E
Primary erythroblast
NT
ZNF410#1
ZNF410#2
FPKM
NBPF11
LIN54
CHD4
TIMELESS
ICAM2
CBX8
SUPT16H
F
Pro
Enh
Scale
chr12:
10 kb
hg38
6,575,000
6,580,000
6,585,000
6,590,000
6,595,000
6,600,000
6,605,000
6,610,000
6,615,000
CHD4
HUDEP-2
primary erythroblast
ZNF410
HA
ZNF410
ChIP-seq
RNA-seq
Cons 100 Verts
Baboon
Green_monkey
Bushbaby
Mouse
Rat
Pika
Pig
Chicken
Fugu
G
Human
Mouse
ZF1
ZF2
ZF3
ZF4
ZF5

Figure S5. ZNF410 binds to the CHD4 locus in an evolutionarily conserved manner, related to Figure 4

(A) Immunoblot analysis confirming overexpression of HA-ZNF410 in HUDEP-2 cells.

(B) Gene track of endogenous ZNF410, HA-ZNF410 and H3K27ac ChIP-seq occupancy at the β-globin locus. The enhancer (LCR) is highlighted with line. LCR: local control region.

(C) Heatmaps of the signal intensities of the 8 ZNF410 bound regions from endogenous ZNF410, HA-ZNF410, H3K27ac ChIP-seq and ATAC-seq in HUDEP-2 or primary erythroblasts. ATAC-seq of primary erythroblast at polychromatic stage were obtained from published data (Ludwig et al., 2019).

(D-E) Expression levels of the 7 ZNF410 bound genes in HUDEP-2 cells transduced with indicated sgRNAs (F) and primary erythroblasts with indicated sgRNAs (G) by RNA-seq analysis. NT: Non-targeting.

(F) PhastCons (from 0 to 1) estimates of evolutionary conservation among 100 vertebrates at the CHD4 gene locus. The CHD4 promoter and enhancer are shaded in orange. CHD4 exons are marked by red * in RNA-seq track. ZNF410 ChIP-seq tracks show ZNF410 binding at the CHD4 promoter and enhancer. Pro: promoter, Enh: enhancer. Cons: conservation, Verts: vertebrates.

(G) Alignment of human and mouse ZNF410 protein sequence. Identical residues are linked by vertical line. ZF (zinc finger) residues are shaded in yellow.

