SUPPLEMENTAL DATA

KAT2A complexes ATAC and SAGA play unique roles in cell maintenance and identity in hematopoiesis and leukemia

Liliana Arede[1,2], Elena Foerner[1*], Selinde Wind[1*], Rashmi Kulkarni[1], Ana Filipa Domingues[1], George Giotopoulos[1,3], Svenja Kleinwaechter[1], Maximilian Mollenhauer-Starkl[4], Holly Davison[4], Aditya Chandru[2], Ryan Asby[1,3], Ralph C. Samarista[4], Shikha Gupta[1,2], Dorian Forte[1,5], Antonio Curti[6], Elisabeth Scheer[7,8,9], Brian J.P. Huntly[1,3], Laszlo Tora[7,8,9], Cristina Pina[4,10#]

[1]Department of Haematology, [2]Department of Genetics and [3]Wellcome Trust Cambridge Stem Cell Institute, University of Cambridge, Cambridge, United Kingdom; [4]College of Health, Medicine and Life Sciences, Division of Biosciences, Brunel University London, United Kingdom; [5]Department of Experimental, Diagnostic and Specialty Medicine – DIMES, Institute of Hematology "Seràgnoli", Bologna, Italy; [6]IRCCS Azienda ospedaliero-universitaria di Bologna, Istituto di Ematologia "Seràgnoli", Bologna, Italy; [7]Institut de Génétique et de Biologie Moléculaire et Cellulaire, 67404 Illkirch, France; [8]Centre National de la Recherche Scientifique (CNRS), UMR7104, 67404 Illkirch, France; [9]Institut National de la Santé et de la Recherche Médicale (INSERM), U1258, 67404 Illkirch, France and [10]Centre for Genome Engineering and Maintenance (CenGEM), Brunel University London, Uxbridge UB8 3PH, United Kingdom

*E.F. and S.W. contributed equally to this study

#Corresponding author (cristina.pina@brunel.ac.uk)

#Correspondence to:

Cristina Pina, MD D.Phil.
College of Health, Medicine and Life Sciences, Division of Biosciences
Brunel University London
Uxbridge UB8 3PH
London
United Kingdom
+44 (0)1895 266676
Email: (cristina.pina@brunel.ac.uk)

SUPPLEMENTAL METHODS

Flow cytometry

CD34+ CB cells or cell lines were stained as described[28]: cell surface antibodies are listed in Supplemental Table 4; quantification of apoptosis used Annexin V (Thermo); cell cycle profiles were obtained using Hoechst 33342 (Thermo). Flow analysis data was acquired on Gallios (Beckman Coulter) and Attune (Thermo) instruments and analysed in Kaluza (Beckman Coulter). Cells were sorted on a FACSAria™ Fusion or Influx instruments (BD).

Cell divisional tracking

For cell divisional tracking and proliferation, 1E6 cells/mL of cell suspension were stained as per manufacturer's instructions using 1μL of 5 μM Tag-it Violet (BioLegend) in PBS. Cells were incubated for 20 min at RT, protected from light. Staining was quenched by adding 5 mL of cell culture medium R20. Pelleted cells were re-suspended in pre-warmed R20 incubated for 10 min at 37°C and analysed daily by flow cytometry on Day0 (loading) to Day3.

Cytospins

Approximately 10E5 transduced cells per construct were centrifuged onto slides for 5 min at 700 rpm, stained with rapid Romanowsky stain pack as per manufacturer's instructions and fixed with Depex mounting medium.

Western Blot

Western blot analysis was performed on whole cell extracts (SPT20, KAT2A, USP22, SGF29), or nuclear extracts (SPT20, ZZZ3) using the antibodies listed in Supplemental Table 2. Blots were resolved by SDS-PAGE. Transfer membranes (Immobilon-P PVDF pore size 0.45 μm, Millipore) were activated with Methanol (Sigma) for 5 minutes. Gels were transferred onto the transfer membrane for about 90 minutes at 100V. Membranes were stained with Ponceau for 1 minute, blocked in blocking buffer (5% skimmed milk in TBS-T (Tris 10 mmol/L, NaCl 50 mmol/L, Tween 0.005%)) for 1h at RT, washed and immunoblotted with primary antibodies diluted in blocking buffer overnight, at 4°C with rotation. The following day, membranes were washed in TBS-T 3x for 10 minutes, and incubated with the appropriate fluorescently-labelled secondary antibody diluted in 1.5% skimmed milk in TBS-T for 1h at RT. Membranes were washed 2x in TBS-T and 1x in TBS

for 10 minutes, subjected to 5 minutes SuperSignal West Pico PLUS Chemiluminescent Subtract (Thermo, Cat. No. 34580) or Immobilon Classico Western HRP Substrate from Merck Millipore (Cat. No. WBLUC0100) for detection of HRP. Membranes were developed on a SRX-101A Processor using the hypercassette method. After drying, membranes were taped onto a hypercassete (Amersham Biosciences) and exposed to X-ray film (exposure dependent on the antibody strength) prior to developing.

Quantitative Real time PCR (Q-RT-PCR)

RNA extraction, cDNA synthesis and Q-PCR analysis were performed as described[7]. Primers and Taqman probes assays (Thermo) are listed in Supplemental Tables 5 and 6, respectively. Relative gene expression calculated by the $2^{-\Delta\Delta Ct}$ method using *HPRT1* as reference.

RNA sequencing

RNA was extracted from *CTRLsh* or *KAT2Ash* HSCs obtained from 2 individual donors using Trizol reagent (Thermo) and linear polyacrylamide (Sigma) as a carrier. RNA-seq libraries were prepared at the Cambridge Stem Cell Institute Genomics Core Facility using the Ovation RNA-seq kit (NuGen) with incorporated DNase treatment, as per manufacturer's instructions. Libraries were sequenced on an Illumina HiSeq4000 instrument at the CRUK Cambridge Research Institute Genomics Core Facility using 50bp single-end reads. The raw fastq files were processed as per the RSEM v1.2.31 workflow and aligned to the reference human genome assembly GRCh37. Differentially expressed genes were obtained at 10% FDR using the R packge edgeR. Gene signatures of down-regulated genes in *KAT2Ash* HSC for gene set enrichment analysis were obtained from MSigDB. Erythroid differentiation (NES= -1.79, q-val=0.018) and platelet biology (NES= -2.05, q-val=0.004).

SUPPLEMENTAL TABLES

Supplemental Table 1. *Sequences of shRNA constructs.*

Oligonucleotide	Sequence (5' to 3')
CTRL	TCAACAAGATGAAGAGCACCAAGGGATCCGTTGGTGCTCTTCATCTTGTTGTTTTTTC
KAT2A	TGCTGAACTTTGTGCAGTACAAGGGATCCGTTGTACTGCACAAAGTTCAGCTTTTTTC
SUPT20Hsh	TCCATCAAGTATTCCTCGGAAAGGGATCCTTTCCGAGGAATACTTGATGGTTTTTT

SUPT20Hsh2	TGCGGATGTGTCATAGCAGAAAGGGATCCTTTCTGCTATGAC ACATCCGCTTTTTT
TADA2Bsh	TCGTGACTGTGAAGACTATTATGGGATCCATAATAGTCTTCA CAGTCACGTTTTTTG
TADA2Ash	TTTGAAGATGACTCGGACATTTGGGATCCAAATGTCCGAGTC ATCTTCAATTTTTTG
TADA2Ash2	TGCACTATATGAAGCATTTCATGGGATCCATGAAATGCTTCA TATAGTGCTTTTTT
USP22	TAGCTACCAGGAGTCCACAAAGGGGATCCCTTTGTGGACTC CTGGTAGCTTTTTTTC
ZZZ3	TGCATCAGATGACGAAAGTATTGGGATCCAATACTTTCGTCA TCTGATGCTTTTTTC

Supplemental Table 2. *Antibodies used in WB and in chromatin immunoprecipitation (ChIP).*

Antibody	**Catalogue #**	**Supplier**
ACTB	ab8227	Abcam
CCDC101	ab204367	Abcam
GCN5 (KAT2A)	sc-20698	Santa Cruz
H3K9ac	07-352	Millipore
Rabbit IgG	12-370	Millipore
SPT20	SPT20 3006	Produced in the Tora Lab[13]
TBP	Ab63766	Abcam
Tubulin	sc-5286	Santa Cruz
USP22	ab195289	Abcam
ZZZ3	ZZZ3 2616	Produced in the Tora Lab[13]

