Cellular origin and pathophysiology of chronic lymphocytic leukemia

Marc Seifert,[1] Ludger Sellmann,[1,2] Johannes Bloehdorn,[3] Frederik Wein,[1] Stephan Stilgenbauer,[3] Jan Dürig,[2] and Ralf Küppers[1]

[1]Institute of Cell Biology (Cancer Research) and [2]Department of Hematology, Medical Faculty, University of Duisburg-Essen, Essen 45122, Germany
[3]Department of Internal Medicine III, University of Ulm, Ulm 89081, Germany

The cellular origin of chronic lymphocytic leukemia (CLL) is still debated, although this information is critical to understanding its pathogenesis. Transcriptome analyses of CLL and the main normal B cell subsets from human blood and spleen revealed that immunoglobulin variable region (IgV) gene unmutated CLL derives from unmutated mature CD5+ B cells and mutated CLL derives from a distinct, previously unrecognized CD5+CD27+ post–germinal center B cell subset. Stereotyped V gene rearrangements are enriched among CD5+ B cells, providing independent evidence for a CD5+ B cell derivation of CLL. Notably, these CD5+ B cell populations include oligoclonal expansions already found in young healthy adults, putatively representing an early phase in CLL development before the CLL precursor lesion monoclonal B cell lymphocytosis. Finally, we identified deregulated proteins, including EBF1 and KLF transcription factors, that were not detected in previous comparisons of CLL and conventional B cells.

CORRESPONDENCE
Ralf Küppers:
ralf.kueppers@uk-essen.de

Abbreviations used: BCR, B cell receptor; CDR, complementarity–determining region; CLL, chronic lymphocytic leukemia; GC, germinal center; GEP, gene expression profiling; Igκ, immunoglobulin κ chain; IgL, Ig light chain; IgV, Ig variable region; PB, peripheral blood; PCA, principal component analysis; sMGZ, splenic marginal zone

Chronic lymphocytic leukemia (CLL) is the most frequent B cell leukemia in elderly patients (Zenz et al., 2010). Approximately half of the cases of CLL carry unmutated Ig variable region (IgV) genes (uCLL), and the remaining cases have somatically mutated IgV genes (mCLL; Damle et al., 1999; Hamblin et al., 1999). This distinction is of biological interest and clinical relevance because uCLL is more aggressive with a significantly shorter time to first treatment (Rassenti et al., 2008). The identification of the cellular origin of CLL is essential to elucidating the pathobiology of a tumor. Only then can the full natural history of the disease be revealed and the dysregulation of gene expression and cellular functions be appreciated (Küppers et al., 1999). For CLL, the consistent expression of CD5 led to initial speculations that CLL might be a malignancy of CD5+ B cells (Caligaris-Cappio et al., 1982; Caligaris-Cappio, 1996), which, in mouse, represents a distinct B cell lineage (B1 B cells; Dorshkind and Montecino-Rodriguez, 2007). However, functional similarities between CLL and splenic marginal zone (sMGZ) B cells led to a proposal that CLL might be derived from such B cells (Chiorazzi and Ferrarini, 2011). Based on a study of specific IgV gene rearrangements, a derivation of uCLL from conventional

naive B cells was proposed (Forconi et al., 2010). About 10 yr ago, detailed gene expression profiling (GEP) of CLL and normal human B cell subsets surprisingly indicated that mCLL and uCLL are similar to memory B cells, but not CD5+ B cells (Klein et al., 2001), indicating that both CLL subsets originate from antigen–experienced B cells (Klein et al., 2001; Rosenwald et al., 2001). This is supported by the finding that ~30% of CLL cases show highly similar IgV genes, which have been grouped into >150 sets of stereotyped receptors (Stamatopoulos et al., 2007; Murray et al., 2008). This strongly suggests that such CLL recognized the same antigens, and hence B cell receptor (BCR) specificity plays a role in CLL pathogenesis.

However, regarding the previous GEP studies (Klein et al., 2001; Rosenwald et al., 2001), there are several caveats. First, none of these studies included sMGZ B cells. Second, in the previous most comprehensive gene expression study of CLL and normal B cells,

The Rockefeller University Press $30.00
J. Exp. Med. 2012 Vol. 209 No. 12 2183–2198
www.jem.org/cgi/doi/10.1084/jem.20120833

2183

memory B cells were isolated as bulk CD27[+] B cells (Klein et al., 2001). However, approximately half of CD27[+] B cells are class-switched, and the remaining cells are mostly IgM[+]IgD[+]CD27[+] B cells (Klein et al., 1998), and few are IgM-only B cells (IgD[low/−]). Importantly, the generation of IgM[+]IgD[+]CD27[+] B cells in germinal center (GC) responses or alternative pathways is discussed (Klein et al., 1998; Kruetzmann et al., 2003; Seifert and Küppers, 2009; Weill et al., 2009). Third, in the previous study including CD5[+] B cells, these were isolated from cord blood, in which practically all B cells are CD5[+] (Klein et al., 2001). However, it was recently reported that a fraction of human peripheral blood (PB) B cells are transitional, but not mature B cells, and that these cells are CD5[+] (Sims et al., 2005). Importantly, at birth the majority of CD5[+] B cells are transitional B cells (Ha et al., 2008; Marie-Cardine et al., 2008; Sims et al., 2005). Hence, in the previous GEP study, mostly transitional B cells and not mature CD5[+] B cells were compared with CLL. Because of these restrictions, we performed a new GEP study of CLL in comparison to normal naive, sMGZ, mature CD5[+] and class-switched cells, as well as IgM[+] memory B cells.

Additionally, we performed an IgV gene analysis from CD5[+] and CD5[−] B cells, to search for the normal B cell subset in which CLL-typical stereotyped BCR can be found. Both independent studies revealed that mCLL and uCLL cells are most closely related to mature CD5[+] B cells. Thus, we conclude that CLL is a malignancy of CD5[+] B cells. Moreover, we identified a small subpopulation of CD5[+] B cells expressing CD27 and carrying somatically mutated IgV genes. These putative post-GC B cells may represent the physiological counterpart of mCLL.

RESULTS

Human naive and CD5[+] B cells show a gene expression pattern highly similar to CLL

For a comprehensive analysis of differential gene expression between CLL and normal human B cells, we isolated PB naive B cells, memory B cell subsets (class-switched, IgM[+]IgD[+]CD27[+], and IgM-only B cells), CD5[+] B cells (excluding transitional B cells), and sMGZ B cells. Global RNA expression from five to seven samples each was analyzed using Affymetrix HGU133 2.0 Plus arrays.

Between 500 and 3,000 transcripts selected according to highest SD were chosen for hierarchical clustering analyses. Fig. 1 shows a representative dendrogram of the 46 samples, based on 2,000 transcripts (Table S1, U133 HC). With the exception of IgM[+]IgD[+]CD27[+] and IgM-only B cells (both referred to as IgM memory B cells), each B cell subset forms a separate branch, supporting the identification of consistent, subset-specific expression patterns.

A separation of the 46 profiles into two major branches is evident. The first major group includes naive and CD5[+] B cells and CLL, the second group consists of the three memory B cell subsets, as well as sMGZ B cells. Thus, the normal B cells arrange according to their IgV gene mutation status and relation
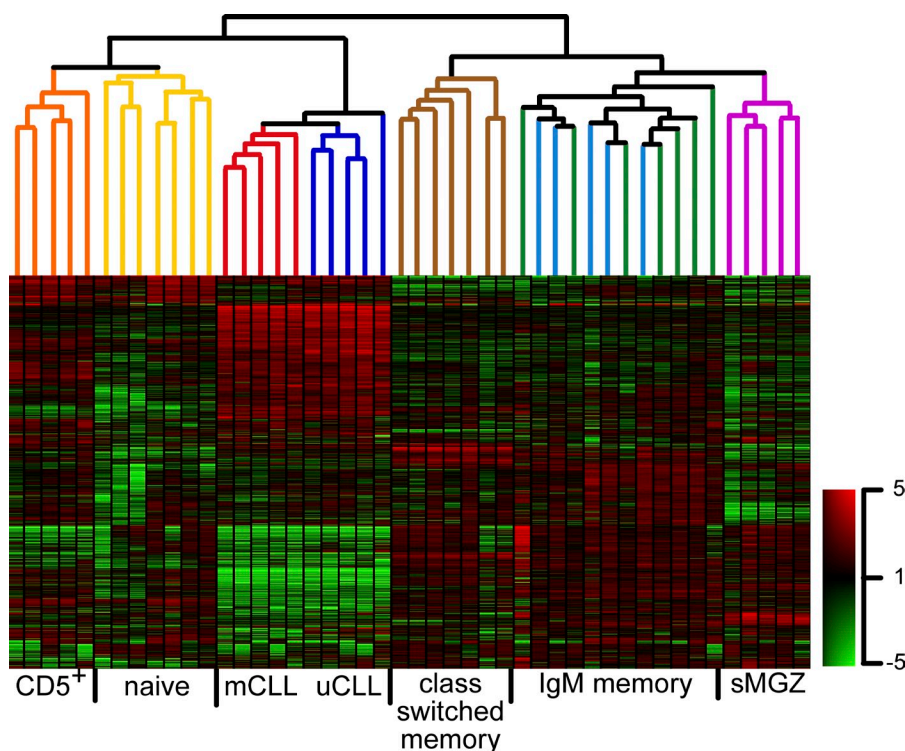

Figure 1. Hierarchical clustering of normal human B cell subsets and CLL. The dendrogram is based on a Spearman ranking of 2000 transcripts with highest SD (Table S1). Subcluster stability was confirmed by bootstrapping procedure (>70%). The color bar depicts normalized intensity values. CD5[+], CD5[+]CD27[−]CD38[low] B cells; naive, conventional naive B cells; mCLL, IgV mutated CLL; uCLL, IgV unmutated CLL; class-switched, IgG[+]CD27[+] and IgA[+]CD27[+] B cells; IgM[+]IgD[+] memory, IgM[+]IgD[+]CD27[+] and IgM-only B cells; sMGZ, splenic marginal zone B cells.

to GC experience (the majority of CD5$^+$ B cells carry unmutated V genes; Brezinschek et al., 1997; Fischer et al., 1997; Dono et al., 2007), although the latter is controversially discussed for sMGZ B cells (Weller et al., 2004). Importantly, above a dominant gene expression signature shared by the CLL subsets, both cluster in the branch of naive and CD5$^+$ B cells, indicating more similar gene expression to these cells than to PB memory B cells or sMGZ B cells.

Principal component analysis (PCA) reveals a strong similarity between CLL and CD5$^+$ but not conventional B cell subsets

In addition to agglomerative hierarchical clustering, we calculated the similarity of transcriptional profiles using PCA. In Fig. 2 a, the distance of individual samples is depicted according to the first and second principal components, explaining 35.2% of the total variance. CD5$^+$ B cells and CLL share a high degree of similarity in their respective expression patterns (Fig. 2 a).

To dissect the relationship of CLL to normal B cell subsets in more detail, we performed PCA based on gene lists derived from pairwise comparisons (Fig. 2, b–f). Relying on the indication of a high similarity of CLL to memory B cells (Klein et al., 2001), we tested whether CLL is more similar to either class-switched or IgM$^+$IgD$^+$CD27$^+$ memory B cells. In this PCA, CLL turned out to be considerably more similar to the IgM memory B cells (Fig. 2 b). However, in line with the hierarchical clustering analysis (Fig. 1), Fig. 2 c suggests a higher similarity of CLL and naive, but not IgM memory B cells. This contradicts a previous publication (Klein et al., 2001), although it should be mentioned that this study included tonsillar B cells.

Importantly, when CD5$^+$ B cells were compared with naive, or any of the CD27$^+$ B cell subsets, mCLL and uCLL were significantly more similar to CD5$^+$ B cells than to the other subsets (Fig. 2 d–f and unpublished data). Notably, this high similarity is already reflected at the level of simple pairwise comparisons, where the amount of differentially expressed genes is minimal between CLL and CD5$^+$ B cells (Table S2). Thus, among the five major B cell subsets included in the analysis, mature CD5$^+$ B cells show the highest concordance in gene expression to CLL, indicating a derivation of CLL from these cells but not from memory or sMGZ B cells.