Figure S6

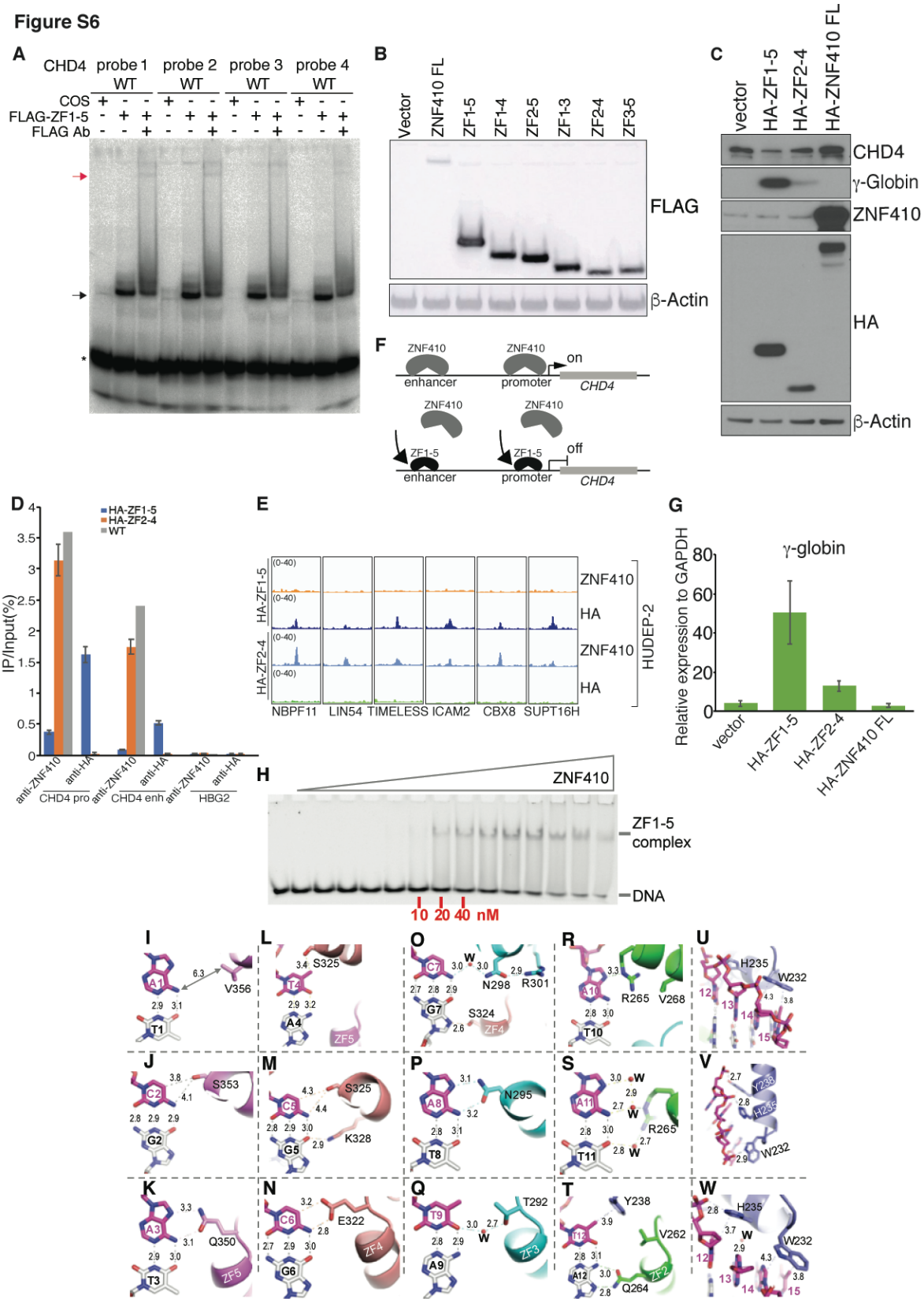
A
CHD4
probe 1
probe 2
probe 3
probe 4
WT
COS
FLAG-ZF1-5
FLAG Ab
B
Vector
ZNF410 FL
ZF1-5
ZF1-4
ZF2-5
ZF1-3
ZF2-4
ZF3-5
FLAG
β-Actin
C
vector
HA-ZF1-5
HA-ZF2-4
HA-ZNF410 FL
CHD4
γ-Globin
ZNF410
HA
β-Actin
F
ZNF410
enhancer
promoter
on
CHD4
ZF1-5
off
D
HA-ZF1-5
HA-ZF2-4
WT
IP/Input(%)
anti-ZNF410
anti-HA
CHD4 pro
CHD4 enh
HBG2
E
HA-ZF1-5
HA-ZF2-4
ZNF410
HA
HUDEP-2
NBPF11
LIN54
TIMELESS
ICAM2
CBX8
SUPT16H
G
γ-globin
Relative expression to GAPDH
H
ZNF410
ZF1-5 complex
DNA
10 20 40 nM
I
L
O
R
U
J
M
P
S
V
K
N
Q
T
W

Figure S6. The ZF domain of ZNF410 is sufficient for DNA binding in vitro and in vivo, related to Figure 5 and Figure 6

(A) The ZF domain of ZNF410 binds to the four motifs from the CHD4 promoter and enhancer sites. Ab, antibody. Black arrow: ZF domain-probe complex; red arrow: FLAG antibody-ZF domain-probe complex. * : free probes.

(B) Western blot analysis of FLAG-ZNF410 constructs expressed in COS-7 cells. Control: empty vector. All constructs were N-terminal FLAG tagged.

(C) Western blot analysis using whole-cell lysates from differentiated HUDEP-2 cells overexpressing HA-ZF1-5, HA-ZF2-4 or HA-ZNF410 FL.

(D)ZNF410 and HA ChIP-qPCR in WT HUDEP-2 cells, HA-ZF1-5 or HA-ZF2-4 overexpressed HUDEP-2 cells. HBG2 region serves as negative control. Results are shown as mean ± SD (n=2 for HA-ZF1-5 and HA-2-4 overexpressed cells)

(E) ChIP-seq tracks of endogenous ZNF410, HA-ZNF410 at the 6 ZNF410 bound regions in HUDEP-2 cells overexpressing indicated constructs.

(F) Diagram of HA-ZF1-5 displacing endogenous ZNF410 binding.

(G) γ-globin levels measured by RT-qPCR in differentiated HUDEP-2 cells overexpressing indicated constructs. Empty vector serves as control. Results are shown as mean ± SD (n=2). GAPDH was used for normalization.

(H) EMSA of ZNF410 ZF1-5. The maximum protein concentration used were 0.5 μM (the right most lane) followed by 2-fold serial dilutions.

(I-K) ZF5 interactions with base pairs 1-3. V356 is too far from A1 to make direct contact. S353 makes van der Waals contacts with C2. Q350 interacts with A3.

