

Supplementary Table 1. Primers for real-time RT-PCR SYBR green analysis.

Gene	RT PCR primers	Length	Annealing temp(°C)
<i>LHX2</i>	(F)- CCAAGGACTTGAAGCAGCTC (R)- GTAAGAGGTTGCGCCTGAAC	176	68
<i>NRP2</i>	(F)- CTCTGAAAAGGGGAATCCAT (R)- GAAGGCTTTTTCGCCTTTCT	110	60
<i>KCNK2</i>	(F)- TAACAAC TATTGGATTTGGTGACTAC (R)- GCCCTACAAGGATCCAGAAC	100	58
<i>HPRT1</i>	(F)- TGACACTGGCAAAAACAATGCA (R)- GGCCTTTTCACCAGCAAGCT	90	58

*F: Forward; R: Reverse; RT: Reverse transcriptase.*

Supplementary Table 2. Primer sequences and PCR conditions.

Gene	Primers	Amplicon location	Product size (bp)	Tm (°C)	Assay	Enzyme
<i>DLG1</i>	(F)-GTTTTTAGTTAGGATATGGT (R)-CTTCTTTTCTACACATCAACA	chr8:13,034,845-13,035,136	292	56	COBRA	BSTU1
<i>KCNK2</i>	(F)-TTTAGTAAAGGGTTTGTGTTGAG (R)-AACCCCTAACTTCTCCAATCTACAC	chr1:213,322,021-213,322,249	230	56	COBRA	Hpy99I
<i>ADAM12</i>	(F)-GTGGATTATTTATAGGTTGGTTTTT (R)-CTAAACTTCTAACCTTTCATTTTAAAA	chr10:128,066,859-128,067,044	186	56	COBRA/bisulfite sequencing	BSTU1
<i>LHX2</i>	(F)-GTTAGTAAAGTTTGTAGGGTGGTG (R)-AAAAAATACTACTACCTTCTCC	chr9:125,834,465-125,834,702	238	63	COBRA	BSTU1
<i>DLEU7</i>	(F)-TGGTTGATGGAGGTTATTAAGG (R)-TTCAAACAACCTTAAATCAAAACAC	chr13:50,315,746-50,316,012	267	62	COBRA/bisulfite sequencing	BSTU1
<i>DMRT2</i>	(F)-TTGGTGAAGGTTAATGATTTTATT (R)-ATACCTAAAACACTCTCCCAACAACC	chr9:1,038,332-1,038,597	266	56	COBRA	BSTU1
<i>SFRP2</i> *	(F)-TTTTTATTTTTAGATTTGTATAAAAA (R)-AACCAAAAACCCCTAACACATC	chr4:154,929,678-154,929,994	316	54	COBRA	BSTU1
<i>POU3F3</i>	(F)-TTTATATTAGGGTTATTTGGGGGT (R)-AATACACCAAACCTTAAACTCCACC	chr2:104,838,934-104,839,170	236	56	COBRA	BSTU1
<i>PCDHGB7</i>	(F)-TGGGGTAGAATAAGGTAGTAGTAAAGGAA (R)-ACAATCCACACAAAACCTCTAAAC	chr5:140,777,593-140,777,963	371	56	COBRA	BSTU1
<i>LRP1B</i>	(F)-TAGGAAAGTTAAGGAAGTTAGGGGA (R)-CACATCTACTCCAAAACAAAAAAA	chr2:142,604,635-142,604,955	319	56	COBRA	Taq <sup>®</sup> 1
<i>HOXC10</i>	(F)-TTTTTTTTGAAAATGATATGTTTT (R)-TACAACAACAACATCTCCTCCTTAAC	chr12:52,665,299-52,665,647	349	56	COBRA	BSTU1
<i>DUOX2</i>	(F)-TTTTAATCGGATTTAAGTCTCGG (M) (R)-AATATCAAACTCCTTAACGACGAA (M)	chr15:43,192,664-43,192,828	164	60	MSP	N/A
	(F)-TTTTAATGGATTTAAGTGTGG (U) (R)-TATCAAACTCCTTAACAACAAA (U)		162	62		
<i>NRP2</i>	(F)-TTTTAGAGATTAGCGTTGTAGTCGA (M) (R)-AAACCGAAACTAAAACCTCCG (M)	chr2:206,259,130-206,259,300	169	60	MSP	N/A
	(F)-TTTTAGAGATTAGTTGTAGTTGA (U) (R)-AAAACCAAACCTAAAACCTCCAC (U)		170	60		
<i>RLN2</i>	(F)-AGATTAGTGGTAGGTGAGAGTTTTC (M) (R)-CACGACTAAATAAAAACCTAAAACGAT (M)	chr9:5,294,231-5,294,382	151	60	MSP	N/A
	(F)-TAGTGGTAGGTGAGAGTTTGT (U) (R)-CCACAACCTAAAATAAAAACCTAAAACAAT (U)		148	60		
<i>APC2(5')</i>	(F)-TGTTTTTGTGTTGAAAATTTTTTA (R)-CAACCAATCCCAACAATCTC	chr19:1,407,976-1,408,199	224	56	COBRA	BSTU1
<i>APC2(3')</i>	(F)-GGGTAAGAAATAGAGTAGGGTTGGAG (R)-TCAACAATAAAAACCTAAAACCTC	chr19:1,418,674-1,419,012	339	56	COBRA	BSTU1

