

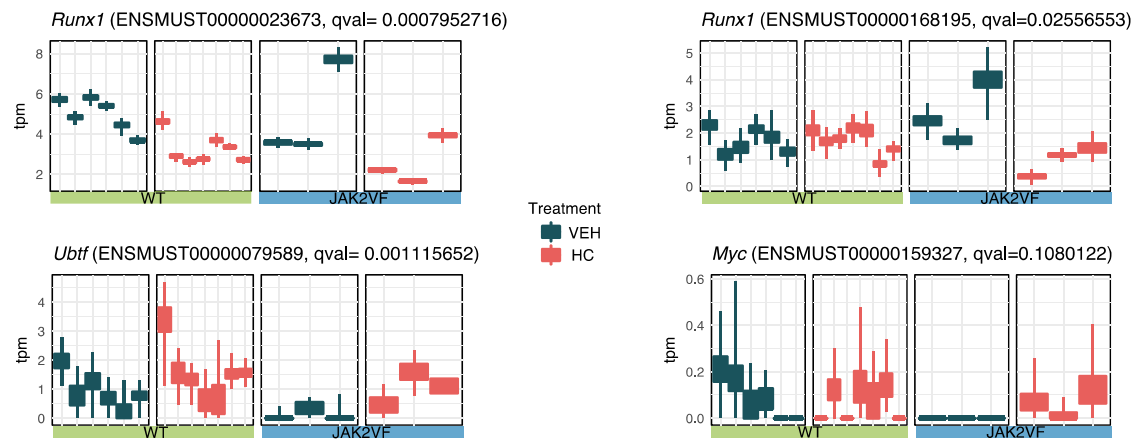
HYDROXYCARBAMIDE EFFECT ON DNA METHYLATION AND GENE EXPRESSION IN MYELOPROLIFERATIVE NEOPLASMS

Stephania Contreras-Castillo, Bertille Montibus, Azucena Rocha, Will Duke, Ferdinand von Meyenn, Donal McLornan, Claire Harrison, Ann Mullally, Reiner Schulz & Rebecca J. Oakey.

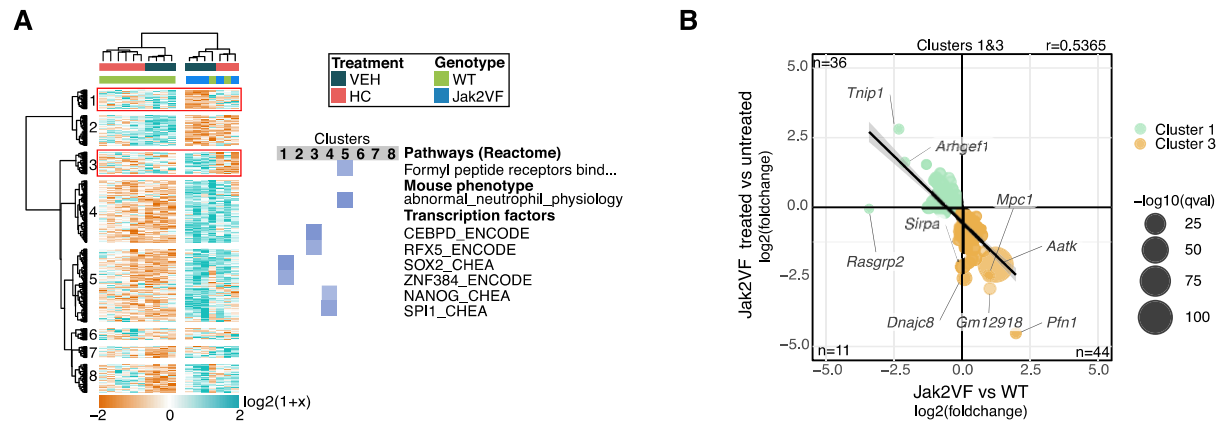
SUPPLEMENTARY FIGURES AND TABLES

TABLE OF CONTENTS

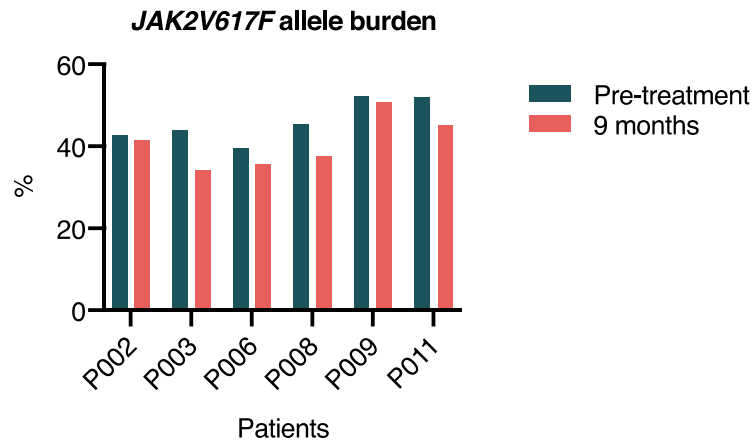
Supplementary Figure S1	2
Supplementary Figure S2	3
Supplementary Figure S3	4
Supplementary Figure S4	5
Supplementary Figure S5	6
Supplementary Figure S6	7
Supplementary Figure S7	8
Supplementary Figure S8	9
Supplementary Figure S9	10
Supplementary Figure S10	11
Supplementary Table S1	12
Supplementary Table S2	13
Supplementary Table S3	14
Supplementary Table S5	15
Supplementary Table S6	16
Supplementary Table S7	17
Supplementary Table S8	18
Supplementary Table S9	19



