SUPPLEMENTAL MATERIAL

Seifert et al., http://www.jem.org/cgi/content/full/jem.20120833/DC1

Table S1, provided as an Excel file, shows the raw signal values of the HG U133 2.0 Plus hierarchical clustering analysis in Fig. 1.

Table S2. Pairwise comparison of normal B cell subsets and CLL

	CD5	Naive	IgM+IgD+CD27+	IgM-only	Class-switched	sMGZ
CLL	587	861	766	907	867	1,274
CD5		728	600	833	814	1,586
naive			645	713	627	551
IgM+IgD+CD27+				5	232	739
IgM-only					228	921
class-switched						757

Numbers of differentially expressed genes between normal B cell subsets and CLL, filtered by differential expression of at least twofold change and statistical significance ($P < 0.05$) according to two-sided T-testing procedure. mCLL and uCLL cases were pooled. sMGZ, splenic marginal zone B cells.

Table S3. VH1 and VH3 rearrangement analysis of human PB CD5+ and conventional B cell subsets

V gene family	Donor	Subset	Sequences	% mutated (range)	Average mutation	Average mutation	Clones with diversity/	average frequency of	# clones shared between	
									Batches[d]	Populations
VH1	1	CD5+CD27−	21	14	0.08	0	3/4	0.18	4	
		CD5+IgM+CD27+	21	76	3.2	3.2	2/3	0.37	0	
		CD5+IgG/IgA+	20	75	1.6	1.2	4/4	0.69	0	
VH1	2	CD5+CD27−	21	0	0	0	0/3	0	0	
		CD5+IgM+CD27+	19	95	2.3	3.3	1/1	0.13	1	2
		CD5+IgG/IgA+	24	75	2.2	1.6	1/3	0.03	2	2
VH1	3	CD5+CD27−	20	0	0	0	0/5	0	0	
		CD5+IgM+CD27+	27	85	2.8	2.7	4/5	0.62	0	
		CD5+IgG/IgA+	27	96	0.5	0.6	1/1	0.74	0	
		CD5− unmutated	14	0	0	0	0	0	0	
		CD5- mutated	11	100	2.2	2.2	0	0	0	
VH1	4	CD5+CD27−	26	0	0	0	0/5	0	5	
		CD5+IgM+CD27+	27	74	2.0	1.9	3/6	0.67	1	
		CD5+IgG/IgA+	13	100	3.8	4.6	2/2	1.06	0	
		CD5− unmutated	18	0	0	0	0	0	0	
		CD5- mutated	17	100	2.9	2.9	0	0	0	
VH3	3	CD5+CD27−	28	7	0.06	0.04	0/4	0	1	
		CD5+IgM+CD27+	22	91	3.4	3.1	2/3	1.33	2	1
		CD5+IgG/IgA+	17	35	0.9	0.4	1/3	0.54	3	1
		CD5− unmutated	19	0	0	0	0	0	0	
		CD5- mutated	12	100	3.5	3.5	0	0	0	
VH3	4	CD5+CD27−	32	9	0.1	0.05	2/10	0.15	5	
		CD5+IgM+CD27+	29	100	2.5	2.5	2/2	4.54	0	
		CD5+IgG/IgA+	25	100	4.0	3.9	2/2	1.26	2	
		CD5− unmutated	22	0	0	0	0	0	0	
		CD5− mutated	12	100	1.9	1.9	0	0	0	
average		CD5+CD27−	148	5 (0–14)	0.04	0.02	5/31	0.06	15	
		CD5+IgM+CD27+	145	87 (76–95)	2.7	2.78	14/20	1.28	4	3
		CD5+IgG/IgA+	126	82 (72–100)	2.17	2.05	11/15	0.72	7	3
		CD5− unmutated[e]	73	0	0	0	0	0	0	
		CD5− mutated[e]	52	100	2.63	2.63	0	0	0	
		total							26	3

[a]For calculation of the average mutation frequency, identical sequences were counted once, as they might derive from one cell. In case of intraclonal diversity, each unique sequence was regarded as derived from an independent cell and counted once. Both in-frame and out-of-frame rearrangements were considered.
[b]For calculation of the average mutation frequency, each sequence was counted.
[c]Calculated as the average number of unique but not shared mutations within clonally related sequences, divided by the number of clonally related sequences.
[d]The term Batches denotes independently processed aliquots in PCR and cloning.
[e]Conventional B cells were isolated and analyzed as CD19+CD5− B cells and afterward the V gene sequences obtained from these cells were separated into mutated and unmutated sequences.