(L-N) ZF4 interactions with base pairs 4-6. S325 forms a weak hydrogen bond with T4, and makes van der Waals contact with C5. E322 interacts with C6.

(O-Q) ZF3 interactions with base pairs 7-9. N298 conducts a water mediated interaction with C7. S324 interacts with G7. N295 interacts with A8. T292 conducts a water mediated interaction with T9.

(R-T) ZF2 DNA interaction with base pairs 10-12. R265 forms a weak hydrogen bond with A10 and mediates a water network connecting with A:T base pair at position 11. Q264 interacts with A12. Y238 of ZF1 forms a π–methyl interaction with T12.

(U-W) ZF1 interactions with base pair 12-15. W232 forms van der Waals contact with A14 and T15. H235, Y238 and W232 form polar interactions with DNA backbone phosphate groups.

Table S2. Primers for RT-qPCR, related to Figures 1-5 and S1-6

Name	Sequence 5'-3'
Gamma_globin_F	TGGCAAGAAGGTGCTGACTTC
Gamma_globin_R	GCAAAGGTGCCCTTGAGATC
Beta_globin_F	TGGGCAACCCTAAGGTGAAG
Beta_globin_R	GTGAGCCAGGCCATCACTAAA
Alpha_globin_F	AAGACCTACTTCCCGCACTTC
Alpha_globin_R	GTTGGGCATGTCGTCCAC
Gamma_globin_PT_F	TTTGTGGCACCTTCTGACTG
Gamma_globin_PT_R	GCCAAAGCTGTCAAAGAA
MBD2_F	AAGTGCTGGCAAGAGCGATGTCTA
MBD2_R	TTTCCCAGGTACCTTGCCAACTGA
GAPDH_F	AGCCACATCGCTCAGACAC
GAPDH_R	GCCCAATACGACCAAATCC
CHD4_F	GCTGCAACCATCCATACCTC
CHD4_R	ACCATCGATGCGTTCGTATT
GATA1_F	CTGTCCCCAATAGTGCTTATGG
GATA1_R	GAATAGGCTGCTGAATTGAGGG
MTA2_F	TTCCCACCTACACTAAGCC
MTA2_R	AGGCCCTTCTGAAATCCAG
BCL11A_F	ACAAACGGAAACAATGCAATGG
BCL11A_R	TTTCATCTCGATTGGTGAAGGG
HDAC2_F	CATGGTGATGGTGTTGAAGAAG
HDAC2_F	TCATTGGAAAATTGACAGCATAGT
GATAD2A_F	ACGAGTTCATCTACCTGGTCGG
GATAD2A_R	ACGTGAAGTCCGTCTTGCACTG
NBPF11_F	GCATGGCTGTTGACATAGGCAG
NBPF11_R	CTATCCAGTGAGTCCTGCAAGAC
LIN54_F	GTCCGACTTGTTACTGCCACATC
LIN54_R	TTGGAACTGACATCCGCACTGG
TIMELESS_F	AAGTGGTCCAGGTGTCGGAGAA
TIMELESS_R	GTGGGCACTATTCTGCTGGTAG
ICAM2_F	ATGACACGGTCCTCCAATGCCA
ICAM2_R	GCACTCAATGGTGAAGGACTTGC
CBX8_F	AACATCCTGGATGCTCGCTTGC
CBX8_R	TTTGAGGAGGAAGGTTTTGGGCT
SUPT16H_F	CATTGGTGACACAGTGCTTGTGG
SUPT16H_R	CCAAAAGGTCCTCTGCCTCATC

Table S3. Top20 differentially expressed transcripts from HUDEP-2 and primary erythroblasts by RNA-seq, related to Figure 3

HUDEP-2 cells		Primary CD34+ cells	
Top10 upregulated genes	log2(Z#2/NT)	**Top10 upregulated genes**	log2(Z#2/NT)
HBG1/2	4.28	HSPB1	2.15
AC104389.5	4.01	IFI27	1.78
RNU5B-1	3.60	**HBG1/2**	1.75
SNORD97	2.29	SELENOW	1.67
CKB	2.24	RN7SL471P	1.60
PRG2	2.17	PKN1	1.43
H1F0	2.13	AC104389.5	1.43
AK1	1.83	STUB1	1.40
RNU5E-1	1.82	PAFAH1B3	1.40
CD63	1.75	CDKN1A	1.38
	log2(Z#1/NT)		log2(Z#1/NT)
HBG1/2	4.26	HSPB1	2.01
AC104389.5	3.95	SELENOW	1.55
CKB	2.44	**HBG1/2**	1.47
H1F0	2.41	STUB1	1.46
PRG2	2.09	RPS19BP1	1.42
AK1	1.99	PAFAH1B3	1.35
MIF	1.92	ISOC2	1.34
CD63	1.90	PKN1	1.32
TMSB10	1.79	COA4	1.30
ATF5	1.65	DYNLT1	1.27
Top10 downregulated genes	log2(Z#2/NT)	**Top10 downregulated genes**	log2(Z#2/NT)
FOS	-1.88	AC130456.3	-2.03
S100A10	-1.59	CBX6	-2.01
AC092490.1	-1.50	TMCC2	-1.69
SLC2A3P1	-1.49	**AC092490.1**	-1.47
SCARNA3	-1.30	ACSL6	-1.47
CHD4	-1.23	KDM7A	-1.32
SCARNA18	-1.20	ZNF410	-1.24
RNU6-11P	-1.11	SMOX	-1.24
TUBA1A	-1.07	PRDX5	-1.09
PLA2G6	-1.01	**CHD4**	-1.08
	log2(Z#1/NT)		log2(Z#1/NT)