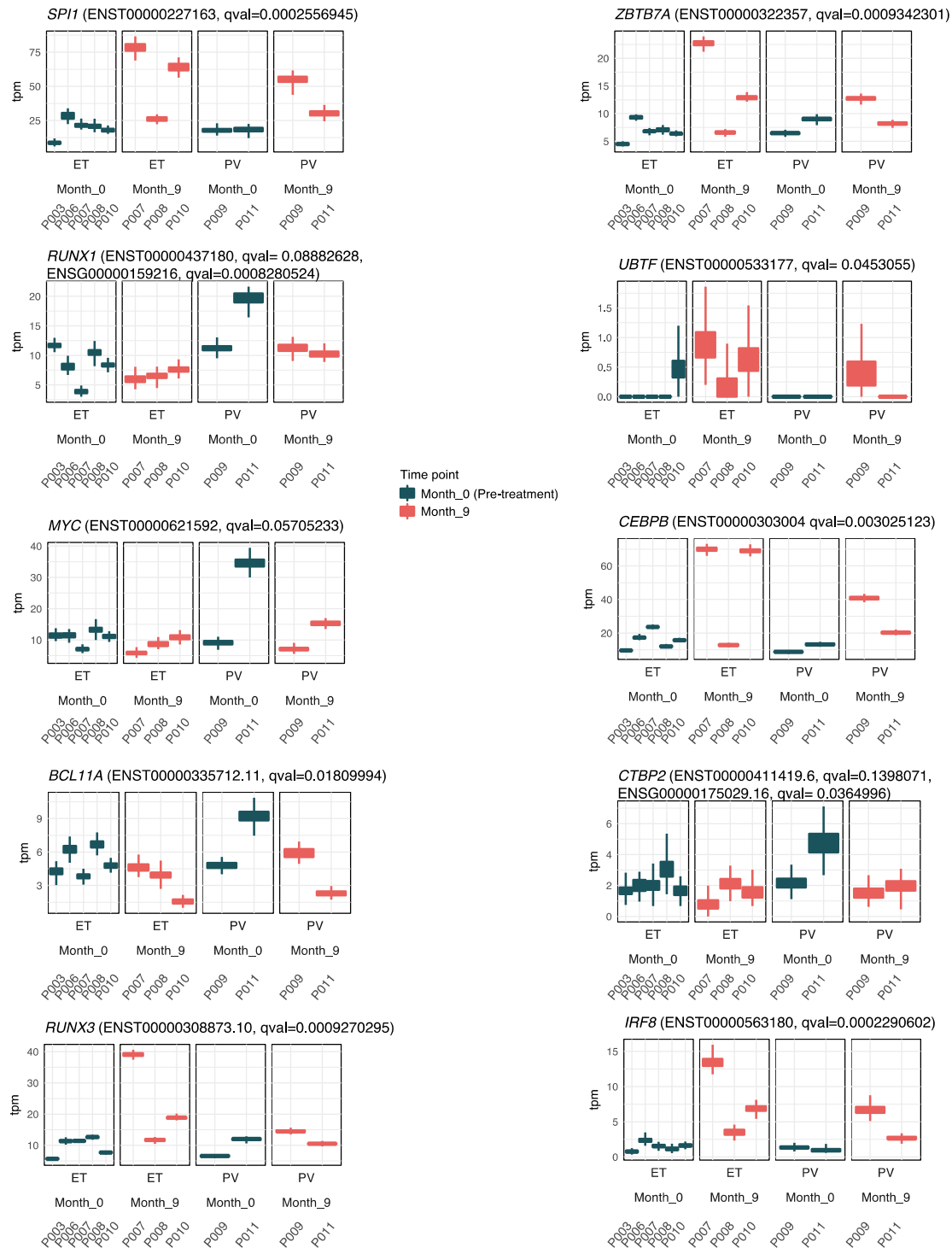
Supplementary Figure S1. Differentially expressed transcripts in mouse LSK cells. Transcripts per million reads (tpms) per each differentially expressed transcript. q-values indicated. VEH: vehicle; HC: hydroxycarbamide.



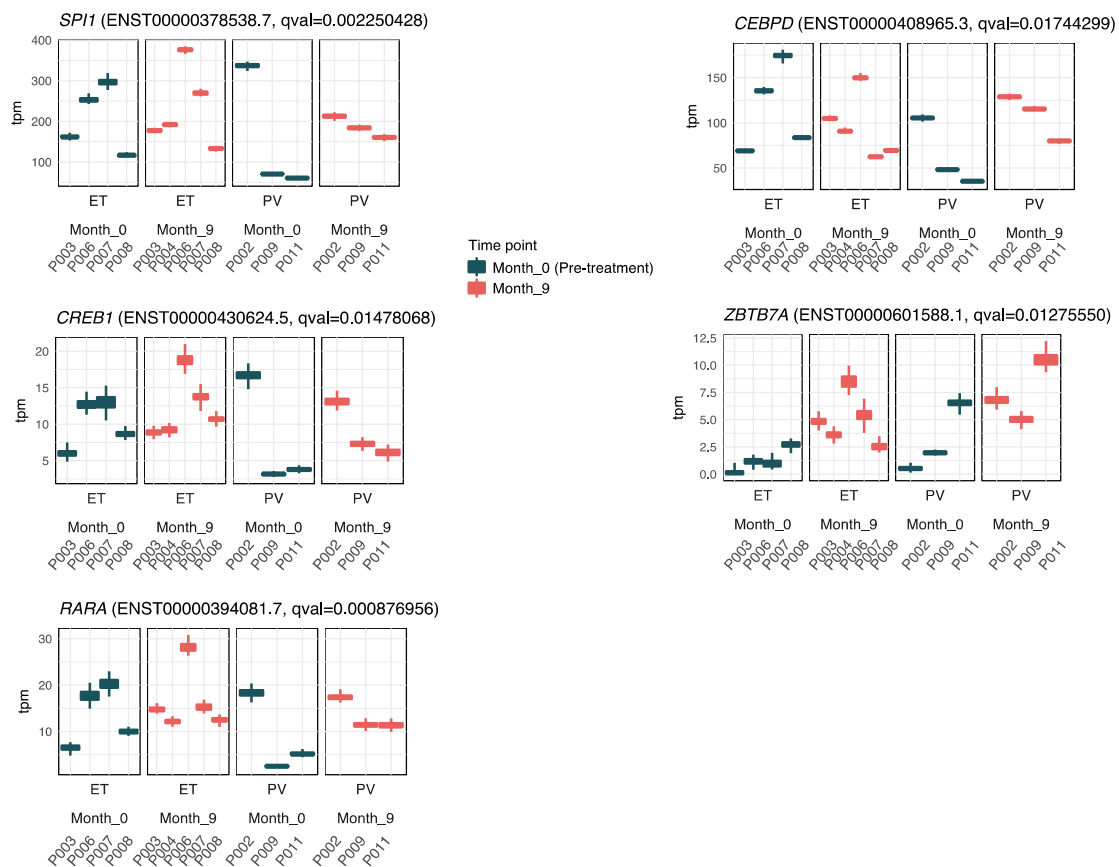
Supplementary Figure S2. Comparison between wild-type vs Jak2VF mutant, and Jak2VF mutant treated vs untreated in neutrophil cells. (A) Heatmap plot of hierarchically clustered DEGs between WT vs Jak2VF mutant and Jak2VF mutant treated vs untreated and top significant enriched terms of clusters identified. (B) Correlation plot of DEGs from Cluster 1 and 3 and their log2-fold change expression. VEH: vehicle; HC: hydroxycarbamide.



