

Characterization and evolutionary origin of novel C2H2 zinc finger protein (ZNF648) required for both erythroid and megakaryocyte differentiation in humans

Daniel C. J. Ferguson,^{1*} Juraidah Haji Mokim,^{1*} Marjolein Meinders,¹ Edmund R. R. Moody,² Tom A. Williams,² Sarah Cooke,¹ Kongtana Trakarnsanga,³ Deborah E. Daniels,^{1,4} Ivan Ferrer-Vicens,¹ Deborah Shoemark,¹ Chartsiam Tipgomut,³ Katherine A. Macinnes,^{1,4} Marieangela C. Wilson,¹ Belinda K. Singleton^{4,5} and Jan Frayne^{1,4}

¹School of Biochemistry, University of Bristol, Bristol, UK; ²School of Biological Sciences, University of Bristol, Bristol, UK; ³Department of Biochemistry, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand; ⁴NIHR Blood and Transplant Research Unit in Red Blood Cell Products, University of Bristol, Bristol, UK and ⁵Bristol Institute for Transfusion Sciences, National Health Service Blood and Transplant (NHSBT), Bristol, UK

**DCJF and JHM contributed equally as co-first authors.*

©2021 Ferrata Storti Foundation. This is an open-access paper. doi:10.3324/haematol.2020.256347

Received: April 22, 2020.

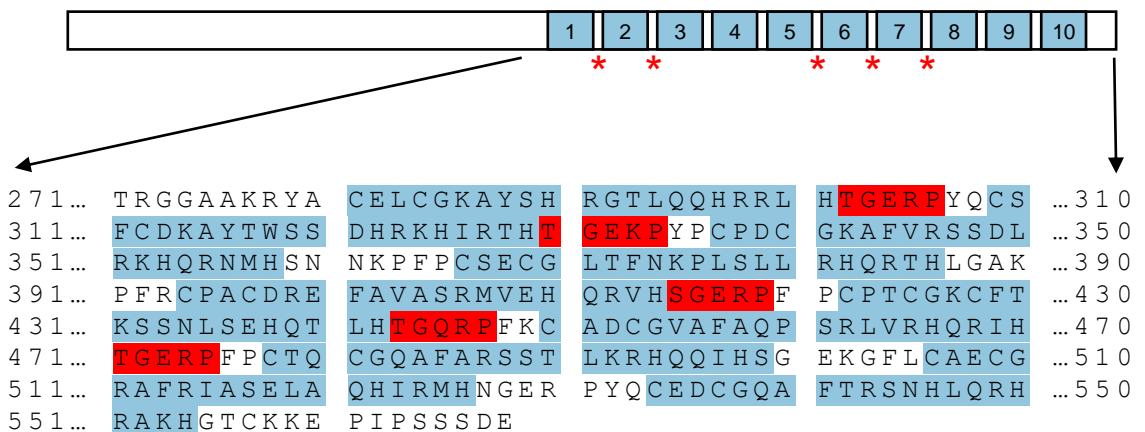
Accepted: September 15, 2020.

Pre-published: October 5, 2020.

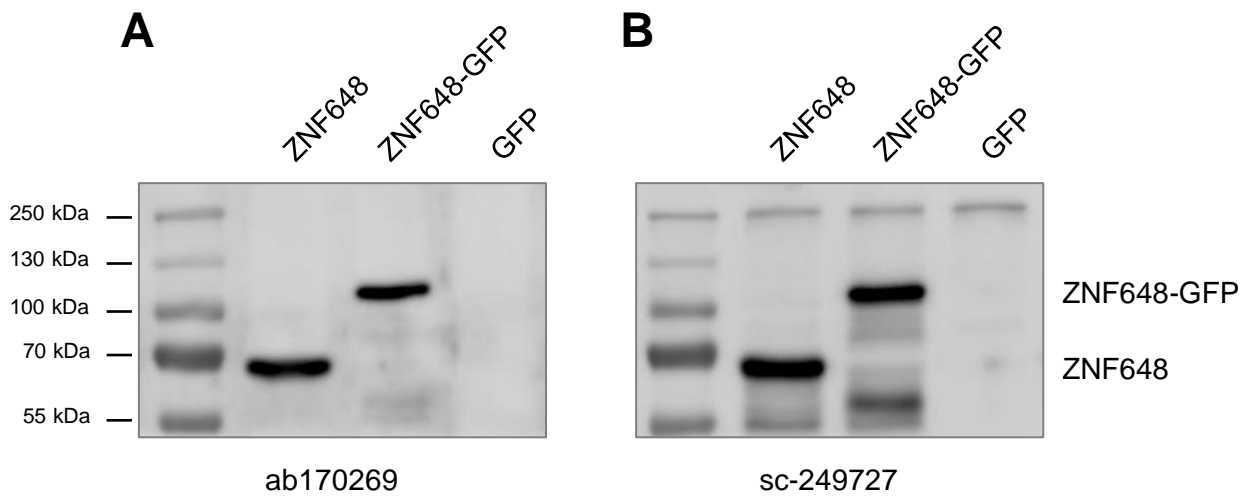
Correspondence: JAN FRAYNE - Jan.Frayne@Bristol.ac.uk

A

Linker sequence	Position
TGERP	Between zinc fingers 1 and 2 Between zinc fingers 7 and 8
TGEKP (consensus)	Between zinc fingers 2 and 3
SGERP	Between zinc fingers 5 and 6
TGQRP	Between zinc fingers 6 and 7

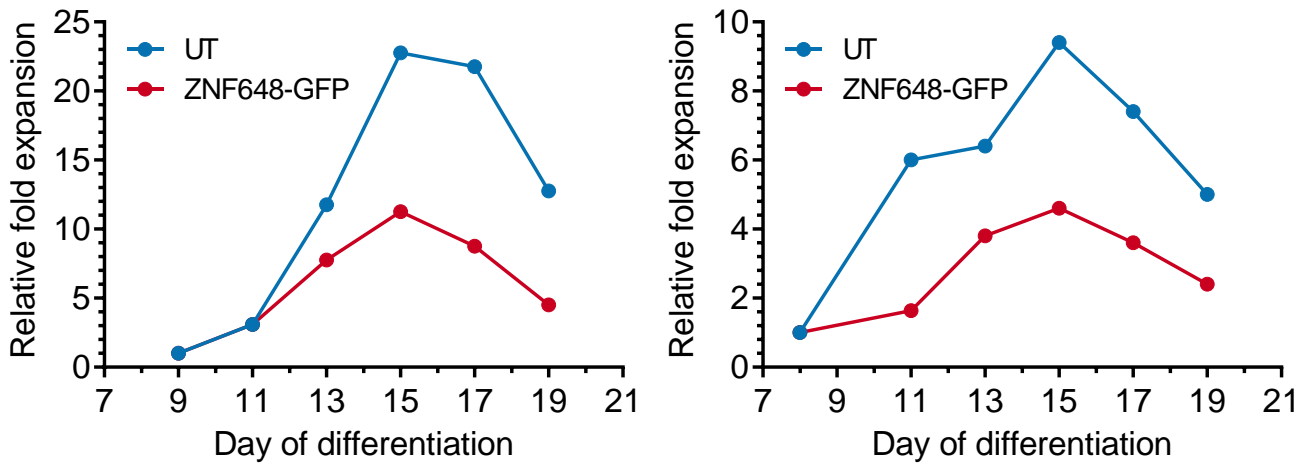
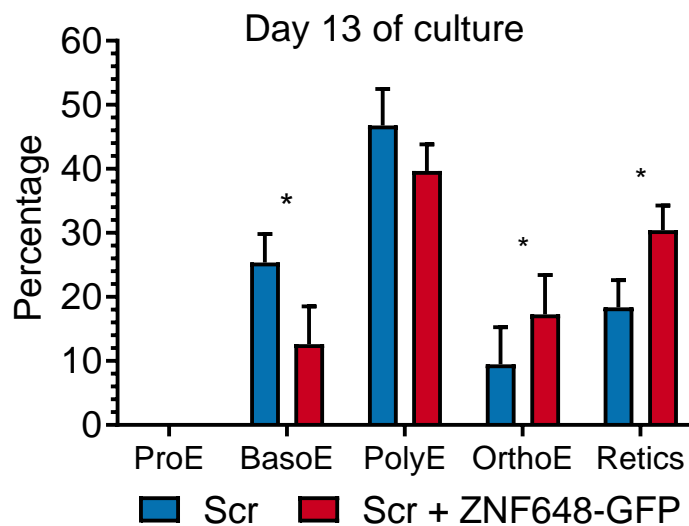
B

Supplementary Figure 1. Linker sequences between Zinc fingers of human ZNF648. A) Linker sequences detected between indicated zinc fingers. B) Sequence of zinc finger domain with linker sequences highlighted in red



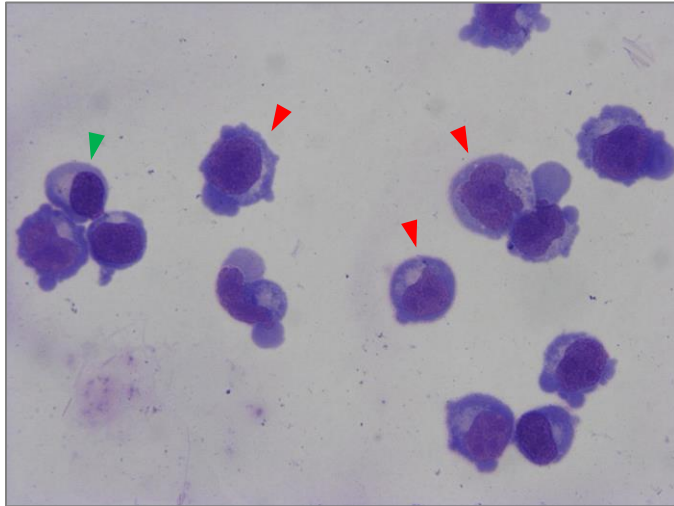
Supplementary Figure 2. Validation of ZNF648 antibodies

Cells were transduced with pXLG3-ZNF648 or pXLG3-ZNF648-GFP, containing the coding region of ZNF648, or a control construct (GFP only). Western blot was probed with ZNF648 antibody A) ab170269 (Abcam) or B) sc-249727 (Santa Cruz)

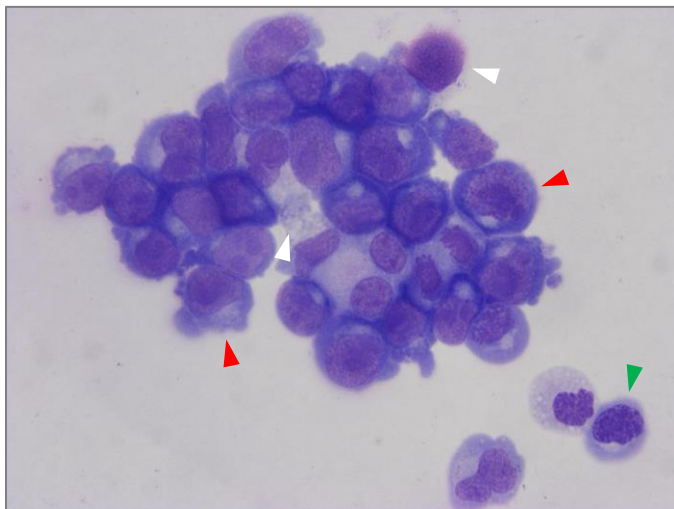
A**B**

Supplementary Figure 3. Exogenous expression of ZNF648 in adult erythroid cells impedes proliferation and advances differentiation. Erythroid cells were differentiated from a pool of CD34⁺ cells isolated from peripheral blood of four adult donors and transduced with ZNF648-GFP on day 3 in culture, with GFP-positive cells isolated on day 8 by FACS. A) Relative fold expansion of ZNF648 over-expressing adult erythroid cells compared to control cells. Graphs are repeats for two separate pools of peripheral blood CD34⁺ cells B) proportion of adult erythroid cells at each stage of differentiation present at day 13 of culture. ProE = proerythroblast, BasoE = basophilic erythroblast, PolyE = polychromatic erythroblast, OrthoE = orthochromatic erythroblast, Retics = reticulocytes. *P < 0.05, n = 2 ± SD one-tailed t-test.