Genes differentially expressed between CLL and other B cell lymphomas reveal a CLL signature predominantly in normal CD5$^+$ B cells

So far, we identified specific gene signatures of normal B cell subsets to reveal their similarity to CLL gene expression patterns. In a complementary approach, we identified "CLL-specific" genes by comparing transcription profiles of CLL to other B cell lymphomas, and then analyzed which normal B cell subset is most similar to CLL for these genes. We generated a list of 27 annotated transcripts that are differentially expressed between CLL and two other major types of mature B cell lymphomas, i.e., follicular and diffuse large B cell
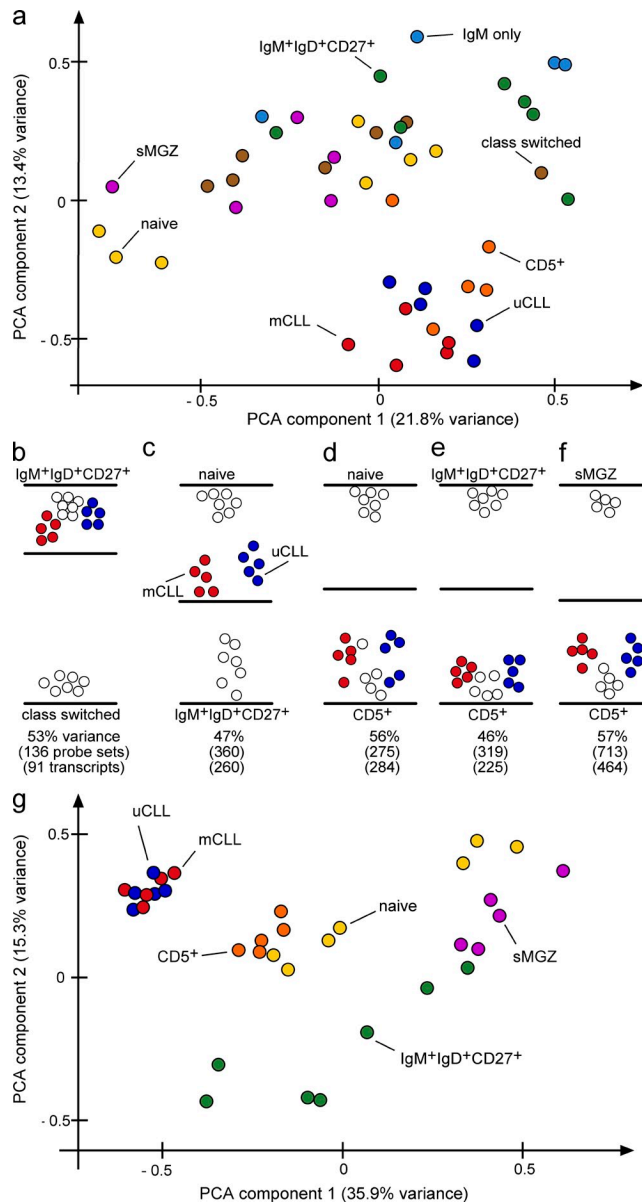

Figure 2. PCA of CD5$^+$ and CD5$^-$ conventional B cell subsets and CLL. (a) Unsupervised PCA shows a high similarity of CLL to CD5$^+$ B cells, but not conventional B cells. The PCA is based on 10,395 annotated transcripts, explaining >35% of total variance. Axis scaling according to mean centering and scaling. Samples belonging to distinct subsets are depicted in the same color. (b–f) Supervised PCA showing mathematical distances of mCLL (red) and uCLL (blue) according to the first principal component of pairwise compared normal B cell subsets (annotated transcripts, >two-fold change; P < 0.05; FDR < 0.05). Subset samples are depicted on top or bottom. Black bars represent maximum distances and baseline according to mean centering and scaling. (g) Supervised PCA of CLL and selected normal B cell samples. Shown are mathematical distances of CLL cases and CD5$^+$ and conventional B cell subsets according to 27 annotated transcripts differentially expressed between CLL and follicular lymphoma or diffuse large B cell lymphoma. Axis scaling according to mean centering and scaling.

lymphomas, including the genes *CTLA4*, *CD200*, and *BCL2* (Rosenwald et al., 2001). The PCA displayed CLL in closer proximity to CD5$^+$ B cells than to conventional naive, IgM memory, and sMGZ B lymphocytes (Fig. 2 g). Similar results were obtained by applying a list of 19 transcription factors with specific expression in CLL, but not in diffuse large B cell lymphoma, hairy cell leukemia, mantle cell lymphoma or follicular lymphoma (Andréasson et al., 2010; unpublished data). This further argues for a close relationship of normal CD5$^+$ B cells and CLL, as CLL-specific genes include typical CD5$^+$ B cell genes.

CD5$^+$ PB B cells are clonally expanded and include a small post–GC B cell population

The indication that CD5$^+$ B cells are CLL precursors involves an important caveat: most, if not all, CD5$^+$ B cells are regarded as pre-GC lymphocytes with unmutated IgV genes (Brezinschek et al., 1997; Fischer et al., 1997), but approximately half of CLL harbor mutated IgV genes. Hence, the question arises whether a distinct subset of CD5$^+$ B cells with mutated IgV genes exists, that may be the specific precursor of mCLL. Indeed, there is a small fraction of CD5$^+$ B cells (4–17%) that coexpress the memory B cell marker CD27 (Klein et al., 1998; Fig. 3). Moreover, a small subpopulation

Figure 3. Flow cytometric analysis of CD5$^+$ human PB B cells of a healthy donor. Depicted are FACS plots of CD19-enriched human lymphocytes stained for CD5, CD27, IgG, and IgA expression. (top left) CD19$^+$ B cells contain a small fraction of CD5$^+$CD27$^+$ B cells. CD5^{high}CD27^{high} events represent residual T cells, as verified by CD3 staining (not depicted), and therefore were excluded by gating procedure. (top right and bottom left) CD5$^+$CD27$^+$ B cells contain minor populations of class-switched IgG$^+$ or IgA$^+$ B cells. These stainings are specific, as there are no IgA$^+$IgG$^+$ double-positive B cells detectable (bottom right). Data are representative of 10 healthy donors.

of these cells (between 0.5 and 2% of CD5$^+$ PB B cells) is class-switched to IgG or IgA (Fig. 3). An average fraction of 18.5% (7.5–42%) of CD5$^+$CD27$^+$ cells expressed CD43 (unpublished data), which has been proposed as a marker for B1 B cells (Griffin et al., 2011).

To clarify whether CD27 expression marks somatically mutated CD5$^+$ B cells, we analyzed V$_H$1 and V$_H$3 family gene rearrangements from CD5$^+$ B cell subsets of four healthy donors. As expected for a pre-GC B cell population, CD5$^+$CD27$^-$ B lymphocytes were consistently unmutated (Table 1 and Table S3). Importantly, between 76 and 95% of V$_H$ genes from CD5$^+$CD27$^+$IgM$^+$ and 72 and 100% from CD5$^+$CD27$^+$ class-switched B cells were mutated with average mutation frequencies of 2.7 and 2.2%, respectively. That is in the range of 2–4%, typical for conventional IgM memory B cells (Klein et al., 1997; Klein et al., 1998). Thus, we have identified a distinct subset of IgV gene mutated human PB CD5$^+$ B cells, characterized by expression of CD27.

Surprisingly, the V$_H$ sequences of the CD5$^+$ B cell analysis revealed a substantial fraction of clonally related sequences within CD5$^+$CD27$^-$, CD5$^+$CD27$^+$IgM$^+$, and CD5$^+$CD27$^+$ class-switched B cells, where 78%, 51%, and 92% of the sequences were assigned to 31, 20, and 15 independent clones, respectively (Table 1 and Table S3). For several reasons, most of these related sequences represent expanded B cell clones. First, clonally related sequences were for 20–48% of the clones detectable within independently processed duplicate aliquots of 10,000 cells from the same CD5$^+$ B cell population of a given donor (Table S3). Second, expanded clones were never detected in CD5$^-$ B cells, which were processed in parallel to the CD5$^+$ B cells (Table 1). Third, the sporadically observed intraclonal diversity among sequences derived from CD5$^+$CD27$^+$ B cells is only compatible with a derivation from distinct cells (Table S3).

Whether somatic hypermutation takes place only in GC B cells is the subject of ongoing discussion (Kruetzmann et al., 2003; Seifert and Küppers, 2009; Weill et al., 2009). We sought to validate that mutated CD5$^+$ B cells are post-GC B cells. *BCL6* mutations are a genetic trait of B cells mutating in the GC and are found in 20–30% of conventional post-GC memory B cells (Pasqualucci et al., 1998; Seifert and Küppers, 2009). Importantly, somatic hypermutation is strictly dependent on target gene transcription (Fukita et al., 1998; Bachl et al., 2001; Yang et al., 2006). Thus, *Bcl6* mutations can only occur in B cells when these cells acquire mutations in a GC reaction, i.e., when *Bcl6* is strongly transcribed (Klein and Dalla-Favera, 2008). There is a low level of *Bcl6* transcription in conventional naive B cells (Ye et al., 1993), but this is reduced when the cells undergo immune responses outside of the GC (Allman et al., 1996; Marshall et al., 2011). Thus, *Bcl6* cannot be targeted by somatic hypermutation in extrafollicular responses. *Bcl6* is transcribed at higher levels for a short time in pre–B cells (Nahar et al., 2011). However, these pre–B cells lack substantial expression of AID (Sitte et al., 2012), and although low-level AID expression has been assigned to another (immature) precursor B cell population, the AID expression level is considerably lower than in GC
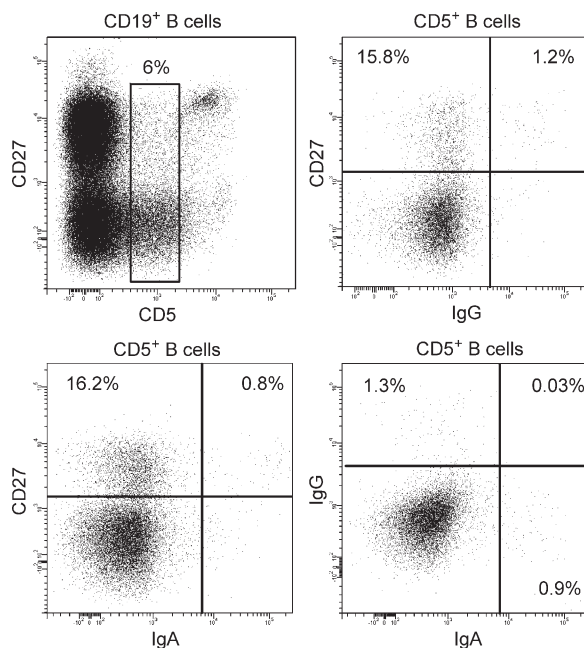
Table 1. VH1 and VH3 gene analysis of human PB CD5[+] and conventional B cell subsets

Subset[a]	Sequences	% mutated (range)	Average mutation frequency[b]	Sequences assigned to clones	Number of clones
CD5[+]CD27[−]	148	5 (0–14)	0.04	115	31
CD5[+]IgM[+]CD27[+]	145	87 (76–95)	2.7	74	20
CD5[+]IgG/IgA[+]	126	82 (72–100)	2.17	116	15
CD5[−] unmutated [c]	73	0	0	0	0
CD5[−] mutated [c]	52	100	2.63	0	0

[a]For CD5[+] and conventional subsets total numbers of four and two healthy donors are given, respectively.
[b]For calculation of the average mutation frequency, identical sequences were counted once, as they might derive from one cell. If all sequences were considered, very similar values were obtained (Table S3). 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.
[c]CD5[−], conventional B cells were isolated and analyzed as CD19[+] B cells and, afterward, separated into mutated and unmutated sequences.

B cells (Meyers et al., 2011) and the effective mutation load of the corresponding Ig genes is negligible (Kuraoka et al., 2011). Hence, an extra-GC derivation of mutations in *Bcl6*, the master regulator of the GC B cell differentiation program (Klein and Dalla-Favera, 2008), can be neglected, and the detection of such mutations is a strong argument for a GC experience of the respective B cells.

BCL6 mutations were indeed detected in CD5[+]CD27[+] B cell subsets in a frequency similar to that found in conventional memory B cells (Table S4). Collectively, we identified CD5[+]CD27[+] B cells as a distinct subset of PB CD5[+] B cells

with somatically mutated V genes, which are GC experienced and partly class-switched. Remarkably, we observed clonal expansion among each of the CD5[+] B cell subsets.

