

Supplementary Figure S1. Validation between $SF3B1^{WT}$ and $SF3B1^{mut}$ CLL patients. Correlation plots between methylation values of CpGs measured with EpiTyper and QSEA beta values for the 16 DMRs chosen for validation. DMRs were sorted by the correlation p-value.

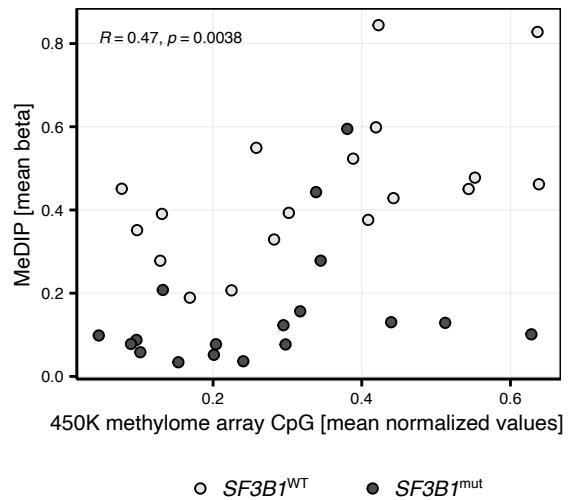


Figure S2. Correlation between mean methylation of 18 DMRs and overlapping CpG from 450K methylome array. Each dot represents a group (CLL $SF3B1^{WT}$ or CLL $SF3B1^{mut}$) mean beta methylation value of a DMR detected in this study and corresponding group mean methylation of a CpG located within same region derived from CLLmethylation study and available from BloodCancerMultiOmics2017 R package [1].

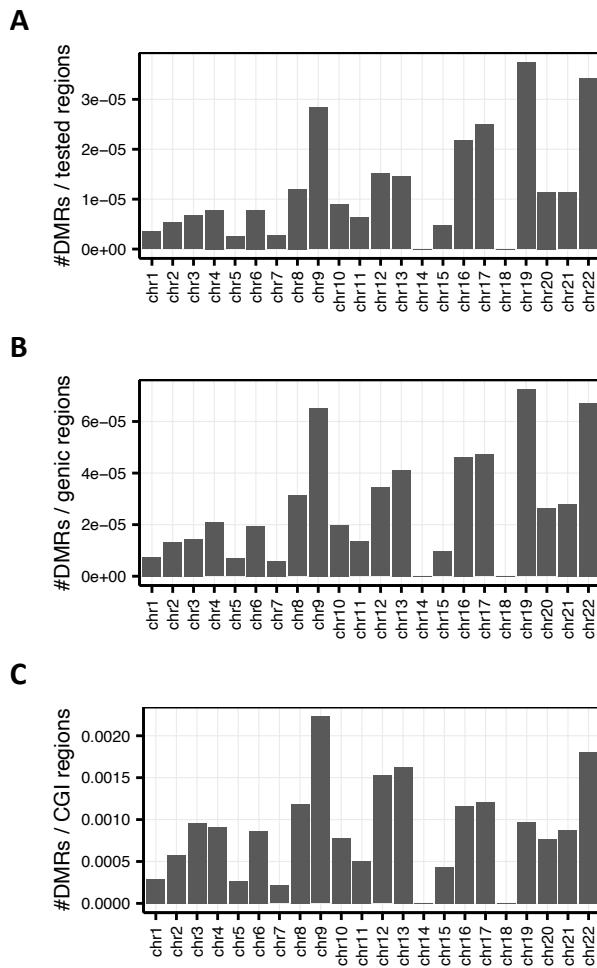


Figure S3. Number of DMRs corrected for A) number of regions tested; B) number of genic regions tested; C) number of regions with a CpG island. The number of differentially methylated regions per chromosome that was identified between *SF3B1^{WT}* and *SF3B1^{mut}* CLL patients' samples divided by the number of A) all regions tested; B) regions that overlapped a gene body; C) regions that overlapped a CpG island at each chromosome.