\*SFRP2 primer sequences were already published.

COBRA: Combined bisulfite restriction analysis; MSP: Methylation-specific PCR; Tm: Melting temperature.

Supplementary Table 3. Statistical analysis of COBRA and MSP results at different levels of CD38.

Genes	1–10% CD38		11–92% CD38		Exact P	Odds ratio	Log odds	95% CI log odds
	M	%	M	%				
DLEU7	12/12	100	7/21	33.33	0.0002	0.021	-3.878	(-6.839, -0.917)
HOXC10	4/10	40	19/19	100	0.0004	56.333	4.031	(0.978, 7.085)
SFRP2	7/14	50	20/22	90.91	0.0144	8.200	2.104	(0.446, 3.762)
POU3F3	5/9	55.56	19/22	86.36	0.1504	4.558	1.517	(-0.17, 3.204)
LHX2	12/14	85.71	13/21	61.9	0.2516	0.318	-1.147	(-2.753, 0.46)
ADAM12	7/11	63.64	14/18	77.78	0.4327	1.933	0.659	(-0.917, 2.235)
LRP1B	7/10	70	11/22	50	0.4461	0.467	-0.762	(-2.271, 0.747)
DMRT2	4/11	36.36	10/19	52.63	0.4664	1.842	0.611	(-0.851, 2.072)
DUOX2	5/12	41.67	6/18	33.33	0.7116	0.709	-0.344	(-1.796, 1.108)
DLC1	9/14	64.29	16/22	72.73	0.7159	1.470	0.385	(-1.003, 1.773)
NRP2	4/10	40	11/23	47.83	0.7220	1.329	0.284	(-1.16, 1.729)
APC(3')	10/14	71.43	16/22	72.73	1.0000	1.088	0.084	(-1.345, 1.514)
APC(5')	7/10	70	15/16	93.75	0.2642	4.822	1.573	(-0.529, 3.675)
KCNK2	4/14	28.57	6/21	28.57	1.0000	0.978	-0.022	(-1.456, 1.413)
RLN2	7/11	63.64	12/19	63.16	1.0000	1.000	0.000	(-1.478, 1.478)
PCDHGB7	14/14	100	22/22	100	NA	1.552	0.439	(-3.536, 4.414)

For each of the 15 genes listed (two CpG islands for APC2), the number and % of cases confirmed as methylated are reported, as well as exact p-values and odds ratios (odds of methylation in the CD38<sup>high</sup> group/odds of methylation in the CD38<sup>low</sup> group). CD38<sup>low</sup> is defined as 1–10% and CD38<sup>high</sup> as 11–92% CD38 expression. The right column of the table contains the 95% CIs of the log odds ratios. In the case of PCDHGB7, since every sample tested was methylated, there is no calculated p-value.