Supplemental Table 3. *Sequences of ChIP-qPCR primers.*

Gene	**Forward**	**Reverse**
HBB	*GCCATCCATTTTTCTTAATTCTGA G*	*TGAGGGCACCATTAGCCAG*
HOXA10	*GTTTATAGCGGCGCATTCCA*	*CGGGTTTGATTTCTGAGCCC*
HOXA9	*CGCTCTCATTCTCAGCATTG*	*TTAAACCTGAACCGCTGTCG*
Intergenic region	*TGGTTTGGAGTGGGTGCT*	*TCCTGCTCTCCGTCACCT*
KRT5	*AGGTTGTAGAGGCTCCGGCT*	*CAGCTTCACCTCCGTGTCCC*

MEIS1	*CCAGAAGAAGACAGAGCGGA*	*CCCTCAGACCCAACTACCAA*
RPS7	*CCTGCTCTCCGACAGAACTT*	*CGGGTAATCGGCTGTATCCC*
EPOR	*TGGCACATAGCGAACATTCCA*	*GGCTGGGAAGAGAATGCTGATT*

Supplemental Table 4. *Antibodies used in flow cytometry analysis and cell sorting.*

Antibody	**Fluorochrome**	**Catalogue #**	**Clone**	**Dilution**	**Supplier**
CD34	PE-Cy7	343515	581	1:200	BioLegend
CD38	PE	356603	HB-7	1:100	BioLegend
CD45RA	APC	304118	HI100	1:100	BioLegend
CD123	PE-Cy5	306008	HI264	1:100	BioLegend
CD71	PE	334105	GY1G4	1:100	BioLegend
CD235a	PB	349103	H1264	1:100	BioLegend
Annexin V	APC	640941	-	1:100	BioLegend
Hoechst 33342	-	H3570	-	1:10000	LifeTechnologies
Tag-it Violet	APC	425101	-	5nM	BioLegend

Supplemental Table 5. *Sequences of qRT-PCR primers.*

Gene	**Forward**	**Reverse**
HBB	*AGGAGAAGTCTGCCGTTACTG*	*CCGAGCACTTTCTTGCCATGA*
HOXA10	*GAGAGCAGCAAAGCCTCGC*	*CCAGTGTCTGGTGCTTCGTG*
HOXA9	*GGTGACTGTCCCACGCTTGAC*	*GAGTGGAGCGCGCATGAAG*
HPRT1	*CCTGGCGTCGTGATTAGTGAT*	*TCGAGCAAGACGTTCAGTCC*
KAT2A	*CCCGCTACGAAACCACTCAT*	*GCATGGACAGGAATTTGGGGA*
RPL13	*CGCAGGAGCCGCAGG*	*CTGCCAGTCCTTGTGGAAGT*
RPS7	*CCCAGGAGCCGTACTCTGA*	*GCCATCTAGTTTGACGCGGA*
SUPT20H	*CCCTTAAATCTACTCCAGCTTCCAG*	*TTGACTGGTTGAACCTTGCTC*
TADA2A	*CGCCTTAAACGCACTATGCTC*	*GGCCGGAATCAATGTCAGCTT*
TADA2B	*GCTACCACGGCTACCAGC*	*AGCCATATCTTCCCAGTTTCCG*
USP22	*GAGGCCATGGACGCCG*	*AGATACAGGACTTGGCCTTGC*

ZZZ3	GGACAGCAAAACAGGTTGCC	GTGCTGTCGTCTGCTTGTTG

Supplemental Table 6. *Taqman probes.*

Gene	Catalogue#	Supplier
CEBPA	Hs00269972_s1	Thermo
EPOR	Hs00959427_m1	Thermo
GATA1	Hs01085823_m1	Thermo
GATA2	Hs00231119_m1	Thermo
HPRT1	Hs02800695_m1	Thermo
ID2	Hs04187239_m1	Thermo
KAT2A	Hs00221499_m1	Thermo
KLF1	Hs00610592_m1	Thermo
MPO	Hs00924296_m1	Thermo
RPL15	Hs04334752_g1	Thermo
RPL3	Hs01581771_g1	Thermo
TAL1	Hs01097987_m1	Thermo

SUPPLEMENTAL FILES

Supplemental File 1: ZZZ3 and SPT20 ChIP-seq peaks in human K562 cells.

Supplemental File 2: RNA-seq differentially expressed genes in CTRLsh *vs* KAT2Ash HSCs (10% FDR).

Supplemental File 3: Genes enriched in Intermediate Erythroblasts (IntE) as defined in a microarray study of human erythroid differentiation by (Merryweather-Clarke et al. (2011)[31]. Cluster 17 obtained from the Human Erythroblast Maturation database (https://cellline.molbiol.ox.ac.uk/eryth/index.html).

Supplemental File 4: Enriched genes in erythroid-basophil-megakaryocyte-biased progenitors (EBMP) as per detailed single-cell profiling of erythroid development by Tusi et al. (2018)[32]. EBMP region selected using SPRING tool (https://kleintools.hms.harvard.edu/paper_websites/tusi_et_al/).

Supplemental File 5: Characteristics of AML patient samples.

SUPPLEMENTAL FIGURE LEGENDS

Supplemental Figure 1 - KAT2A-containing ATAC and SAGA complexes have unique functional associations in K562 cells.
(A, B) Specificity of ATAC (ZZZ3) and SAGA (SPT20) ChIP-seq targets against ENCODE; data as retrieved by EnrichR online annotation tool (Kuleshov et al., 2016)[34]. **(C, D)** Top Gene Ontology (GO) associations of ZZZ3 and SPT20 ChIP-seq targets on biological processes as calculated by EnrichR (Kuleshov et al., 2016)[34]. **(E)** Western blot analysis of ZZZ3 (left) and SPT20 (right) nuclear protein abundance in K562 cells transduced with *CTRLsh*, *ZZZ3sh* and *SUPT20Hsh*. Observed ZZZ3, SPT20 and SPT20 variant as per Krebs et al. 2011[13] using in-house purified anti-sera. TBP (38 kDa) was used as loading control. **(F)** Western blot analysis of KAT2A protein abundance in erythroid-affiliated K562 cells transduced with *CTRLsh* and *KAT2Ash*. Predicted MW for KAT2A is 95 kDa. ACTB (42 kDa) was used as loading control. **(G)** Quantitative RT-PCR analysis of gene expression in K562 cells transduced with *KAT2Ash*. Mean ± SEM of 4 individual experiments; gene expression relative to *CTRLsh*, normalized to *HPRT1* housekeeping gene. Two-tailed t-test for significance *p<0.05. **(H)** Growth curve of K562 cells transduced with shRNA constructs against ZZZ3, SUPT20H and KAT2A. Mean ± SEM of 3 independent experiments. ANOVA for mixed effects analysis significance *p<0.05.

Supplemental Figure 2 - Investigation of ATAC and SAGA element knockdown in human cord blood (CB) cell specification. (A) Photographs of representative colony morphologies from transduced human CB HSC and progenitors. BFU-E: Boost-forming unit – erythrocyte colony. CFU-GM: Colony-forming unit - granulocyte monocyte colony; CFUGEMM or mixed: Colony-forming unit – granulocyte erythrocyte monocyte megakaryocyte colony. **(B)** Proportion of *ZZZ3sh* transduced CB HSC and progenitors. Mean ± SEM of >3 individual sorting experiments. Two tailed paired t-test for significance; no significant differences. **(C)** Proportion of *SUPT20Hsh* transduced CB HSC and progenitors. Mean ± SEM of >3 individual sorting experiments. Two-tailed paired t-test for significance; no significant differences. **(D)** Western blot analysis of USP22 protein abundance in K562 cells transduced with *CTRLsh* and *USP22sh*. Predicted USP22 MW is 60 kDa. ACTB (42 kDa) was used as loading control. **(E)** Quantitative RT-PCR validation of USP22 knockdown in CB HSCs. Mean ± SEM of 3 individual experiments; gene expression relative to *CTRLsh*, normalized to *HPRT1* housekeeping gene. Paired two-tailed t-test for significance **p<0.01.

(F) Frequency of colony-forming cells in the HSC/MPP, MEP (top), GMP and MLP (bottom) compartments transduced with *USP22sh*. Mean ± SEM of 4 individual CB samples. Two-tailed paired t-test for significance; *p<0.05. **(G)** Proportion of *USP22sh* transduced CB HSC and progenitors. Mean ± SEM of >4 individual sorting experiments. Two-tailed paired t-test for significance; no significant differences.