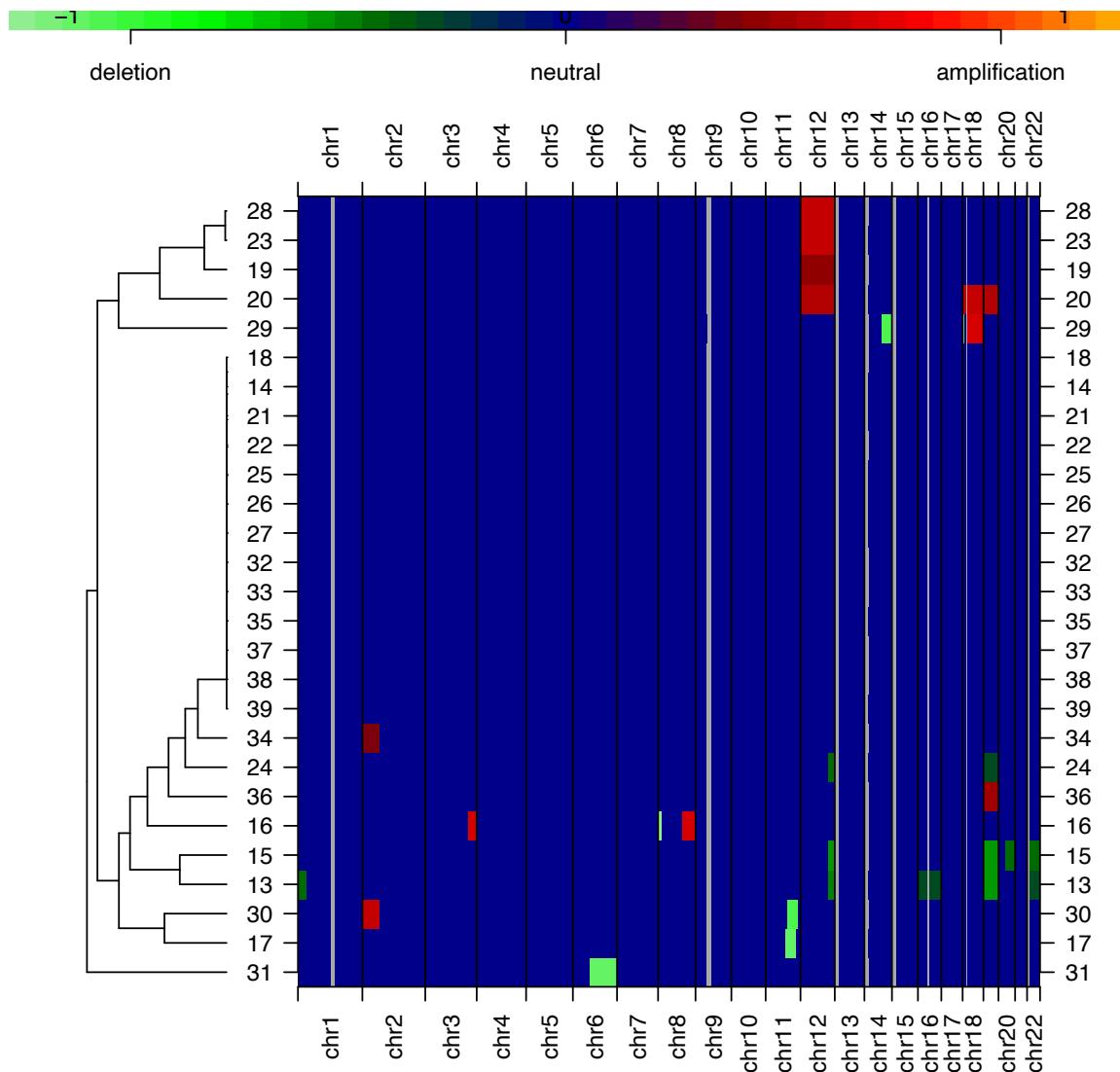
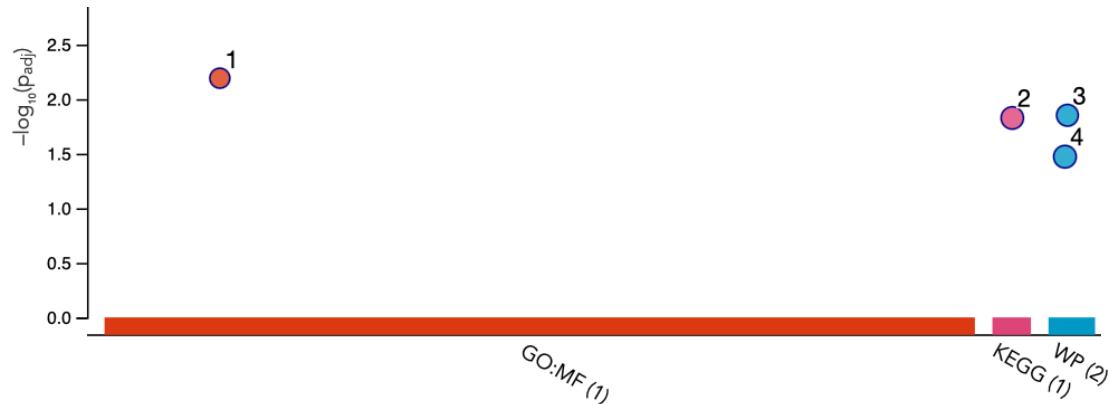