Derivation of uCLL from CD5[+]CD27[−] and mCLL from CD5[+]IgM[+]CD27[+] B cells

Because CD5[+] B cells consist of pre- and post-GC lymphocytes, we aimed to clarify whether uCLL and mCLL derive from these normal B cell populations, respectively. We performed an additional GEP analysis of CD5[+]IgM[+]CD27[+] and CD5[+]CD27[−] B cells, depleted for transitional and CD43[+]
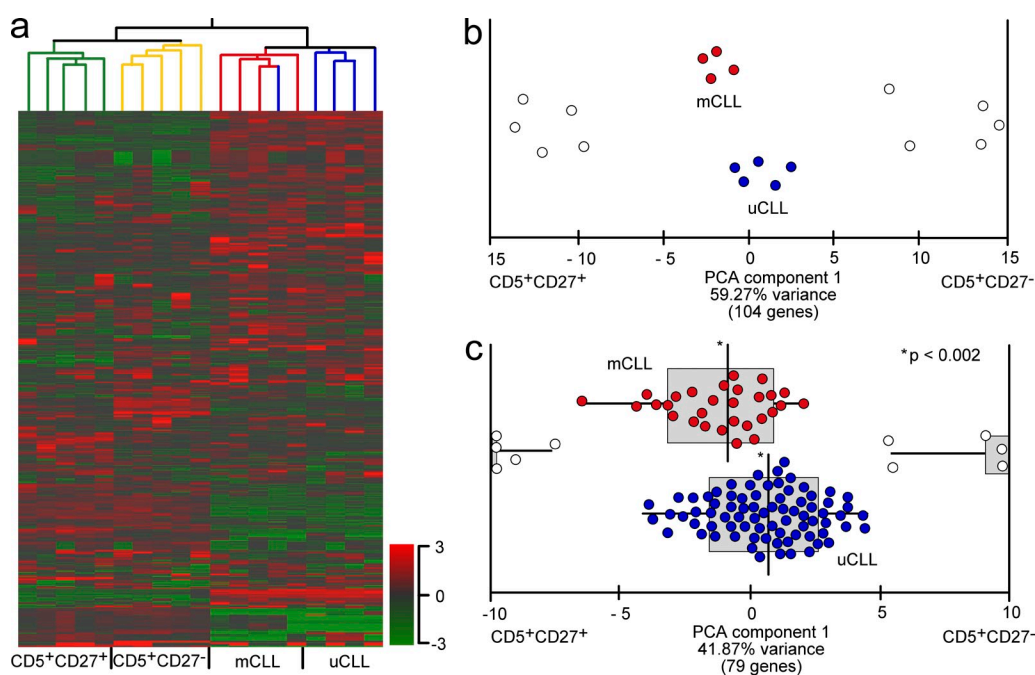

Figure 4. Unsupervised hierarchical clustering and supervised PCA of normal mature PB CD5[+] B cell subsets and CLL samples. (a) Highly purified CD5[+]CD27[−]CD38[low] naive B cells (yellow), CD5[+]CD27[+]CD38[−]CD43[−] memory B cells (green), and mCLL (red) and uCLL (blue) were included in a HuGene-1_0-st-v1 GEP study. Hierarchical clustering is based on 500 transcripts with the highest SD, according to Manhattan clustering and average linkage method. This clustering is representative for dendrograms based on 250–3,500 transcripts. (b) The PCA is based on 104 genes (twofold change; P < 0.05; FDR< 0.05) differentially expressed between CD5[+]CD27[−] and CD5[+]CD27[+] B lymphocytes. mCLL (red) and uCLL (blue) are displayed along the first principal component, covering >59% of total variance. (c) PCA with 107 CLL cases from an independent exon expression study (Table S6). Only probe sets scored with "best match" in a HuGene-HuExon array comparison (Affymetrix) were considered, resulting in 79 genes present on both platforms. The abscissa depicts the Eigenvector values of the similarity matrix associated with the dataset according to mean centering (mean zero) and scaling (to unit SD). The distribution of the data points along the ordinate was chosen arbitrarily to display all data points separately.

B cells, from five healthy donors, and 5 mCLL and 4 uCLL (a fifth case was excluded because of technical failure) on HuGene-1_0-st-v1 arrays. Hierarchical clusterings based on 500–5,000 genes with highest SD showed that normal CD5+CD27− and CD5+CD27+ samples were stably arranged in one and CLL cases in another branch (Fig. 4 a and Table S5). Besides one uCLL arranging with mCLL, all cell types clustered separately. Thus, the distance tree mainly reflected differences between normal and CLL cells, as expected from previous publications (Klein et al., 2001; Rosenwald et al., 2001).

However, a PCA based on 104 genes significantly separating CD5+CD27− from CD5+CD27+ B lymphocytes revealed that mCLL tend to be more similar to CD5+CD27+ cells and uCLL more similar to CD5+CD27− B cells (Fig. 4 b). As expected, the transcriptionally very similar CLL samples were not clearly separated (see Discussion). Therefore, we sought to ascertain if the weak association of the CLL subtypes to either mutated or unmutated CD5+ B cells was statistically significant. A larger cohort of 107 CLL was subjected to PCA according to the same algorithm covering 41.9% of the total variance (Table S6). Importantly, although mCLL and uCLL samples were overlapping, the distributions of the two CLL subsets were significantly different (P < 0.002), and numerous uCLL and mCLL were positioned closer to CD5+CD27− and CD5+CD27+ B cells, respectively (Fig. 4 c).

Collectively, CD5+CD27− and CD5+CD27+ B cells show consistent differences in their gene expression pattern, and uCLL tend to be more similar to IgV unmutated CD5+ B cells, whereas mCLL show a higher similarity to post-GC CD5+ B cells.

CD5+CD27− and CD5+CD27+ B cells, but not CD43+ B cells, preferentially express stereotyped BCR

Up to 30% of CLL express stereotyped BCR. These are defined by highly homologous heavy and light chain complementarity–determining region 3 (CDR3) amino acid

sequences (>60% homology), as well as confined V-gene usage, and either presence or absence of somatic mutations (Messmer et al., 2004; Stamatopoulos et al., 2007). If mature CD5+ B cells are precursors of CLL, preferential expression of such stereotyped BCR in these cells can be postulated. We identified 12 stereotyped IgV genes among 160 unique sequences derived from CD5+ B cells (7.4%) and a single one among 107 sequences derived from conventional B cells (<1%) of four healthy donors (Table 2 and Table S7). This difference is statistically significant (P < 0.018).

To further validate these findings we designed a PCR specific for rearrangements using *IGHV* segments frequently used in stereotyped BCR and analyzed two additional individuals (Table S8). We included CD43+ B cells, as they were recently proposed as human B1 B cells with phenotypic similarity to CLL (Griffin et al., 2011). Again, the frequency of stereotyped rearrangements was significantly higher among CD5+ B cells as compared with CD5− B cells (18/145 versus 6/141, respectively; P < 0.018; Table 2 and Table S8). Importantly, among the stereotyped receptors from both approaches, eight of nine VH1-69 rearrangements were unmutated and four of seven VH3-21, four of four VH3-23, and five of seven VH4-34 rearrangements were mutated (unpublished data). This correlation of V_H gene usage and mutational status among stereotyped BCR is in line with published data (Murray et al., 2008). Furthermore, 19 out of the 30 stereotyped IGHV rearrangements detected among mature CD5+ B cells could be assigned to 6 of the 10 main CLL categories (Stamatopoulos et al., 2007; Table S7). However, three stereotyped IGHV rearrangements from class-switched CD5+ B cells did not belong to typically class-switched CLL stereotypes (unpublished data).

The CD43+ B cells frequently showed identical rearrangements, indicating that these cells rarely use the *IGHV* segments tested, and/or encompass expanded B cell clones. None of the 72 *IGHV* region genes from CD43+ B cells included in

Table 2. Stereotyped BCR in human PB conventional, CD5+, and CD43+ B cells

Donors analyzed	Sample	V genes obtained	V genes included in analysis[a]	Unique sequences[b]	Stereotyped receptors[b]	Fisher's exact test[c]
4[d]	conventional	143	125	107	1	
4[d]	CD5+	460	419	160	12	P < 0.018
2[e]	conventional	196	169	141	6	
2[e]	CD5+	180	158	145	18	P < 0.018
2[e]	CD43+	115	72	12	0	P < 1

[a]Only full-length V gene rearrangements and correct V-gene, i.e., VH1-69, VH3-21, VH3-23, VH3-48, and VH4-34 in stereotyped BCR-specific PCR approach
[b]Identical sequences counted as one. When all sequences obtained are counted individually, the statistical significance is below P < 0.001. Stereotypy was determined based on a 60% homology of the CDR3.
[c]P-values calculated by Fisher's exact test versus unique conventional B cell sequences.
[d]Total numbers of V_H1 and V_H3 gene analysis of four independent healthy blood donors are given, for details see Table S3.
[e]Total numbers of V genes amplified by PCR for V_H gene segments frequently used by stereotyped BCR of two independent healthy blood donors are given, for details see Table S7.

the analysis (representing 12 different rearrangements) was stereotypic (Table 2 and Table S8).

We conclude that stereotyped *IGHV* rearrangements are predominantly found in mature CD5$^+$ B cells at a frequency similar to and reflecting the *IGHV* gene mutation pattern in CLL. Moreover, CD43$^+$ B cells do not fulfill the criteria to be the specific precursors of CLL.

Similarities and differences between CD5$^+$ PB B cell subsets and conventional B cells

Having identified two subsets of mature CD5$^+$ PB B cells based on CD27 expression, we wondered how these relate to conventional CD27$^-$ naive and CD27$^+$ memory B cells in their gene expression patterns. We generated a list of 519 transcripts that were differentially expressed with at least a twofold change (P < 0.05 and FDR < 0.05) between conventional naive and IgM memory B cells, including typical naive (ABCB1, IL4R, and FCER2) and memory cell genes (CD27, ITGB2, CD80, CD95, and CD32). A heatmap of the expression pattern of these genes for the two CD5$^+$ B cell subsets revealed that CD5$^+$CD27$^-$ B cells resembled conventional naive B cells and CD5$^+$CD27$^+$ B cells conventional IgM memory B cells (Fig. 5 a and Table S9). This further argues for a pre-GC differentiation stage of CD5$^+$CD27$^-$ B cells and a post-GC differentiation stage of CD5$^+$CD27$^+$ B cells. However, regarding the overlap of genes differentially expressed between conventional naive and memory B cells (Table S9) and those differentially expressed between CD5$^+$CD27$^-$ and CD5$^+$CD27$^+$ B cells (1.5 fold-change; P < 0.05; false discovery rate < 0.05; Table S9), there was an overlap of only ~18% of these genes, although the majority of the remaining transcripts showed similar expression tendencies (i.e., up- or down-regulated in CD27$^-$ versus CD27$^+$ B cells; unpublished data). This indicates specific differences between conventional and CD5$^+$ B cells, in line with their separation in the unsupervised hierarchical clustering (Fig. 4 a). To further analyze these differences, we performed gene set enrichment analysis (GSEA) of CD5$^+$CD27$^-$ and conventional naive B cells. Our main focus was on activated B cell signatures because CD5 is also considered as a B cell activation marker (Berland and Wortis, 2002) and we wanted to clarify whether CD5$^+$ B cells in human PB may simply represent activated conventional B cells. Importantly, immediate early and delayed early activation genes were consistently down-regulated in CD5$^+$ versus conventional naive, IgM memory, or MGZ B cells (Fig. 6, g and h). Moreover, CD5$^+$ B cells did not show signs of increased NF-κB activity or CD40 signaling (Fig. 6 i, Table S10, and not depicted). This finding is in line with a previous publication showing that CD5$^+$ B cells do not express typical B cell activation markers (Damle et al., 2002). Besides lacking an early activation or increased NF-κB target gene signature (Table S10), CD5$^+$ B cells exhibited reduced cytokine, interleukin, TNF, STAT3, and MAPK signaling according to GSEA when compared with conventional naive B cells (unpublished data).

Finally, according to GSEA, CD5$^+$ B cells showed signs of homeostatic proliferation (Fig. 6 j). As the majority of

CD5$^+$ B cells is resting, with >99% of the cells in G0/G1 phase (unpublished data), this could mean that a small fraction of CD5$^+$ B cells proliferates. In line with this view, we detected an increased expression of MYC target genes in CD5$^+$ B cells (Fig. 6 k), especially in the CD27$^+$ subset (unpublished data). Fluorescence microscopic analysis of CD5$^+$ B cells revealed that a small fraction of CD5$^+$ B cells (<1%) does show nuclear MYC expression (Fig. 6 a).

Collectively, CD5$^+$ B cells do not display an activation signature in comparison to conventional naive B cells, but there is a signature for preferential homeostatic proliferation in CD5$^+$ B cells. Determining a full picture of the biological properties of CD5$^+$ B cells will require additional detailed studies of mature conventional and CD5$^+$ B cell subsets, which is beyond the scope of this study.

Evaluation of genes with similar expression in CD5$^+$ B cells and CLL

We wondered whether there are genes that were previously regarded as deregulated in CLL in comparison to conventional B cells, but that do not show deregulation in comparison of CLL to CD5$^+$ B cells, which we now consider as their normal counterpart. A list of selected genes with previously proposed functional relevance for CLL is shown in Fig. 5 b and Table S11.

FOXP1, a transcription factor involved in lymphocyte development and with a potential role of an oncogene in lymphomas (Koon et al., 2007), is increased in CLL in comparison with conventional B cells (Korać et al., 2009). Similarly, the transcription factor LEF1, which has important functions in lymphopoiesis, promotes CLL survival and is not expressed in conventional mature B cells (Gutierrez et al., 2010; Tandon et al., 2011). Both transcription factors are transcribed in CD5$^+$ B cells at a level similar to that found in CLL cells, suggesting they already have functions in normal CD5$^+$ B cells and remain active after transformation into CLL. In line with this view, GSEA shows that LEF1 target genes are up-regulated in CD5$^+$ versus conventional naive B cells (unpublished data).

The tyrosine kinase LCK is involved in S phase transition and apoptosis (Paterson et al., 2006) and tetraspanin family member CD9 modulates adhesion and migration of the neoplastic B cells (Barrena et al., 2005). These genes were previously reported to be up-regulated in CLL. However, they are expressed at similar levels in normal CD5$^+$ B cells, indicating that their expression pattern reflects unique signaling potential and migration properties of the CD5$^+$ precursor population.

IL-24 is expressed in CLL and contributes to tumor survival (Sainz-Perez et al., 2006). Normal CD5$^+$ B cells show similarly high transcript levels of IL-24 (Fig. 5 b). Interestingly, the role of IL-24 in normal B cells is to inhibit the plasma cell differentiation program in GC B cells (Maarof et al., 2010), thus IL-24 may also contribute to a restrained differentiation potential in CLL.