Supplemental Figure 3 - KAT2A is required for human CB progenitor survival and erythroid specification with no impact on lymphoid and myeloid potential. (A) Relative representation of colony-forming cells present in total CB CD34+ cells transduced with *KAT2Ash*. Mean ± SEM of 3 individual CB samples. Paired two-tailed t-test for significance *p<0.05. **(B)** Relative representation of colony-forming cells present in total CB CD34+ cells treated with MB-3 200μM and vehicle DMSO. Data are represented relative to DMSO. Mean ± SEM of 3 individual CB samples. Paired two-tailed t-test for significance *p<0.05. **(C)** Relative representation of colony-forming cells present in total CB CD34+ cells transduced with *CCDC101sh*. Mean ± SEM of 3 individual CB samples. Paired two-tailed t-test for significance *p<0.05. **(D)** Western blot analysis of SGF29 (*CCDC101* gene) protein abundance in erythroid-affiliated K562 cells transduced with *CTRLsh* and *CCDC101sh*. Predicted SGF29 MW is 36 kDa. ACTB (42 kDa) was used as loading control. **(E)** Flow cytometry analysis of CB CD34+ cells transduced with *CTRLsh* or *KAT2Ash* after 7 days of liquid culture in mixed lineage differentiation conditions (see Methods). Cells plotted were gated as live GFP+ (transduced) CD34-. Data are representative of 2 independent experiments. **(F)** Flow cytometry analysis of CB CD34+ cells treated with MB-3 100μM and 0.1% DMSO after 6 days of liquid culture in mixed lineage differentiation conditions (see Methods). Cells plotted were gated as live GFP+ (transduced) CD34-. Data are representative of 2 independent experiments. **(G)** Flow cytometry analysis of cell cycle in HSC/MPP, MEP, GMP and MLP transduced cells with *KAT2Ash* vs *CTRLsh*. N=3 individual cord blood samples; mean ± SEM. Two-tailed paired t-test not significant. **(H)** Flow cytometry analysis of apoptosis by Annexin-V staining. N=3 individual cord blood samples; mean ± SEM. Two-tailed paired t-test for significance *p<0.05.

Supplemental Figure 4 – Investigation of erythroid differentiation in CB and K562 cells upon knockdown of SAGA and ATAC – KAT2A complexes. (A) Analysis of cell death in *CTRLsh* and *KAT2A*sh CB HSC/MPP in erythroid differentiation cultures. Mean ± SD of 4

independent differentiation experiments. Two-tailed paired t-test for significance *p<0.05. **(B)** Time course of erythroid marker profiling by flow cytometry in K562 cells transduced with *CTRLsh* treated with 1.5% DMSO. Proportion of cells expressing early (CD71) and late (CD235a) erythroid lineage markers is shown at each time-point. Data summarise mean ± SD of 6 independent experiments. **(C)** Quantitative RT-PCR analysis of erythroid gene expression progression in K562 cells transduced with *CTRLsh* and treated with 1.5% DMSO for erythroid molecular induction. Mean ± SD of N>3 independent experiments; data are represented relative to day 0 normalised to *HPRT1* housekeeping gene. Two-tailed paired t-test for significance *p<0.05, **p<0.01, ***p<0.001. **(D)** Quantitative RT-PCR validation of SAGA-specific *SUPT20Hsh2* and ATAC-specific *TADA2Ash* knockdown in K562 cells. N=3 independent experiments, mean ± SEM of gene expression relative to *CTRLsh*, normalised to *HPRT1* housekeeping gene. Two-tailed t-test for significance *p<0.05, **p<0.01, ***p<0.001. **(E)** Western blot analysis of SPT20 protein abundance in erythroid-affiliated K562 cells transduced with *CTRLsh* (lines 1-3), *SUPT20Hsh2* (lines 4-6) and *TADA2Ash* (lines 7-9). Observed SPT20 as per Krebs et al. 2011[13] using an in-house purified anti-serum. α-Tubulin (50 kDa) was used as loading control. **(F)** Quantitative RT-PCR analysis of erythroid gene expression progression in K562 cells transduced with *CTRLsh*, *SUPT20Hsh2*, or *TADA2Ash* and treated with 1.5% DMSO for erythroid molecular induction. Mean ± SD of N>3 independent experiments, mean ± SEM of gene expression relative to day 0, normalised to *HPRT1* housekeeping gene. Two-tailed paired t-test for significance *p<0.05, **p<0.01. **(G)** SAGA-specific elements peak at the Intermediate (IntE) phase of late erythroid differentiation. Representation of Cluster 17 extracted from visualisation database: https://cellline.molbiol.ox.ac.uk. Details of individual genes can be found in Supplemental File 4. **(H)** Visualisation of gene expression (selected region) pattern of *Kat2a* and *Zzz3* using SPRING tool (https://kleintools.hms.harvard.edu/paper_websites/tusi_et_al/). As per Supplemental File 3, other genes in this region include regulators of E/Meg cell fate commitment including *Gata2*, *Zfpm1* and *Myb*. MPP: multipotent progenitors. EBMP: erythroid basophil-megakaryocyte-biased progenitors. Meg: megakaryocyte. Ba: basophil. CEP: committed erythroid progenitors. ETD: erythroid terminal differentiation. **(I)** ChIP-seq peak track of the *EPOR* locus bound by ZZZ3 replicate 2 in K562 cells. **(J)** Quantitative RT-PCR analysis of *EPOR* expression in undifferentiated K562 cells transduced with and ATAC *ZZZ3sh* and *TADA2Ash*. Mean ± SD of N=3 independent experiments each for *ZZZ3sh* and *TADA2Ash* against the respective *CTRLsh*. Experiments are represented separately, as they reflect distinct K562 cell cultivars.

Supplemental Figure 5 - Investigation of ATAC and SAGA element knockdown in MOLM13 AML cells. (A) Growth curve of MOLM13 cells transduced with shRNA constructs against *KAT2A, SUPT20H* and *ZZZ3*. Mean ± SEM of 3 independent experiments. **(B)** Growth curve of MOLM13 cells transduced with shRNA constructs against *TADA2B* and *TADA2A*. Mean ± SEM of 3 independent experiments. **(C)** Representative flow cytometry plots of cell cycle analysis of MOLM-13 cells transduced with *CTRLsh*, *KAT2Ash*, SAGA-specific *SUPT20Hsh* and *TADA2Bsh*, and ATAC-specific *ZZZ3sh* and *TADA2Ash*. **(D)** Growth curve of MOLM13 cells transduced with *CTRLsh* and shRNA constructs against *SUPT20H* and *ZZZ3* and stained with Tag-it Violet cell tracking dye. Mean ± SEM of 3 individual experiments. **(E)** Flow cytometry analysis of median fluorescence intensity of MOLM13 cells stained with Tag-it Violet proliferation and cell tracking dye. As expected, the fluorescence is halved at each division. **(F)** Quantitative RT-PCR validation of *TADA2Ash2* knockdown in lentiviral-producing cells. Representative experiment shown. **(G)** Quantification of blast-like and monocytic differentiated cells in MOLM13 cultures transduced with *CTRLsh*, *KAT2Ash, SUPT20Hsh* and *ZZZ3sh* by an independent observer. Scoring of 3 randomly selected fields of >100 cells; 2-way ANOVA with Tukey's multiple comparisons test; adjusted p-values for significance of pairwise comparisons *$p<0.05$, **$p<0.01$, ***$p<0.001$. **(H)** Flow cytometry analysis of apoptosis by Annexin-V staining in MOLM13 cells transduced with *CTRLsh*, *KAT2Ash*, *SUPT20Hsh*, *ZZZ3sh* (left) and *CTRLsh*, *TADA2Bsh* and *TADA2Ash* (right). The combination of Annexin V and DAPI staining allows for a distinction between viable cells (double negative), cells in early apoptosis (Annexin V positive), cells in late apoptosis (double positive) and dead cells (DAPI positive). N ≥3 individual samples; mean ± SEM. Two-tailed t-test for significance; no significant differences.