Figure S4. Copy number variation estimated based on MeDIP-seq data. The copy number variation per sample (rows) for each chromosome was calculated with addCNV function of QSEA R package [2] with the following options: window size = 2Mb and fragment size =250bp. The blue color indicates single copy, whereas green indicates a deletion and red an amplification.

A.

ID	Source	Term ID	Term Name	p _{adj} (query_1)
1	GO:MF	GO:0005112	Notch binding	6.430×10 ⁻³
2	KEGG	KEGG:04330	Notch signaling pathway	1.487×10 ⁻²
3	WP	WP:WP268	Notch Signaling	1.407×10 ⁻²
4	WP	WP:WP61	Notch Signaling Pathway Netpath	3.369×10 ⁻²

version
date
organism

e103_eg50_p15_68c0e33

g:Profiler

13/05/2021, 15:21:35

hsapiens

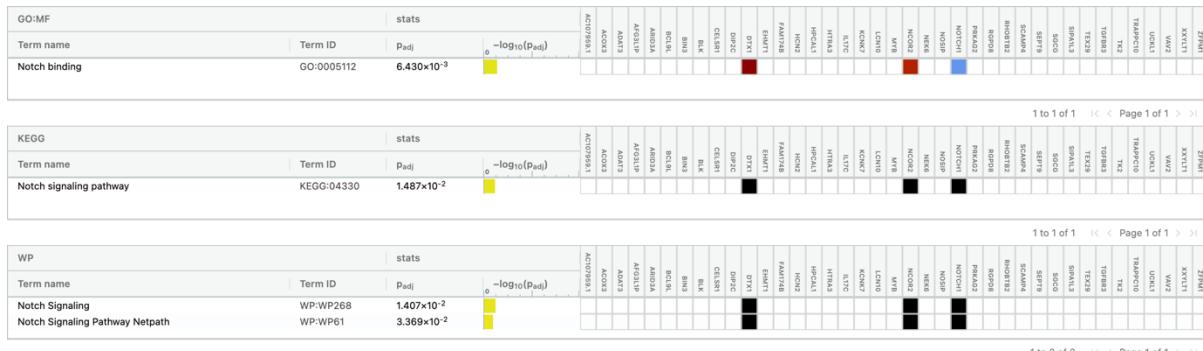
B.

Figure S5. Biological annotation of genes with DMRs. The 39 unique genes that contain a DMR and were annotated were used for enrichment analysis with gProfiler. A) Graphical representation of the enrichment test results with significant terms plotted on the adjusted p-value logarithmic scale. B) Detailed table with the results of the enrichment analysis. All genes taken for the analysis are listed and the boxes are filled if a gene is annotated for the term listed in the first column.