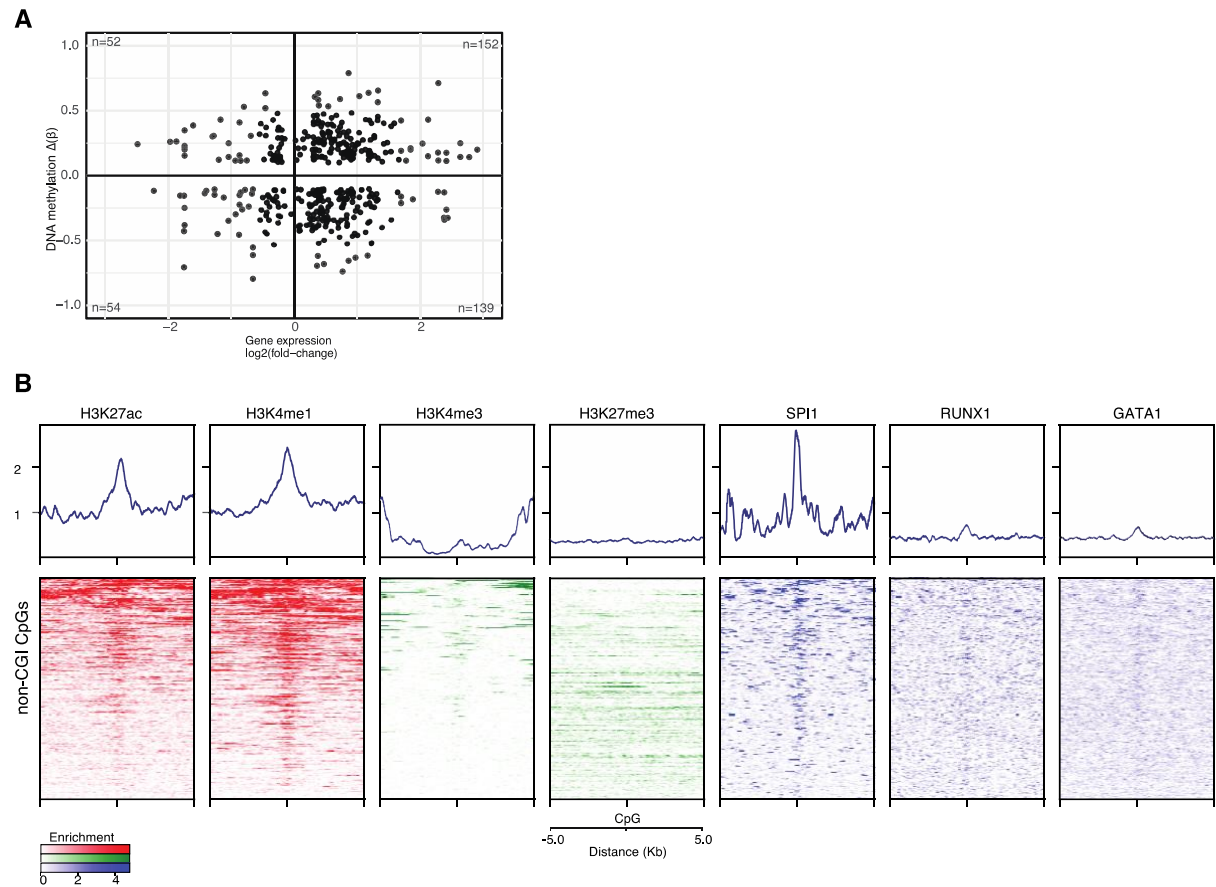
Supplementary Figure S3. Percentage of *JAK2V617F* burden in pre-treatment and 9-months samples. DNA from neutrophil cells were used to measure the percentage of *JAK2V617F* allele burden. Samples from pre-treatment and 9 months after HC treatment were used.