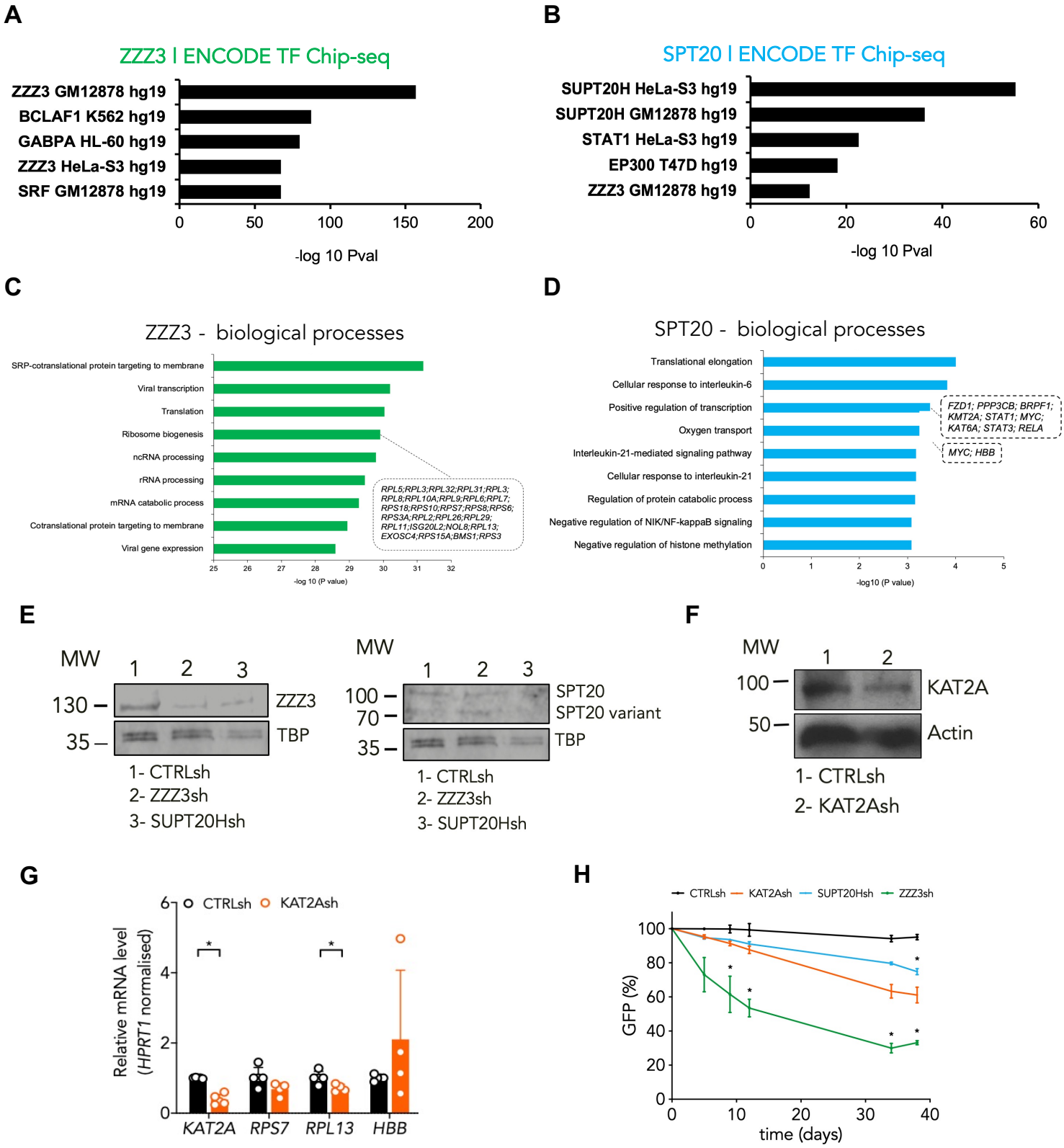
Supplemental Figure 6 - Investigation of ATAC and SAGA element knockdown in CD34+ AML cell lines KG1a and Kasumi-1. (A) Flow cytometry analysis of cell cycle in KG1a AML cells transduced cells with shRNA constructs against KAT2A, SUPT20H, TADA2B, ZZZ3 and TADA2A N=3 individual experiments; mean ± SEM. Two-tailed t-test for significance *$p<0.05$. **(B)** Flow cytometry analysis of apoptosis by Annexin-V staining in KG1a cells. N3 ≥3 individual samples; mean ± SEM. Two-tailed t-test for significance; no significant differences. **(C)** Flow cytometry analysis of apoptosis by Annexin-V staining in

Kasumi-1 cells. N3 ≥3 individual samples; mean ± SEM. Two-tailed t-test for significance; no significant differences.

Supplemental Figure 7 - KAT2A chemical inhibition with MB3 suppresses growth and expansion of primary AML cells for up to 3 weeks in MS5 co-culture. (A) Flow cytometry gating strategy used for analysis of primitive AML CD34+ and GMP-like (L-GMPs) cell populations. Cells gated for live cells (SSC-A vs FSC-A), and singlets (SSC-A vs SSC-H) considered. Singlets were then gated for GFP+ to establish the global level of GFP against which individual sub-populations were analyzed. Sub-populations analyzed were CD34+ (left), and L-GMPs (right), gated as CD34+38+CD123+CD45RA+ cells. GFP levels within CD34+ cells and L-GMP were obtained and compared with global GFP levels for each construct to determine relative preservation of each sub-population upon gene expression knockdown. The result of this analysis is presented in Fig. 7G-H, directly comparing *CTRLsh* with *KAT2Ash*, *SUPT20Hsh* or *ZZZ3sh*. **(B)** Growth of human primary CD34+ AML cells treated with KAT2A inhibitor MB3 (100μM) vs vehicle DMSO in the MS5 coculture system for up to 3 weeks. 3 individual patient samples are shown. **(C)** Percentage of L-GMPs at week 2 of MS5 co-culture for AML2 cells treated with KAT2A inhibitor MB3 (100μM) vs vehicle DMSO.

SUPPLEMENTAL FIGURES

Supplemental Figure 1



Supplemental Figure 1 - KAT2A-containing ATAC and SAGA complexes have unique functional associations in K562 cells.

(A, B) Specificity of ATAC (ZZZ3) and SAGA (SPT20) ChIP-seq targets against ENCODE; data as retrieved by EnrichR online annotation tool (Kuleshov et al., 2016)[34].

(C, D) Top Gene Ontology (GO) associations of ZZZ3 and SPT20 ChIP-seq targets on biological processes as calculated by EnrichR (Kuleshov et al., 2016)[34].

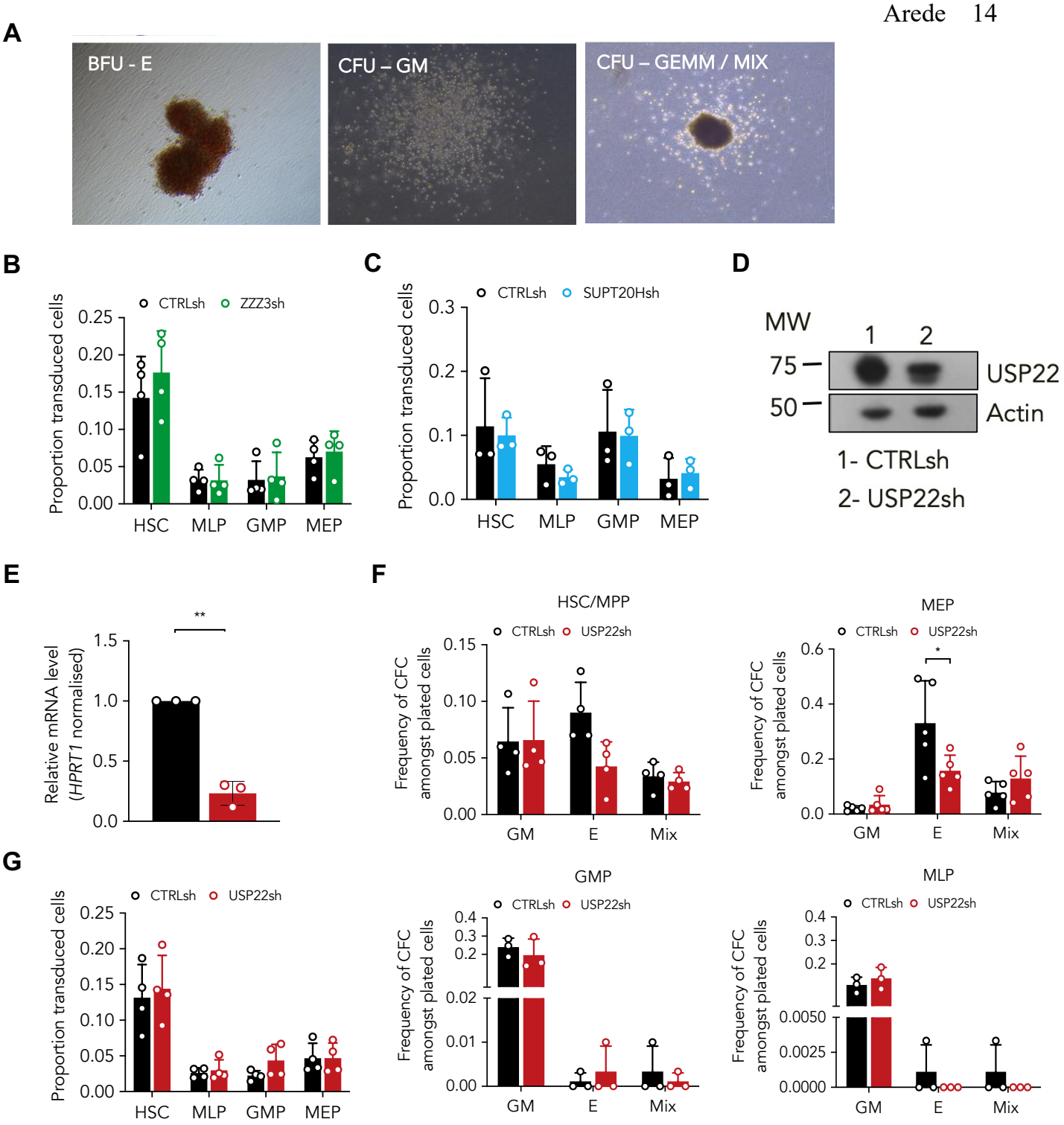
(E) Western blot analysis of ZZZ3 (left) and SPT20 (right) nuclear protein abundance in K562 cells transduced with *CTRLsh*, *ZZZ3sh* and *SUPT20Hsh*. Observed ZZZ3, SPT20 and SPT20 variant as per Krebs et al. 2011[13] using in-house purified anti-sera. TBP (38 kDa) was used as loading control.