Members of the Rho family subclass of small GTPases showed similar expression levels in CD5$^+$ B cells and CLL,
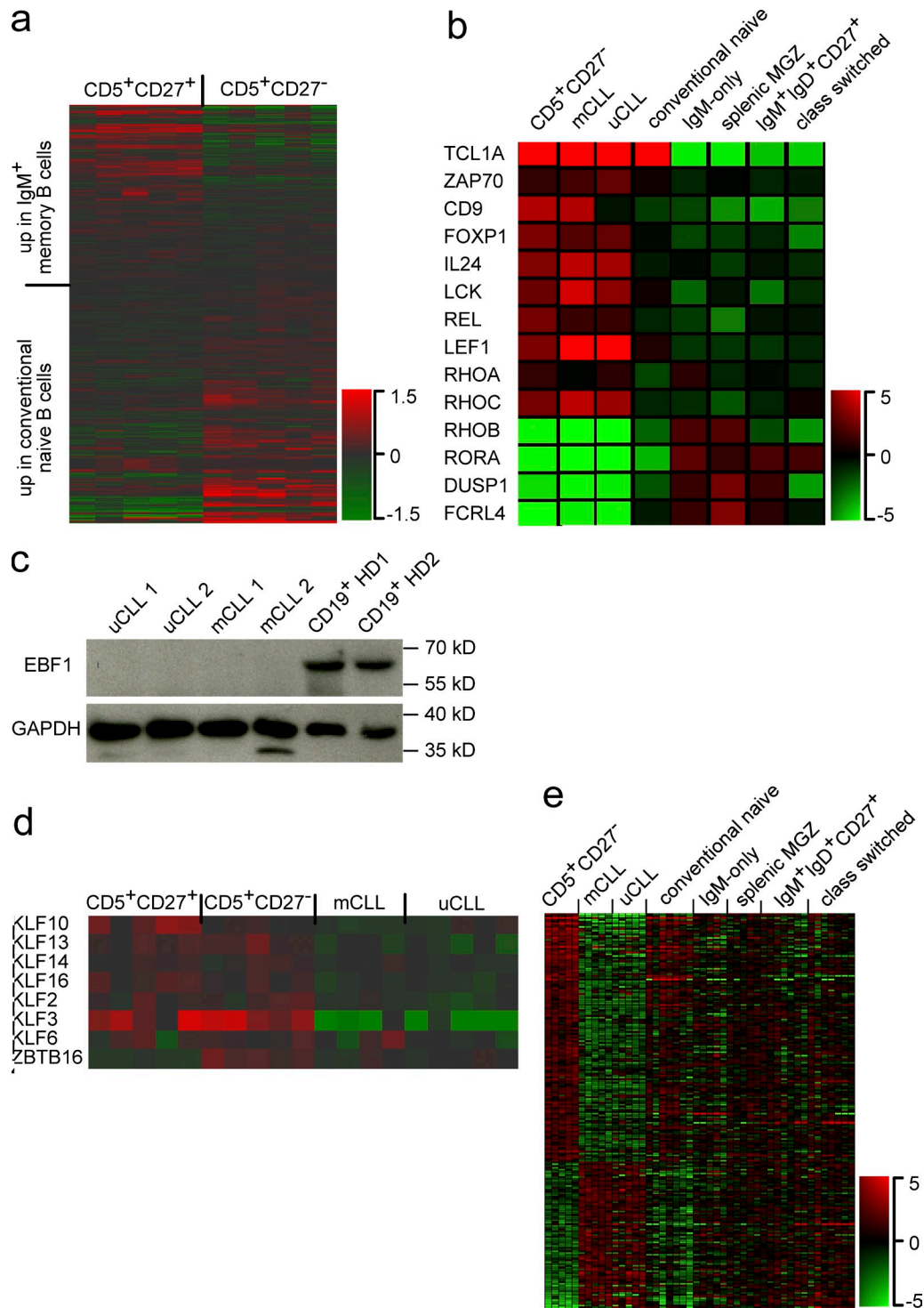
Figure 5. Evaluation of CD5+ B cell and CLL gene expression profiles. (a) Heatmap of CD5+ B cell subset transcription patterns of genes with differential expression between conventional naive and IgM+ memory B cells (Table S9). The color bar depicts normalized intensity values. (b) Heatmap of selected genes with similar transcription in CD5+ B cells and CLL (Table S11). (c) Immunoblotting of EBF1. Protein lysates of CLL and CD19+ B cells from healthy donors were analyzed for EBF1 and GAPDH content. Data are representative of 10 CLL analyzed. (d) Expression heatmap of selected KLF family members in CLL and CD5+ B cells. GEP of normal CD5+ B cell subsets and CLL were filtered for KLF family members with differential expression within the four cell types (Table S12). Depicted are normalized signal intensities. (e) Differentially expressed transcripts (>twofold-change; P < 0.05; FDR < 0.05) between uCLL and CD5+CD27− B cells (HG U133 array), but not (<twofold change; P < 0.05; FDR < 0.05) between uCLL and bulk conventional B cells (naive and memory B cells combined). All transcripts in this list (Table S14) are considered as deregulated genes in CLL that were not detectable (or underestimated) in analyses including conventional CD19+ B cells. Normalized signal intensities are shown.

but not conventional B cells. RHOA, RHOB, and RHOC proteins promote reorganization of the actin cytoskeleton and regulate cell shape, attachment, and motility (Bustelo et al., 2007). *rhoB*-deficient animals indicate tumor suppressing properties of this small GTPase, whereas *rhoC*-deficient tumor cells showed reduced migration and lower invasiveness, thus indicating a beneficial role for RHOC in tumor development (Bustelo et al., 2007). Remarkably, low expression of RHOB, but high levels of RHOA and RHOC are already detectable in normal CD5$^+$ B cells, suggesting that the CLL tumor cells may profit from taking over this expression pattern. In line with this result, GSEA revealed that shared characteristics of CD5$^+$ B cells and CLL, for example, reduced G protein–coupled receptor activity and interaction with extracellular matrix receptors (unpublished data) in contrast to conventional B cells.

Several genes show similarly low transcription levels in CLL and CD5$^+$ but not conventional B cell subsets (Fig. 5 b and Table S11). FCRL4 was previously reported to be down-regulated in CLL (Kazemi et al., 2008), with the potential to dampen BCR signaling and enhance TLR signaling (Sohn et al., 2011). Similarly, RORA, a potent negative regulator of NF-κB signaling (Delerive et al., 2001) and DUSP1, a negative regulator of MAPK signaling, are already down-regulated in mature CD5$^+$ B cells, contributing to the idea of unique signaling characteristics in CD5$^+$ B cells that are (at least partially) inherited in CLL.

Finally, TCL1A and ZAP70, two molecules characteristically expressed in CLL, preferentially in uCLL cases with poor clinical outcome, are already transcribed at similarly high levels not only in conventional naive B cells, but also in mature CD5$^+$ B cells (Fig. 5 b and Table S11).

Evaluation of deregulated genes and implications for pathogenetic mechanisms in CLL

As we propose mature CD5$^+$ B cells as the cells of origin of CLL, we screened the respective transcription patterns for deregulated genes in CLL that were not identified in previous studies, including memory or bulk B cells. 215 annotated transcripts with consistent differential expression (on both array platforms) between CD5$^+$ B cell subsets and CLL included 72 genes (30%), known to be deregulated in CLL, as well as 143 (70%) to the best of our knowledge so far unreported genes (unpublished data). From the 72 previously published genes, 71 showed an expression tendency that was in line with published data. Thus, the GEPs of CLL and CD5$^+$ B cells are of high quality and neatly reproduce a large number of well described CLL features.

Among newly identified deregulated genes, the B cell lineage transcription factor EBF1 was significantly down-regulated in CLL compared with CD5$^+$ or conventional B cells. This observation was supported by increased expression of EBF1-induced genes in normal B cells, but not in CLL (Fig. 6, l). EBF1-repressed genes were not decreased at statistically significant levels. Whereas EBF1 was detectable in conventional B cells by immunoblotting, it was below the detection limit

in mCLL and uCLL (Fig. 5 c). The low expression of EBF1 may lead to reduced levels of numerous B cell signaling factors, thereby contributing to an anergic signature of CLL cases (Mockridge et al., 2007; Muzio et al., 2008) and low susceptibility to host immune recognition (Schultze et al., 1996).

Similarly, gene sets that were up-regulated in plasma cells instead of B cells (Fig. 6 m) are significantly enriched in CD5$^+$ B cells. Moreover, the flow cytometric analysis of intracellular Ig verifies a low content of cytoplasmic Ig in CLL (unpublished data). However, the reduced potential to differentiate into antibody-secreting cells may also be more directly mediated by down-regulation of plasma cell factors. BACH2, a transcriptional repressor inhibiting plasma cell differentiation, was expressed in CLL and genes containing a predicted BACH2-binding site in their promoter region were significantly decreased in CLL (unpublished data).

Members of the KLF family of transcription factors are considered to have tumor-suppressive properties. We observed a down-regulation of several KLF factors in CLL when compared with CD5$^+$ B cells (Fig. 5 d). Although the fold-change of down-regulation of some of these KLF factors is low (Table S12), the decrease of these transcriptional regulators is consistent and may therefore add up to an important impact on CLL biology. Indeed, KLF2-induced genes (Haaland et al., 2005) were significantly increased in CD5$^+$ B cells versus CLL (Fig. 6 n). This includes genes like CDKN2D, an inhibitor of CDK4, which in turn has been proposed to keep CLL cells blocked in the G1 cell cycle stage and contribute to tumor accumulation (Wolowiec et al., 2001). Moreover, the CLL growth inhibiting TGFB1 (DeCoteau et al., 1997) and caspases 8 and 10, both important regulators of programmed cell death in CLL (Enjuanes et al., 2008), are KLF2 targets, found to be down-regulated in CLL. KLF2-repressed genes were not decreased at statistically significant levels. Immunofluorescence analysis of KLF2 and KLF3 confirmed that nuclear expression of both factors was consistently detectable in a fraction of normal CD5$^+$ B cells but virtually absent in CLL (Fig. 6 and Table S13). Although the transcription level in CD5$^+$ B cells was only marginally higher than in CLL (1.5-fold-change), KLF3 expression was more pronounced on protein level, with bright nuclear stainings in 28 to 40% of CD5$^+$ B cells, whereas KLF2 was detected in ~1% of CD5$^+$ B cells analyzed (Fig. 6, c–f).

In addition, we filtered for genes with differential expression between CLL and CD5$^+$ but not conventional B cells (Fig. 5 e and Table S14), i.e., deregulated genes in CLL that were not detectable (or underestimated) in analyses, including only conventional B cells for comparison to CLL. A substantial number of these transcripts included typical B cell genes like CD20 (MS4A1), CD21 (CR2), CD40, CD79B, and IGHD, affirming the idea of a decreased B cell phenotype in CLL, mediated by EBF1 down-regulation (Fig. 5 c and unpublished data).