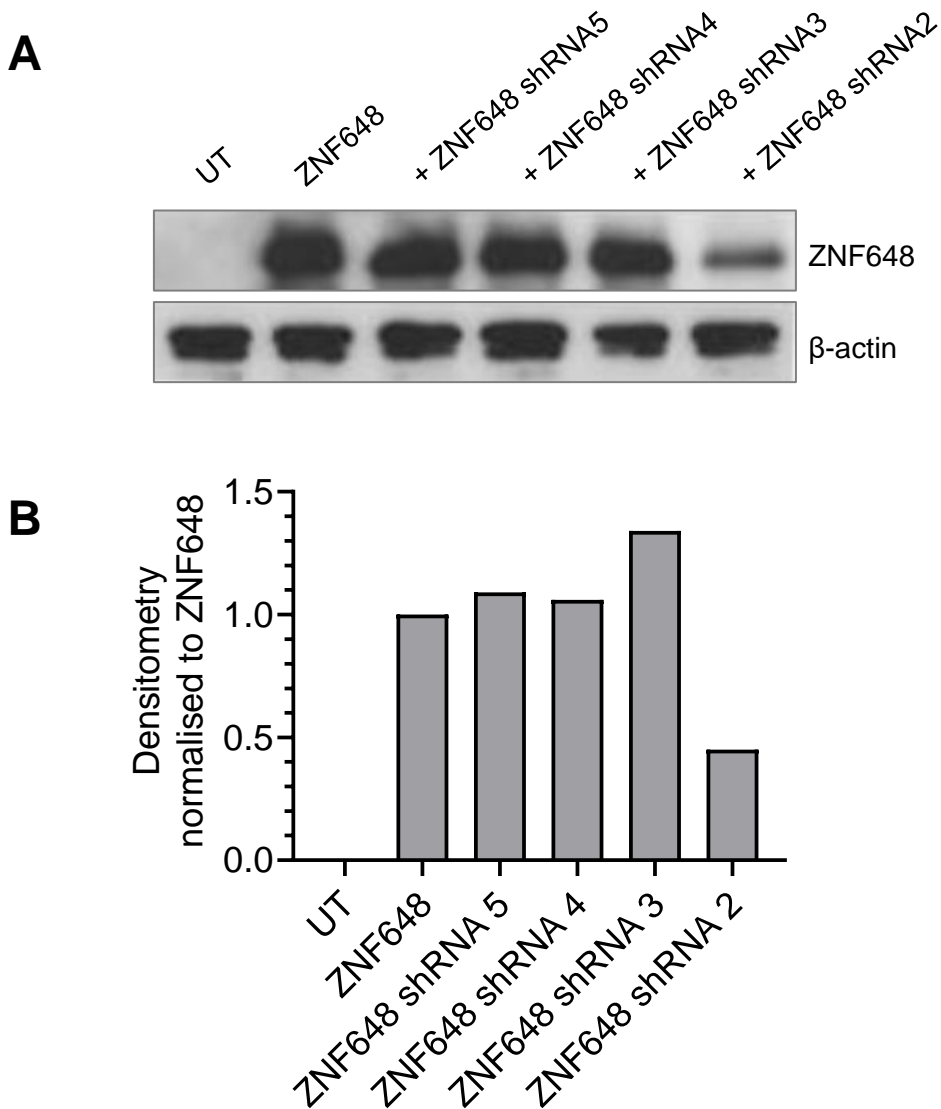
Scr shRNA control



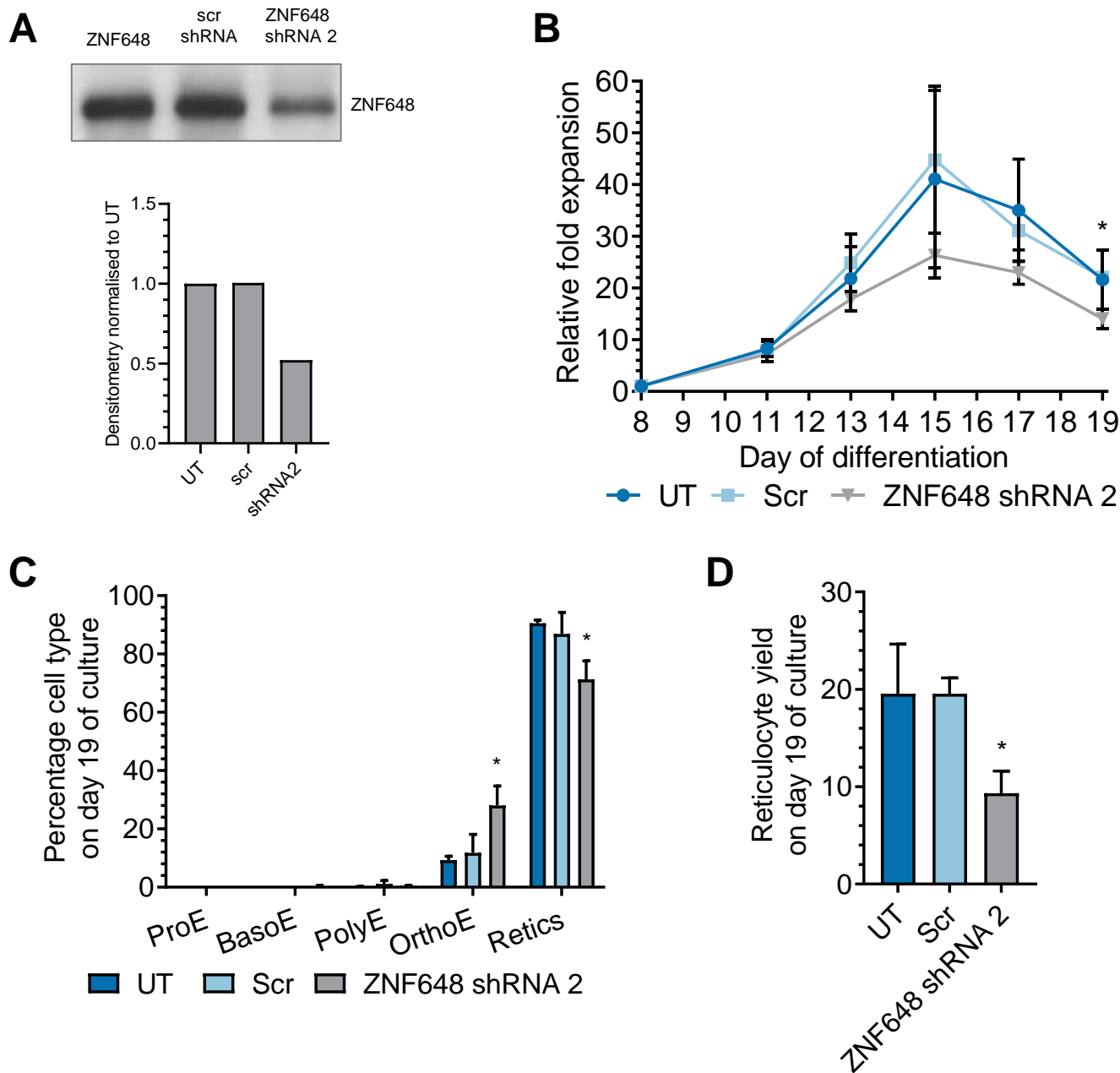
ZNF648 shRNA



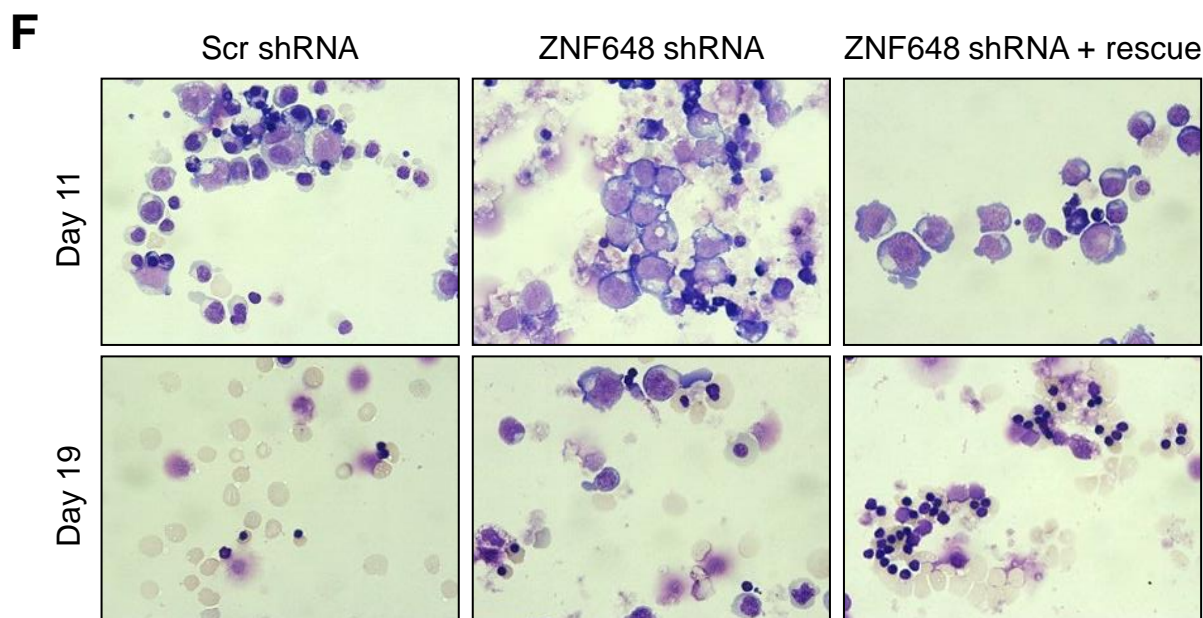
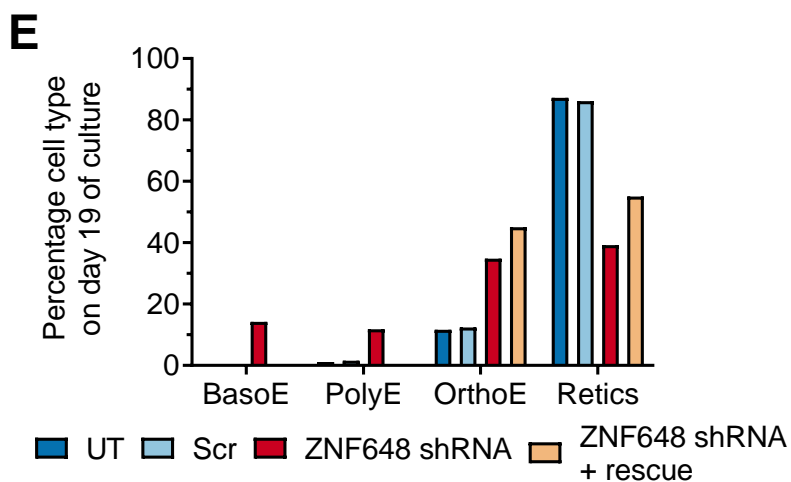
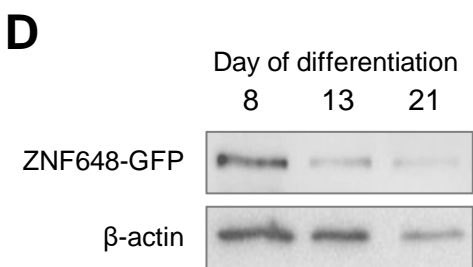
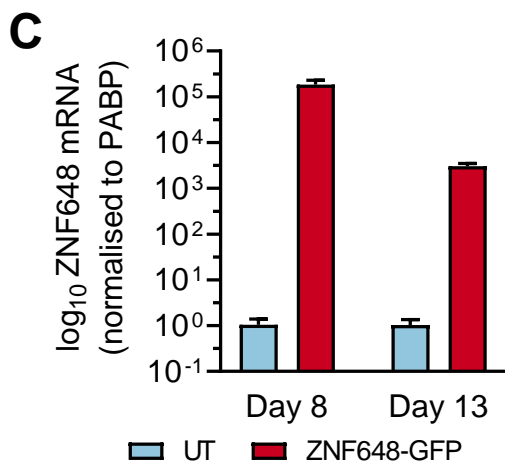
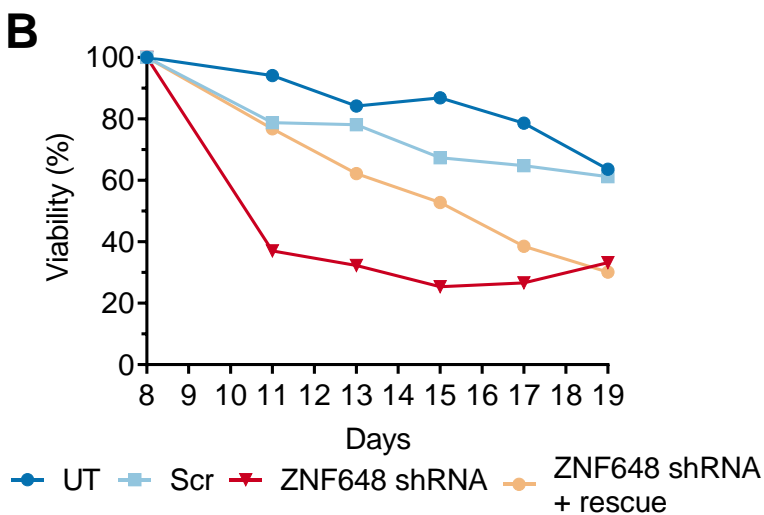
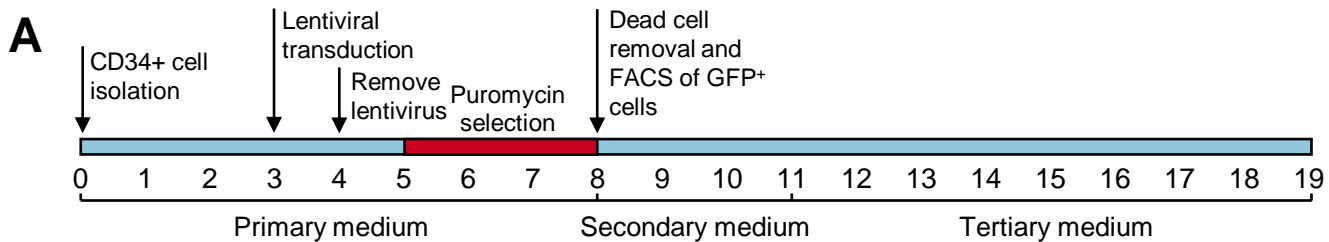
Supplementary Figure 4. Morphology of erythroid cells differentiated from adult PB CD34⁺ cells transduced with scrambled (Scr) or ZNF648 shRNA at day 8 in culture. Cell stained with May-Grünwald-Giemsa. Red arrow = pro-erythroblast, green arrow = basophilic erythroblast, white arrow = cell debris.



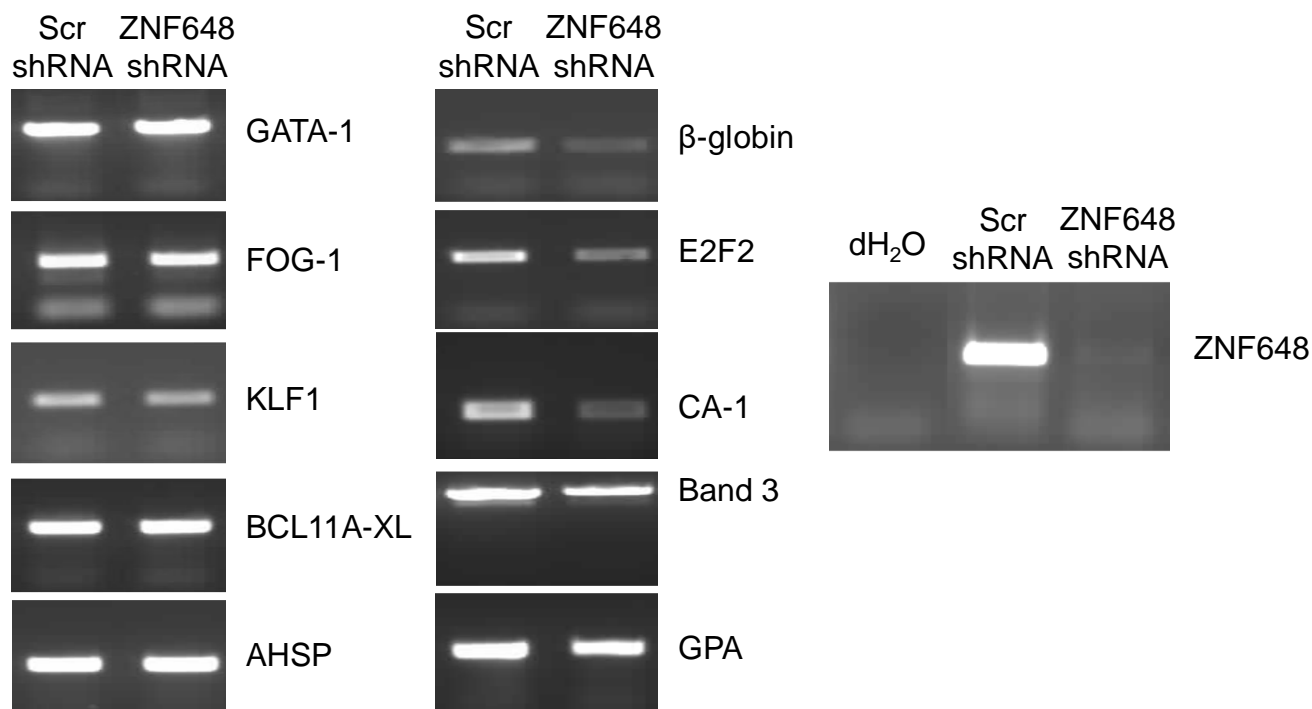
Supplementary Figure 5. Evaluating additional ZNF648 shRNAs. The erythroid cell line K562 was first transduced with pXLG3-ZNF648, containing the coding region of ZNF648, and allowed to recover for 3 days. Cells were then transduced with a control scrambled shRNA (Scr) or with one of four ZNF648 shRNAs. A) Western blot of whole cell lysates probed with ZNF648 antibody. B) Densitometry values from western blot



Supplementary Figure 6. Knockdown of ZNF648 in adult erythroid cells with ZNF648 shRNA2. A) effect of ZNF648 shRNA2 on levels of ectopic ZNF648 in transduced K562 cells. Western blot probed with ZNF648 antibody and associated densitometry values. Adult PB CD34⁺ cells at day 3 in culture were transduced with scrambled shRNA (Scr) or ZNF648 shRNA2, puromycin selected and seeded at same cell density along with untransduced cells (UT). B) Relative fold expansion during differentiation compared to cell number at day 8 in culture. C) Proportion of cells at different stages of differentiation at day 19 in culture, ProE = proerythroblast, BasoE= basophilic erythroblast, PolyE = polychromatic erythroblast, OrthoE = orthochromatic erythroblast, Retics = reticulocytes. D) Reticulocyte yield at day 19 of culture. *p<0.05, t-test, n=3 ± SD.