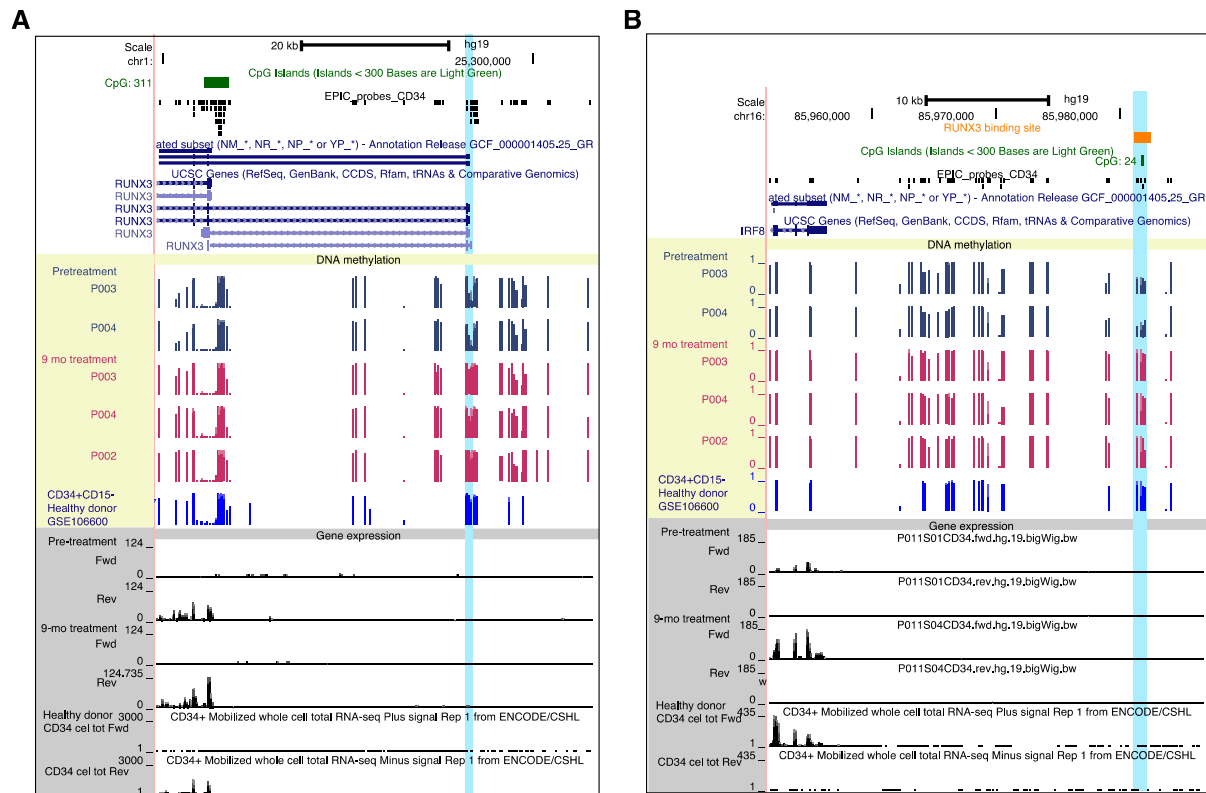
Supplementary Figure S4. Differentially expressed transcripts in human CD34⁺ cells. Transcripts per million reads (tpms) per each differentially expressed transcript. Q-values indicated.



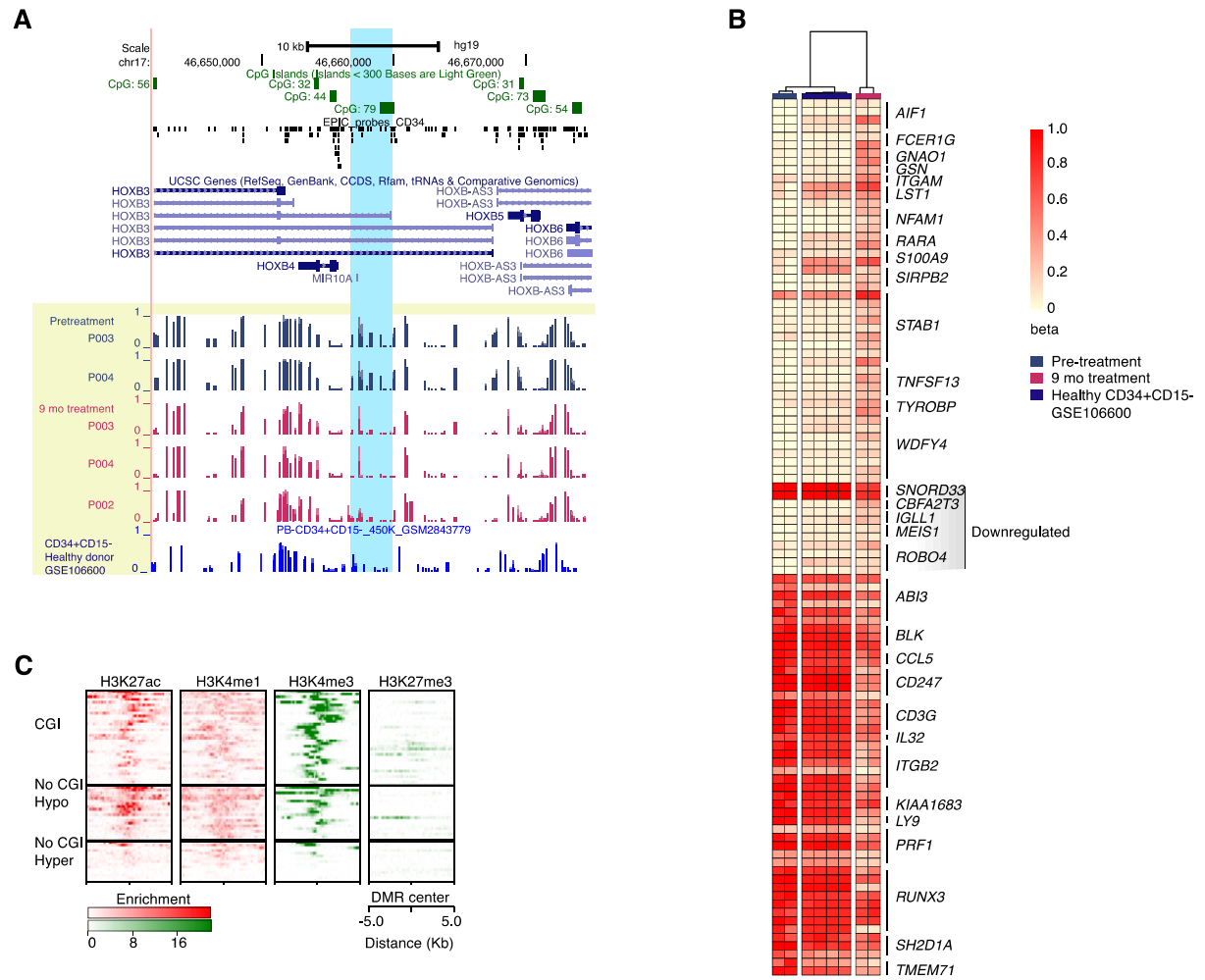
Supplementary Figure S5. Differentially expressed transcripts in human neutrophils. Transcripts per million reads (tpms) per each differentially expressed transcript. Q-values indicated.