(F) Western blot analysis of KAT2A protein abundance in erythroid-affiliated K562 cells transduced with *CTRLsh* and *KAT2Ash*. Predicted MW for KAT2A is 95 kDa. ACTB (42 kDa) was used as loading control.

(G) Quantitative RT-PCR analysis of gene expression in K562 cells transduced with *KAT2Ash*. Mean ± SEM of 4 individual experiments; gene expression relative to *CTRLsh*, normalized to *HPRT1* housekeeping gene. Two-tailed t-test for significance *p<0.05.

(H) Growth curve of K562 cells transduced with shRNA constructs against ZZZ3, SUPT20H and KAT2A. Mean ± SEM of 3 independent experiments. ANOVA for mixed effects analysis significance *p<0.05.

Supplemental Figure 2



Supplemental Figure 2 - Investigation of ATAC and SAGA element knockdown in human cord blood (CB) cell specification.

(A) Photographs of representative colony morphologies from transduced human CB HSC and progenitors. BFU-E: Boost-forming unit – erythrocyte colony. CFU-GM: Colony-forming unit - granulocyte monocyte colony; CFUGEMM or mixed: Colony-forming unit – granulocyte erythrocyte monocyte megakaryocyte colony.

(B) Proportion of *ZZZ3sh* transduced CB HSC and progenitors. Mean ± SEM of >3 individual sorting experiments. Two tailed paired t-test for significance; no significant differences.

(C) Proportion of *SUPT20Hsh* transduced CB HSC and progenitors. Mean ± SEM of >3 individual sorting experiments. Two-tailed paired t-test for significance; no significant differences.

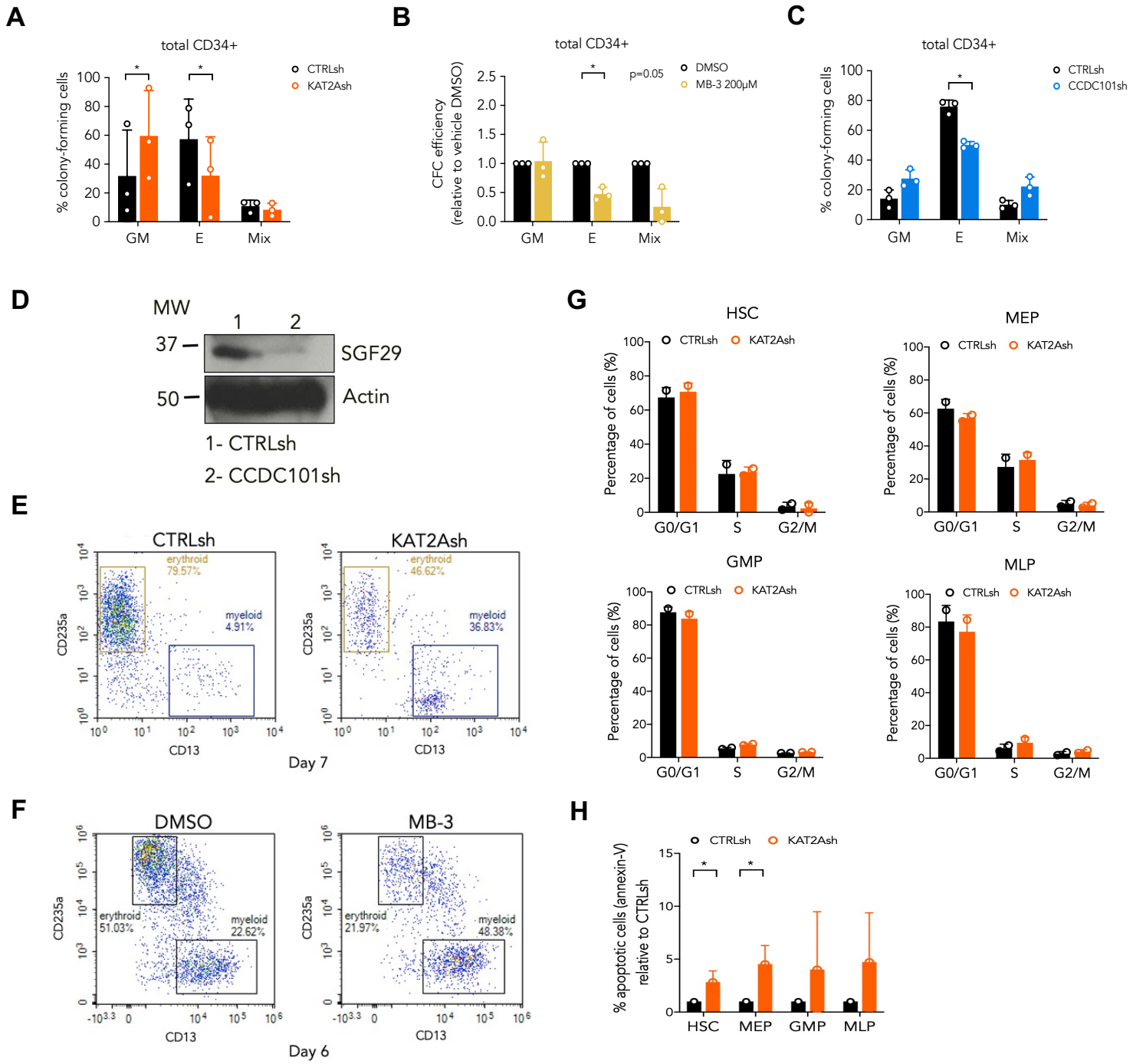
(D) Western blot analysis of USP22 protein abundance in K562 cells transduced with *CTRLsh* and *USP22sh*. Predicted USP22 MW is 60 kDa. ACTB (42 kDa) was used as loading control.

(E) Quantitative RT-PCR validation of USP22 knockdown in CB HSCs. Mean ± SEM of 3 individual experiments; gene expression relative to *CTRLsh*, normalized to *HPRT1* housekeeping gene. Paired two-tailed t-test for significance **p<0.01.

(F) Frequency of colony-forming cells in the HSC/MPP, MEP (top), GMP and MLP (bottom) compartments transduced with *USP22sh*. Mean ± SEM of 4 individual CB samples. Two-tailed paired t-test for significance; *p<0.05.

(G) Proportion of *USP22sh* transduced CB HSC and progenitors. Mean ± SEM of >4 individual sorting experiments. Two-tailed paired t-test for significance; no significant differences.

Supplemental Figure 3



Supplemental Figure 3 - KAT2A is required for human CB progenitor survival and erythroid specification with no impact on lymphoid and myeloid potential.