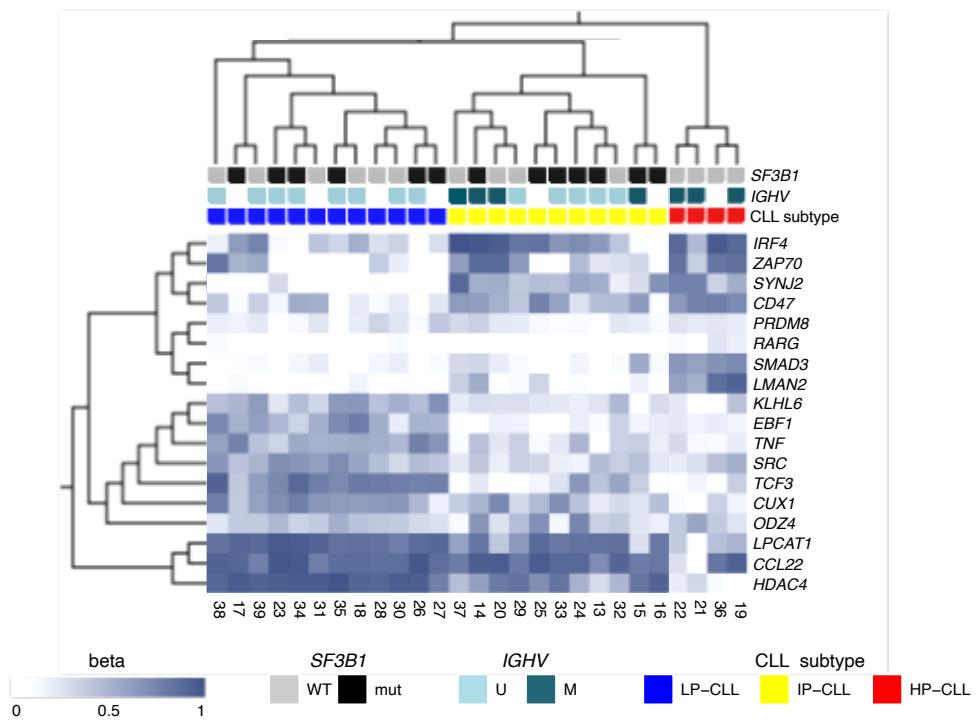


Figure S6. *SF3B1* mutation affects B cells in at least two out of three CLL subtypes. Heat map showing the methylation beta values for the eighteen most variable regions that were used to differentiate CLL subtypes as defined in Oakes et al. [3]. The samples were categorized based on the clusters shown in Figure 4A.