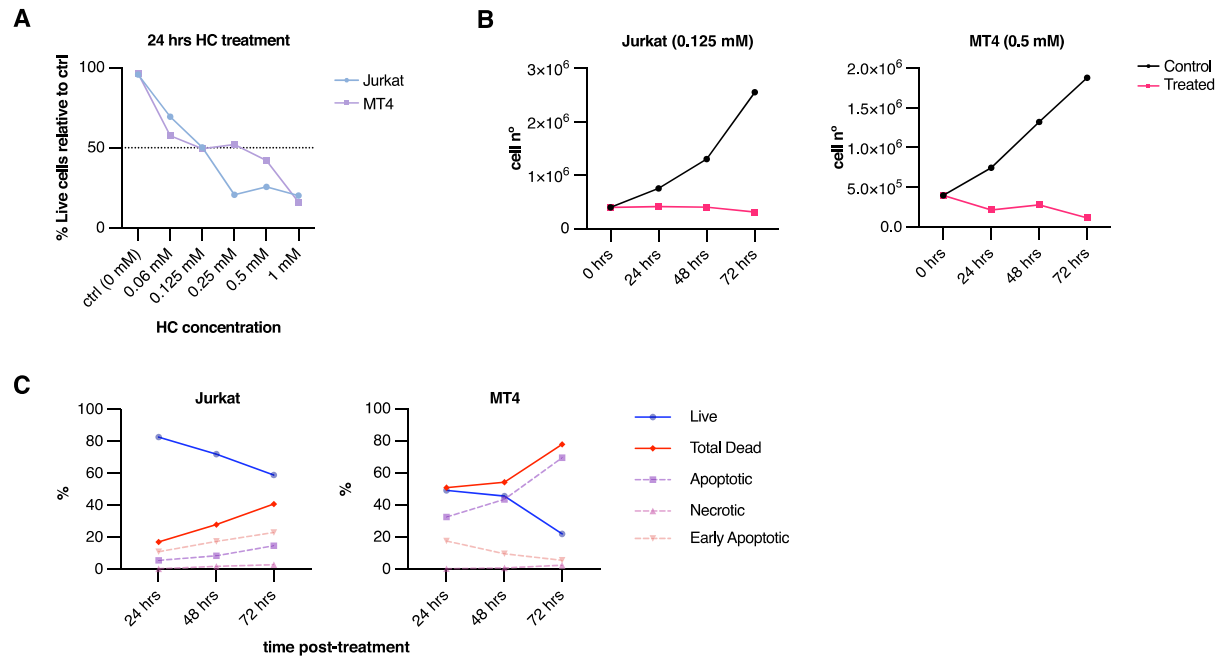
Supplementary Figure S6. Non-CpG sites associated to differentially expressed genes (A) Scatterplot of each CpG site with its DNA methylation change (y-axis) and gene expression change (x-axis) of its associated gene. (B) heatmap plots of each CpG sites mapped to ChIP-seq datasets from CD34⁺ healthy cells.



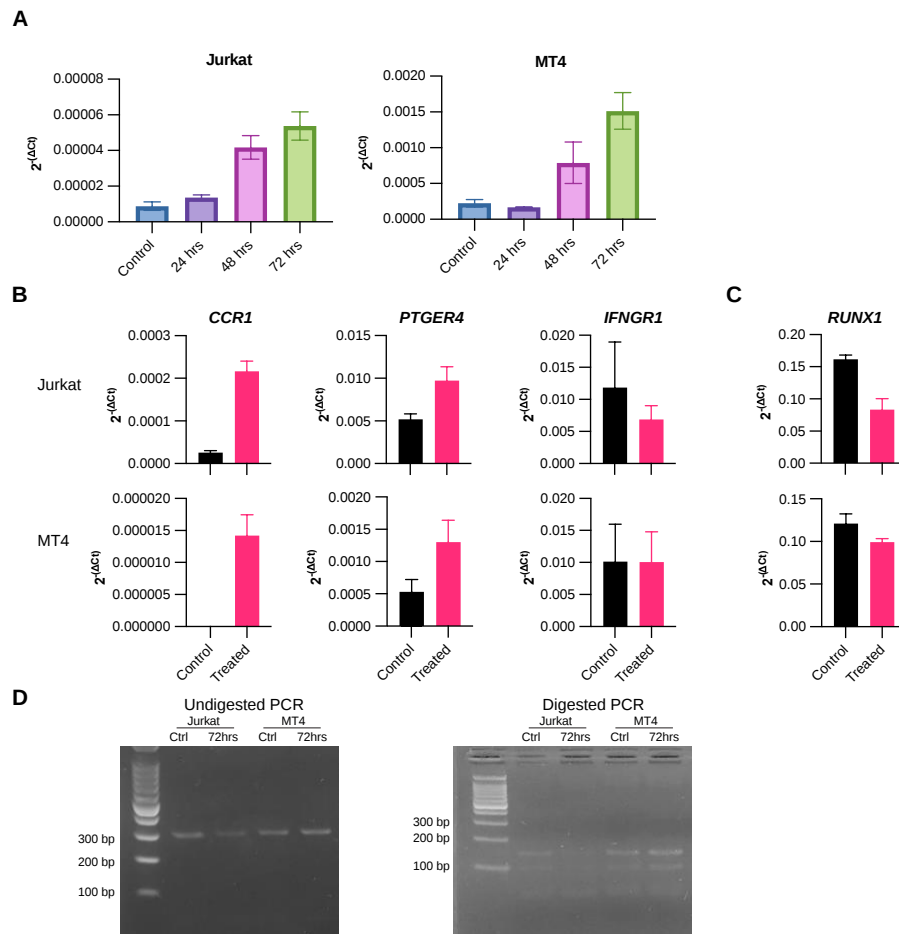
Supplementary Figure S7. DNA methylation and gene expression in human CD34⁺ cells at selected loci. (A) *RUNX3* locus and DMR at gene promoter. (B) *IRF8* locus with DMR3 where *RUNX3* binds.



Supplementary Figure S8. DMRs analysis of CD34⁺ cells. (A) *HOXB4* locus with DMR overlapping CGI positively correlated with *HOXB4* expression after 9-months of HC treatment (B) Heatmap of individual differentially methylated CpGs located in DMRs that do not overlap CGIs at pre-treatment, 9 months of HC treatment and healthy individuals. Genes associated to the differentially methylated CpGs indicated. (C) Heatmaps of publicly available ChIP-seq data from CD34⁺ cells overlapping DMRs associated to CGI or not.