(A) Relative representation of colony-forming cells present in total CB CD34+ cells transduced with *KAT2Ash*. Mean ± SEM of 3 individual CB samples. Paired two-tailed t-test for significance *p<0.05.

(B) Relative representation of colony-forming cells present in total CB CD34+ cells treated with MB-3 200µM and vehicle DMSO. Data are represented relative to DMSO. Mean ± SEM of 3 individual CB samples. Paired two-tailed t-test for significance *p<0.05.

(C) Relative representation of colony-forming cells present in total CB CD34+ cells transduced with *CCDC101sh*. Mean ± SEM of 3 individual CB samples. Paired two-tailed t-test for significance *p<0.05.

(D) Western blot analysis of SGF29 (*CCDC101* gene) protein abundance in erythroid-affiliated K562 cells transduced with *CTRLsh* and *CCDC101sh*. Predicted SGF29 MW is 36 kDa. ACTB (42 kDa) was used as loading control.

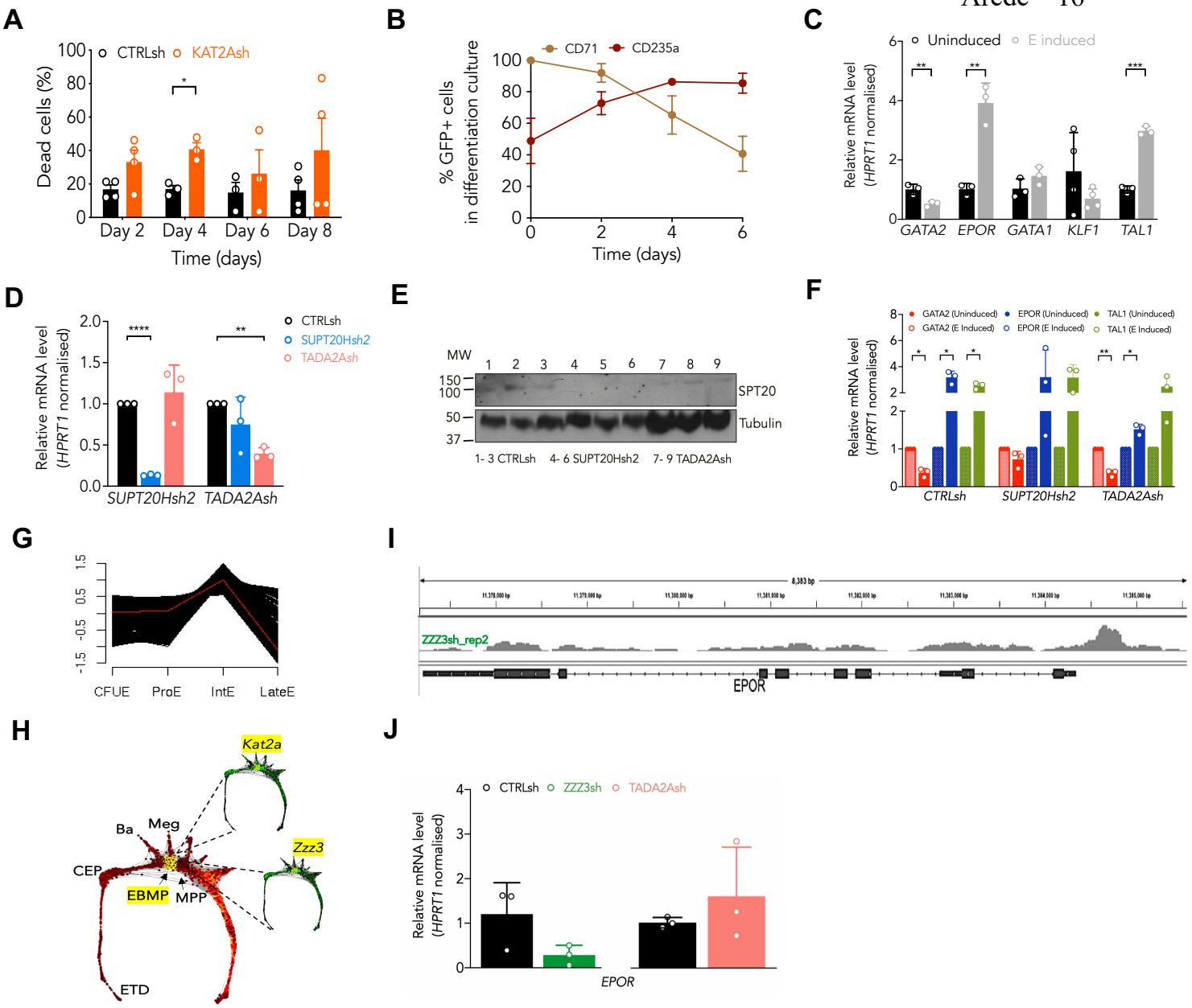
(E) Flow cytometry analysis of CB CD34+ cells transduced with *CTRLsh* or *KAT2Ash* after 7 days of liquid culture in mixed lineage differentiation conditions (see Methods). Cells plotted were gated as live GFP+ (transduced) CD34-. Data are representative of 2 independent experiments.

(F) Flow cytometry analysis of CB CD34+ cells treated with MB-3 100µM and 0.1% DMSO after 6 days of liquid culture in mixed lineage differentiation conditions (see Methods). Cells plotted were gated as live GFP+ (transduced) CD34-. Data are representative of 2 independent experiments.

(G) Flow cytometry analysis of cell cycle in HSC/MPP, MEP, GMP and MLP transduced cells with *KAT2Ash* vs *CTRLsh*. N=3 individual cord blood samples; mean ± SEM. Two-tailed paired t-test not significant.

(H) Flow cytometry analysis of apoptosis by Annexin-V staining. N=3 individual cord blood samples; mean ± SEM. Two-tailed paired t-test for significance *p<0.05.

Supplemental Figure 4



Supplemental Figure 4 - Investigation of erythroid differentiation in CB and K562 cells upon knockdown of SAGA and ATAC – KAT2A complexes.

(A) Analysis of cell death in *CTRLsh* and *KAT2A*sh CB HSC/MPP in erythroid differentiation cultures. Mean ± SD of 4 independent differentiation experiments. Two-tailed paired t-test for significance *p<0.05.

(B) Time course of erythroid marker profiling by flow cytometry in K562 cells transduced with *CTRLsh* treated with 1.5% DMSO. Proportion of cells expressing early (CD71) and late (CD235a) erythroid lineage markers is shown at each time-point. Data summarise mean ± SD of 6 independent experiments.

(C) Quantitative RT-PCR analysis of erythroid gene expression progression in K562 cells transduced with *CTRLsh* and treated with 1.5% DMSO for erythroid molecular induction. Mean ± SD of N>3 independent experiments; data are represented relative to day 0 normalised to *HPRT1* housekeeping gene. Two-tailed paired t-test for significance *p<0.05, **p<0.01, ***p<0.001.

(D) Quantitative RT-PCR validation of SAGA-specific *SUPT20Hsh2* and ATAC-specific *TADA2Ash* knockdown in K562 cells. N=3 independent experiments, mean ± SEM of gene expression relative to *CTRLsh*, normalised to *HPRT1* housekeeping gene. Two-tailed t-test for significance *p<0.05, **p<0.01, ***p<0.001.

(E) Western blot analysis of SPT20 protein abundance in erythroid-affiliated K562 cells transduced with *CTRLsh* (lines 1-3), *SUPT20Hsh2* (lines 4-6) and *TADA2Ash* (lines 7-9). Observed SPT20 as per Krebs et al. 2011[13] using an in-house purified anti-serum. α-Tubulin (50 kDa) was used as loading control.

(F) Quantitative RT-PCR analysis of erythroid gene expression progression in K562 cells transduced with *CTRLsh*, *SUPT20Hsh2*, or *TADA2Ash* and treated with 1.5% DMSO for erythroid molecular induction. Mean ± SD of N>3 independent experiments, mean ± SEM of gene expression relative to day 0, normalised to *HPRT1* housekeeping gene. Two-tailed paired t-test for significance *p<0.05, **p<0.01.

