JCI The Journal of Clinical Investigation

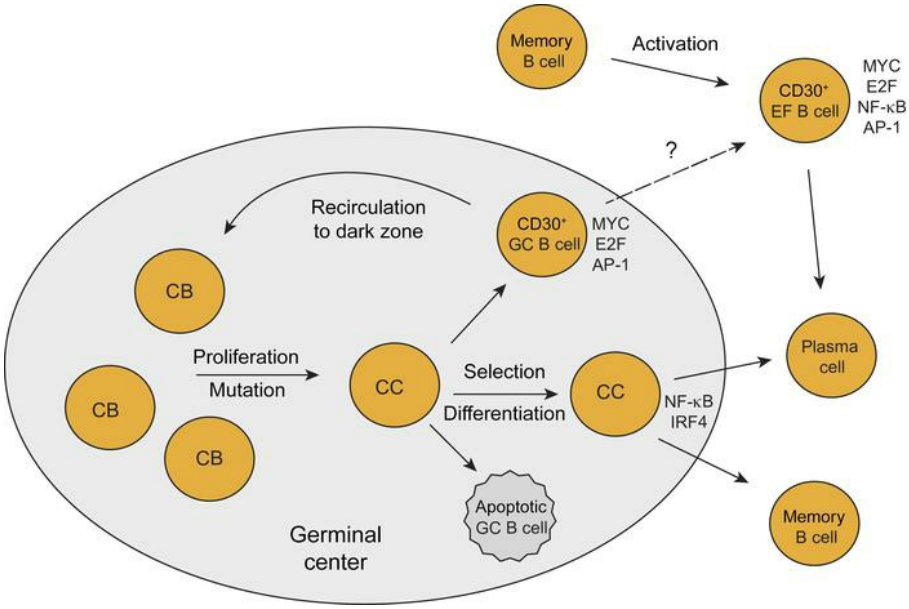
Human CD30+ B cells represent a unique subset related to Hodgkin lymphoma cells

Marc A. Weniger, … , Martin-Leo Hansmann, Ralf Küppers

J Clin Invest. 2018;128(7):2996-3007. https://doi.org/10.1172/JCI95993.

Research Article | Hematology | Immunology

Graphical abstract



Human CD30$^+$ B cells represent a unique subset related to Hodgkin lymphoma cells

Marc A. Weniger,[1] Enrico Tiacci,[1] Stefanie Schneider,[1] Judith Arnolds,[2] Sabrina Rüschenbaum,[1] Janine Duppach,[1] Marc Seifert,[1] Claudia Döring,[3] Martin-Leo Hansmann,[3,4] and Ralf Küppers[1]

[1]Institute of Cell Biology (Cancer Research), and [2]Department of Otorhinolaryngology, University of Duisburg-Essen, Essen, Germany. [3]Dr. Senckenberg Institute of Pathology, University of Frankfurt/Main, Medical School, Frankfurt, Germany. [4]Frankfurt Institute for Advanced Studies, Frankfurt, Germany.

Very few B cells in germinal centers (GCs) and extrafollicular (EF) regions of lymph nodes express CD30. Their specific features and relationship to CD30-expressing Hodgkin and Reed/Sternberg (HRS) cells of Hodgkin lymphoma are unclear but highly relevant, because numerous patients with lymphoma are currently treated with an anti-CD30 immunotoxin. We performed a comprehensive analysis of human CD30$^+$ B cells. Phenotypic and IgV gene analyses indicated that CD30$^+$ GC B lymphocytes represent typical GC B cells, and that CD30$^+$ EF B cells are mostly post-GC B cells. The transcriptomes of CD30$^+$ GC and EF B cells largely overlapped, sharing a strong MYC signature, but were strikingly different from conventional GC B cells and memory B and plasma cells, respectively. CD30$^+$ GC B cells represent MYC$^+$ centrocytes redifferentiating into centroblasts; CD30$^+$ EF B cells represent active, proliferating memory B cells. HRS cells shared typical transcriptome patterns with CD30$^+$ B cells, suggesting that they originate from these lymphocytes or acquire their characteristic features during lymphomagenesis. By comparing HRS to normal CD30$^+$ B cells we redefined aberrant and disease-specific features of HRS cells. A remarkable downregulation of genes regulating genomic stability and cytokinesis in HRS cells may explain their genomic instability and multinuclearity.