Supplementary Figure S9. HC dose selection and response to treatment. (A) % of live cells relative to control on Jurkat and MT-4 cell lines after 24 hrs of HC treatment at different HC concentration. (B) Number of live cells on control and HC treated cell lines at 24, 48 and 72 hrs post HC-treatment. (C) Percentage of live, total dead, apoptotic, necrotic and early apoptotic cells at 24, 48 and 72 hrs post HC-treatment.



Supplementary Figure S10. Gene expression analysis of cell lines treated with HC. (A) Gene expression of *SPI1* on two cell lines, Jurkat and MT-4, at different time points of HC treatment. (B) Gene expression of control and treated cells (72 hrs treatment) of genes associated to *SPI1* binding sites which region is differentially methylated in $CD34^+$ cells. (C) *RUNX1* gene expression in control and treated cells (72 hrs treatment). (D) Agarose gel pictures showing the PCR product of *SPI1* promoter locus before and after enzyme restriction digestion. Picture shows no change of DNA methylation after 72 hrs of HC treatment.

Supplementary Table S1. Patients' clinical parameters per collection time-point and genome-wide data included in the study

Sample ID	Time Point	Age	Sex	Diagnosis	Molecular Diagnosis	Other mutations	Hct (%)	MVC (fL)	Platelets (x10 ⁹ /L)	WBC (x10 ⁹ /L)	Neutrophil (x10 ⁹ /L)	Hb (g/L)	RNA seq data		DNA meth data	
													CD34	NEU	CD34	NEU
P002	Pre-tto	50	F	PV	JAK2	DNMT3A (p.Arg771Ter, VAF 17%)	0.43	92	336	5.3	3.6	143	Yes*	Yes	Yes*	Yes
	6 mo	51					0.427	94	267	5.4	3.6	141				Yes
	9 mo	51					0.389	94	219	4.3	2.6	130	Yes*	Yes	Yes	Yes
P003	Pre-tto	55	F	ET	JAK2		0.424	91	661	10.9	7.3	135	Yes*	Yes	Yes	Yes
	3 mo	55					0.399	98	369	7.6	4.7	134				Yes
	6 mo	55					0.396	103	350	8	5	131				Yes
	9 mo	56					0.387	105	259	6.8	4.4	131	Yes*	Yes	Yes	Yes
P004	Pre-tto	42	F	ET	TN		0.451	90	628	11.3	6.8	144	Yes*	Yes*	Yes	Yes
	3 mo	43					0.43	93	597	9.1	5.3	147				Yes
	9 mo	43					0.409	93	592	8.4	4.6	137	Yes*	Yes*	Yes	Yes
P006	Pre-tto	67	F	ET	JAK2		0.422	85	534	7.6	4.4	136	Yes*	Yes		Yes
	3 mo	67					0.397	88	479	5.5	3.4	133				Yes
	6 mo	68					0.435	92	439	5.4	3.1	139				Yes
	9 mo	68					0.415	91	430	6.6	4.2	144	Yes*	Yes		Yes
P007	Pre-tto	77	F	ET	JAK2		0.403	83	529	8.6	2.8	129	Yes	Yes		Yes
	3 mo	78					0.409	88	427	7.1	2.7	125				Yes*
	6 mo	78					0.389	95	487	7.7	3.4	130				Yes
	9 mo	78					0.394	94	466	6.6	2.2	134	Yes	Yes		Yes
P008	Pre-tto	88	F	ET	JAK2	DNMT3A (p.Arg899Cys, VAF 18%)	0.416	82	728	10.6	7	133	Yes	Yes		Yes
	3 mo	88					0.384	92	359	6.9	4.4	127				Yes
	6 mo	88					0.38	93	351	6.3	3.5	123				Yes
	9 mo	89					0.406	93	389	6.6	4	134	Yes	Yes		Yes
P009	Pre-tto	68	M	PV	JAK2		0.455	82	635	16.6	13.6	148	Yes	Yes		Yes
	3 mo	69					0.45	83	641	13.1	10.7	144				Yes
	6 mo	69					0.453	87	423	9.8	7.8	145				Yes
	9 mo	69					0.447	98	369	8.2	6.4	153	Yes	Yes		Yes
P010	Pre-tto	77	F	ET	TN		0.386	92	787	6.4	4.1	127	Yes			Yes
	3 mo	77					0.389	101	563	4.5	2.3	128				Yes
	6 mo	78					0.388	105	642	3.5	2.2	127				Yes
	9 mo	78					0.376	107	522	8.3	5.9	126	Yes			Yes
P011	Pre-tto	52	M	PV	JAK2	TET2 (p.Gln659Ter, VAF 6% c.3955-2A>G, VAF 48%)	0.427	67	699	12.6	8.7	136	Yes	Yes		Yes
	3 mo	52					0.448	70	355	9	6	144				Yes
	9 mo	53					0.486	76	209	6.6	3.8	158	Yes	Yes		Yes

TN: triple negative; Hct: hematocrit; MVC: Mean corpuscular volume; WBC: White blood cells; Hb: Hemoglobin; VAF: variant allele frequency; *Low quality samples excluded from analysis

Supplementary Table S2. Common differentially expressed genes encoding TFs between human and mouse