COBRA: Combined bisulfite restriction analysis; MSP: Methylation-specific PCR.

Supplementary Table 4. Statistical analysis of COBRA and MSP results at different levels of CD38.

Genes	1–30% CD38		31–92% CD38		Exact P	Odds ratio	Log odds	95% CI log odds
	M	%	M	%				
DLEU7	18/21	85.71	1/12	8.33	<0.0001	0.025	-3.702	(-5.751, -1.652)
NRP2	5/20	25	10/13	76.92	0.0052	8.455	2.135	(0.582, 3.687)
SFRP2	14/23	55.56	13/13	100	0.0136	17.690	2.873	(-0.066, 5.812)
ADAM12	11/19	57.9	10/10	100	0.0265	15.522	2.742	(-0.23, 5.715)
APC(3')	13/22	59.1	13/14	92.86	0.0536	6.333	1.846	(-0.034, 3.726)
HOXC10	12/18	66.67	11/11	100	0.0576	11.960	2.482	(-0.504, 5.467)
DMRT2	7/19	36.84	7/11	63.64	0.2568	2.778	1.022	(-0.457, 2.5)
DLC1	14/23	60.87	11/13	84.62	0.2586	3.014	1.103	(-0.491, 2.697)
LRP1B	12/18	66.67	6/14	42.86	0.2831	0.398	-0.922	(-2.315, 0.471)
POU3F3	12/17	70.59	12/14	85.71	0.4117	2.200	0.788	(-0.9, 2.477)
DUOX2	8/19	42.11	3/11	27.27	0.4661	0.557	-0.585	(-2.113, 0.943)
RLN2	13/19	68.42	6/11	54.55	0.6956	0.569	-0.564	(-2.035, 0.908)
KCNK2	7/22	31.82	3/13	23.07	0.7094	0.689	-0.373	(-1.864, 1.118)
APC(5')	14/17	82.35	8/9	88.89	1.0000	1.368	0.313	(-1.779, 2.405)
LHX2	16/22	72.72	9/13	69.23	1.0000	0.832	-0.184	(-1.627, 1.259)
PCDHGB7	23/23	100	13/13	100	NA	0.574	-0.554	(-4.531, 3.422)

For each of the 15 genes listed (two CpG islands for APC2), the number and % of cases confirmed as methylated are reported, as well as exact p-values and odds ratios (odds of methylation in the CD38<sup>high</sup> group/odds of methylation in the CD38<sup>low</sup> group). CD38<sup>low</sup> is defined as 1–30% and CD38<sup>high</sup> as 31–92% CD38 expression. The right column of the table contains the 95% CIs of the log odds ratios. In the case of PCDHGB7, since every sample tested was methylated, there is no calculated p-value.

COBRA: Combined bisulfite restriction analysis; MSP: Methylation-specific PCR.

Supplementary Table 5. Candidate genes/regions preferentially methylated in groups expressing high or low CD38.