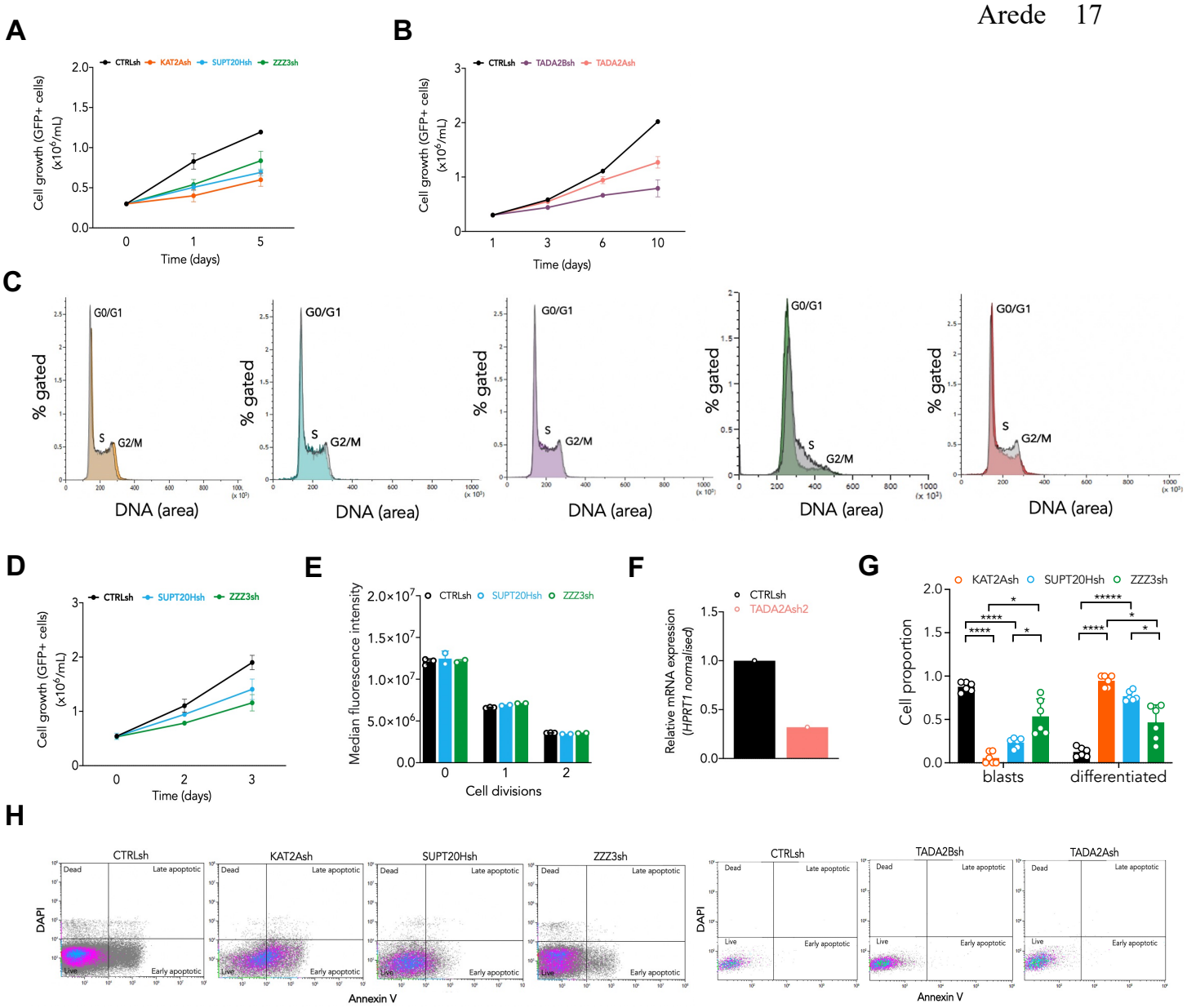
(G) SAGA-specific elements peak at the Intermediate (IntE) phase of late erythroid differentiation. Representation of Cluster 17 extracted from visualisation database: https://cellline.molbiol.ox.ac.uk. Details of individual genes can be found in Supplemental File 4.

(H) Visualisation of gene expression (selected region) pattern of *Kat2a* and *Zzz3* using SPRING tool (https://kleintools.hms.harvard.edu/paper_websites/tusi_et_al/). As per Supplemental File 3, other genes in this region include regulators of E/Meg cell fate commitment including *Gata2*, *Zfpm1* and *Myb*. MPP: multipotent progenitors. EBMP: erythroid basophil-megakaryocyte-biased progenitors. Meg: megakaryocyte. Ba: basophil. CEP: committed erythroid progenitors. ETD: erythroid terminal differentiation.

(I) ChIP-seq peak track of the *EPOR* locus bound by ZZZ3 replicate 2 in K562 cells.

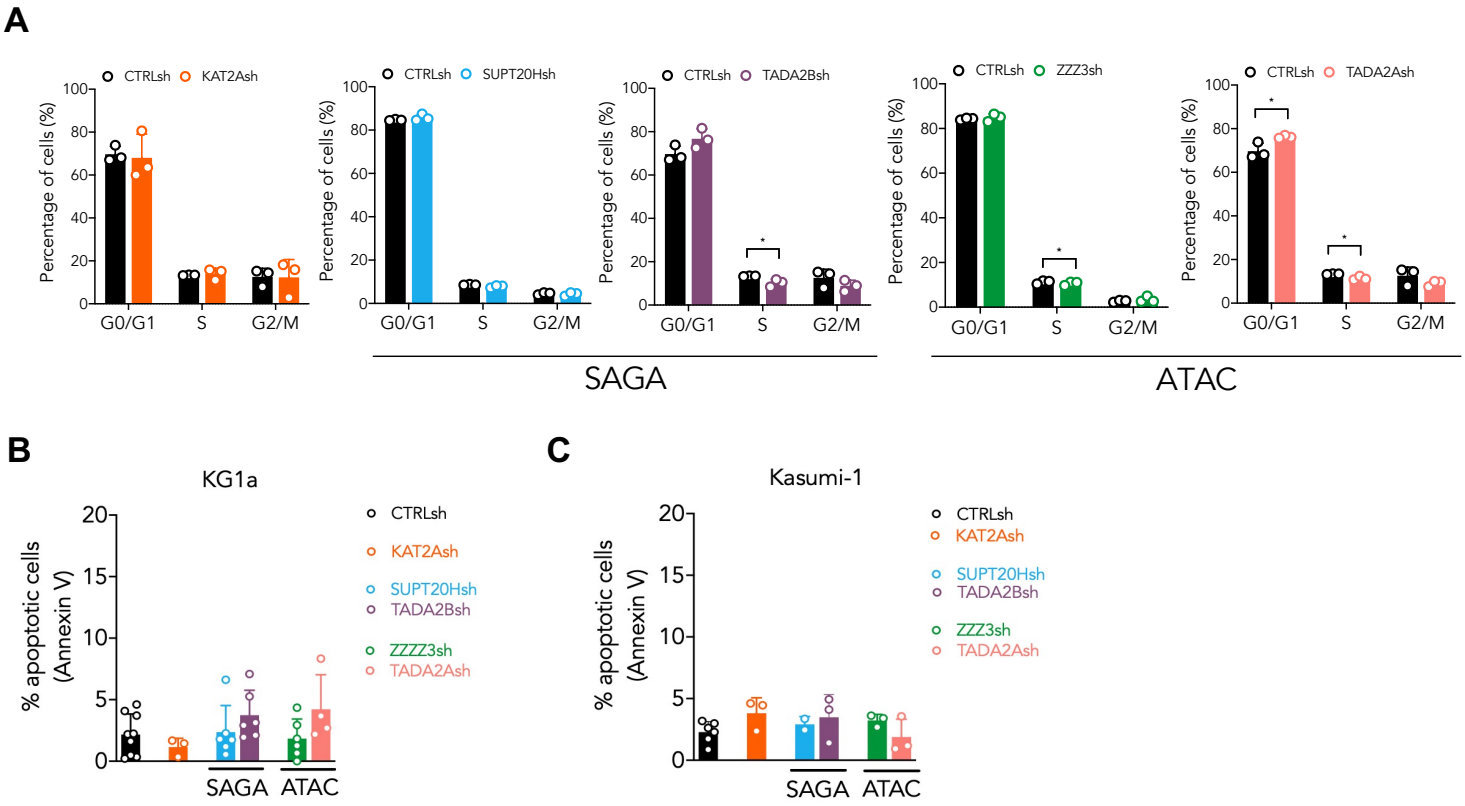
(J) Quantitative RT-PCR analysis of *EPOR* expression in undifferentiated K562 cells transduced with and ATAC *ZZZ3sh* and *TADA2Ash*. Mean ± SD of N=3 independent experiments each for *ZZZ3sh* and *TADA2Ash* against the respective *CTRLsh*. Experiments are represented separately, as they reflect distinct K562 cell cultivars.