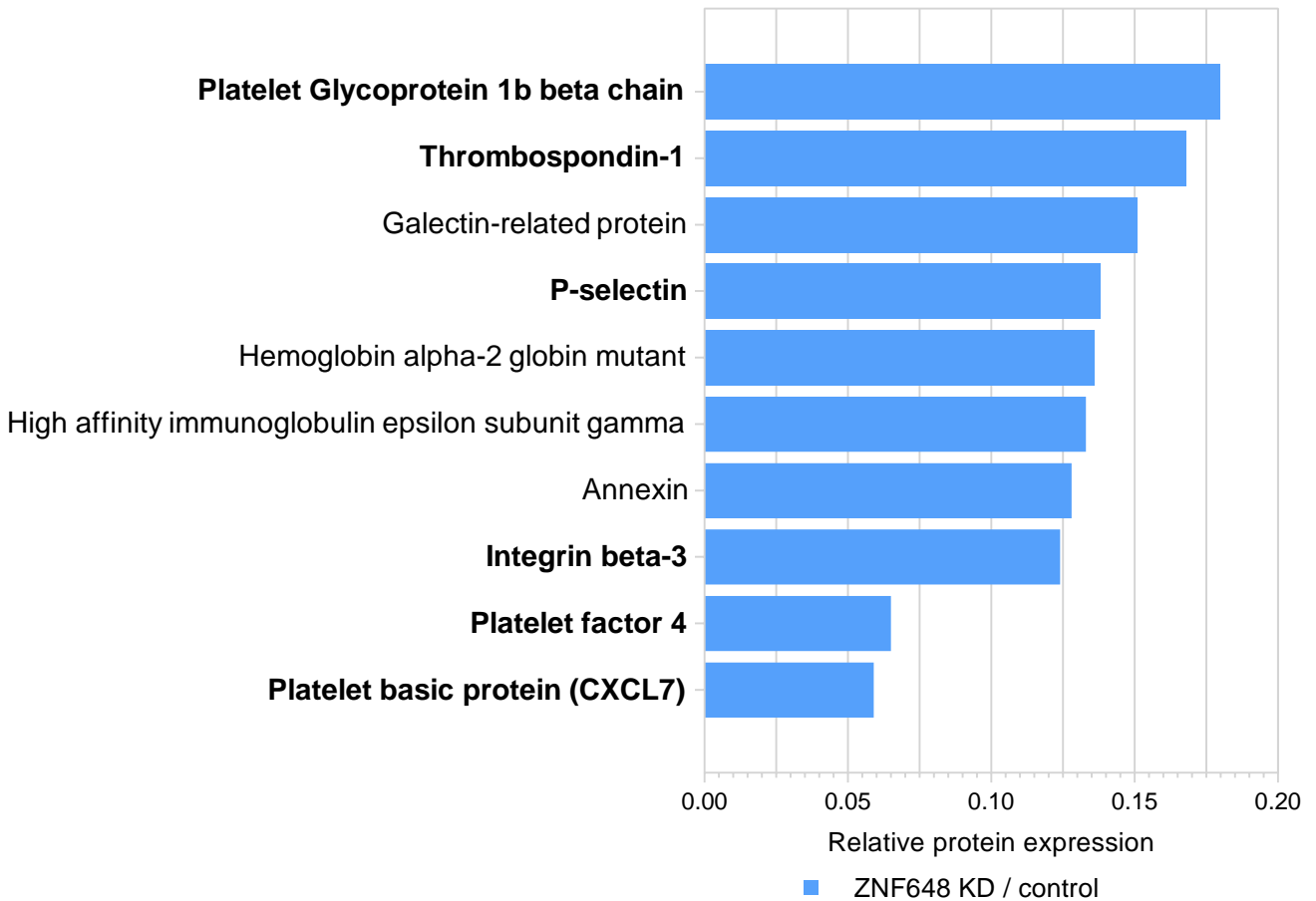


Supplementary Figure 7. Rescue of ZNF648 in erythroid cells following knockdown recovers phenotype. Erythroblasts differentiated from adult PB CD34⁺ cells at day 3 in culture were transduced with scrambled shRNA (Scr) as a control, ZNF648 shRNA, or co-transduced with ZNF648 shRNA and ZNF648-GFP (rescue). Untransduced cells (UT) served as a further control. A) Schematic of experimental design. B) Viability of cells during differentiation assessed by trypan blue exclusion. C) qPCR of ZNF648-GFP transcript levels normalised to UT controls at day 8 and day 13 of differentiation. D) Western blot of whole cell lysates from cells at day 8, 13 and 21 of differentiation probed with ZNF648 antibody along with antibody to actin as a protein loading control. E) Proportion of cells at different stages of differentiation at day 19 of differentiation, Baso = basophilic erythroblast, PolyE = polychromatic erythroblast, OrthoE = orthochromatic erythroblast, Retics = reticulocytes. F) Morphology of cells at day 11 and 19 in culture stained with May-Grünwald-Giemsa. Magnification x500.

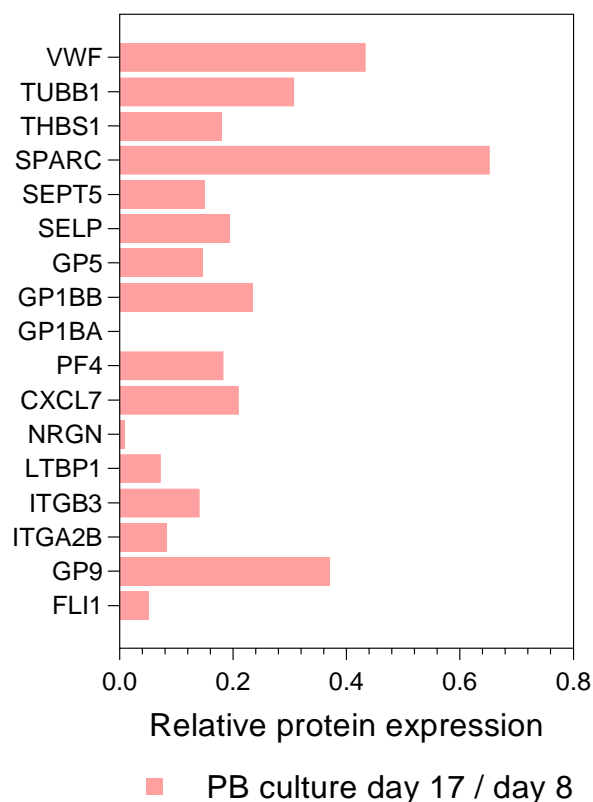
A**B**

GENE	Forward primer	Reverse primer
GATA1	5'-TTCAGCAGCCTATTCTCTCC-3'	5'-CCTTGGTAGAGATGGGCAGTA-3'
FOG-1	5'-AAGGACAGGAACCAGAACC-3'	5'-CTCTGCTGGCTCCTTCTTCACA-3'
KLF1	5'-GCAAGAGCTACACCAAGAGC-3'	5'-GCTGTCTATGGGTCCGTGTT-3'
BCL11A-XL	5'-AGATCCCTTCCTTAGCTTCG-3'	5'-TCAAACTCGATCACTGTGC-3'
AHSP	5'-GTTAGACCTGAAGGCAGATGGC-3'	5'-CTAAGGACATGAAGGTCAGAG-3'
HBB	5'-CTTTAGTGATGGCCTGGCTC-3'	5'-GGCAGAATCCAGATGCTCAA-3'
E2F2	5'-CTCGGTATGACACTTCGCTGG-3'	5'-GCACTTCCAGTCTCGTCTGCG-3'
CA1	5'-TCCTACAACCCAGCCACAG-3'	5'-CTTTGAGGCAGCTTCAGC-3'
SLC4A1 (Band3)	5'-CAGAACTCTCGGTGCCTG-3'	5'-CCACTTCGTCGTATTCATCCC-3'
GPA	5'-GTGAGCATATCAGCATCAAGTACC-3'	5'-TAAAGGCACGTCTGTGTCAGG-3'
ZNF648	5'-GTATGGAAGCCTACGTGCTC-3'	5'-AGAGGTCACACCCTAGTTCC-3'

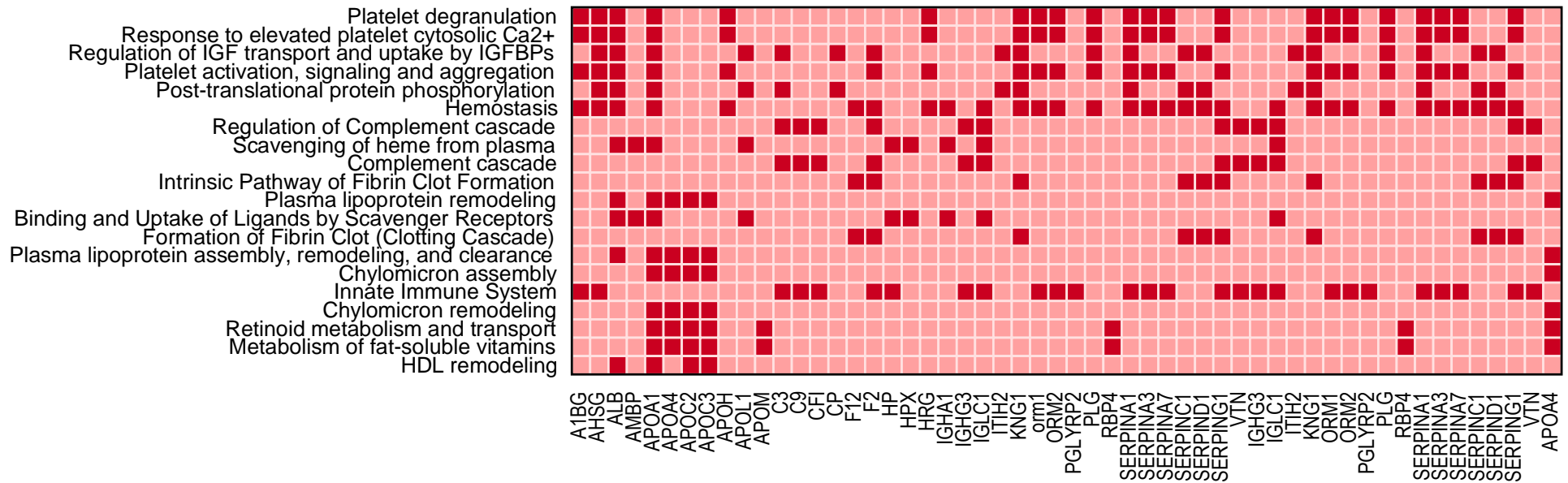
Supplementary Figure 8. Effect of ZNF648 knockdown on gene expression by adult erythroid cells. A) Expression of indicated genes measured by RT-PCR using RNA isolated from erythroid cells differentiated from adult peripheral blood CD34⁺ cells transduced with scrambled (Scr) shRNA as control or ZNF648 shRNA, at day 8 in culture. B) Primer sequences for RT-PCR.



Supplementary Figure 9. Six of the ten proteins most decreased in level following ZNF648 knock down are megakaryocyte associated. Tandem Mass Tag labelling and nano-LC MS/MS was performed on whole cell lysates of erythroid cells transduced with ZNF648 or src (control) shRNA at day 8 in culture. The relative fold decrease in abundance of the 10 proteins displaying greatest magnitude decrease following ZNF648 knockdown are shown. Megakaryocyte-associated proteins are shown in bold text.