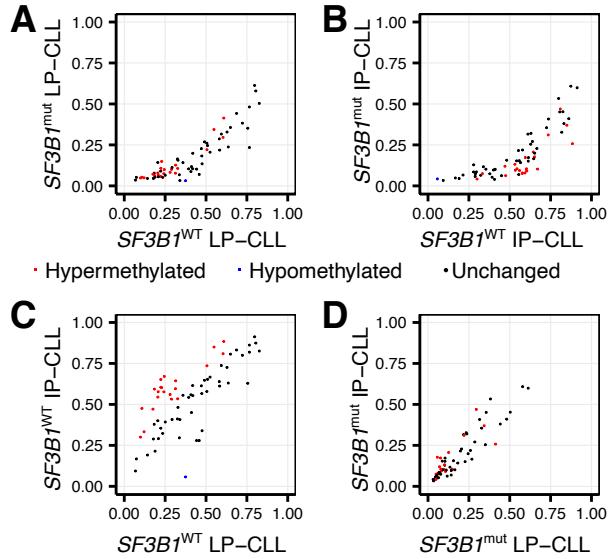


Figure S7. DMRs during physiological B cell maturation. For the 67 differentially methylated regions (DMRs) detected between all *SF3B1*^{mut} and *SF3B1*^{WT} samples we plotted their mean methylation levels (beta normalized) for LP-CLL and IP-CLL samples separately. Every dot represents a DMR. The DMRs are colored by physiological changes during LP- to IP-CLL maturation in *SF3B1*^{WT} - with blue color indicating hypomethylated and red hypermethylated regions ($\geq 20\%$ methylation level difference). (A) Comparison between LP-CLL- *SF3B1*^{mut} and LP-CLL- *SF3B1*^{WT}; (B) Comparison between IP-CLL- *SF3B1*^{mut} and IP-CLL- *SF3B1*^{WT}; (C) Comparison between IP-CLL- *SF3B1*^{WT} and LP-CLL- *SF3B1*^{WT}; D) Comparison between IP-CLL- *SF3B1*^{mut} and LP-CLL- *SF3B1*^{mut}.

Table S1. Information about the CLL patients used within the study.

MeDIP ID	MeDIP batch	Binet* disease stage	White blood cells [1/ml]	SF3B1	IGHV	Sex	Age at sampling [years]	SF3B1 mutation**	TP53 mutation**	ATM mutation**	XPO1 mutation**	NOTCH1 mutation**
18	1	C	49.29	WT	U	male	50					
19	1	B/C	98.89	WT	M	male	64					p.F1606L (5%)
20	1	n.a.	24.3	WT	M	male	48					
21	1	n.a.	24.39	WT	M	female	57					
22	1	A/B	31.62	WT	M	male	60					
28	2	A/B	49.26	WT	N/A	male	74					
29	2	A/B	45.19	WT	U	male	68					
30	2	A/B	52.46	WT	U	male	67					
31	2	A/B	58.5	WT	N/A	male	47					
32	2	n.a.	78.42	WT	U	male	70					
36	3	n.a.	27.64	WT	N/A	male	65					
37	3	A/B	48.71	WT	M	male	57					
38	3	A/B	48.87	WT	U	female	76					
39	3	n.a.	17.85	WT	U	male	64					
13	1	A/B	124.02	mut	U	male	52	H662Q (45%)				
14	1	A/B	149.42	mut	M	male	64	K700E (44%)				
15	1	A/B	34.91	mut	M	male	49	G742D (40%)				
16	1	A/B	483.77	mut	N/A	male	60	E622D (49%)				
17	1	A/B	45.81	mut	N/A	male	57	K666M (48%)		p.E571K (49%)		
23	2	A/B	59.92	mut	U	male	75	Q699E (51%)				p.P2514fs (12%)
24	2	C	140.33	mut***	U	male	66	A284T (49%)***				
25	2	C	33.3	mut	N/A	female	74	K700E (48%)				
26	2	A/B	26.69	mut	U	male	69	N626Y (48%)				
27	2	C	51.15	mut	N/A	male	49	D894G (54%); I704N (19%)				
33	3	C	110.26	mut	U	male	72	Y623C (42%)				
34	3	A/B	73.62	mut	U	male	59	I704F (48%)	p.L43**** (6%)			
35	3	A/B	32.76	mut	U	male	64	K666E (51%)			p.C1736Y (66%)	

* Binet's disease stage as defined in Binet *et al.* [4]** Mutation identified in the patient by Vollbrecht *et al.* 2015 [5]*** The mutation was identified in the N-terminal, outside the highly conserved HEAT repeats 5–8, see Fig. 2 legend in Vollbrecht *et al.* 2015 [5]**** Deletion on the second allele (see Vollbrecht *et al.* 2015 [5])