Gene abbreviation	Gene name	Gene functions
<b>Candidates preferentially methylated in CD38<sup>high</sup></b>		
ABTB2	Ankyrin repeat and BTB (POZ) domain containing 2	Regulation of cellular growth, and nucleosome assembly
ADAMTS17	ADAM metalloproteinase with thrombospondin type 1 motif, 17	Metalloproteinase, proteolysis
ARHGAP5	Rho GTPase activating protein 5	Regulates RHO GTPases, a family which may mediate cytoskeleton changes and influence cell migration and the WNT planar cell polarity pathway
ARID2	AT rich interactive domain 2 (ARID, RFX-like)	Chromatin remodeling. Required for stability of the SWI/SNF chromatin remodeling complex
ASTN2	Astroctactin 2, member of astroctactin family;	Neuronal adhesion molecule required for glial-guided migration of young postmitotic neuroblasts in developing brain
ATP11B	ATPase, class VI, type 11B	Drives uphill transport of ions across membranes
BARHL1	BarH-like homeobox 1	Homeodomain transcription factor of the BarH class expressed by developing inner ear hair cells, cerebellar granule cells, precerebellar neurons, and collicular neurons
BHLHB4	Basic helix-loop-helix domain containing, class B, 4	Transcriptional regulator implicated in retinal functions
CACNB3	calcium channel, voltage-dependent, $\beta$ 3 subunit	Contributes to the function of the calcium channel by increasing peak calcium current
CCDC28A	Coiled-coil domain containing 28A	Uncharacterized function
EBF3	Early B-cell factor 3	Transcription factor involved in B-cell development
EFEMP1	EGF-containing fibulin-like extracellular matrix protein 1	Extracellular matrix protein important in vision
EFNA5	Ephrin-A5	Involved in cell migration and induces compartmentalized signaling within a caveolae-like membrane microdomain when bound to its cognate receptor
ELOVL4	Elongation of very long chain fatty acids (FEN1/Elo2, SUR4/Elo3, yeast)-like 4	Membrane-bound protein which is a member of the ELO family, proteins which participate in the biosynthesis of fatty acids
EN1	Engrailed homeobox 1	Mutations in the mouse homologs, En1 and En2, are frequently lethal. The human engrailed homologs 1 and 2 encode homeodomain-containing proteins implicated in the control of pattern formation during development of the CNS.
FOXE1	Forkhead box E1 (thyroid transcription factor 2)	The forkhead family of transcription factors, which is characterized by a distinct forkhead domain. This gene functions as a thyroid transcription factor which likely plays a crucial role in thyroid morphogenesis
GATA6	GATA binding protein 6	Important in early tissue development. GATA6-regulated WNT signaling is required for lung development
HAND1	Heart and neural crest derivatives expressed 1	Plays an essential role in early trophoblast differentiation and in cardiac morphogenesis. HAND transcription factors have a crucial function in sustaining the survival of neonatal sympathetic neurons
HSA-MIR-615	MicroRNA615	MicroRNA of uncharacterized function
KCNJ3	Potassium inwardly-rectifying channel, subfamily J, member 3	This potassium channel is controlled by G proteins. Inward rectifier potassium channels are characterized by a greater tendency to allow potassium to flow into the cell rather than out of it
MPPED2	Metallophosphoesterase domain containing 2	Found in fetal brain. Currently uncharacterized function
NETO1	Neuropilin (NRP) and tolloid (TLL)-like 1	May play a role in the development and/or maintenance of neuronal circuitry. note that nrp2 was methylated in 50% of CLL

CLL: Chronic lymphocytic leukemia.

Supplementary Table 5. Candidate genes/regions preferentially methylated in groups expressing high or low CD38.