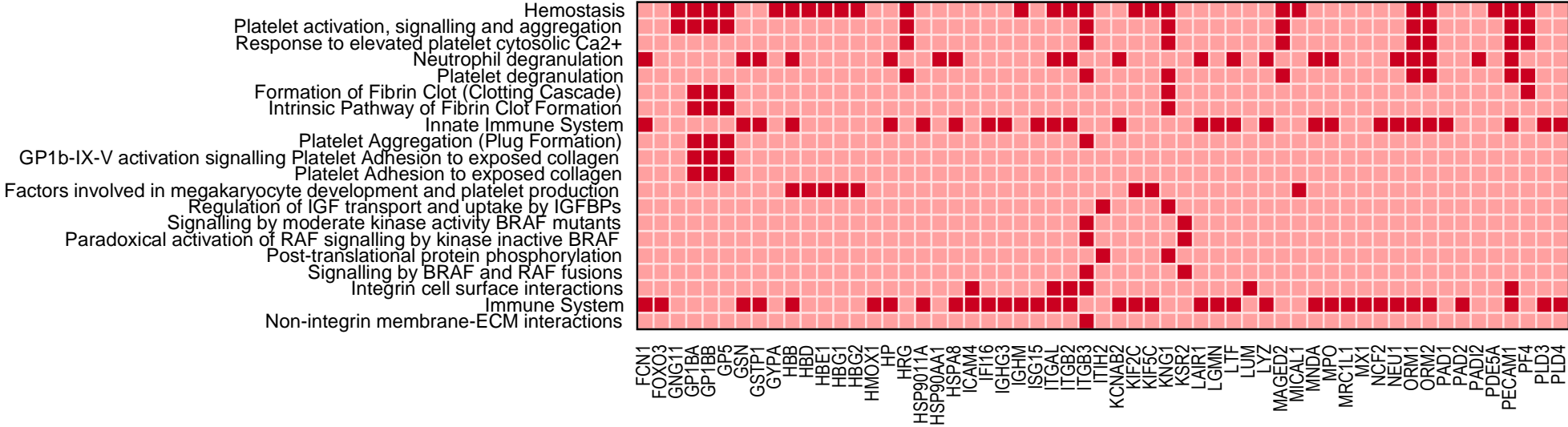
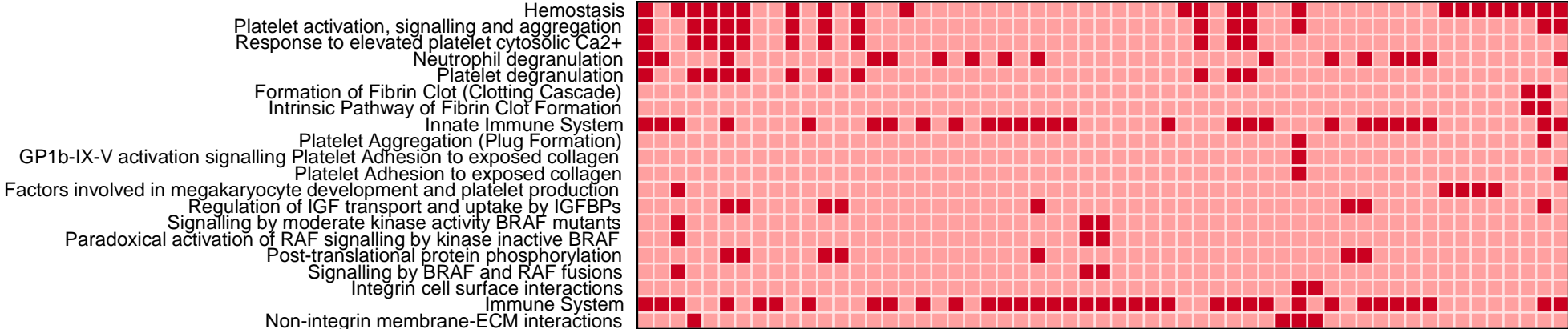


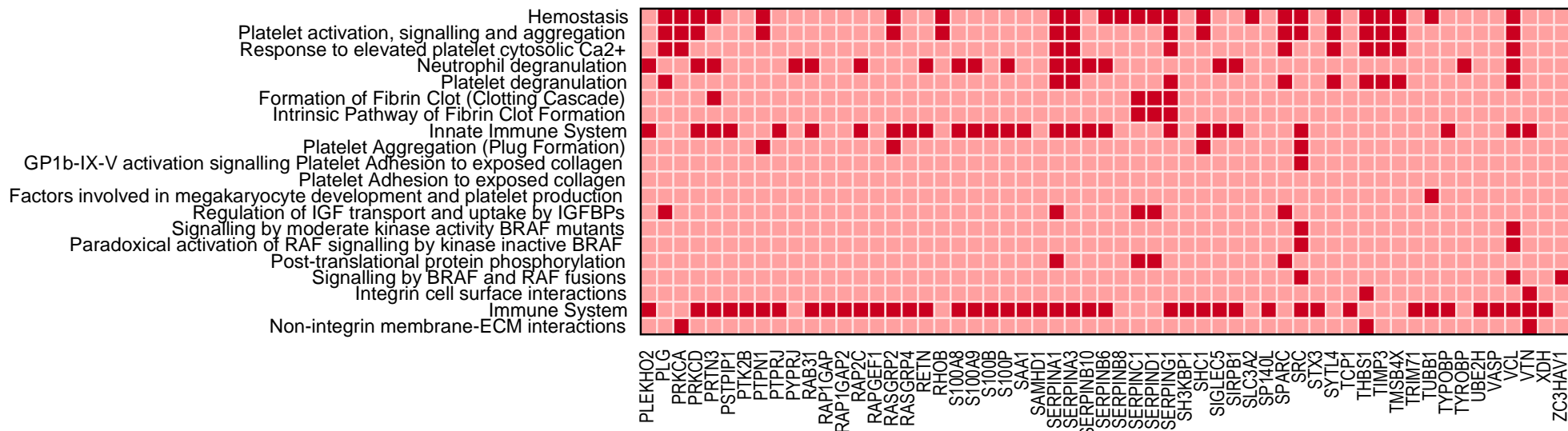
Supplementary Figure 10. Megakaryocyte associated proteins decrease in level during erythroid differentiation. Tandem Mass Tag labelling and nano-LC MS/MS was performed on whole cell lysates of erythroid cells at day 8 and 17 in culture. The relative decrease in level of megakaryocyte associated proteins at day 17 (majority of cells orthochromatic normoblasts and reticulocytes) compared to day 8 (majority of cells pro-erythroblasts) are shown.



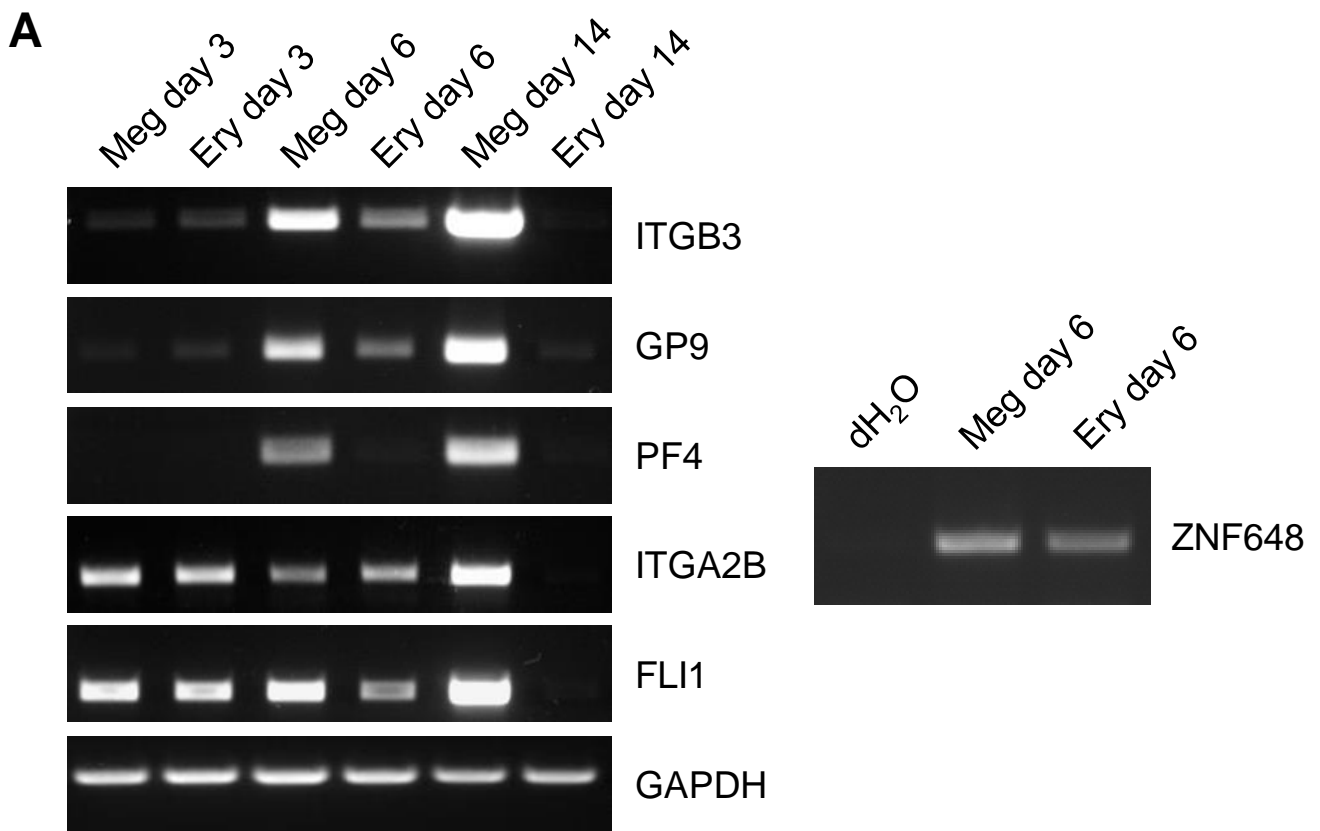
Supplementary Figure 11. Pathway classification of those proteins increased in level following ZNF648 KD also increased in erythroid cells differentiated from three iPSC lines. Proteins increased in level by four-fold or more following ZNF648 knockdown in adult erythroid cells, and also increased in level in erythroid cells differentiated from all three iPSC lines (compared to normal adult erythroid cells; Trakarnsanga *et al* 2014¹²) at day 8 in culture were entered into Reactome. Dark red indicates protein apportioned to pathway shown on left. Top 20 pathways shown with minimum probability of over representation in the dataset, corrected for false discovery rate, of $p < 2.39 \times 10^{-5}$