Table S3. Primers designed for the DMR validation with EpiTYPER MassARRAY

Name	Primer plus tag (small letters)	Anealing temperature [°C]	gene	length of Epityper target	genomic coordinate of Epityper assay	genomic coordinate of DMR from MeDIP-seq
DMR_01_10F	aggaaagagagGAGTAGTTGGGATTATAGGTATTGTT	56		257	chr12:1602716-1602972	chr12:1602751-1603000
DMR_01_T7R	cgttaatacgactcaatagggagaaggctCATAAAAACCCCTCAAATTATTC					
DMR_02_10F	aggaaagagAGATTTTATTGTAGGGTAAAGG	60	UCKL1	264	chr20:63950750-63951013	chr20:63950751-63951000
DMR_02_T7R	cgttaatacgactcaatagggagaaggctACTACCCACAAAAATAATCCTACCC					
DMR_03_10F	aggaaagaggAGATTTTTAGAATTGGGTTA	56	ACOX3	290	chr4:8384511-8384800	chr4:8384501-8384750
DMR_03_T7R	cgttaatacgactcaatagggagaaggctTCCAAACCTTACTTCCAAAACCTTA					
DMR_04_10F	aggaaagaggGTTTTATTAGGGTGGGGTATT					
DMR_04_T7R	cgttaatacgactcaatagggagaaggctAACTCTAACCTCTACTAACACACC	56	BCL9L	480	chr11:118910355-118910834	chr11:118910501-118910750
DMR_05_10F	aggaaagaggGTGGGTTGGTGTAGGGTATT	56	ZFPM1	374	chr16:88478366-88478739	chr16:88478501-88478750
DMR_05_T7R	cgttaatacgactcaatagggagaaggctAACTAAACCTCCTCATCCCTCAA					
DMR_10_10F	aggaaagaggTTTTAGTTGTGTAGGTAGGG	56	SEPTIN9	422	chr17:77435444-77435865	chr17:77435501-77435750
DMR_10_T7R	cgttaatacgactcaatagggagaaggctAACTCAAAAACAAAACCTCCAAT					
DMR_11_10F	aggaaagaggGATTTTGTAAATTGTTGGT	60	HTRA3	290	chr4:8270104-8270393	chr4:8270001-8270250
DMR_11_T7R	cgttaatacgactcaatagggagaaggctAAAAAAACCTCACTCCCTCAA					
DMR_12_10F	aggaaagaggGTTTTAAAGTTGGTGAAGGG	56	SGCG	391	chr13:23324392-23324782	chr13:23324501-23324750
DMR_12_T7R	cgttaatacgactcaatagggagaaggctTCCAATAAAAAACACATTACTTCC					
DMR_13_10F	aggaaagaggAGAGTTGGGGTTTAAATTAGTTTT	56	SCAMP4; ADAT3	465	chr19:190389-1909853	chr19:1909501-1909750
DMR_13_T7R	cgttaatacgactcaatagggagaaggctTCCTTAAACCACATTCTACAAC					
DMR_14_10F	aggaaagaggAAAGAGTAGGGTTAGGAGAGTTG	56	RHOBTB2	344	chr8:23003905-23004248	chr8:23004001-23004250
DMR_14_T7R	cgttaatacgactcaatagggagaaggctCAAAAACCCATTCAAACAAAAAA					
DMR_15_10F	aggaaagaggTATGTGTTAGTGTAAGATTGTTG	60	CELSR1	218	chr22:46420444-46420661	chr22:46420251-46420500
DMR_15_T7R	cgttaatacgactcaatagggagaaggctACATACTACCCAAATACACTCACC					
DMR_20_10F	aggaaagaggTTTTGTTAGGTTGGAGTAGT	60	TK2	413	chr16:66526917-66527329	chr16:66527001-66527250
DMR_20_T7R	cgttaatacgactcaatagggagaaggctCAAAATAAAACCTCTACCCAAAT					
DMR_21_10F	aggaaagaggAGATTTTGTGGTTATTGTT	56	HPCAL1	433	chr2:10325839-10326271	chr2:10326001-10326250
DMR_21_T7R	cgttaatacgactcaatagggagaaggctAAACAAAAAACTCTCCATCTACC					
DMR_22_10F	aggaaagaggTTTTAGGAGGATAGGGTAGGGTTA	56	IL17C	198	chr16:88639754-88639951	chr16:88639751-88640000
DMR_22_T7R	cgttaatacgactcaatagggagaaggctACAAAACCCAAATATAAAAACC					
DMR_24_10F	aggaaagaggTTAGTTTGTGTTATTGAGGGAGG	56	BLK	474	chr8:11517930-11518403	chr8:11518001-11518250
DMR_24_T7R	cgttaatacgactcaatagggagaaggctTCAAAACCCAATATCTAAC					
DMR_25_10F	aggaaagaggGTGTGGGATTAGGTTAGGGTTA	60	BIN3	353	chr8:22627450-22627802	chr8:22627501-22627750
DMR_25_T7R	cgttaatacgactcaatagggagaaggctCTATCTAACCTACCCCCACCTAAC					