SIPA1, a RAP1 GTPase-activating protein is significantly down-regulated in CLL compared with CD5$^+$, but not conventional B cells. This is remarkable, as SIPA1-deficient mice mostly develop myeloproliferative disorders, but in few
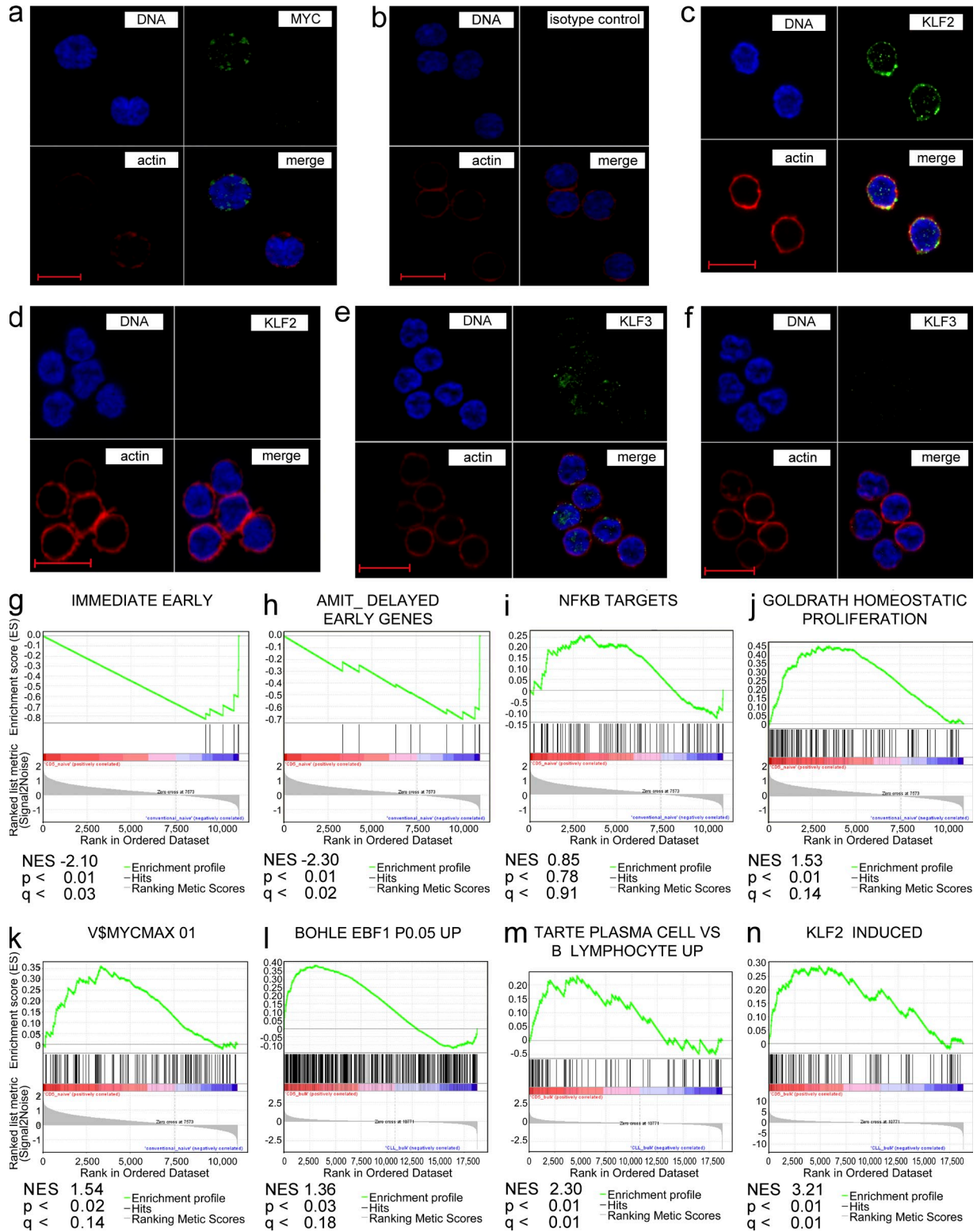
Figure 6. Fluorescence microscopic analysis and GSEA of CD5+ B cells and CLL. CD5+ B cells and CLL cells were stained for intracellular MYC, KLF2, or KLF3 expression (green). DNA and actin were stained with Hoechst 33258 (blue) and Phalloidin-TRITC (red), respectively. MYC was expressed in <1% of CD5+ B cells (a), as compared with isotype negative control stainings (b). Sporadic KLF2 expression was detectable in CD5+ B cells (c) but never in CLL (d). KLF3 was expressed in 28–40% of CD5+ B cells (e) and in 1–5% of CLL cells (f) as determined by two independent blind studies of two normal CD5+ B cell samples and two CLL each (Table S13). Pictures are representative of four healthy donors and five CLL analyzed. (g–n) shows selected plots from a GSEA based on 24,000 probe sets of 5 CD5+CD27− and 5 CD5+CD527+ B cell GEP combined with 7 conventional naive B cell GEP or 9 CLL GEP,

instances accumulate self-reactive CD5$^+$ B cells in the peritoneal cavity (Ishida et al., 2006).

Several genes involved in G protein signaling were found to be deregulated by a factor of more than twofold, including GNG7, GNG11, RAP1GDS1, RASGRP2, RASSF2, VAV1, VAV3, and RAB31. To what extent these findings contribute to CLL pathobiology remains unresolved. However, it is obvious that a comparison of CLL to its mature CD5$^+$ B cell precursor gives rise to several previously unreported aspects.

Similarly, we identified by GSEA several gene sets with preferential deregulation between CLL and CD5$^+$ B cells: FOXO1 was described to be expressed in CLL at levels similar to that of conventional B cells (Xie et al., 2012). However, in comparison to mature CD5$^+$ B cells, FOXO1 target genes are significantly enriched in CLL (unpublished data). Also enriched in CLL are ABC family transporters, which are responsible for membrane permeability for a broad range of nutrients and compounds and genes with a PITX2-binding site in the promoter region (unpublished data). Gene sets significantly enriched in CD5$^+$ B cells (and thus down-regulated upon transition into CLL) are, for example, genes with POU6F1-binding site in the promoter region, a known proliferation-driving homeobox transcription factor in adenocarcinoma, and HIF1- and SRF-driven genes (unpublished data).

DISCUSSION

We aimed to determine the cellular origin of CLL by comparing the global gene expression of mCLL and uCLL to the major human mature B cell subsets. Unsupervised multiparametric analyses of >10,000 transcripts, pointed to CD5$^+$ B cells as the normal B cell subset with the most similar gene expression to CLL. This finding was confirmed by refined assays based on normal B cell subset-specific expression patterns and notably, was already reflected by pairwise comparisons on single gene level. In a complementary approach, the presence of genes that are differentially expressed between CLL and other B cell lymphomas identified gene expression patterns typical for CLL in CD5$^+$ B cells. Thus, we conclude from our transcriptional characterization that mature CD5$^+$ B cells are in all probability the cell of origin of this leukemia. This refers to the cell that was the direct precursor of the tumor clone. As tumorigenesis is a multi-step process, first transforming events may have happened in earlier differentiation stages, perhaps even in hematopoietic stem cells of CLL patients (Kikushige et al., 2011).

An unexpected finding was that not only uCLL but also mCLL was highly similar to CD5$^+$ B cells, which are mostly IgV unmutated (Brezinschek et al., 1997; Fischer et al., 1997; Dono et al., 2007). We specified a distinct subset of CD5$^+$CD27$^+$ B cells (∼1% of total B cells), with the vast

majority of these cells carrying somatically mutated *IGHV* genes. Addressing the current discussion on somatic hypermutation outside the GC (Kruetzmann et al., 2003; Weill et al., 2009), we showed that CD5$^+$CD27$^+$ B cells carry mutations in *BCL6* as a specific hallmark of B cells undergoing hypermutation in the GC (because only in GC B cells *BCL6* is highly transcribed and strong transcription of a gene is essential for somatic hypermutation; Pasqualucci et al., 1998; Seifert and Küppers, 2009). Thus, we identified a distinct subset of somatically mutated post-GC CD5$^+$ memory B cells. The high similarity of the two CD5$^+$ B cell subsets in terms of their gene expression supports the idea that CD5$^+$CD27$^+$ B cells derive from CD5$^+$CD27$^-$ B cells that in rare instances can undergo GC reactions.

A separate GEP analysis revealed that mCLL tend to be more similar to CD5$^+$CD27$^+$ B cells, and uCLL tend to be more similar to CD5$^+$CD27$^-$ B cells. This was also seen when we investigated the CD5$^+$ B cell gene signatures of >100 CLL from an independent cohort. As consistent gene expression differences discriminating mCLL from uCLL are scarce, and CLL subsets are more similar to each other than to any other normal B cell subset, the potential of CD5$^+$ B cell transcription patterns to translate such minor differences into significant CLL subtype predictions is the more impressive. Thus, it appears that uCLL derive from CD5$^+$CD27$^-$ B cells, whereas mCLL derive from CD5$^+$CD27$^+$ B cells. This interpretation is supported by the shared pattern of *IGHV* gene mutations – uCLL and unmutated CD5$^+$CD27$^-$ B cells versus mCLL and mutated CD5$^+$CD27$^+$ B cells – and the distribution of mutated and unmutated stereotyped *IGHV* rearrangements in the two CD5$^+$ B cell subsets, respectively. Moreover, a derivation of mCLL from post-GC CD5$^+$ memory B cells is further supported by the fact that not only normal CD5$^+$CD27$^+$ B cells, but also mCLL carry *BCL6* mutations as a genetic trait of a GC passage (Pasqualucci et al., 2000; Jantus-Lewintre et al., 2009). The potential relationship between a rare, CD5-expressing B cell population from human tonsils, partially located in GCs, and CLL, has been previously described (Caligaris-Cappio et al., 1982). Notably, the expression of CD27 on uCLL is not contradictory to a derivation from CD5$^+$CD27$^-$ B cells, as CD27 is also up-regulated upon T cell-independent B cell activation (Huggins et al., 2007), and expressed on a variety of leukemias and lymphomas (van Oers et al., 1993), independent of their cellular origin (Dong et al., 2002).

In two separate *IGHV* gene PCR analyses, stereotyped BCR were significantly enriched in CD5$^+$ B cells as compared with conventional B cells. This independently supports the conclusion from the transcriptome studies that CD5$^+$ B cells are the precursor population of CLL. This conclusion is not

irrespective of the mutation status. Normalized enrichment score (NES), nominal p-value (p), and FDR (q) are given for each plot. (g–k) CD5$^+$CD27$^-$ B cells are shown on the left side (red) of the plots, conventional naive B cells on the right side (blue). (g) Immediate early genes, (h) delayed early genes, (i) NF-κB target genes, (j) homeostatic proliferation genes, (k) MYC/MAX target genes. (l–n) pools of CD5$^+$ B cell subsets are shown on the left side (red), CLL on the right side (blue). (l) EBF1 target genes, (m) genes up-regulated in plasma cells versus B cells, (n) KLF2-induced genes.

contradicted by another study of stereotyped VH1-69 rearrangements among unmutated V genes (Forconi et al., 2010), because that study did not separate B cell subsets and CD5$^+$ B cells account for ~15% of unmutated B cells in adult PB. Our second *IGHV* gene analysis included CD43$^+$ B cells, as these putative B1 cells were recently proposed to be CLL precursors (Griffin et al., 2011). About two third of CD43$^+$ B cells are CD5$^+$, but unexpectedly, they are also defined as CD27$^+$ B cells (Griffin et al., 2011). As the *IGHV* gene PCR with VH primers for *IGHV* segments frequently used by stereotyped BCR did not amplify many different rearrangements from the isolated CD43$^+$CD27$^+$ B cells, this population is either oligoclonal or it does not frequently use these *IGHV* genes of dominant stereotyped groups, or both. Nevertheless, no stereotyped rearrangements were found among CD43$^+$ B cells. Although we do not exclude that in some instances CD43$^+$CD5$^+$ B cells can transform into CLL, these cells do not represent the specific CLL precursors. It should also be noted that in our analyses, CD43$^+$CD27$^+$ B cells are much rarer (on average 3% of PB B cells from 7 donors) than reported by Griffin et al. (2011; 13% of PB B cells).

The transcriptome of human mature CD5$^+$ B cells is clearly distinct from conventional naive and memory B cells. Moreover, the observed clonality and the increased frequency of stereotyped receptors in CD5$^+$ B cell subsets is not detectable in CD5$^-$ B cells. This autonomy of human CD5$^+$ B cells is supported by independent studies showing marginal but distinct phenotypic features of these cells, e.g., bias in V gene repertoire (Brezinschek et al., 1997) or polyreactivity to nuclear and cytoplasmic antigens (Hervé et al., 2005). Furthermore, normal CD5$^+$ B cells lack a clear activation or NF-κB signature, arguing against the idea that these cells simply represent activated conventional B cells. Hence, CD5 expression defines a distinct differentiation stage, or perhaps even a separate lineage of human B cells.

The derivation of CLL from CD5$^+$ B cells has major implications for our understanding of its pathogenesis. First, CLL often express polyreactive and autoreactive BCR specificities (Catera et al., 2008; Chu et al., 2008; Chu et al., 2010), and it is an important question whether this holds true for normal mature CD5$^+$ B cells as well. Initial studies suggest that CD5$^+$ B cells may possibly be reactive only to selected autoantigens (Hervé et al., 2005). As the autoreactivity of antibodies from CLL is partly linked to stereotyped BCR, the finding of stereotyped *IGHV* gene rearrangements in normal CD5$^+$ B cells supports the idea that some CD5$^+$ B cells express autoreactive receptors and through chronic antigenic stimulation might be at risk to undergo malignant transformation. In this regard, the relatively frequent occurrence of IgV mutated CLL (ca. 50% of cases) as opposed to the rarity of mutated normal CD5$^+$ B cells might indicate that in those rare instances when an (autoreactive?) CD5$^+$ B cell is driven into a GC reaction, it is at increased risk for malignant transformation. This may be due to the extensive clonal expansion and perhaps the mutagenic environment in the GC (Küppers et al., 1999). Furthermore, the fact that mCLL on average have a higher mutation load than normal CD5$^+$CD27$^+$ B cells indicates that a prolonged GC experience (and hence higher mutation load) of these cells leads to a higher propensity of malignant transformation. Second, there is convincing evidence that CLL may be preceded by monoclonal B cell lymphocytosis, which is defined by expansions of mostly CD5$^+$ B cells, and found in ~3% of healthy elderly individuals. Interestingly, these clones often carry already some of the genetic lesions typical for CLL and hence can be regarded as premalignant conditions (Landgren et al., 2009). Although the present analysis was not designed to clarify the clonal composition of human mature CD5$^+$ B cells, it revealed a surprising oligoclonality of these cells. Therefore, we extend the chain of events in CLL development by identifying oligoclonally expanded normal CD5$^+$ B cells, already detectable in young healthy adults, as potential precursors of monoclonal B cell lymphocytosis in the elderly. Finally, we provide evidence that comparison of CLL to mature CD5$^+$ B cells identifies deregulated genes with increased sensitivity. Indeed, we found numerous novel expression patterns that were not recognized in previous studies when CLL were compared with bulk or memory B cells, e.g., diminished EBF1 expression and reduced levels of tumor suppressor genes of the KLF family in CLL. Moreover, we provide a selection of genes with important functions in lymphocytes, deregulated in CLL, which might explain numerous pathophysiological aspects of this disease. These include signaling properties, migration potential and metabolic features.