RF00019	-2.43	TMCC2	-2.01
S100A10	-1.60	ACSL6	-1.67
AC092490.1	-1.45	CBX6	-1.56
FOS	-1.32	**AC092490.1**	-1.50
TMCC2	-1.18	ZNF410	-1.49
GLIPR2	-1.14	SMOX	-1.45
TUBA1A	-1.06	PRDX5	-1.00
CHD4	-0.99	**CHD4**	-0.93
PPP4R1	-0.95	SPC25	-0.93
SCARNA3	-0.89	GPCPD1	-0.93

Z#1: ZNF410 sgRNA#1; Z#2: ZNF410 sgRNA#2; NT: Non-targeting

Table S4. Analysis of H3K27ac ChIP-seq and ATAC signal from primary human erythroid cells at ZNF410 motifs, related to Figure 4

Motif	# of motif in human genome	H3K27ac only	ATAC signal only	Both
CATCCCATAATA	434	11	13	15
TATCCCATAATA	561	17	9	3
CAGCCCATAATA	529	20	15	5
CAACCCATAATA	464	13	9	2
CATCCCATAATT	484	15	7	4
CAGACCATAATA	351	9	9	2
CATACCATAATA	422	11	6	2
CATCCCGTAATA	33	2	2	1
CATCCCATAATG	399	13	11	2

Distinct nucleotides are highlighted in “red”.

Table S5. Summary of X-ray data collection and refinement statistics, related to Figure 6

PDB ID code	6WMI
ZNF410	ZF1-5 (residues 217-366)
DNA oligo (5'-3') (3'-5')	CACATCCCATAATAATG GTGTAGGGTATTATTAC
X-ray Source (wavelength)	SETCAT 22-ID (1.0Å)
Space group	$P6_2$
Unit cell (Å)	186.26, 186.26, 46.95
α, β, γ (°)	90, 90, 120
Resolution (Å)	41.92-2.75 (2.85-2.75) *
[a] R_{merge}	0.22 (0.50)
R_{pim}	0.07 (0.82)
$CC_{1/2}$, CC	0.991, 0.998 (0.498, 0.815)
[b] $\langle I/\sigma I \rangle$	12.3 (1.3)
Completeness (%)	100 (100)
Redundancy	11.1 (10.0)
Observed reflections	522,370
Unique reflections **	47,241 (4,697)
Phase Determination	
Bijvoet Pairs	22,583
Mean FOM (SAD)	0.482 (5 Å), 0.276 (2.75 Å)
Density modification R-Factor	0.34 (5 Å), 0.35 (2.75 Å)
Refinement	
Resolution (Å)	2.75
No. reflections **	47,202
[c] R_{work} / [d] R_{free}	0.159 / 0.194
No. atoms: protein	1140
DNA	350
Zn	10
Solvent	137
B-factor (Å^2): protein	67.4
DNA	58.8
Zn	66.2
Solvent	45.5
R.m.s deviations	
Bond Length(Å)	0.005
Bond angles (°)	0.7

* Values in parenthesis correspond to highest resolution shell. ** Friedel mates kept separately.

[a] $R_{merge}=\Sigma|I-\langle I\rangle|/\Sigma I$, where I is the observed intensity and $\langle I\rangle$ is the averaged intensity from multiple observations.

[b] $\langle I/\sigma I\rangle$=averaged ratio of the intensity (I) to the error of the intensity (σI).

[c] $R_{work}=\Sigma|Fobs-Fcal|/\Sigma|Fobs|$, where Fobs and Fcal are the observed and calculated structure factors.

[d] R_{free} was calculated using a randomly chosen subset (5%) of the reflections not used in refinement.