Gene abbreviation	Gene name	Gene functions
<b>Candidates preferentially methylated in CD38<sup>high</sup> (cont.)</b>		
<i>OLIG2</i>	Oligodendrocyte lineage transcription factor 2	Required for oligodendrocyte and motor neuron specification in the spinal cord. Cooperates with <i>OLIG1</i> to establish the pMN domain of the embryonic neural tube
<i>PAP2D</i>	Phosphatidic acid phosphatase type 2	All members of this protein family convert phosphatidic acid to diacylglycerol, and function in <i>de novo</i> synthesis of glycerolipids as well as in receptor-activated signal transduction mediated by phospholipase D
<i>PTCHD1</i>	Patched domain containing 1	Likely member of the Patched family, but currently of uncharacterized function
<i>PXDN</i>	Peroxidase homolog	Likely member of the Peroxidase family, but currently of uncharacterized function
<i>RTKN</i>	Rhotekin	Encodes a scaffold protein that inhibits the GTPase activity of Rho proteins and activates NF- $\kappa$ -B. Rho proteins and thus regulates many important cellular processes, including cell growth and transformation
<i>ST6GAL2</i>	ST6 $\beta$ -galactosamide $\alpha$ -2,6-sialyltransferase 2	Transfers sialic acid from the donor of substrate CMP-sialic acid to galactose containing acceptor substrates
<i>THBD</i>	Thrombomodulin	A specific endothelial cell receptor that forms a 1:1 stoichiometric complex with thrombin. This complex is responsible for the conversion of protein C to the activated protein C (protein Ca)
<i>UNCX</i>	UNC homeobox	Transcription factor involved in somitogenesis and neurogenesis.
<i>VSX1</i>	Visual system homeobox 1	Binds to the 37-bp core of the locus control region (LCR) of the red/green visual pigment gene cluster
<i>ZFP37</i>	Zinc finger protein 37 homolog (mouse)	May be involved in transcriptional regulation
<b>Candidates preferentially methylated in CD38<sup>low</sup></b>		
<i>C16ORF30</i>	Transmembrane protein 204	Can influence paracellular permeability. Appears to be involved in cell–cell interactions through adherens junctions
<i>TMEM204</i>		
<i>C19ORF29</i>	Chromosome 19 open reading frame 29	May be involved in pre-mRNA splicing
<i>CLSPN</i>	Claspin homolog	Required for checkpoint mediated cell-cycle arrest in response to DNA damage. May play a role as a sensor that monitors the integrity of DNA replication forks
<i>WNT9A</i>	Wingless-type MMTV integration site family, member 9A	Ligand for members of the frizzled family of receptors. May be a signaling molecule that affects the development of discrete regions of tissues
<i>ZFX1B</i>	Zinc finger E-box binding homeobox 2	Transcriptional inhibitor that binds to DNA sequence 5'-CACCT-3' in different promoters. Represses transcription of E-cadherin

CLL: Chronic lymphocytic leukemia.

Supplementary Table 6. Candidate genes/loci methylated across all CD38 expression levels.