Supplementary Figure 12





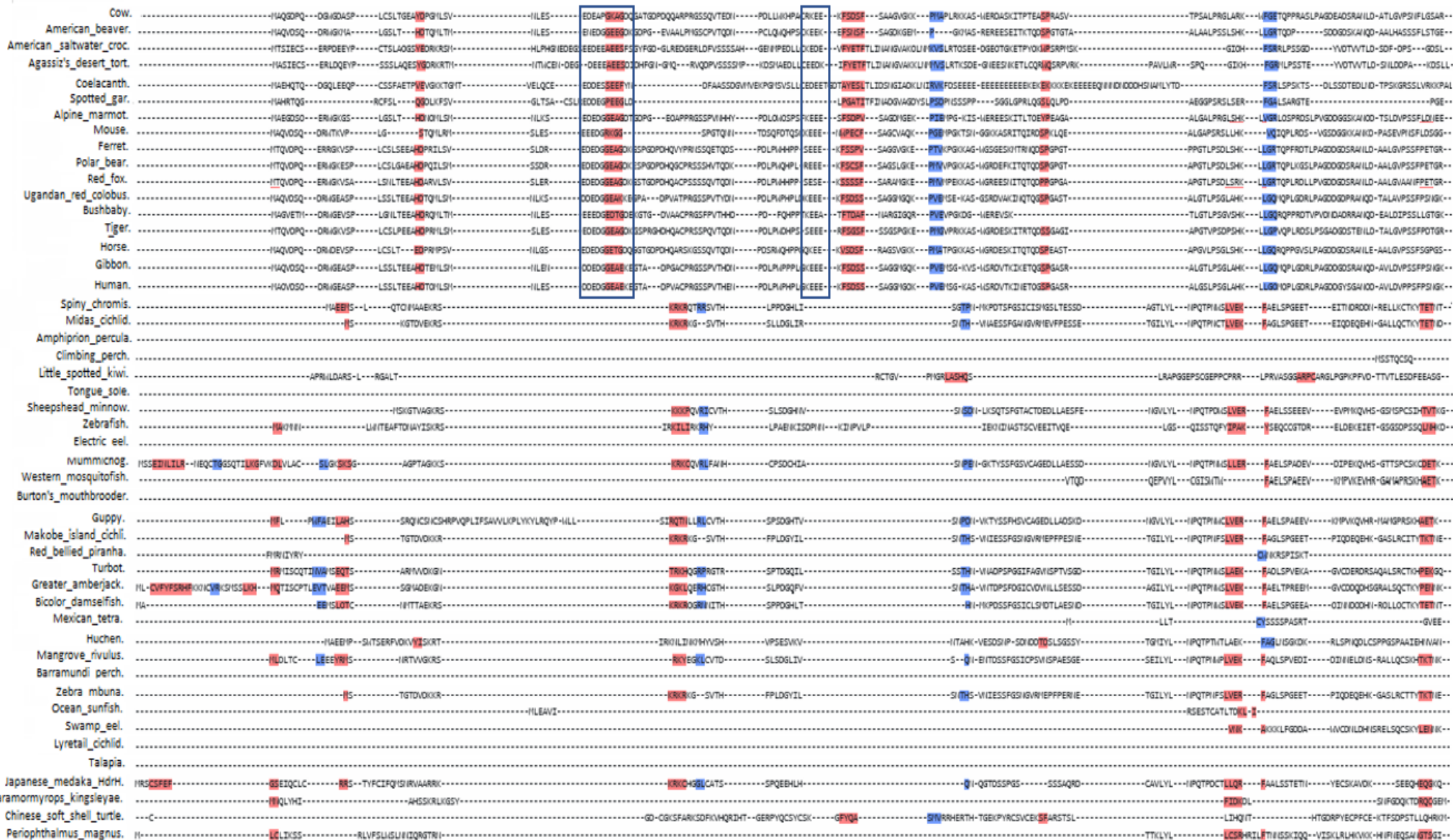
Supplementary Figure 12. Pathway classification of proteins increased in erythroid cells differentiated from three iPSC lines compared to normal adult erythroid cells. Common proteins increased in level by two-fold or more in proteomic dataset (Trakarnsanga *et al* 2014¹²) from three iPSC lines compared to normal adult erythroid cells at day 8 in culture were entered into Reactome. Dark red indicates protein apportioned to pathway shown on left. Top 20 pathways shown with minimum probability of over representation in the dataset, corrected for false discovery rate, of $p < 1 \times 10^{-5}$



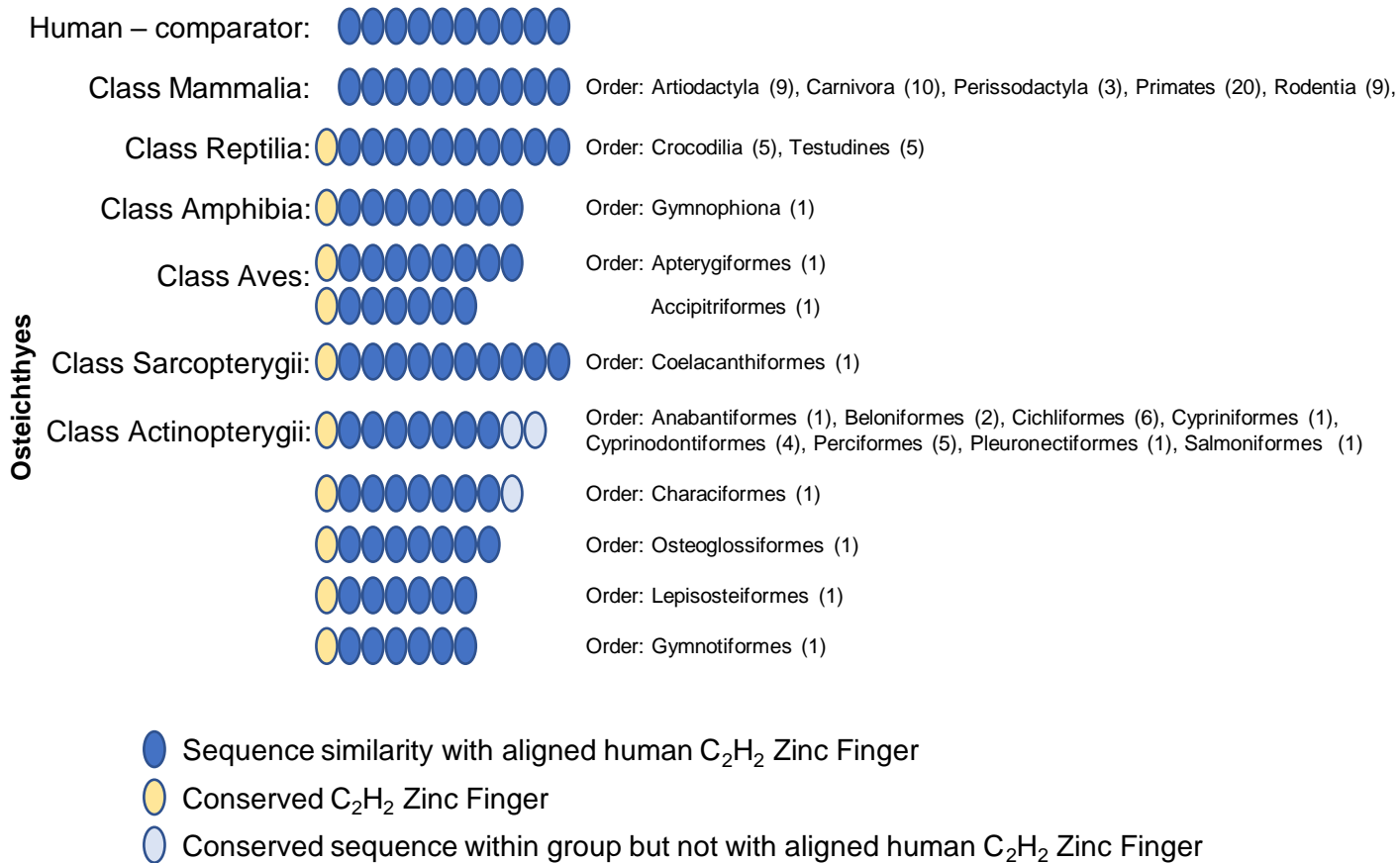
B

GENE	Forward primer	Reverse primer
ZNF648	5'-GTATGGAAGCCTACGTGCTC-3'	5'-AGAGGTCACACCCTAGTTCC-3'
ITGB3	5'-GAGTTCCCAGTGAGTGAGGC-3'	5'-CTCATTGAAGCGGGTCACCT-3'
GP9	5'-ACACTCCACCTTCCCTAGTCA-3'	5'-GCGTCACATCGAGGGTCTG-3'
PF4	5'-CCAACTGATAGCCACGCTGA-3'	5'-CCAAGCATAACCAGTATTCAC-3'
ITGA2B	5'-TTACTTGGGGTTCCAGTTTGA-3'	5'-GCTACCTACGGGCGTCTTC-3'
FLI1	5'-CCTCAACAAAAGTCCTCCCCTT-3'	5'-GGGAGGGACAAAGTTCACCT-3'
GAPDH	5'-ACCACAGTCCATGCCATCAC-3'	5'-TCCACCACCCTGTTGCTGTA-3'

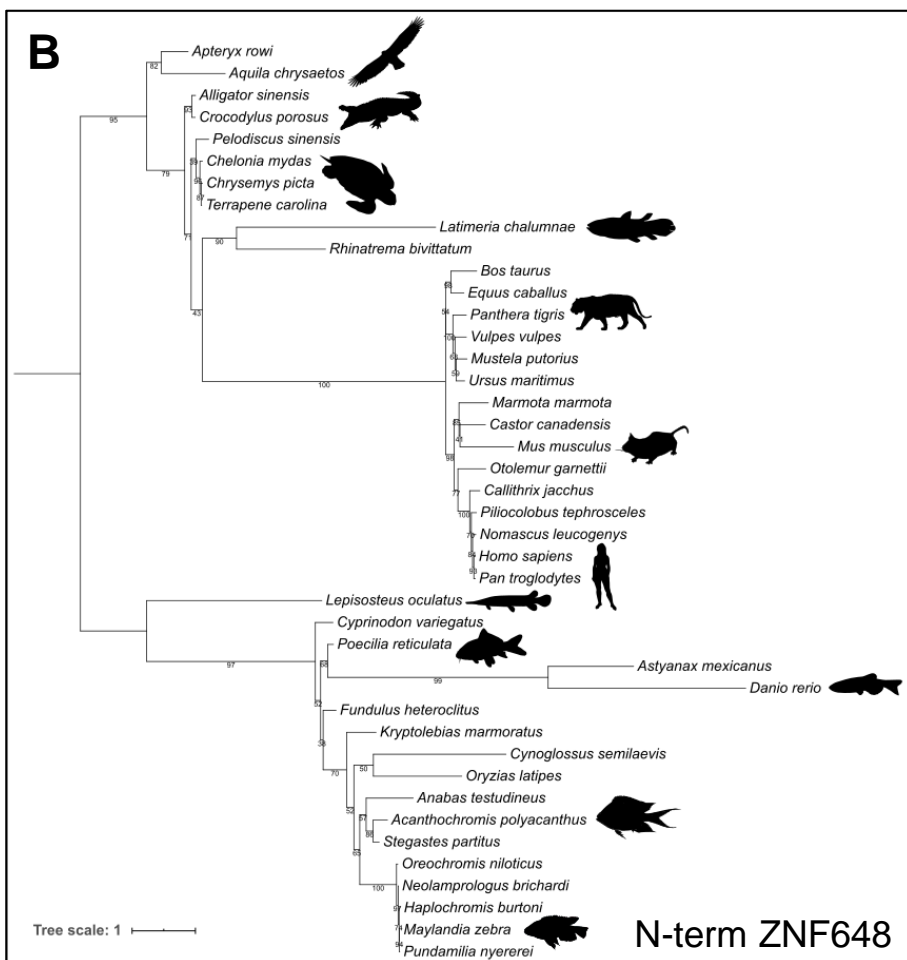
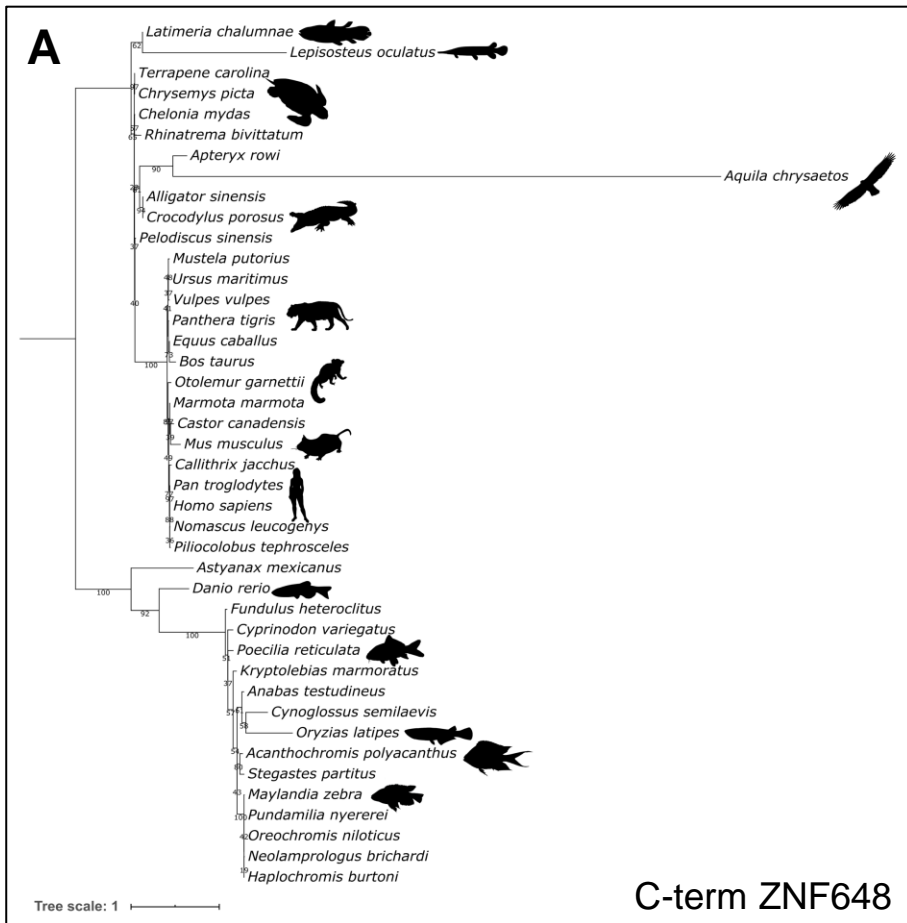
Supplementary Figure 13. Comparison of ZNF648 expression and expression of key megakaryocyte proteins in erythroblasts and megakaryoblasts. A) Expression of ZNF648, ITGB3, GP9, PF4, ITGA2B and FLI1 was analysed by RT-PCR in erythroblasts (Ery) and megakaryoblasts (Meg) differentiated from adult peripheral blood CD34⁺ cells at day in culture indicated. RT-PCR of GAPDH was used as a control for cDNA concentration. Note on day six 32% of the population are megakaryoblasts (CD61⁺/CD41⁺) and 14% of the population on day fourteen are mature megakaryocytes (CD61⁺/CD42b⁺). In contrast all cells in the erythroid cultures are erythroid. B) Primer sequences for RT-PCR.