Introduction

CD30 (TNFRSF8) is a transmembrane glycosylated protein and member of the tumor necrosis factor receptor superfamily (1). Within the hematopoietic system, this receptor is expressed by only a few activated T and B cells in lymphoid tissues (2). CD30 is also expressed by some EBV-infected B cells in infectious mononucleosis, but is virtually absent on peripheral blood B cells (1–3). Little is known about the specific features of CD30$^+$ B cells. These cells are large mononuclear cells that are found within germinal centers (GCs), as well as in extrafollicular (EF) regions of lymph nodes (2, 4). CD30$^+$ EF B cells express the B cell transcription factor PAX5, the plasmablast/plasma cell factor IRF4, and JUNB, a member of the AP-1 transcription factor family (4, 5). A fraction of them expresses activation-induced cytidine deaminase (AID), a key factor for somatic hypermutation and class switch recombination (4, 5).

Being a hallmark of Hodgkin and Reed/Sternberg (HRS) tumor cells in classical Hodgkin lymphoma (cHL), CD30 is an important marker in lymphoma diagnosis (6). CD30-expressing tumor cells are also characteristic for primary mediastinal B cell lymphoma, a fraction of conventional diffuse large B cell lymphomas (DLBCLs), and anaplastic large cell lymphoma (1, 6–8). The expression of CD30 by malignant B and T cells in cHL and other lymphomas is not only of biological and diagnostic relevance, but has also recently gained major impact for the treatment of these tumors. An anti-CD30 antibody coupled to a microtubule inhibitor (brentuximab vedotin) shows promising results in the eradication of CD30$^+$ lymphoma cells and is currently being tested in numerous clinical studies (9–11).

We isolated human tonsillar CD30$^+$ B cells, characterized the cells phenotypically, analyzed their rearranged *IGHV* genes, and compared their global gene expression to that of the main subsets of normal mature B cells and of cHL HRS cells. We aimed to clarify the differentiation stage and specific features of normal CD30$^+$ B cells and their relationship to cHL HRS cells.

Results

Normal CD30$^+$ GC and EF B cells are mostly CD27$^+$ and class-switched. Previous immunohistochemical analyses recognized large CD30$^+$ B cells inside GCs and outside of follicles (2, 4). Accordingly, we distinguished CD30$^+$ GC B cells (CD20^{hi}CD38$^+$) and CD30$^+$ EF B cells (CD20$^+$CD38$^{lo/-}$) by flow cytometry (Figure 1A). Typically, only 0.1%–1.7% (mean 0.7%) of tonsillar mononuclear cells are CD30$^+$ B cells (Supplemental Table 1; supplemental material available online with this article; https://doi.org/10.1172/JCI95993DS1). We analyzed CD30$^+$ B cells for the expression of CD27, a marker for memory B cells, GC B cells, and plasma cells (12, 13). Most cells of both CD30$^+$ B cell subsets express CD27 levels similar to those in conventional GC and memory B cells (Supplemental Figure 1). The Ig isotype distribution of CD30$^+$ GC and EF B cells was largely similar (Supplemental Table 2): on average, about 50% of CD30$^+$ GC and EF B cells expressed IgG, and about 20% of both subsets are IgA$^+$ (Figure 1 and Supplemental Table 2). On average, IgM was expressed in

Authorship note: MAW and ET are co-first authors. CD, MLH, and RK are co-senior authors.
Conflict of interest: The authors have declared that no conflict of interest exists.
Submitted: June 27, 2017; **Accepted:** April 17, 2018.
Reference information: *J Clin Invest*. 2018;128(7):2996–3007.
https://doi.org/10.1172/JCI95993.