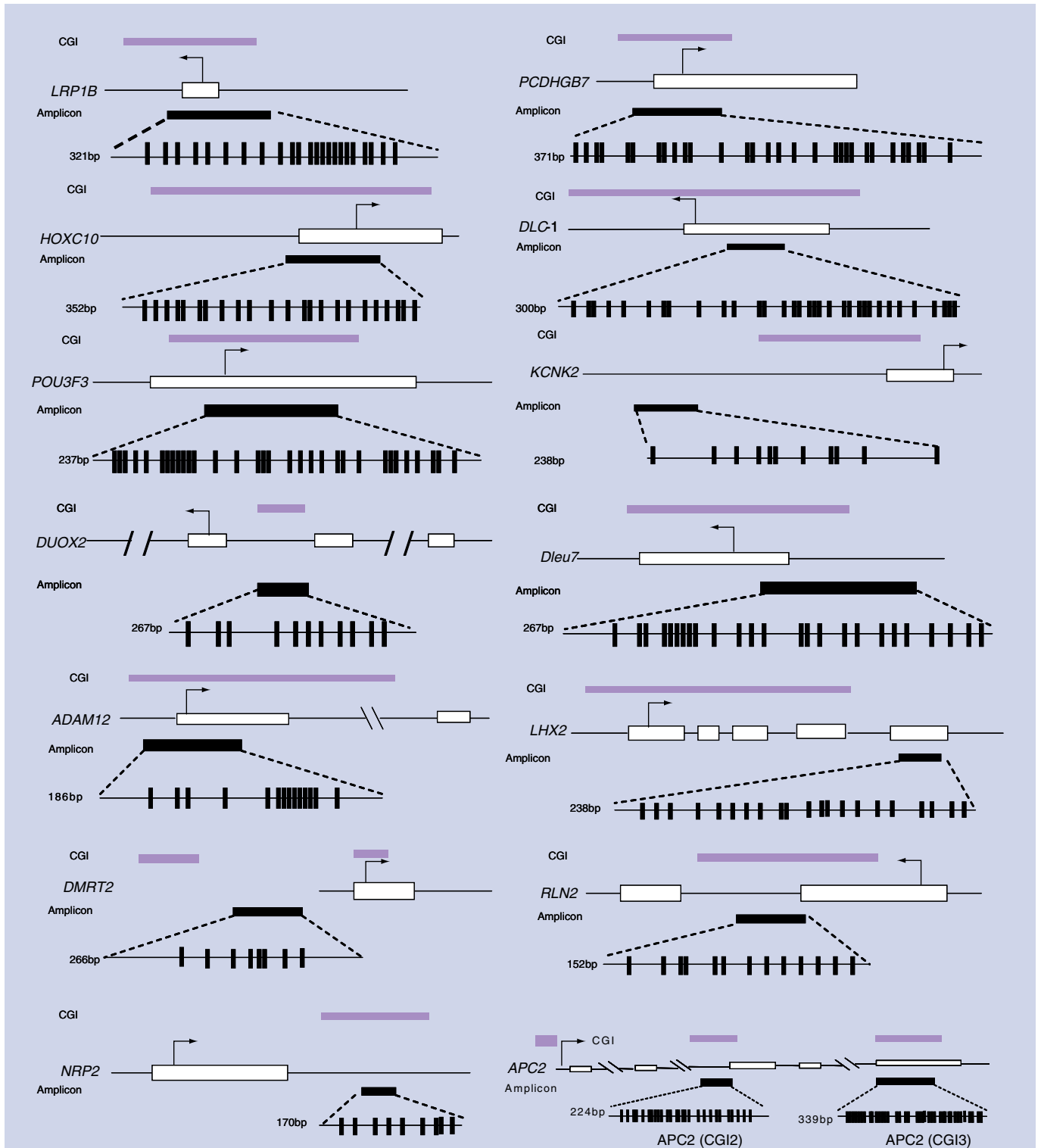
Gene symbol	DAVID gene name
AATF	Apoptosis antagonizing transcription factor
ADARB2	Adenosine deaminase, RNA-specific, B2 (red2 homolog rat)
ASPH	Aspartate $\beta$ -hydroxylase
CDX2	Caudal type homeobox transcription factor 2
CEP152	Centrosomal protein 152 kda
CORO2B	Coronin, actin binding protein, 2b
CPT1C	Carnitine palmitoyltransferase 1c
CTBP2	C-terminal binding protein 2
CYP46A1	Cytochrome P450, family 46, subfamily A, polypeptide 1
DBC1	Deleted in bladder cancer 1
DLX1	Distal-less homeobox 1
EN2	Engrailed homolog 2
EPHA4	EPH receptor A4
FGF9	Fibroblast growth factor 9 (glia-activating factor)
FLJ46347	Hypothetical loc389064
FOXA1	Forkhead box A1
FOXC2	Forkhead box C2 (Mfh-1, mesenchyme forkhead 1)
FOXD1	Forkhead box D1
FTMT	Ferritin mitochondrial
GBX2	Gastrulation brain homeobox 2
HAND2	Heart and neural crest derivatives expressed 2
HES4	Hairy and enhancer of split 4 ( <i>Drosophila</i> )
HOXD3	Homeobox D3