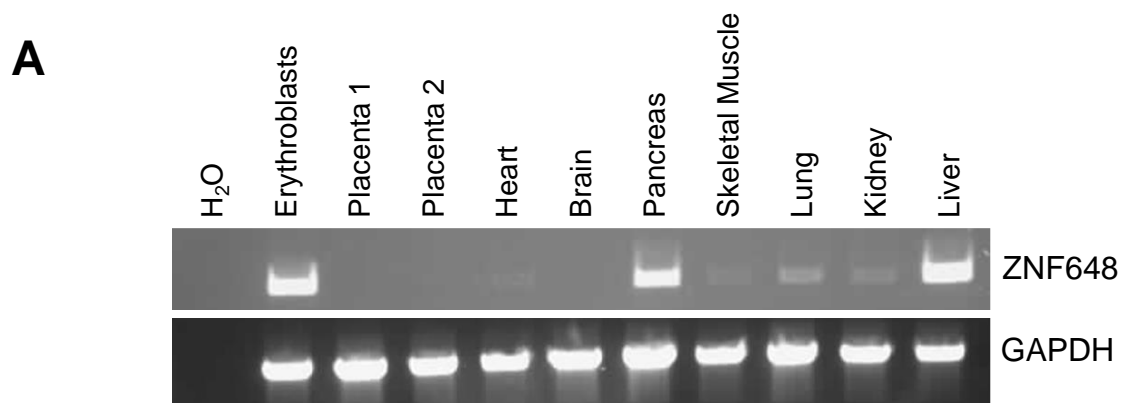
Supplementary Figure 14. N-terminal regions of ZNF648 from range of species analysed by predicted secondary structure (PSIPRED) alignment. Sequences shown aligned from residue 1 to 226 of human sequence. The cyan boxes surround putative patches of conserved charge residues. Secondary structure: α -helix highlighted in red, β -sheet highlighted in blue



Supplementary Figure 15. Schematic of ZNF648 Zinc Finger domain structure in range of animals. Structure of human ZNF648 zinc finger domain is comparator. ZNF648 was identified in family members of all orders shown. Values in brackets are the number of members of each order for which sequences were compared.



Supplementary Figure 16. Phylogeny of ZNF648 orthologue
A) ZNF648 C-terminal ZnF domain and B) ZNF648 N-terminal region in Osteichthyes. See legend to Figure 8 for details.



B

GENE	Forward primer	Reverse primer
ZNF648	5'-GTATGGAAGCCTACGTGCTC-3'	5'-AGAGGTCACACCCTAGTTCC-3'
GAPDH	5'-ACCACAGTCCATGCCATCAC-3'	5'-TCCACCACCCTGTTGCTGTA-3'

Supplementary Figure 17. Expression of ZNF648 in range of human tissues. A) ZNF648 transcript level analysed by PCR in range of tissues using Human MTC™ Panel I from Clontech along with day 8 adult erythroid cells. PCR of GAPDH was used as a loading control, along with a no cDNA control (H₂O). B) Primer sequences for RT-PCR.

Supplementary Table 1. Genes up and down regulated following ZNF648 knockdown.

Table 1A. genes down regulated

ABCC4	ACER3	AGPAT4	AGPAT9	AK7
ALOX12	ALOX5	ANGPT1	ANKRD27	ANXA3
APOBEC3A	ARG1	ARHGEF3	ARPP-19	ASAP2
ATP6V1A	AZIN1	BANK1	BCAN	BDNF
BMP2K	BPGM	BTK	C12orf39	C19orf33
C19orf59	C21orf45	C2orf88	C5orf4	CA13
CALD1	CALM3	CAPRIN1	CCDC68	CCND3
CD226	CD9	CDC42EP3	CDKN2D	CEACAM1
CLEC1B	CLU	CMTM5	COL18A1	COMMD6
CRIPT	CTSE	CTTN	CXCL5	DAB2
DAPP1	DKK1	DMC1	DNM3	EAF2
EHD3	ELK1	ELOVL7	ENDOD1	F13A1
F2R	F2RL2	FAM150B	FAM40B	FAM65A
FCER1A	FCER1G	FCGR2A	FERMT3	FGFR1OP2
FGFR3	FHL1	FLCN	FLI1	FRMD4B
FYB	FYN	GAS2L1	GFPT1	GJA4
GNAZ	GNG11	GOLSYN	GP1BA	GP5
GP6	GP9	GPX7	GSN	GUCY1B3
GYPA	H1F0	HBB	HBE1	HEMGN
HGD	HIPK2	HMB5	HPSE	HSPC159
HTR2A	ICK	IGF2BP3	IGSF3	ITGA2B
ITGA6	ITGB3	JAZF1	KALRN	KCNE1
KCNE3	KIAA0776	KIF26A	KLHL5	L2HGDH
LAMP2	LAT2	LCN2	LCP2	LOC100127983
LOC100130938	LOC339803	MAML3	MAPK8	MEF2C
MEIS1	MFAP3L	MGLL	MMRN1	MOSPD1
MPL	MS4A2	MS4A3	MYCT1	MYLK
MYO9A	NCF1C	NFIB	NRGN	OGFRL1
OSBP2	PATL2	PDE5A	PDE7A	PDLIM1
PEAR1	PECAM1	PF4	PF4V1	PLEK
PLS1	PPBP	PREP	PRICKLE1	PRTFDC1
PTGS1	QPCT	RAB21	RAB27B	RAB6A
RASGRP3	RCN2	RGS18	RHD	RTN1
SCD5	SDPR	SELENBP1	SELP	SENP1
SEPT5	SERPINE2	SF3B2	SFRP1	SH3BGRL3
SH3BP5	SH3TC2	SHISA5	SLC22A17	SLC35D3
SLC37A1	SNCA	SPARC	ST3GAL6	STARD8
SUCNR1	SYTL3	TBXA2R	TBXAS1	TC2N
TCF3	TGFB1	TMEM163	TMEM40	TMEM55A
TMSB15A	TUBO3	TRPC6	TUBA4A	TUBB1
UBASH3B	VCL	VDR	VWF	WFDC1
ZNF185	ZNF25	ZNF479		

Table 1B. genes up regulated

ABHD14A	ACPL2	ANKRD5	ANKRD7	AP1S3
ARNTL2	ASB9	ATP8B1	BLCAP	C12orf5
C1orf53	C21orf2	CALB1	DACH1	DDAH1
DDX31	DENND5B	DOCK9	DYX1C1	EFCAB2
EME2	EPB41L5	FAM135A	FAM83H	FMNL2
FST	HBS1L	HFE	HOMER1	HOXA9
HTR1F	IKZF1	KIAA1598	LEPR	LGALS3BP
LOC642852	LOC729991	LRRC7	LTB	MAGI3
MARCKSL1	MFSD3	MIOS	MIPOL1	MOBKL2A
MYO6	NAT14	NBPF1	NMU	NTS
PANK1	PDE3B	PLEKHA1	PNMAG6A	PPAP2A
PRTN3	RABEP2	RARS	RDH10	RUFY3
SEPP1	SGPL1	SLC26A2	SLC38A1	SORBS1
SPAG1	STAMBPL1	STX4	STXBP6	TANC1
TJP1	TMCO4	TMEM117	TMEM176B	TRIM26
TSPAN6	TULP4	UBL3	USP44	WT1
ZNF165	ZNF521	ZNF625	ZSCAN16	

Supplementary Table 2. Gene names for proteins increased and decreased in level following ZNF648 knock down in adult erythroid cells.