Table S4. Bcl6 mutation analysis of human PB CD5+ B cell subsets

Donor	cell type	sequences	Mutated seq. (%)	number of mutations [a]	mutation frequency
1	$CD5^+CD27^-$	20	7 (28)	6 × 1, 1 × 2	0.05
1	$CD5^+IgM^+CD27^+$	20	9 (45)	3 × 1, 3 × 2, 1 × 3, 1 × 4, 1 × 5	0.12
1	$CD5^+IgG/IgA^+$	23	10 (43)	7 × 1, 1 × 2, 2 × 3	0.09
2	$CD5^+CD27^-$	26	5 (19)	5 × 1	0.02
2	$CD5^+IgM^+CD27^+$	30	8 (27)	5 × 1, 3 × 2	0.05
2	$CD5^+IgG/IgA^+$	23 [b]	23 (100)	4 × 1, 4 × 2, 2 × 3, 2 × 4, 1 × 5	0.32 [b]
3	$CD5^+CD27^-$	24	2 (8)	2 × 1	0.01
3	$CD5^+IgM^+CD27^+$	24	7 (29)	5 × 1, 2 × 2	0.05
3	$CD5^+IgG/IgA^+$	12	8 (67)	5 × 1, 1 × 2, 1 × 3, 1 × 4	0.16
4	$CD5^+CD27^-$	18	2 (11)	2 × 1	0.01
4	$CD5^+IgM^+CD27^+$	19	9 (47)	5 × 1, 2 × 2, 1 × 4, 1 × 6	0.1
4	$CD5^+IgG/IgA^+$	10	4 (40)	1 × 1, 3 × 2	0.09
4	T cells[c]	20	3 (15)	3 × 1	0.02
5	T cells[c]	14	3 (21)	2 × 1, 1 × 2	0.04
6	T cells[c]	16	5 (31)	4 × 1, 1 × 2	0.05
average	$CD5^+CD27^-$	88	16 (18)	15 × 1, 1 × 2	0.03
	$CD5^+IgM^+CD27^+$	93	33 (35)	17 × 1, 11 × 2, 1 × 3, 2 × 4, 1 × 5, 1 × 6	0.11
	$CD5^+IgG/IgA^+$	68	44 (65)	17 × 1, 9 × 2, 5 × 3, 3 × 4, 1 × 5	0.22
	T cells [c]	50	11 (22)	9 × 1, 2 × 2	0.04

[a]Identically mutated sequences counted as one. Indicated are number of sequences with number of mutations per sequence.
[b]In this population, several sequences harbored shared mutations, leading to a total number of 13 unique sequences when identical sequences were counted as one.
[c]T cells were analyzed to determine the experimental background of mutation.

Table S5, provided as an Excel file, shows the raw signal values of the Gene ST array hierarchical clustering analysis in Fig. 4 a.

Table S6, provided as an Excel file, shows the raw signal values from the PCA of $CD5^+$ B cell subsets and 107 CLL samples in Fig. 4 c.

Table S7, provided as an Excel file, shows the IGHV rearrangements assigned to stereotyped receptor families in Table 2.

Table S8. Stereotyped BCR in human PB CD5+, CD43+, and conventional B cells, as determined by PCR for selected IGHV gene segments

PCR	Donor	Sample [a]	V genes obtained	V genes included in analysis [b]	Sequences mutated (mut. freq.) [c]	Unique sequences[d]	Stereotyped receptors[d]	
VH1-69	1	conventional	24	22	10 (3.2)	20	1	
		CD5+	20	16	5 (1.0)	13	2	
		CD43+	24	24	0	0	0	
VH1-69	2	conventional	22	22	14 (3.3)	12	0	
		CD5+	26	26	6 (2.6)	22	2	
		CD43+	21	21	21 (12.6)	4	0	
VH4-34	1	conventional	23	19	7 (2.4)	19	2	
		CD5+	14	11	9 (1.2)	11	1	
		CD43+	0	0	0	0	0	
VH4-34	2	conventional	19	13	10 (3.7)	12	1	
		CD5+	19	14	8 (1.8)	12	3	
		CD43+	0	0	0	0	0	
VH3-23	1	conventional	21	20	14 (3.6)	20	0	
		CD5+	28	24	18 (2.8)	24	1	
		CD43+	0	0	0	0	0	
VH3-23	2	conventional	23	15	10 (1.9)	14	1	
		CD5+	22	20	15 (1.9)	20	2	
		CD43+	22	20	20 (3.6)	5	0	
VH3-48	1	conventional	10	9	8 (1.2)	9	0	
		CD5+	10	9	4 (1.7)	9	0	
		CD43+	24	0	0	0	0	
VH3-48	2	conventional	5	5	1 (0.4)	4	0	
		CD5+	7	5	5 (1.1)	5	1	
		CD43+	7	7	7 (1.5)	3	0	
VH3-21	1	conventional	28	24	12 (1.5)	21	1	
		CD5+	16	16	7 (1.2)	15	4	
		CD43+	8	0	0	0	0	
VH3-21	2	conventional	21	20	11 (0.5)	10	0	
		CD5+	18	17	3 (0.6)	14	2	
		CD43+	10	0	0	0	0	
								Fisher's exact test [e]
Total		conventional	196	169	97 (2.2)	141	6	
		CD5+	180	158	80 (1.6)	145	18	P < 0.018
		CD43+	116	72	48	12	0	P < 1