HOXD8	Homeobox D8
HSPA4L	Heat shock 70kda protein 4-like
IGF2BP1	Insulin-like growth factor 2 mRNA binding protein 1
IRX4	Iroquois homeobox protein 4
JDP2	JUN dimerization protein 2
JMJD1A	Jumonji domain containing 1a
LHX1	LIM homeobox 1
MAGI1	Membrane associated guanylate kinase, Ww and PdZ domain containing 1
NELL2	Nel-like 2 (chicken)
NKX2-2	NK2 transcription factor related, locus 2 ( <i>Drosophila</i> )
NKX2-3	NK2 transcription factor homolog C ( <i>Drosophila</i> )
NPAS2	Neuronal PAS domain protein 2
NPY2R	Neuropeptide Y receptor Y2
NR2F2	Nuclear receptor subfamily 2, group F, member 2
NRG1	Neuregulin 1
NTRK3	Neurotrophic tyrosine kinase, receptor, type 3
PAX6	Paired box gene 6 (aniridia, keratitis)
PCDH19	Protocadherin 19
PDE10A	Phosphodiesterase 10a
PENK	Proenkephalin
PIK3R2	Phosphoinositide-3-kinase, regulatory subunit 2 (P85 $\beta$ )
PRAC	Small nuclear protein PRAC
PRDM12	PR domain containing 12
SALL1	SAL-like 1 ( <i>Drosophila</i> )