On the other hand, we identified genes which appeared differentially expressed between CLL and conventional B cells but turned out to be similarly expressed by normal CD5$^+$ B cells and CLL, so that they were not deregulated during cellular transformation (Fig. 5 e and Table S14). Thus, we reveal a highly similar expression pattern of the normal and malignant cells, which is characteristic for CLL. This implicates that the tumor inherited or adopted normal B cell properties. CLL pathobiology may be better understood, if the role of these genes in normal CD5$^+$ B cells is clarified.

Importantly, comparing CLL cells to their specific cell of origin is also of relevance when deregulated miRNA expression in CLL is studied, or epigenetic changes during the transformation process in CLL are being evaluated. We conclude that the identification of two distinct CD5$^+$ B cell subsets as cellular origin of CLL significantly contributes to a better understanding of CLL pathobiology.

MATERIALS AND METHODS

CLL samples and healthy blood donors. PB samples of CLL patients and healthy donors and splenic tissue were analyzed with the approval of the ethical review committee of the University of Duisburg-Essen and with informed consent of the donors according to our institutional guidelines. At the time of sample collection, patients were either untreated or treatment-free for >3 mo. Clinical and laboratory data are shown in Table S15. All CLL expressed IgM and IgD. Normal PB B cell subsets for GEP were taken from healthy donors (age, 23–59 yr; mean, 35 yr), samples for other experiments were prepared from buffy coats (age, 18–66 yr). Splenic tissue was derived from surgery to repair traumatic rupture or tumor surgery, but without direct tumor affection (donor age, 49–76 yr; mean, 59 yr).

Flow cytometry and cell sorting. B and CLL cells were isolated by Ficoll density gradient centrifugation (GE Healthcare) and CD19-MACS strategy (Miltenyi Biotec) to a purity of >98%. Samples were FACS-sorted and analyzed by staining with anti–CD5-APC (UCHT2; BioLegend), CD43-FITC (1G10), CD38-APC (HIT2), CD27-FITC (MT271), CD23-PE (M-L233), CD24-PE (ML5), IgG-FITC (G18-145), IgA-FITC (F0316; Dako), CD21-PE (1048), IgM-FITC (G20-127), IgD-PE (IA6-2), IgD-PE-Cy7 (IA6-2), and CD27-APC (MT271). If not stated otherwise, all antibodies were purchased from BD Biosciences (BD). Samples for HG U133 2.0 plus arrays were sorted as: $CD5^+CD27^-CD23^+CD24^-$ ("$CD5^+$"), $CD23^+IgD^{high}CD27^-$ ("naive"), $IgM^+IgD^+CD27^+$, $IgM^+IgD^{low/-}CD27^+$ ("IgM-only"), IgG^+/IgA^+CD27^+ ("class-switched"), $IgM^+CD21^{high}CD27^+$ ("sMGZ"). B cells for HuGene-1_0-st-v1 arrays, *BCL6* and *IGHV* analysis were negatively enriched by the EasySep Human B Cell Enrichment kit 19054 (STEMCELL), thus excluding $CD43^+$ B cells. $CD5^+$ B cells were subsequently enriched by anti–CD5-APC (UCHT2) and anti-APC MicroBeads (Miltenyi Biotec). Samples were sorted as $CD5^+IgD^+CD38^{low}CD27^-$ ($CD5^+CD27^-$) and $CD5^+CD38^-CD27^+$ ($CD5^+$ memory) B cells. $CD5^+$ class-switched B cells were sorted as $CD5^+CD38^-CD27^+$ and IgG^+ or IgA^+. Transitional B cells ($CD38^{high}CD24^{high}$) were excluded. For the IgV gene PCR, $CD43^+$ B1 B cells ($CD43^+CD27^+$) and B2 B cell subsets were isolated after CD19-MACS. CLL samples were sorted as $CD5^+CD23^+$ cells. After saturating staining of extracellular Ig light chain (IgL), intracellular IgL was stained with the BD Cytofix/Cytoperm kit according to the manufacturer's instructions. Antibodies used were anti–Ig κ chain (Igκ)-PE and anti–Igκ-FITC (both G20-193; BD). FACS data were acquired with a FACSCanto cytometer (BD).

GEP sample preparation. RNA was extracted from 10,000 cells by the Gentra Purescript protocol (Gentra). RNA integrity was assessed by Agilent 2100 Bioanalyzer (Agilent). Samples with RNA integrity number >9.0 were processed by MessageAmp II aRNA amplification kit and MessageAmp II Biotin Enhanced kit (Ambion). For the HG U133 2.0 Plus GEP analysis data were generated in three batches, with the first two equally composed of naive, IgM-only, $IgM^+IgD^+CD27^+$, class-switched, and sMGZ B cells, and a third batch containing the $CD5^+$ B cell and CLL samples and two naive, $IgM^+IgD^+CD27^+$, and class-switched B cell samples each for batch correction. Vsn-normalized data were corrected for batch effect by ComBat software (Johnson et al., 2007). For the HuGene-1_0-st-v1 GEP analysis, 50 ng RNA was processed with the OVATION Pico WTA System, the WT Ovation Exon Module, and the Encore Biotin Module (NuGen). Arrays were scanned with a GeneChip Scanner 3000 7G (Affymetrix). GeneChip data have been submitted to the GEO database under accession no. GSE36907.

IGHV gene rearrangement analysis and BCL6 PCR. For *IGHV*-PCR analysis, cells were sorted in duplicate or triplicate aliquots. Genomic DNA was extracted by Gentra Puregene Blood kit (QIAGEN). *IGHV* gene rearrangements of the V_H1 and V_H3 family were amplified in a seminested, two-rounded multiplex PCR assay with the Expand High Fidelity PCR system (Roche; Küppers, 2004).

For the stereotyped BCR-specific PCR, DNA from aliquots of 2,500–5,000 cells was amplified while seminested with *IGHV* gene-specific leader exon and two sets of *IGHJ* primers for 30 cycles, two times. Stereotypy was determined by CDR3 amino acid homology of ≥60% to the consensus sequence of one of 48 stereotype subsets (Stamatopoulos et al., 2007). The calculation of CDR3 homology with respect to physicochemical properties of the amino acids used was performed in a blinded fashion using the DNASIS MAX software. Although other algorithms exist to determine stereotypy (Darzentas et al., 2010), the approach used here is sufficient, as it aims to compare two cell types with the same measure.

The *BCL6* major mutation cluster was amplified from 5,000 cell aliquots by seminested PCR (Pasqualucci et al., 1998) by the High Fidelity PCR System (Roche).

All PCR products were subcloned using the TOPO TA Cloning kit (Invitrogen) and XL1-Blue–competent cells (Agilent). Sequences were obtained with the BigDye Deoxy sequencing kit (Applied Biosystems) and an automated sequencer (ABI3100; Applied Biosystems). Sequences were analyzed by the international ImMunoGeneTics (IMGT) information system and Lasergene 8 (DNASTAR) software. Sequence data have been submitted to GenBank database under accession nos. JX432019–JX432961.

Fluorescence microscopy. Expression of KLF2 was determined by intracellular staining with anti-KLF2 (665333; R&D Systems) and anti–mouse-Cy2 (Jackson ImmunoResearch Laboratories). $CD5^+$ B cells were isolated by T cell depletion (EasySep Human B Cell Enrichment kit 19054), $CD10^+$ transitional B cell depletion (anti–human CD10 MicroBeads; Miltenyi Biotec), and subsequent enrichment by rabbit anti-CD5 (SAB4503585; Sigma-Aldrich) and anti–rabbit IgG MicroBeads (Miltenyi Biotec). CLL cells were purified by CD19-MACS. Expression of KLF3 was assessed by intracellular stainings with anti-KLF3 (ab49221; Abcam) and anti–rabbit-Cy2 (Jackson ImmunoResearch Laboratories). Expression of Myc was analyzed by intracellular staining with anti–c-MYC (9E70; Invitrogen) and anti–rabbit-Cy2 (Jackson ImmunoResearch Laboratories). All stainings were combined with Phalloidin-TRITC (Sigma-Aldrich) and Hoechst 33258 (Roche). Fluorescence microscopy was performed on a Zeiss Axio Observer.Z1 fluorescence microscope equipped with the respective filter sets and an Apotome. Image acquisition was performed via a Plan-Apochromat 63×/1,40 oil objective lense (1.46 numerical aperture) and an AxioCam MRm camera from cell suspensions in fluorescent mounting medium (S3023; Dako) at 23°C. Images were processed with AxioVision Rel. 4.8 software (Carl Zeiss).

Immunoblotting. Equal amounts of protein (15 µg) of CLL and $CD19^+$ B cells were loaded onto 6% acrylamide Tris-Glycine gels (Invitrogen) and transferred to PVDF membrane (Millipore). Anti-EBF1 (H00001879-M01; Abnova) and anti-GAPDH (sc-31915; Santa Cruz Biotechnology, Inc.) primary antibodies were used at 1:1,000 dilution. Protein detection was performed by HRP-conjugated secondary antibodies (115–036-062; Jackson ImmunoResearch Laboratories; sc-2350; Santa Cruz Biotechnology, Inc.) and the ECL Plus chemiluminescence detection kit (GE Healthcare).

Oligonucleotides. *IGHV-leader*–specific primers for *IGHV* genes frequently used by stereotyped receptors are as follows: IGHV1-2 5′-TCT-TCT-TGG-TGG-CAG-CAG-CCA-CAG-GT-3′; IGHV1-69, 5′-GGA-CTG-GAC-CTG-GAG-GTT-CCT-CTT-TG-3′; IGHV3-11, 5′-TGC-TAT-AAT-AAA-AGG-TGT-CCA-GTG-TC-3′; IGHV3-21, 5′-CGA-GGA-TTC-ACC-ATG-GAA-CTG-GGG-CTC-C-3′; IGHV3-48, 5′-TGC-TGG-GTT-TTC-CTT-GTT-GCT-ATT-TTA-G-3′; IGHV4-34, 5′-CAG-GTG-CAG-CTA-CAG-CAG-TGG-GGC-G-3′; IGHV4-39, 5′-TGT-CTC-TGG-TGG-CTC-CAT-CAG-CAG-TAG-3′.

Statistical analysis. Data were analyzed with GeneSpring GX software (Affymetrix) and R software for statistical computing (R Development Core Team [2008]; http://www.R-project.org) Probe sets with a minimum raw signal of 50 and at least 4 present calls in at least one condition, according to MAS5 software, were used for further analysis. Multivariate data analysis was performed with ANOVA and Tukey post-hoc testing procedures, and pairwise comparisons were tested for statistical significance by Student's t test ($P < 0.05$). The Benjamini-Hochberg method was used for multiple testing correction. Gene set enrichment analysis was performed with the GSEA2 software (Subramanian et al., 2005). The significance of association of stereotyped receptor usage with one of the B cell subsets analyzed was determined by two-tailed Fisher's exact tests.

Online supplemental material. Table S1 lists the transcripts for the hierarchical clustering in Fig. 1. Table S2 gives the numbers of differentially expressed genes between CLL and mature B cell subsets. Table S3 lists the details of the V gene analysis of $CD5^+$ and conventional B cell subsets. Table S4 summarizes the Bcl6 analysis of $CD5^+$ B cell subsets. Table S5 lists the transcripts for the hierarchical clustering in Fig. 4 a. Table S6 lists the transcripts and mutation status of the samples used in the PCA in Fig. 4 c. Table S7 gives the stereotypic IGHV rearrangements from our study. Table S8 gives a detailed overview on the PCR analysis of $CD5^+$, $CD43^+$, and conventional B cells

for stereotyped IGHV rearrangements. Table S9 lists the transcripts for the heatmap in Fig. 5 a. Table S10 is a list of NF-κB target genes in human B cells. Table S11 lists the transcripts for the heatmap in Fig. 5 b. Table S12 gives the raw signal values for the heatmap in Fig. 5 d. Table S13 gives a detailed overview on the KLF3 nuclear expression pattern of CLL and CD5+ B cells. Table S14 lists the genes with differential expression between CLL and CD5+, but not conventional B cells. Table S15 summarizes the patient characteristics. Online supplemental material is available at http://www.jem.org/cgi/content/full/jem.20120833/DC1.

REFERENCES

Allman, D., A. Jain, A. Dent, R.R. Maile, T. Selvaggi, M.R. Kehry, and L.M. Staudt. 1996. BCL-6 expression during B-cell activation. *Blood*. 87:5257–5268.

Andréasson, U., P. Edén, C. Peterson, C.M. Högerkorp, M. Jerkeman, N. Andersen, M. Berglund, C. Sundström, R. Rosenquist, C.A. Borrebaeck, and S. Ek. 2010. Identification of uniquely expressed transcription factors in highly purified B-cell lymphoma samples. *Am. J. Hematol*. 85:418–425.

Bachl, J., C. Carlson, V. Gray-Schopfer, M. Dessing, and C. Olsson. 2001. Increased transcription levels induce higher mutation rates in a hypermutating cell line. *J. Immunol*. 166:5051–5057.