Table 2A. Proteins decreased in level

ALAS2	ANKRD27	ANXA2	ANXA4	AQP1
AQP4	ASXL2	BCAM	BPGM	C15orf41
C2orf88	CA1	CCND3	CCPG1	CD9
CDRT4	CENPV	CKB	CLASP2	CMTM5
DDRKG1	DHX57	DKFZp434P232	DLG4	DNAJB9
DOK3	EIF2AK1	ENDOD1	EPB42	EPN2
ERICH6B	ESAM	F13A1	F2R	FAM65A
FBXO28	FCER1G	FCGR2A	FLI1	FYB
GLYR1	GP1BA	GP1BB	GP5	GP9
GUCY1B3	GYP A	HBA1	HBA2	HBB
HBD	HBG1	HBMS	HBZ	HEMGN
HMBS	HPSE	HVCN1	IKBK G	IMPA2
ISG20	ITGA2B	ITGB3	KALRN	LGALS3
LGALS L	LINC00176	LTBP1	MARC	MFF
MGLL	MTSS1L	NCAPG	NCF2	NEK6
NOGOC	NRGN	OSBP2	PDE5A	PDLIM1
PF4	PPBP	PPP1R14A	PRC1	PREP
QARS	RAB27B	RAB3IL1	RAB5B	RCN2
RGCC	RGS14	RGS18	RGS3	RHCE
RHOC	RNF115	RTN1	RTN4	S100A10
SDPR	SELENBP1	SELP	SERPINE1	SERPINE2
SH3BGRL3	SLC2A1	SLC4A1	SOX6	SPARC
SQRDL	SRPK2	STON1-GTF2A1L	SULT1A1	SULT1A3
TBXAS1	TCF3	THBS1	TUBB1	UROS
VWF	ZNF14	ZNF397		

Table 2B. Proteins increased in level

A1BG	ABCA12	ABI2	ACTR1A	ADAR2d
AFM	AGKC	AGT	AHSG	ALB
AMBP	AMBP	ANKLE1	APC2	APOA1
APOA2	APOA4	APOC2	APOC3	APOE
APOH	APOL1	APOM	ATP2B1	ATRN
ATXN2	AZGP1	BSDC1	C19orf47	C1QB
C1S	C3	C4B	C4orf29	C5
C6	C7	C9	CALB1	CBX3
CCAR1	CCNL1	CD5L	CD5L	CDC42BPB
CF1	CFHR1	CFI	COL23A1	CP
CPEB3	CSNK1D	DACH1	DKFZp434K0831	DKFZp667N107
DKFZp686K03196	DKFZp686N02209	DOCK9	EIF2C3	ENTHD2
EPCAM	EPS8L1	F12	F2	F2
FAM127A	FBXW9	FLJ37549	FRMPD3	GC
GLUL	GML	GPT2	GSKIP	HAT1
hCG	hCG_40889	HNRNPA3	HNRNPC	HNRNPL
HP	HPR	HPX	HPX	HRG
HSP90AB3P	IDH3A	IGHA1	IGHG1	IGHG1
IGHG2	IGHG3	IGHM	IGKC	IGKV1-17
IGKV3-20	IGKV4-1	IGLC1	IGM	ITIH1
ITIH2	ITIH2	KANSL3	KIAA1731	KIT
KNG1	KRT6A	LAMA2	LRG1	LRRD1
LUM	MFS D3	MGAT5	MTG1	MYO5A
NBL1	NCOA5	NONO	NRXN3	orm1
ORM2	PANK1	PANK3	PGLYRP2	PKD1
PLG	PON1	PPAP2A	PPIG	PRPF40A
PTOV1	PURG	RAB6A	RALY	RAP2C
RBBP6	RBM1	RBM12B	RBM15	RBM25
RBMX	RBP4	RGPD4	RNF135	SAFB
SAFB2	SERPINA1	SERPINA3	SERPINA6	SERPINA7
SERPINC1	SERPIND1	SERPINF1	SERPINF2	SERPING1
SFPQ	SFRS10	SLTM	SON	SPIN1
SRSF10	SRSF8	SRSF9	SUB1	SYNCRIP
TF	TRA2A	TSPAN6	TTR	VTN
ZBED9	ZNF106	ZNF326	ZNF638	ZNF814

Supplementary Table 3. Gene names for proteins increased in level following ZNF648 knock down in adult erythroid cells and in erythroid cells differentiated from iPSC lines.

A1BG	C3	IGHA1	RF-IP9
AFM	C9	IGHG1	SERPINA1
AGT	CD5L	IGHG3	SERPINA3
AHSG	CF1	IGL	SERPINA7
AMBP	CP	IGL@	SERPINC1
APOAI	F2	KNG1	SERPIND1
APOA4	F12	LUM	SERPINF1
APOC2	GC	ORM1	SERPING1
APOC3	HP	ORM2	VTN
APOH	HPX	PGLYRP2	ZBED9
APOL1	HRG	PLG	
APOM	IDH3A	RBP4	

Supplementary Table 4. Sequence similarity of ZNF648 across a range of mammals.

Organism Subject Human	C-terminal zinc finger domain homology (%)	N-terminal region homology (%)	Whole protein homology (%)
Great Ape: Bonobo/Pygmy Chimpanzee (<i>Pan paniscus</i>)	100	98	99
Lesser Ape: White-Cheeked Gibbon (<i>Nomascus leucogenys</i>)	100	94	97
Old World Monkey: Drill (<i>Mandrillus leucophaeus</i>)	99	89	94
New World Monkey: White-headed capuchin (<i>Cebus capucinus imitator</i>)	98	87	92
Cat (<i>Felis catus</i>)	97	70	83
Dog (<i>Canis lupus familiaris</i>)	97	71	84
Horse (<i>Equus caballus</i>)	97	69	83
Polar Bear (<i>Ursus maritimus</i>)	97	72	84
Orca (<i>Orcinus orca</i>)	96	64	79
Mouse (<i>Mus musculus</i>)	94	50	73

Supplementary Table 5. ZNF648 sequence conservation between closely related species.

Great Ape Subject: Human	C-terminal zinc finger domain homology (%)	N-terminal region homology (%)	Whole protein homology (%)
Bonobo/pygmy chimp	100	98	99
Gorilla	100	98	99
Common chimpanzee	100	97	98
Sumatran Orang-utan	99	97	98
Old World Monkey Subject: Drill (<i>Mandrillus leucophaeus</i> ; family Cercopithecidae)	C-terminal zinc finger domain homology (%)	N-terminal region homology (%)	Whole protein homology (%)
Family CERCOPTHECINAE			
<i>Macaca nemestrina</i> (Southern pig-tailed macaque)	99	97	98
<i>Theropithecus gelada</i> (Gelada baboon)	100	97	98
<i>Macaca fascicularis</i> (Crab-eating macaque)	99	97	98
<i>Papio anubis</i> (Olive baboon)	100	97	98
<i>Macaca mulatta</i> (Rhesus macaque)	99	97	98
<i>Cercocebus atys</i> (Sooty mangabey)	100	96	98
<i>Chlorocebus sabaeus</i> (Green monkey)	99	94	96
Family COLOBINAE			
<i>Rhinopithecus bieti</i> (Black snub-nosed monkey)	100	93	96
<i>Rhinopithecus roxellana</i> (Golden snub-nosed monkey)	100	93	96
<i>Colobus angolensis</i> (Angolan colobus)	100	92	96
Old World Monkey Subject: Angolan colobus (<i>Colobus angolensis</i> ; family Colobus)	C-terminal zinc finger domain homology (%)	N-terminal region homology (%)	Whole protein homology (%)
Family COLOBINAE			
<i>Piliocolobus tephrosceles</i> (Ugandan red colobus)	100	97	98
<i>Rhinopithecus bieti</i> (Black snub-nosed monkey)	100	96	98
<i>Rhinopithecus roxellana</i> (Golden snub-nosed monkey)	100	96	98
Family CERCOPTHECINAE			
<i>Chlorocebus sabaeus</i> (Green monkey)	99	95	97
<i>Macaca nemestrina</i> (Southern pig-tailed macaque)	99	94	96
<i>Macaca fascicularis</i> (Crab-eating macaque)	99	94	96
<i>Papio anubis</i> (Olive baboon)	100	93	96
<i>Theropithecus gelada</i> (Gelada baboon)	100	93	96
<i>Macaca mulatta</i> (Rhesus macaque)	99	93	96
<i>Cercocebus atys</i> (Sooty mangabey)	100	92	96
New World Monkey Subject: White-headed capuchin (<i>Cebus capucinus imitator</i> ; family Cebidae)	C-terminal zinc finger domain homology (%)	N-terminal region homology (%)	Whole protein homology (%)
Cebidae			
<i>Saimiri boliviensis</i>	98	92	96
Aotidae			
<i>Aotus nancymae</i>	98	92	95
Callitrichidae			
<i>Callithrix jacchus</i>	98	90	94

Felidae Subject: Cat (<i>Felis catus</i>)	C-terminal zinc finger domain homology (%)	N-terminal region homology (%)	Whole protein homology (%)
<i>Acinonyx jubatus</i> (Cheetah)	100	98	99
<i>Puma concolor</i> (Cougar)	100	98	99
<i>Panthera tigris altaica</i> (Siberian Tiger)	100	97	98
<i>Panthera pardus</i> (Leopard)	100	97	98
Carnidae Subject: Dog (<i>Canis lupus familiaris</i>)	C-terminal zinc finger domain homology (%)	N-terminal region homology (%)	Whole protein homology (%)
<i>Canis lupus dingo</i> (Dingo)	100	100	100
<i>Vulpes vulpes</i> (red fox)	100	96	98
Equine Subject: Horse (<i>Equus caballus</i>)	C-terminal zinc finger domain homology (%)	N-terminal region homology (%)	Whole protein homology (%)
<i>Equus przewalskii</i> (Przewalski's horse)	100	99	99
<i>Equus asinus</i> (Donkey)	100	99	99
Ursidae Subject: Polar bear (<i>Ursus maritimus</i>)	C-terminal zinc finger domain homology (%)	N-terminal region homology (%)	Whole protein homology (%)
<i>Ursus arctos horribilis</i> (grizzly bear)	100	100	100
<i>Ailuropoda melanoleuca</i> (giant panda)	100	93	96
Muroidea Subject: House Mouse (<i>Mus musculus</i>)	C-terminal zinc finger domain homology (%)	N-terminal region homology (%)	Whole protein homology (%)
Ryukyu Mouse (<i>Mus caroli</i>)	99	95	97
Gairdner's Shrewmouse (<i>Mus pahari</i>)	98	86	93
Brown Rat (<i>Rattus norvegicus</i>)	98	79	89
Desert Woodrat (<i>Neotoma lepida</i>)	97	71	85
Deer Mouse (<i>Peromyscus maniculatus bairdii</i>)	97	72	85
Thicket rat (<i>Grammomys surdaster</i>)	98	80	90
Chinese Hamster (<i>Cricetulus griseus</i>)	98	70	85
Mongolian gerbil (<i>Meriones unguiculatus</i>)	96	71	84
Cetaceans Subject: Orca (<i>Orcinus orca</i>)	C-terminal zinc finger domain homology (%)	N-terminal region homology (%)	Whole protein homology (%)
<i>Delphinapterus leucas</i> (Beluga whale)	98	97	98
<i>Eschrichtius robustus</i> (Grey Whale)	99	95	97
<i>Balaenoptera acutorostrata scammoni</i> (Minke Whale)	99	93	96
<i>Monodon Monoceros</i> (Narwhal)	98	98	98
<i>Lagenorhynchus obliquidens</i> (Pacific white-sided dolphin)	98	99	98
<i>Tursiops truncatus</i> (Common bottlenose dolphin)	98	96	98
<i>lipotes vexillifer</i> (Baiji)	98	96	97
<i>Neophocaena asiaeorientalis asiaeorientalis</i> (Finless porpoise)	98	96	97

'Subject' is the ZNF648 sequence used to compare with that of the organisms in respective table. There was a very limited number of ZNF648 sequences available for many family and sub family members.