DMR	sample	SF3B1	MeDIP beta	EpiTYPER mean	CpG position in amplicon		
					404		
DMR_24	19	WT	0.82	0.88	0.88		
DMR_24	21	WT	0.76	0.78	0.78		
DMR_24	28	WT	0.77	0.78	0.78		
DMR_24	30	WT	0.73	0.73	0.73		
DMR_24	31	WT	0.44	0.38	0.38		
DMR_24	36	WT	0.75	0.52	0.52		
DMR_24	14	Mut	0.13	0.16	0.16		
DMR_24	17	Mut	0.33	0.11	0.11		
DMR_24	24	Mut	0.37	0.50	0.50		
DMR_24	25	Mut	0.16	NA	NA		
DMR_24	26	Mut	0.14	0.10	0.10		
DMR_24	33	Mut	0.34	0.07	0.07		

DMR	sample	SF3B1	MeDIP beta	EpiTYPER mean	CpG position in amplicon		
					75	336	
DMR_25	19	WT	0.83	0.39	0.78	0.00	
DMR_25	21	WT	0.83	0.45	0.82	0.07	
DMR_25	28	WT	0.78	0.15	0.29	0.00	
DMR_25	30	WT	0.67	0.10	0.15	0.04	
DMR_25	31	WT	0.74	0.47	0.93	0.00	
DMR_25	36	WT	NA	0.31	0.48	0.13	
DMR_25	14	Mut	0.48	0.32	0.55	0.09	
DMR_25	17	Mut	0.34	0.05	0.10	0.00	
DMR_25	24	Mut	0.26	0.12	0.22	0.02	
DMR_25	25	Mut	0.17	NA	NA	NA	
DMR_25	26	Mut	0.40	NA	NA	NA	
DMR_25	33	Mut	0.56	0.17	0.19	0.14	

Table S5. Validation of MeDIP-seq results by EpiTYPER MASSarray. Given are the mean EpiTYPER methylation values and the mean beta-values calculated from MeDIP-seq by QSEA. The six samples analysed by EpiTYPER are a subset of the samples used for MeDIP-seq.

DMR	EpiTYPER means		MeDIP beta means	
	SF3B1 WT (6 samples)	SF3B1 mut (6 samples)	SF3B1 WT (14 samples)	SF3B1 mut (13 samples)
DMR_1	0.21	0.09	0.48	0.11
DMR_2	0.57	0.27	0.84	0.46
DMR_3	0.76	0.39	0.79	0.40
DMR_4	0.28	0.11	0.40	0.10
DMR_5	0.20	0.02	0.44	0.06
DMR_10	0.49	0.32	0.46	0.13
DMR_11	0.07	0.04	0.28	0.06
DMR_12	0.32	0.06	0.46	0.11
DMR_13	0.50	0.30	0.51	0.15
DMR_14	0.33	0.04	0.39	0.07
DMR_15	0.35	0.10	0.88	0.59
DMR_20	0.33	0.03	0.62	0.22
DMR_21	0.53	0.25	0.71	0.26
DMR_22	0.22	0.03	0.34	0.04
DMR_24	0.68	0.19	0.71	0.26
DMR_25	0.31	0.16	0.77	0.41

Table S6. Motif enrichment analysis of the 67 hypomethylated regions in *SF3B1*^{mut} patients.