[a]Conventional and CD5+ bulk B cell subsets were analyzed and not separated in terms of CD27 expression.
[b]Only in-frame rearrangements with correct V gene.
[c]The number of mutated in-frame rearrangements and their average mutation frequency (total number of mutations divided by number of nucleotides) is given.
[d]Identical sequences counted as one. When any sequence obtained is counted individually, the statistical significance is even below P < 0.001.
[e]P-values calculated by Fisher´s exact test versus conventional B cells. Only unique sequences were considered.

Table S9, provided as an Excel file, shows the transcripts differentially regulated between CD5+ and conventional naive and CD27+ B cells.

Table S10, provided as an Excel file, shows NF-κB target genes in B cells used in the GSEA from Fig. 6 i.

Table S11, provided as an Excel file, is a list of genes with similar transcription in CD5+ B cells and CLL from Fig. 5 b.

Table S12, provided as an Excel file, lists the raw signal values of the KLF factor expression heatmap in Fig. 5 d.

Table S13. KLF3 expression pattern in CD5+ B cells and CLL[a]

			Blinded study 1	Blinded study 2	
Donor	Cell type	Cells analyzed	% cells with nuclear KLF3 expression	% cells with nuclear KLF3 expression	Fisher's exact test (CD5+ vs. CLL)
1	CD5+	110	31	45	
2	CD5+	109	26	37	
mean[b]			28	40	P < 0.001
1	CLL	98	1	5	
2	CLL	96	0	3	
mean[b]			1	5	P < 0.001

[a]Evaluation was done in a blinded fashion.
[b]Mean values are calculated as weighted mean

Table S14, provided as an Excel file, is a list of genes with differential expression between CD5+ B cells and CLL from Fig. 5 e.

Table S15. Patient characteristics

Patient	Sex	Age	V gene status	Binet Stage	Treatment	TTFT (mo)	CD38	Lymphocyte count (per nl)	Initial diagnosis	I-FISH	Progression	HG U133 2.0 plus	HuGene-1_0-st-v1
A	m	79	mutated	A	CLB	38	negative	49.9	Feb 2007	n.d.	stable	mCLL_1	mCLL_1
B	m	55	mutated	A	none		negative	49.9	Apr 2009	13q-	stable	mCLL_2	mCLL_2
C	f	73	mutated	A	none		negative	9.9	Dec 2008	13q-	stable	mCLL_3	mCLL_3
D	f	70	mutated	A	none		negative	75.1	Dec 2005	13q-	stable	mCLL_4	mCLL_4
E	m	75	mutated	A	none		negative	74.3	June 2005	13q-	stable	mCLL_5	mCLL_5
F	f	54	unmutated	A	none		positive	45	June 2009	n.d.	progressive	uCLL_1	uCLL_1
G	f	58	unmutated	A	none		negative	20.7	May 2007	13q-	stable	uCLL_2	uCLL_2
H	f	69	unmutated	A	R-Bendamustine	39	positive	18.2	June 2006	11q-	progressive	uCLL_3	uCLL_3
I	m	72	unmutated	B	FC, Ofatumumab	45	positive	90.1	Nov 1997	11q-, 6q-13q-	progressive	uCLL_4	
J	m	58	unmutated	C	FC-Ofatumumab	22	negative	17.4	Jan 2006	13q-	progressive	uCLL_5	uCLL_4
K	m	55	unmutated	C	CLB, R-CHOP, R-FC	41	positive	12	Nov 2000	13q-	progressive		uCLL_5

TTFT, time to first treatment; mo, months; I-FISH, interphase fluorescence in situ hybridization; CLB, chlorambucil; R, rituximab; FC, fludarabine and cyclophosphamide; CHOP, cyclophosphamide, doxorubicin, vincristine, and prednisone.