Supplemental Figure 5



Supplemental Figure 5 - Investigation of ATAC and SAGA element knockdown in MOLM13 AML cells.
(A) Growth curve of MOLM13 cells transduced with shRNA constructs against *KAT2A, SUPT20H* and *ZZZ3*. Mean ± SEM of 3 independent experiments.
(B) Growth curve of MOLM13 cells transduced with shRNA constructs against *TADA2B* and *TADA2A*. Mean ± SEM of 3 independent experiments.
(C) Representative flow cytometry plots of cell cycle analysis of MOLM-13 cells transduced with *CTRLsh*, *KAT2Ash*, SAGA-specific *SUPT20Hsh* and *TADA2Bsh*, and ATAC-specific *ZZZ3sh* and *TADA2Ash*.
(D) Growth curve of MOLM13 cells transduced with *CTRLsh* and shRNA constructs against *SUPT20H* and *ZZZ3* and stained with Tag-it Violet cell tracking dye. Mean ± SEM of 3 individual experiments.
(E) Flow cytometry analysis of median fluorescence intensity of MOLM13 cells stained with Tag-it Violet proliferation and cell tracking dye. As expected, the fluorescence is halved at each division.
(F) Quantitative RT-PCR validation of *TADA2Ash2* knockdown in lentiviral-producing cells. Representative experiment shown.
(G) Quantification of blast-like and monocytic differentiated cells in MOLM13 cultures transduced with *CTRLsh*, *KAT2Ash, SUPT20Hsh* and *ZZZ3sh* by an independent observer. Scoring of 3 randomly selected fields of >100 cells; 2-way ANOVA with Tukey's multiple comparisons test; adjusted p-values for significance of pairwise comparisons *p<0.05, **p<0.01, ***p<0.001.
(H) Flow cytometry analysis of apoptosis by Annexin-V staining in MOLM13 cells transduced with *CTRLsh*, *KAT2Ash*, *SUPT20Hsh*, *ZZZ3sh* (left) and *CTRLsh*, *TADA2Bsh* and *TADA2Ash* (right). The combination of Annexin V and DAPI staining allows for a distinction between viable cells (double negative), cells in early apoptosis (Annexin V positive), cells in late apoptosis (double positive) and dead cells (DAPI positive). N ≥3 individual samples; mean ± SEM. Two-tailed t-test for significance; no significant differences.

Supplemental Figure 6



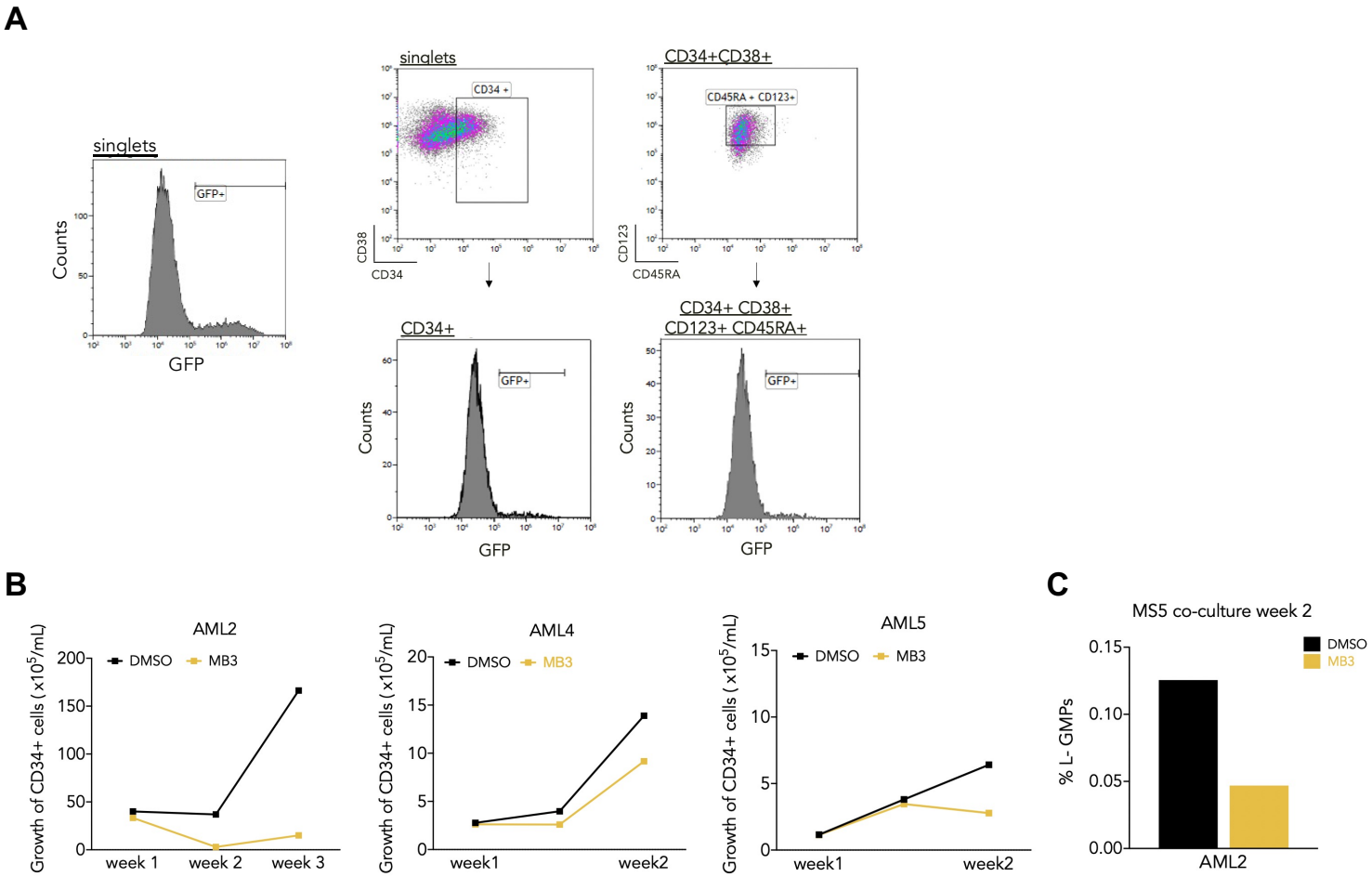
Supplemental Figure 6 - Investigation of ATAC and SAGA element knockdown in CD34+ AML cell lines KG1a and Kasumi-1.

(A) Flow cytometry analysis of cell cycle in KG1a AML cells transduced cells with shRNA constructs against KAT2A, SUPT20H, TADA2B, ZZZ3 and TADA2A N=3 individual experiments; mean ± SEM. Two-tailed t-test for significance *p<0.05.

(B) Flow cytometry analysis of apoptosis by Annexin-V staining in KG1a cells. N3 ≥3 individual samples; mean ± SEM. Two-tailed t-test for significance; no significant differences.

(C) Flow cytometry analysis of apoptosis by Annexin-V staining in Kasumi-1 cells. N3 ≥3 individual samples; mean ± SEM. Two-tailed t-test for significance; no significant differences.

Supplemental Figure 7



Supplemental Figure 7 - KAT2A chemical inhibition with MB3 suppresses growth and expansion of primary AML cells for up to 3 weeks in MS5 co-culture.

(A) Flow cytometry gating strategy used for analysis of primitive AML CD34+ and GMP-like (L-GMPs) cell populations. Cells gated for live cells (SSC-A vs FSC-A), and singlets (SSC-A vs SSC-H) considered. Singlets were then gated for GFP+ to establish the global level of GFP against which individual sub-populations were analyzed. Sub-populations analyzed were CD34+ (left), and L-GMPs (right), gated as CD34+38+CD123+CD45RA+ cells. GFP levels within CD34+ cells and L-GMP were obtained and compared with global GFP levels for each construct to determine relative preservation of each sub-population upon gene expression knockdown. The result of this analysis is presented in Fig. 7G-H, directly comparing *CTRLsh* with *KAT2Ash*, *SUPT20Hsh* or *ZZZ3sh*.

(B) Growth of human primary CD34+ AML cells treated with KAT2A inhibitor MB3 (100µM) vs vehicle DMSO in the MS5 coculture system for up to 3 weeks. 3 individual patient samples are shown.

(C) Percentage of L-GMPs at week 2 of MS5 co-culture for AML2 cells treated with KAT2A inhibitor MB3 (100µM) vs vehicle DMSO.