DAVID: Database for annotation, visualization and integrated discovery.

**Supplementary Table 6. Candidate genes/loci methylated across all CD38 expression levels.**

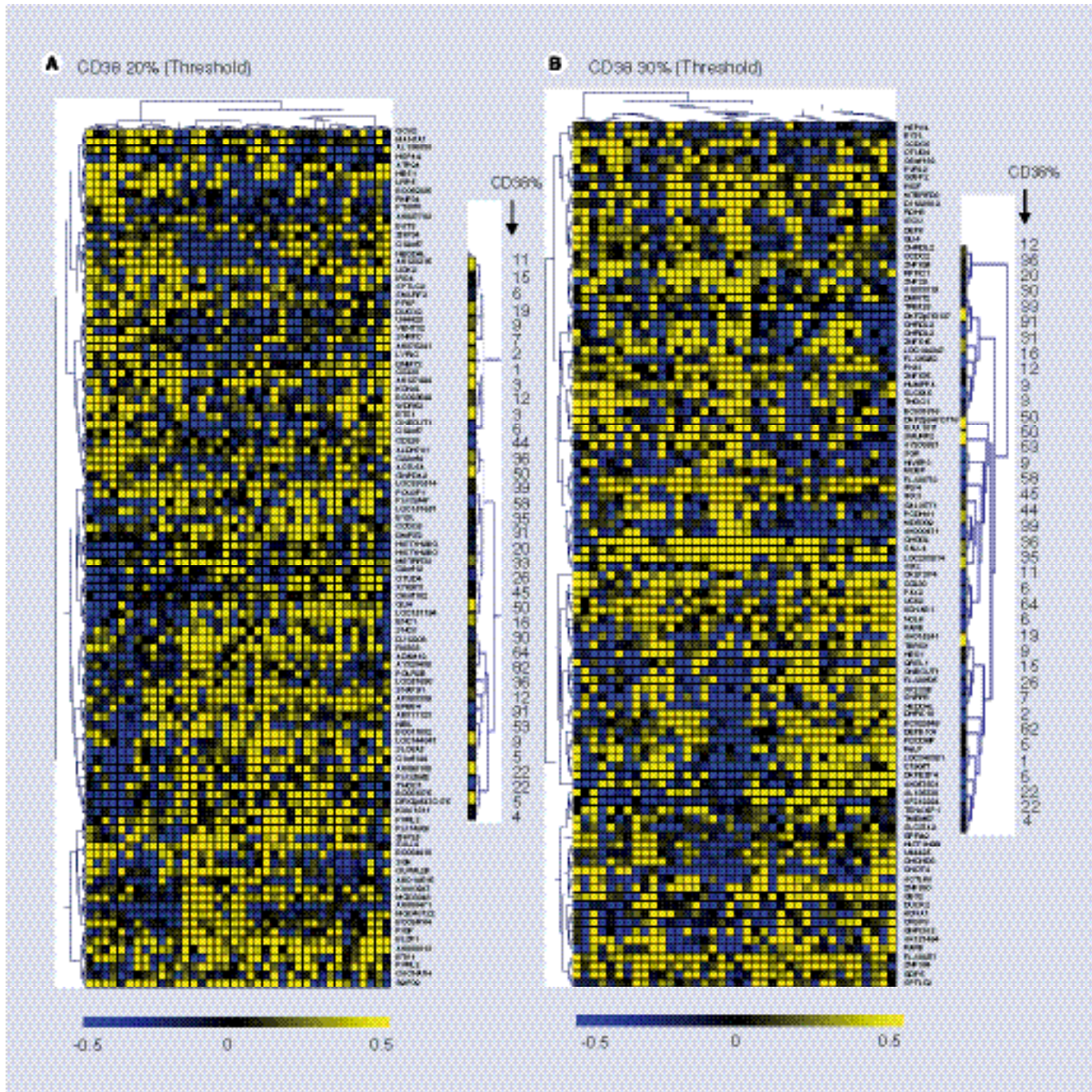
<b>Gene symbol</b>	<b>DAVID gene name</b>
<i>SALL3</i>	SAL-like 3 ( <i>Drosophila</i> )
<i>SIM2</i>	Single-minded homolog 2 ( <i>Drosophila</i> )
<i>SIN3A</i>	SIN3 homolog A, transcription regulator (yeast)
<i>SLC16A12</i>	Solute carrier family 16 (monocarboxylic acid transporters), member 12
<i>SLC26A1</i>	Solute carrier family 26 (sulfate transporter), member 1
<i>SLC6A2</i>	Solute carrier family 6 (neurotransmitter transporter, noradrenalin), member 2
<i>SLC6A5</i>	Solute carrier family 6 (neurotransmitter transporter, glycine), member 5
<i>SMC2</i>	SMC2 structural maintenance of chromosomes 2-like 1 (yeast)
<i>TBX18</i>	T-box 18
<i>TBX20</i>	T-box 20
<i>TBX3</i>	T-box 3 (ulnar mammary syndrome)
<i>TBX4</i>	T-box 4
<i>TRAF3IP3</i>	TRAF3 interacting protein 3
<i>WBSCR17</i>	Williams-Beuren syndrome chromosome region 17
<i>WNT7A</i>	Wingless-type mmtv integration site family, member 7a
<i>ZNF503</i>	Zinc finger protein 503
<i>ZNF667</i>	Zinc finger protein 667
<i>ZNRD1</i>	Zinc ribbon domain containing, 1

DAVID: Database for annotation, visualization and integrated discovery.



**Supplementary Figure 1. Relative locations of CpG islands and gene regulatory regions.**  
 CGI: CpG island.





**Supplementary Figure 2. Hierarchical clustering analysis of DNA methylation. (A)** This cluster is based on a 20% CD38 threshold similar to that in [FIGURE 1](#). This illustrates a measure of relatedness of DNA methylation across all loci for each sample. Each column represents a patient sample and each row represents a clone/locus on microarray chip. The fluorescence ratios of cy3/cy5 are measures of DNA methylation and are depicted as a color intensity (-0.5 to +0.5) in log base 2; yellow indicates loci that have a higher level of DNA methylation in chronic lymphocytic leukemia compared with normal controls, blue indicates a lower level of methylation and black indicates no change. Graded colors across the spectrum represent various levels of methylation. The dendrogram from the top of the cluster (rotated 90° and enlarged on the right) represents the CD38 expression level of each sample. Not every patient sample clustered with the expected group. **(B)** This cluster is based on a 30% CD38 threshold similar to that in **(A)**.