Symbol	Human			Mouse		
	b	q-val transcript	q-val gene	b	q-val transcript	q-val gene
<i>NR4A2</i>	0.841	0.046	0.039	0.559	0.008	0.024
<i>MAFB</i>	1.591	0.003	0.001	1.467	0.005	0.003
<i>MEF2D</i>	0.372	0.008	0	0.076	0.396	0.029
<i>ATF3</i>	1.392	0.005	0	1.521	0.007	0.005
<i>ZBTB7A</i>	0.432	0.001	0	0.134	0.529	0
<i>FOXJ2</i>	0.229	0.071	0.028	0.111	1	0.008
<i>MAFG</i>	0.297	0.084	0.019	0.084	0.408	0.026
<i>ETS1</i>	0.486	0.013	0.001	0.08	1	0
<i>ELK4</i>	0.28	0.101	0.05	0.118	0.532	0.003
<i>FOXN3</i>	0.127	0.169	0.023	0.651	0.11	0
<i>ARID5A</i>	0.578	0.006	0.003	0.051	1	0
<i>ID2</i>	1.245	0.003	0	0.591	0.01	0.014
<i>PLAG1</i>	-1.04	0.002	0	-1.317	0	0
<i>HOXA5</i>	-0.322	0.063	0.046	-0.299	0.018	0.013
<i>RUNX1</i>	-0.225	0.089	0.001	-0.305	0.001	0.001
<i>BCL11A</i>	-0.516	0.018	0	-0.298	0	0
<i>MXD1</i>	-0.595	1	0.008	-0.434	0	0
<i>STAT6</i>	-1.268	0.342	0.005	-0.129	0.012	0.008
<i>ERG</i>	-0.549	0.005	0	-0.465	0.001	0.001
<i>ZNF652</i>	-0.22	0.223	0.002	-0.3	0	0
<i>HOXB4</i>	-0.472	0.033	0.023	-0.226	0.058	0.042
<i>IKZF2</i>	-0.33	0.128	0.029	-0.297	0.002	0
<i>MGA</i>	-0.226	0.144	0.023	-0.192	0.007	0.045
<i>TFDP2</i>	-0.588	0.003	0	-0.279	0.001	0.013
<i>KAT6A</i>	-1.101	0.21	0.011	-0.174	0.024	0.048
<i>KAT6B</i>	-0.56	0.057	0.029	-0.351	0.006	0.029
<i>KDM5B</i>	-0.29	0.018	0.001	-0.495	0	0

Dataset of transcription factor from Chawla, *et. al.*, 2013 (Chawla et al., 2013); Bold indicates TF significantly enriched; b: log(foldchange) computed by Sleuth

Supplementary Table S3. Enrichment results from i-cisTarget

Enriched TFs that bind differentially methylated CpGs located in intergenic regions with the total number of genes associated to the TF and number of genes that were also DEGs in our dataset.

Type of CpG	T. Factor	N° of genes associated found by i-cisTarget	DEGs		Dataset used by i-cisTarget
			up	down	
Non-CGI CpGs hypo (7,791 CpGs)	GATA2	225	29	14	GATA2 ChIP-seq protocol PCR1x on human K562
	GATA1	133	15	3	GATA1 ChIP-seq on human PBDE Fetal
	SPI1	435	87	12	SPI1 ChIP-seq protocol PCR1x on human K562
Non-CGI CpGs hyper (8,034 CpGs)	CEBPB	172	42	5	CEBPB ChIP-seq protocol v042211.1 on human GM12878
	TCF12	544	117	19	TCF12 ChIP-seq protocol PCR1x on human GM12878
	BCL11A	477	99	18	BCL11A ChIP-seq protocol PCR1x on human GM12878
CGI hypo (3,049 CpGs)	SUZ12	212	9	2	SUZ12 ChIP-seq on human H1-hESC
	EZH2	329	10	5	ChIP-seq on human H1-hESC EZH2
	CTBP2	432	38	8	CTBP2 ChIP-seq on human H1-hESC
CGI hyper (1,974 CpGs)	SUZ12	73	6	0	SUZ12 ChIP-seq on human H1-hESC
	EZH2	159	8	0	ChIP-seq on human H1-hESC EZH2
	CTBP2	127	16	1	CTBP2 ChIP-seq on human H1-hESC

CGI: CpG island; T. Factor: transcription factor; hyper: hypermethylated; hypo: hypomethylated; DEGs: differentially expressed gene; bold indicates that the gene encoding the TF is differentially expressed.

Supplementary Table S5. Differentially methylated intergenic CpG islands and genes associated identified by i-cisTarget.

Position of the DMR			Width (bp)	N° CpGs	q-value	Mean methylation change (β -value)	Gene associated (i-cisTarget)
Chr	Start	End					
chr4	174428790	174430614	1825	12	1.15E-41	-0.072	<i>HAND2</i>
chr5	139088316	139089549	1234	7	5.68E-38	-0.203	
chr5	43000303	43001210	908	8	1.14E-33	-0.282	
chr2	241562085	241562650	566	5	9.23E-33	-0.218	
chr6	156717406	156718546	1141	6	3.94E-31	0.283	<i>ARID1B</i>
chr2	85640762	85641985	1224	13	4.21E-31	-0.158	
chr21	45575014	45576085	1072	5	2.63E-30	0.443	<i>C21orf33</i>
chr16	85981336	85981947	612	5	4.49E-26	0.387	<i>IRF8</i>
chr11	32008659	32009163	505	5	5.34E-24	-0.081	<i>RCN1</i>
chr6	27647843	27648605	763	7	2.64E-23	-0.059	<i>HIST1H2BL</i>
chr19	57149436	57149813	378	7	1.03E-22	-0.028	
chr22	36806001	36806655	655	6	2.80E-22	-0.197	
chr11	47416109	47417205	1097	6	4.26E-21	-0.126	
chr1	38941882	38942644	763	8	8.53E-20	-0.086	
chr4	85403167	85403915	749	5	4.17E-19	-0.066	
chr5	77253544	77253990	447	5	7.31E-19	-0.142	<i>TBCA</i>
chr2	223185049	223185848	800	11	1.10E-18	-0.125	
chr6	29521499	29521803	305	16	1.59E-18	-0.032	<i>UBD</i>
chr8	8820844	8821258	415	5	3.81E-18	-0.129	
chr11	20618001	20618651	651	6	4.39E-16	-0.028	<i>SLC6A5</i>
Bold indicates differentially expressed genes							

Supplementary Table S6. Top enrichment results from i-cisTarget of intergenic and intragenic CpG islands that are differentially methylated regions. Genes associated with the feature identified by i-cisTarget were overlapped with the list of differentially expressed genes (DEGs). The number of the associated genes that were differentially expressed is indicated.