Barrena, S., J. Almeida, M. Yunta, A. López, N. Fernández-Mosteirín, M. Giralt, M. Romero, L. Perdiguer, M. Delgado, A. Orfao, and P.A. Lazo. 2005. Aberrant expression of tetraspanin molecules in B-cell chronic lymphoproliferative disorders and its correlation with normal B-cell maturation. *Leukemia*. 19:1376–1383. http://dx.doi.org/10.1038/sj.leu.2403822

Berland, R., and H.H. Wortis. 2002. Origins and functions of B-1 cells with notes on the role of CD5. *Annu. Rev. Immunol*. 20:253–300. http://dx.doi.org/10.1146/annurev.immunol.20.100301.064833

Brezinschek, H.P., S.J. Foster, R.I. Brezinschek, T. Dörner, R. Domiati-Saad, and P.E. Lipsky. 1997. Analysis of the human VH gene repertoire. Differential effects of selection and somatic hypermutation on human peripheral CD5(+)/IgM+ and CD5(-)/IgM+ B cells. *J. Clin. Invest*. 99:2488–2501. http://dx.doi.org/10.1172/JCI119433

Bustelo, X.R., V. Sauzeau, and I.M. Berenjeno. 2007. GTP-binding proteins of the Rho/Rac family: regulation, effectors and functions in vivo. *Bioessays*. 29:356–370. http://dx.doi.org/10.1002/bies.20558

Caligaris-Cappio, F. 1996. B-chronic lymphocytic leukemia: a malignancy of anti-self B cells. *Blood*. 87:2615–2620.

Caligaris-Cappio, F., M. Gobbi, M. Bofill, and G. Janossy. 1982. Infrequent normal B lymphocytes express features of B-chronic lymphocytic leukemia. *J. Exp. Med*. 155:623–628. http://dx.doi.org/10.1084/jem.155.2.623

Catera, R., G.J. Silverman, K. Hatzi, T. Seiler, S. Didier, L. Zhang, M. Hervé, E. Meffre, D.G. Oscier, H. Vlassara, et al. 2008. Chronic lymphocytic leukemia cells recognize conserved epitopes associated with apoptosis and oxidation. *Mol. Med*. 14:665–674. http://dx.doi.org/10.2119/2008-00102.Catera

Chiorazzi, N., and M. Ferrarini. 2011. Cellular origin(s) of chronic lymphocytic leukemia: cautionary notes and additional considerations and possibilities. *Blood*. 117:1781–1791. http://dx.doi.org/10.1182/blood-2010-07-155663

Chu, C.C., R. Catera, K. Hatzi, X.J. Yan, L. Zhang, X.B. Wang, H.M. Fales, S.L. Allen, J.E. Kolitz, K.R. Rai, and N. Chiorazzi. 2008. Chronic lymphocytic

leukemia antibodies with a common stereotypic rearrangement recognize nonmuscle myosin heavy chain IIA. *Blood*. 112:5122–5129. http://dx.doi.org/10.1182/blood-2008-06-162024

Chu, C.C., R. Catera, L. Zhang, S. Didier, B.M. Agagnina, R.N. Damle, M.S. Kaufman, J.E. Kolitz, S.L. Allen, K.R. Rai, and N. Chiorazzi. 2010. Many chronic lymphocytic leukemia antibodies recognize apoptotic cells with exposed nonmuscle myosin heavy chain IIA: implications for patient outcome and cell of origin. *Blood*. 115:3907–3915. http://dx.doi.org/10.1182/blood-2009-09-244251

Damle, R.N., T. Wasil, F. Fais, F. Ghiotto, A. Valetto, S.L. Allen, A. Buchbinder, D. Budman, K. Dittmar, J. Kolitz, et al. 1999. Ig V gene mutation status and CD38 expression as novel prognostic indicators in chronic lymphocytic leukemia. *Blood*. 94:1840–1847.

Damle, R.N., F. Ghiotto, A. Valetto, E. Albesiano, F. Fais, X.J. Yan, C.P. Sison, S.L. Allen, J. Kolitz, P. Schulman, et al. 2002. B-cell chronic lymphocytic leukemia cells express a surface membrane phenotype of activated, antigen-experienced B lymphocytes. *Blood*. 99:4087–4093. http://dx.doi.org/10.1182/blood.V99.11.4087

Darzentas, N., A. Hadzidimitriou, F. Murray, K. Hatzi, P. Josefsson, N. Laoutaris, C. Moreno, A. Anagnostopoulos, J. Jurlander, A. Tsaftaris, et al. 2010. A different ontogenesis for chronic lymphocytic leukemia cases carrying stereotyped antigen receptors: molecular and computational evidence. *Leukemia*. 24:125–132. http://dx.doi.org/10.1038/leu.2009.186

DeCoteau, J.F., P.I. Knaus, H. Yankelev, M.D. Reis, R. Lowsky, H.F. Lodish, and M.E. Kadin. 1997. Loss of functional cell surface transforming growth factor beta (TGF-beta) type 1 receptor correlates with insensitivity to TGF-beta in chronic lymphocytic leukemia. *Proc. Natl. Acad. Sci. USA*. 94:5877–5881. http://dx.doi.org/10.1073/pnas.94.11.5877

Delerive, P., D. Monté, G. Dubois, F. Trottein, J. Fruchart-Najib, J. Mariani, J.C. Fruchart, and B. Staels. 2001. The orphan nuclear receptor ROR alpha is a negative regulator of the inflammatory response. *EMBO Rep*. 2:42–48. http://dx.doi.org/10.1093/embo-reports/kve007

Dong, H.Y., A. Shahsafaei, and D.M. Dorfman. 2002. CD148 and CD27 are expressed in B cell lymphomas derived from both memory and naïve B cells. *Leuk. Lymphoma*. 43:1855–1858. http://dx.doi.org/10.1080/1042819021000006385

Dono, M., V.L. Burgio, M. Colombo, S. Sciacchitano, D. Reverberi, V. Tarantino, G. Cutrona, N. Chiorazzi, and M. Ferrarini. 2007. CD5+ B cells with the features of subepithelial B cells found in human tonsils. *Eur. J. Immunol*. 37:2138–2147. http://dx.doi.org/10.1002/eji.200636887

Dorshkind, K., and E. Montecino-Rodriguez. 2007. Fetal B-cell lymphopoiesis and the emergence of B-1-cell potential. *Nat. Rev. Immunol*. 7:213–219. http://dx.doi.org/10.1038/nri2019

Enjuanes, A., Y. Benavente, F. Bosch, I. Martín-Guerrero, D. Colomer, S. Pérez-Alvarez, O. Reina, M.T. Ardanaz, P. Jares, A. García-Orad, et al. 2008. Genetic variants in apoptosis and immunoregulation-related genes are associated with risk of chronic lymphocytic leukemia. *Cancer Res*. 68:10178–10186. http://dx.doi.org/10.1158/0008-5472.CAN-08-2221

Fischer, M., U. Klein, and R. Küppers. 1997. Molecular single-cell analysis reveals that CD5-positive peripheral blood B cells in healthy humans are characterized by rearranged Vkappa genes lacking somatic mutation. *J. Clin. Invest*. 100:1667–1676. http://dx.doi.org/10.1172/JCI119691

Forconi, F., K.N. Potter, I. Wheatley, N. Darzentas, E. Sozzi, K. Stamatopoulos, C.I. Mockridge, G. Packham, and F.K. Stevenson. 2010. The normal IGHV1-69-derived B-cell repertoire contains stereotypic patterns characteristic of unmutated CLL. *Blood*. 115:71–77. http://dx.doi.org/10.1182/blood-2009-06-225813

Fukita, Y., H. Jacobs, and K. Rajewsky. 1998. Somatic hypermutation in the heavy chain locus correlates with transcription. *Immunity*. 9:105–114. http://dx.doi.org/10.1016/S1074-7613(00)80592-0

Griffin, D.O., N.E. Holodick, and T.L. Rothstein. 2011. Human B1 cells in umbilical cord and adult peripheral blood express the novel phenotype CD20+CD27+CD43+CD70-. *J. Exp. Med*. 208:67–80. http://dx.doi.org/10.1084/jem.20101499

Gutierrez, A. Jr., R.C. Tschumper, X. Wu, T.D. Shanafelt, J. Eckel-Passow, P.M. Huddleston III, S.L. Slager, N.E. Kay, and D.F. Jelinek. 2010. LEF-1

is a prosurvival factor in chronic lymphocytic leukemia and is expressed in the preleukemic state of monoclonal B-cell lymphocytosis. *Blood.* 116:2975–2983. http://dx.doi.org/10.1182/blood-2010-02-269878

Ha, Y.J., Y.C. Mun, C.M. Seong, and J.R. Lee. 2008. Characterization of phenotypically distinct B-cell subsets and receptor-stimulated mitogen-activated protein kinase activation in human cord blood B cells. *J. Leukoc. Biol.* 84:1557–1564. http://dx.doi.org/10.1189/jlb.0706457

Haaland, R.E., W. Yu, and A.P. Rice. 2005. Identification of LKLF-regulated genes in quiescent CD4+ T lymphocytes. *Mol. Immunol.* 42:627–641. http://dx.doi.org/10.1016/j.molimm.2004.09.012

Hamblin, T.J., Z. Davis, A. Gardiner, D.G. Oscier, and F.K. Stevenson. 1999. Unmutated Ig V(H) genes are associated with a more aggressive form of chronic lymphocytic leukemia. *Blood.* 94:1848–1854.

Hervé, M., K. Xu, Y.S. Ng, H. Wardemann, E. Albesiano, B.T. Messmer, N. Chiorazzi, and E. Meffre. 2005. Unmutated and mutated chronic lymphocytic leukemias derive from self-reactive B cell precursors despite expressing different antibody reactivity. *J. Clin. Invest.* 115:1636–1643. http://dx.doi.org/10.1172/JCI24387

Huggins, J., T. Pellegrin, R.E. Felgar, C. Wei, M. Brown, B. Zheng, E.C. Milner, S.H. Bernstein, I. Sanz, and M.S. Zand. 2007. CpG DNA activation and plasma-cell differentiation of CD27- naive human B cells. *Blood.* 109:1611–1619. http://dx.doi.org/10.1182/blood-2006-03-008441

Ishida, D., L. Su, A. Tamura, Y. Katayama, Y. Kawai, S.F. Wang, M. Taniwaki, Y. Hamazaki, M. Hattori, and N. Minato. 2006. Rap1 signal controls B cell receptor repertoire and generation of self-reactive B1a cells. *Immunity.* 24:417–427. http://dx.doi.org/10.1016/j.immuni.2006.02.007

Jantus-Lewintre, E., C. Reinoso Martín, C. García Ballesteros, J. Pendas, C. Benet Campos, J.R. Mayans Ferrer, and J. García-Conde. 2009. BCL6: somatic mutations and expression in early-stage chronic lymphocytic leukemia. *Leuk. Lymphoma.* 50:773–780. http://dx.doi.org/10.1080/10428190902842626

Johnson, W.E., C. Li, and A. Rabinovic. 2007. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics.* 8:118–127. http://dx.doi.org/10.1093/biostatistics/kxj037

Kazemi, T., H. Asgarian-Omran, M. Hojjat-Farsangi, M. Shabani, A. Memarian, R.A. Sharifian, S.M. Razavi, M. Jeddi-Tehrani, H. Rabbani, and F. Shokri. 2008. Fc receptor-like 1-5 molecules are similarly expressed in progressive and indolent clinical subtypes of B-cell chronic lymphocytic leukemia. *Int. J. Cancer.* 123:2113–2119. http://dx.doi.org/10.1002/ijc.23751

Kikushige, Y., F. Ishikawa, T. Miyamoto, T. Shima, S. Urata, G. Yoshimoto, Y. Mori, T. Iino, T. Yamauchi, T. Eto, et al. 2011. Self-renewing hematopoietic stem cell is the primary target in pathogenesis of human chronic lymphocytic leukemia. *Cancer Cell.* 20:246–259. http://dx.doi.org/10.1016/j.ccr.2011.06.029

Klein, U., and R. Dalla-Favera. 2008. Germinal centres: role in B-cell physiology and malignancy. *Nat. Rev. Immunol.* 8:22–33. http://dx.doi.org/10.1038/nri2217

Klein, U., R. Küppers, and K. Rajewsky. 1997. Evidence for a large compartment of IgM-expressing memory B cells in humans. *Blood.* 89:1288–1298.

Klein, U., K. Rajewsky, and R. Küppers. 1998. Human immunoglobulin (Ig)M+IgD+ peripheral blood B cells expressing the CD27 cell surface antigen carry somatically mutated variable region genes: CD27 as a general marker for somatically mutated (memory) B cells. *J. Exp. Med.* 188:1679–1689. http://dx.doi.org/10.1084/jem.188.9.1679

Klein, U., Y. Tu, G.A. Stolovitzky, M. Mattioli, G. Cattoretti, H. Husson, A. Freedman, G. Inghirami, L. Cro, L. Baldini, et al. 2001. Gene expression profiling of B cell chronic lymphocytic leukemia reveals a homogeneous phenotype related to memory B cells. *J. Exp. Med.* 194:1625–1638. http://dx.doi.org/10.1084/jem.194.11.1625

Koon, H.B., G.C. Ippolito, A.H. Banham, and P.W. Tucker. 2007. FOXP1: a potential therapeutic target in cancer. *Expert Opin. Ther. Targets.* 11:955–965. http://dx.doi.org/10.1517/14728222.11.7.955

Korać, P., M.G. Vintar, R. Ajdukovi , M.M. Kardum Paro, B. Jaksić, and M. Dominis. 2009. FOXP1 and BCL2 show similar immunoenzymatic pattern in bone marrow trephines of chronic lymphocytic leukemia patients.