Rank	Motif Name	Motif	P-value	q-value (Benjamini)	# of Target Sequences with Motif (of 67)	% of Target Sequences with Motif	# of Background Sequences with Motif (of 42456)	% of Background Sequences with Motif
1	ZNF165 (Zf) WHIM12-ZNF165-ChIP-Seq (GSE65937)		1.00E-03	0.398	9	13.43%	1600.6	3.77%
2	BHLHA15 (bHLH) NIH3T3-BHLHB8.HA-ChIP-Seq (GSE119782)		1.00E-02	0.5857	25	37.31%	9228.3	21.73%
3	Foxf1 (Forkhead) Lung-Foxf1-ChIP-Seq (GSE77951)		1.00E-02	0.5857	8	11.94%	1610.1	3.79%
4	Tcf12 (bHLH) GM12878-Tcf12-ChIP-Seq (GSE32465)		1.00E-02	0.5857	23	34.33%	8411.3	19.81%
5	Olig2 (bHLH) Neuron-Olig2-ChIP-Seq (GSE30882)		1.00E-02	0.5857	25	37.31%	9571.4	22.54%
6	Foxo3 (Forkhead) U2OS-Foxo3-ChIP-Seq (E-MTAB-2701)		1.00E-02	0.5857	7	10.45%	1437.6	3.39%
7	Ascl1 (bHLH) NeuralTubes-Ascl1-ChIP-Seq (GSE55840)		1.00E-02	0.5857	31	46.27%	13364.1	31.47%
8	AMYB (HTH) Testes-AMYB-ChIP-Seq (GSE44588)		1.00E-02	0.5857	17	25.37%	5851.9	13.78%
9	FOXA1 (Forkhead) LNCAP-FOXA1-ChIP-Seq (GSE27824)		1.00E-02	0.5857	9	13.43%	2242.5	5.28%

Bibliography

1. Dietrich, S.; Oleś, M.; Lu, J.; Sellner, L.; Anders, S.; Velten, B.; Wu, B.; Hüllein, J.; da Silva Liberio, M.; Walther, T.; et al. Drug-perturbation-based stratification of blood cancer. *J. Clin. Invest.* **2018**, *128*, 427–445, doi:10.1172/JCI93801.
2. Lienhard, M.; Grasse, S.; Rolff, J.; Frese, S.; Schirmer, U.; Becker, M.; Börno, S.; Timmermann, B.; Chavez, L.; Sültmann, H.; et al. QSEA-modelling of genome-wide DNA methylation from sequencing enrichment experiments. *Nucleic Acids Res.* **2017**, doi:10.1093/nar/gkw1193.
3. Oakes, C.C.; Seifert, M.; Assenov, Y.; Gu, L.; Przekopowitz, M.; Ruppert, A.S.; Wang, Q.; Imbusch, C.D.; Serva, A.; Koser, S.D.; et al. DNA methylation dynamics during B cell maturation underlie a continuum of disease phenotypes in chronic lymphocytic leukemia. *Nat. Genet.* **2016**, doi:10.1038/ng.3488.
4. Binet, J.L.; Auquier, A.; Dighiero, G.; Chastang, C.; Piguet, H.; Goasguen, J.; Vaugier, G.; Potron, G.; Colona, P.; Oberling, F.; et al. A new prognostic classification of chronic lymphocytic leukemia derived from a multivariate survival analysis. *Cancer* **1981**, *48*, 198–206, doi:10.1002/1097-0142(19810701)48:1<198::aid-cncr2820480131>3.0.co;2-v.
5. Vollbrecht, C.; Mairinger, F.D.; Koitzsch, U.; Peifer, M.; Koenig, K.; Heukamp, L.C.; Crispatzu, G.; Wilden, L.; Kreuzer, K.A.; Hallek, M.; et al. Comprehensive analysis of disease-related genes in chronic lymphocytic leukemia by multiplex PCR-based next generation sequencing. *PLoS One* **2015**, doi:10.1371/journal.pone.0129544.