Type of DMR	Feature	N° of genes associated	DEGs		Dataset used by i-cisTarget
			up	down	
Hypo intergenic	CTBP2	8	0	0	CTBP2 ChIP-seq on human H1-hESC
Hyper intergenic	RUNX3	3	1	0	RUNX3 ChIP-seq protocol v042211.1 on human GM12878
Bold indicates differentially expressed genes					

Supplementary Table S7. Differentially methylated regions overlapping CpG islands in gene promoters whose genes are differentially expressed

Symbol	Gene expression details				Differentially methylated region					
	q-val (transcript)	q-val (gene)	b	Chr	Start	End	Width	N° CpGs	DNA meth difference	N° sig CpGs (limma)
CD302	0.043	0.784	0.35	chr2	160653686	160655179	1494	11	-0.217	4
CD63	0.018	0.002	0.434	chr12	56121099	56122120	1022	7	-0.203	5
CDKN1C	0.001	0.001	0.768	chr11	2907670	2908471	802	13	-0.16	4
CLCN5	0.021	0.019	0.518	chrX	49687006	49687331	326	11	-0.174	3
CTIF	0.038	0.027	0.422	chr18	46064542	46065292	751	9	-0.142	2
CTS2	0	0	0.898	chr20	57582006	57583709	1704	27	-0.076	11
ELF4	0.001	0.004	0.599	chrX	129244725	129244816	92	5	-0.147	2
FGR	0.002	0	1.023	chr1	27960788	27962692	1905	13	-0.267	7
G0S2	0.001	0	1.943	chr1	209848479	209849006	528	8	-0.028	1
GK	0.028	0.037	0.597	chrX	30670939	30671488	550	12	-0.09	1
GPSM3	0.003	0	0.494	chr6	32164124	32165200	1077	23	-0.047	4
IFNGR2	0.006	0.001	0.693	chr21	34774627	34775459	833	8	-0.112	4
JMJD1C	0.034	0.166	0.308	chr10	65225648	65226682	1035	9	-0.125	2
KCNE3	0.004	0.002	0.612	chr11	74178114	74179208	1095	11	-0.1	3
MAFG	0.084	0.019	0.297	chr17	79880647	79882042	1396	9	-0.35	5
MAP4K4	0.006	0.001	0.69	chr2	102313069	102313888	820	6	-0.132	3
MMP14	0.056	0.041	0.781	chr14	23305153	23305957	805	9	-0.245	2
RARA	0.002	0	0.548	chr17	38498077	38499096	1020	10	-0.124	4
RBM47	0.002	0	1.131	chr4	40632362	40633572	1211	10	-0.249	5
SORT1	0.002	0.203	0.976	chr1	109940754	109941201	448	7	-0.18	4
SPI1	0	0	0.825	chr11	47398575	47401027	2453	15	-0.284	12
ZNF641	0.106	0.029	0.323	chr12	48744457	48745694	1238	13	-0.036	3
CD34	0.001	0	-0.444	chr1	208084071	208085889	1819	18	-0.082	3
CRHBP	0.03	0.005	-0.505	chr5	76248637	76249502	866	7	-0.265	5
DAPK1	0.055	0.033	-2.245	chr9	90112086	90113998	1913	19	-0.065	5
HOXA5	0.063	0.046	-0.322	chr7	27183591	27185393	1803	34	-0.232	6
HOXB4	0.033	0.023	-0.472	chr17	46656777	46660097	3321	19	-0.201	9
ITGA6	0.004	1	-3.202	chr2	173292978	173294093	1116	9	-0.125	2
MECP2	0.234	0.035	-0.111	chrX	153362589	153363472	884	6	-0.087	3
MORC2	0.017	1	-0.558	chr22	31317914	31318373	460	8	-0.102	5
MSRB3	0.253	0.032	-0.633	chr12	65671924	65672696	773	9	-0.122	3
SLC17A9	0.032	0.029	-0.395	chr20	61583686	61584850	1165	11	-0.153	5
TRIM9	0.006	0	-2.683	chr14	51562381	51563131	751	12	-0.094	2
UCK2	0.022	0.009	-0.358	chr1	165796134	165796797	664	7	-0.136	2
ZNF577	0.033	0.035	-0.632	chr19	52390810	52391789	980	14	-0.104	7
b: log(foldchange) computed by Sleuth										

Supplementary Table S8. List of antibodies used in flow cytometry

Antibody	Clone	Cat #	Manufacturer
CD5	53-7.3	48-0051-82	Bioscience
Gr1	RB6-8C5	108430	Biolegend
B220	RA3-6B2	103227	Biolegend
CD3	17A2	100214	Biolegend
CD11b	M1/70	101224	Biolegend
Ter119	TER-119	116232	Biolegend
CD117	2B8	105812	Bioscience
Sca-1	D7	108106	Bioscience
CD16b	CLB-gran11.5	550868	BD Biosciences
CD14	M5E2	301808	Biolegend
CD3	UCHT1	300452	Biolegend
CD19	H1B19	302216	Biolegend

Supplementary Table S9. List of primers used for RT-qPCR and bisulphite PCR

Gene	Sequence
	RT-qPCR
<i>SPI1</i>	GTT ACA GGC GTG CAA AAT GG
	TGC GTT TGG CGT TGG TAT AG
<i>CCR1</i>	GTC CCT TGG AAC CAG AGA GAA GC
	TAC CAA GGA GTA CAG AGG GGG
<i>PTGER4</i>	TCT GAC CTC GGT GTC CAA AAA
	AGT GCA AGG CTG GGT CTG TAG
<i>IFNGR1</i>	GGT AGC AGC ATG GCT CTC CT
	AAC ATT AGT TGG TGT AGG CAC TGA
<i>RUNX1</i>	CCC ATC GCT TTC AAG GTG GT
	TCA TTG CCA GCC ATC ACA GT
<i>ACTB</i>	CAC CAT TGG CAA TGA GCG GTT C
	AGG TCT TTG CGG ATG TCC ACG T
<i>GADPH</i>	GTC TCC TCT GAC TTC AAC AGC G
	ACC ACC CTG TTG CTG TAG CCA A
Bisulphite PCR	
<i>SPI1</i>	TTA GGA GGG TAG TGG GTG GG
	ACT CAC CCA AAA CTC CTA TAA CTC A