Appl. Immunohistochem. Mol. Morphol. 17:500–504. http://dx.doi.org/10.1097/PAI.0b013e3181a20307

Kruetzmann, S., M.M. Rosado, H. Weber, U. Germing, O. Tournilhac, H.H. Peter, R. Berner, A. Peters, T. Boehm, A. Plebani, et al. 2003. Human immunoglobulin M memory B cells controlling Streptococcus pneumoniae infections are generated in the spleen. *J. Exp. Med.* 197:939–945. http://dx.doi.org/10.1084/jem.20022020

Küppers, R. 2004. Molecular single-cell PCR analysis of rearranged immunoglobulin genes as a tool to determine the clonal composition of normal and malignant human B cells. *Methods Mol. Biol.* 271:225–238.

Küppers, R., U. Klein, M.-L. Hansmann, and K. Rajewsky. 1999. Cellular origin of human B-cell lymphomas. *N. Engl. J. Med.* 341:1520–1529. http://dx.doi.org/10.1056/NEJM199911113412007

Kuraoka, M., T.M. Holl, D. Liao, M. Womble, D.W. Cain, A.E. Reynolds, and G. Kelsoe. 2011. Activation-induced cytidine deaminase mediates central tolerance in B cells. *Proc. Natl. Acad. Sci. USA.* 108:11560–11565. http://dx.doi.org/10.1073/pnas.1102571108

Landgren, O., M. Albitar, W. Ma, F. Abbasi, R.B. Hayes, P. Ghia, G.E. Marti, and N.E. Caporaso. 2009. B-cell clones as early markers for chronic lymphocytic leukemia. *N. Engl. J. Med.* 360:659–667. http://dx.doi.org/10.1056/NEJMoa0806122

Maarof, G., L. Bouchet-Delbos, H. Gary-Gouy, I. Durand-Gasselin, R. Krzysiek, and A. Dalloul. 2010. Interleukin-24 inhibits the plasma cell differentiation program in human germinal center B cells. *Blood.* 115:1718–1726. http://dx.doi.org/10.1182/blood-2009-05-220251

Marie-Cardine, A., F. Divay, I. Dutot, A. Green, A. Perdrix, O. Boyer, N. Contentin, H. Tilly, F. Tron, J.P. Vannier, and S. Jacquot. 2008. Transitional B cells in humans: characterization and insight from B lymphocyte reconstitution after hematopoietic stem cell transplantation. *Clin. Immunol.* 127:14–25. http://dx.doi.org/10.1016/j.clim.2007.11.013

Marshall, J.L., Y. Zhang, L. Pallan, M.C. Hsu, M. Khan, A.F. Cunningham, I.C. MacLennan, and K.M. Toellner. 2011. Early B blasts acquire a capacity for Ig class switch recombination that is lost as they become plasmablasts. *Eur. J. Immunol.* 41:3506–3512. http://dx.doi.org/10.1002/eji.201141762

Messmer, B.T., E. Albesiano, D.G. Efremov, F. Ghiotto, S.L. Allen, J. Kolitz, R. Foa, R.N. Damle, F. Fais, D. Messmer, et al. 2004. Multiple distinct sets of stereotyped antigen receptors indicate a role for antigen in promoting chronic lymphocytic leukemia. *J. Exp. Med.* 200:519–525. http://dx.doi.org/10.1084/jem.20040544

Meyers, G., Y.S. Ng, J.M. Bannock, A. Lavoie, J.E. Walter, L.D. Notarangelo, S.S. Kilic, G. Aksu, M. Debré, F. Rieux-Laucat, et al. 2011. Activation-induced cytidine deaminase (AID) is required for B-cell tolerance in humans. *Proc. Natl. Acad. Sci. USA.* 108:11554–11559. http://dx.doi.org/10.1073/pnas.1102600108

Mockridge, C.I., K.N. Potter, I. Wheatley, L.A. Neville, G. Packham, and F.K. Stevenson. 2007. Reversible anergy of sIgM-mediated signaling in the two subsets of CLL defined by VH-gene mutational status. *Blood.* 109:4424–4431. http://dx.doi.org/10.1182/blood-2006-11-056648

Murray, F., N. Darzentas, A. Hadzidimitriou, G. Tobin, M. Boudjogra, C. Scielzo, N. Laoutaris, K. Karlsson, F. Baran-Marzak, A. Tsaftaris, et al. 2008. Stereotyped patterns of somatic hypermutation in subsets of patients with chronic lymphocytic leukemia: implications for the role of antigen selection in leukemogenesis. *Blood.* 111:1524–1533. http://dx.doi.org/10.1182/blood-2007-07-099564

Muzio, M., B. Apollonio, C. Scielzo, M. Frenquelli, I. Vandoni, V. Boussiotis, F. Caligaris-Cappio, and P. Ghia. 2008. Constitutive activation of distinct BCR-signaling pathways in a subset of CLL patients: a molecular signature of anergy. *Blood.* 112:188–195. http://dx.doi.org/10.1182/blood-2007-09-111344

Nahar, R., P. Ramezani-Rad, M. Mossner, C. Duy, L. Cerchietti, H. Geng, S. Dovat, H. Jumaa, B.H. Ye, A. Melnick, and M. Müschen. 2011. Pre-B cell receptor-mediated activation of BCL6 induces pre-B cell quiescence through transcriptional repression of MYC. *Blood.* 118:4174–4178. http://dx.doi.org/10.1182/blood-2011-01-331181

Pasqualucci, L., A. Migliazza, N. Fracchiolla, C. William, A. Neri, L. Baldini, R.S.K. Chaganti, U. Klein, R. Küppers, K. Rajewsky, and R. Dalla-Favera. 1998. BCL-6 mutations in normal germinal center

B cells: evidence of somatic hypermutation acting outside Ig loci. *Proc. Natl. Acad. Sci. USA*. 95:11816–11821. http://dx.doi.org/10.1073/pnas.95.20.11816

Pasqualucci, L., A. Neri, L. Baldini, R. Dalla-Favera, and A. Migliazza. 2000. BCL-6 mutations are associated with immunoglobulin variable heavy chain mutations in B-cell chronic lymphocytic leukemia. *Cancer Res*. 60:5644–5648.

Paterson, J.C., S. Tedoldi, A. Craxton, M. Jones, M.L. Hansmann, G. Collins, H. Roberton, Y. Natkunam, S. Pileri, E. Campo, et al. 2006. The differential expression of LCK and BAFF-receptor and their role in apoptosis in human lymphomas. *Haematologica*. 91:772–780.

Rassenti, L.Z., S. Jain, M.J. Keating, W.G. Wierda, M.R. Grever, J.C. Byrd, N.E. Kay, J.R. Brown, J.G. Gribben, D.S. Neuberg, et al. 2008. Relative value of ZAP-70, CD38, and immunoglobulin mutation status in predicting aggressive disease in chronic lymphocytic leukemia. *Blood*. 112:1923–1930. http://dx.doi.org/10.1182/blood-2007-05-092882

Rosenwald, A., A.A. Alizadeh, G. Widhopf, R. Simon, R.E. Davis, X. Yu, L. Yang, O.K. Pickeral, L.Z. Rassenti, J. Powell, et al. 2001. Relation of gene expression phenotype to immunoglobulin mutation genotype in B cell chronic lymphocytic leukemia. *J. Exp. Med*. 194:1639–1647. http://dx.doi.org/10.1084/jem.194.11.1639

Sainz-Perez, A., H. Gary-Gouy, A. Portier, F. Davi, H. Merle-Beral, P. Galanaud, and A. Dalloul. 2006. High Mda-7 expression promotes malignant cell survival and p38 MAP kinase activation in chronic lymphocytic leukemia. *Leukemia*. 20:498–504. http://dx.doi.org/10.1038/sj.leu.2404073

Schultze, J., L.M. Nadler, and J.G. Gribben. 1996. B7-mediated costimulation and the immune response. *Blood Rev*. 10:111–127. http://dx.doi.org/10.1016/S0268-960X(96)90040-5

Seifert, M., and R. Küppers. 2009. Molecular footprints of a germinal center derivation of human IgM+(IgD+)CD27+ B cells and the dynamics of memory B cell generation. *J. Exp. Med*. 206:2659–2669. http://dx.doi.org/10.1084/jem.20091087

Sims, G.P., R. Ettinger, Y. Shirota, C.H. Yarboro, G.G. Illei, and P.E. Lipsky. 2005. Identification and characterization of circulating human transitional B cells. *Blood*. 105:4390–4398. http://dx.doi.org/10.1182/blood-2004-11-4284

Sitte, S., J. Gläsner, J. Jellusova, F. Weisel, M. Panattoni, R. Pardi, and A. Gessner. 2012. JAB1 is essential for B cell development and germinal center formation and inversely regulates Fas ligand and Bcl6 expression. *J. Immunol*. 188:2677–2686. http://dx.doi.org/10.4049/jimmunol.1101455

Sohn, H.W., P.D. Krueger, R.S. Davis, and S.K. Pierce. 2011. FcRL4 acts as an adaptive to innate molecular switch dampening BCR signaling and enhancing TLR signaling. *Blood*. 118:6332–6341. http://dx.doi.org/10.1182/blood-2011-05-353102

Stamatopoulos, K., C. Belessi, C. Moreno, M. Boudjograh, G. Guida, T. Smilevska, L. Belhoul, S. Stella, N. Stavroyianni, M. Crespo, et al. 2007. Over 20% of patients with chronic lymphocytic leukemia carry stereotyped receptors: Pathogenetic implications and clinical correlations. *Blood*. 109:259–270. http://dx.doi.org/10.1182/blood-2006-03-012948

Subramanian, A., P. Tamayo, V.K. Mootha, S. Mukherjee, B.L. Ebert, M.A. Gillette, A. Paulovich, S.L. Pomeroy, T.R. Golub, E.S. Lander, and J.P. Mesirov. 2005. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. USA*. 102:15545–15550. http://dx.doi.org/10.1073/pnas.0506580102

Tandon, B., L. Peterson, J. Gao, B. Nelson, S. Ma, S. Rosen, and Y.H. Chen. 2011. Nuclear overexpression of lymphoid-enhancer-binding factor 1 identifies chronic lymphocytic leukemia/small lymphocytic lymphoma in small B-cell lymphomas. *Mod. Pathol*. 24:1433–1443. http://dx.doi.org/10.1038/modpathol.2011.103

van Oers, M.H., S.T. Pals, L.M. Evers, C.E. van der Schoot, G. Koopman, J.M. Bonfrer, R.Q. Hintzen, A.E. von dem Borne, and R.A. van Lier. 1993. Expression and release of CD27 in human B-cell malignancies. *Blood*. 82:3430–3436.

Weill, J.C., S. Weller, and C.A. Reynaud. 2009. Human marginal zone B cells. *Annu. Rev. Immunol*. 27:267–285. http://dx.doi.org/10.1146/annurev.immunol.021908.132607

Weller, S., M.C. Braun, B.K. Tan, A. Rosenwald, C. Cordier, M.E. Conley, A. Plebani, D.S. Kumararatne, D. Bonnet, O. Tournilhac, et al. 2004. Human blood IgM "memory" B cells are circulating splenic marginal zone B cells harboring a prediversified immunoglobulin repertoire. *Blood*. 104:3647–3654. http://dx.doi.org/10.1182/blood-2004-01-0346

Wolowiec, D., L. Ciszak, A. Kosmaczewska, D. Bocko, R. Teodorowska, I. Frydecka, and K. Kuliczkowski. 2001. Cell cycle regulatory proteins and apoptosis in B-cell chronic lymphocytic leukemia. *Haematologica*. 86:1296–1304.

Xie, L., A. Ushmorov, F. Leithäuser, H. Guan, C. Steidl, J. Färbinger, C. Pelzer, M.J. Vogel, H.J. Maier, R.D. Gascoyne, et al. 2012. FOXO1 is a tumor suppressor in classical Hodgkin lymphoma. *Blood*. 119:3503–3511. http://dx.doi.org/10.1182/blood-2011-09-381905

Yang, S.Y., S.D. Fugmann, and D.G. Schatz. 2006. Control of gene conversion and somatic hypermutation by immunoglobulin promoter and enhancer sequences. *J. Exp. Med*. 203:2919–2928. http://dx.doi.org/10.1084/jem.20061835

Ye, B.H., F. Lista, F. Lo Coco, D.M. Knowles, K. Offit, R.S. Chaganti, and R. Dalla-Favera. 1993. Alterations of a zinc finger-encoding gene, BCL-6, in diffuse large-cell lymphoma. *Science*. 262:747–750. http://dx.doi.org/10.1126/science.8235596

Zenz, T., D. Mertens, R. Küppers, H. Döhner, and S. Stilgenbauer. 2010. From pathogenesis to treatment of chronic lymphocytic leukaemia. *Nat. Rev. Cancer*. 10:37–50.