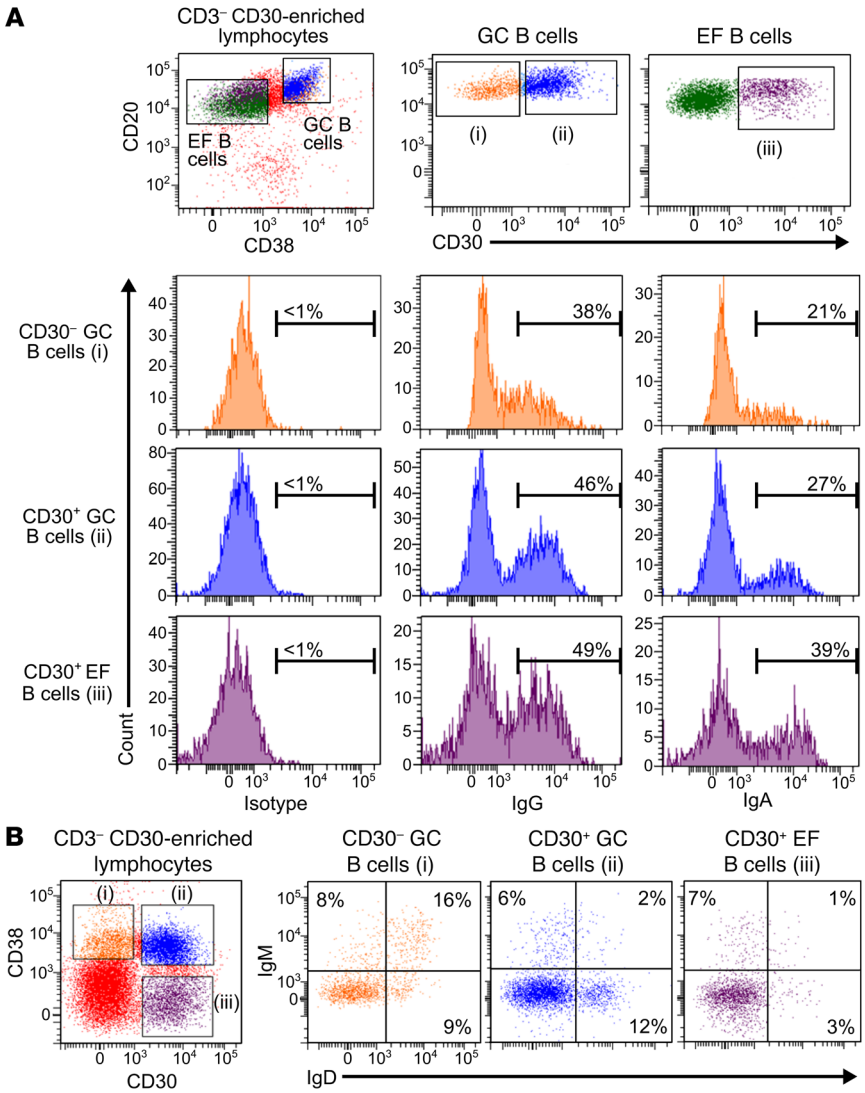


Figure 1. Phenotypic characterization of CD30⁺ B cells. Tonsillar mononuclear cells were depleted of CD3⁺ T cells and enriched for CD30⁺ B cells by consecutive MACS isolation steps. (**A**) CD3⁻ CD30–enriched B cells were stained for CD20, CD30, CD38, and either for IgG, IgA, or an isotype control. Gates defining CD30⁻ GC B cells (i), CD30⁺ GC B cells (ii), and CD30⁺ EF B cells (iii) are given. Histograms show fractions of IgG⁺, IgA⁺ , and isotype control–positive cells. (**B**) IgM and IgD expression on CD30⁺ B cell subsets. The percentages of IgM⁺ and/or IgD⁺ cells are given. Gates defining CD30⁻ GC B cells (i), CD30⁺ GC B cells (ii), and CD30⁺ EF B cells (iii) are shown on the left. The expression pattern of IgM and IgD for these 3 B cell subpopulations are depicted in the plots on the right.

9% of CD30⁺ GC and 22% of CD30⁺ EF B cells (Figure 1B). Most IgM⁺CD30⁺ B cells showed low levels of IgD. IgE was not detectable. The Ig isotype distribution of CD30⁺ GC B cells was similar to that of CD30⁻ GC B cells (Supplemental Table 2).

Taken together, CD27 and Ig isotype expression of CD30⁺ GC and EF B cells is very similar to that of conventional GC B cells and memory B cells, respectively.

Normal CD30⁺ B cells carry mutated IGHV genes. We sequenced rearranged *IGVH* genes from CD30⁺ GC and EF B cells ($n = 4$ each). Nearly all sequences obtained from CD30⁺ GC B cells were somatically mutated, with average mutation frequencies between 4.6% and 8% (Table 1). This is slightly higher than mutation frequencies typically observed for tonsillar GC B cells (14). The Ig framework region replacement-to-silent (R/S) ratio of mutations was lower than 2 in 3 of the 4 samples (2.4 in donor 3), in line with positive selection of a functional B cell receptor (BCR) (14). We found several clonally related V_H gene sequences in 3 of the 4 samples, indicating that CD30⁺ GC B cells can be members of expanded clones.

The *IGVH* gene analysis of CD30⁺ EF B cells from 3 donors showed between 79% and 92% mutated sequences, whereas cells from 1 donor had only 33% mutated *IGVH* gene sequences. Average mutation frequencies (2.9%–8.2%) and R/S ratios in the framework regions (1.3–1.8) were typical for memory B cells (Table 1) (14). Moreover, we identified 2 pairs of clonally related sequences among the CD30⁺ EF B cells, one of which belonged to a clone present among CD30⁺ GC B cells of the same tonsil, demonstrating its GC derivation.

Taking together, *IGHV* gene analysis revealed that CD30⁺ GC B cells represent typical, selected GC B cells, and the vast majority of CD30⁺ EF B cells represent post-GC (memory) B cells, although a minor fraction (or sometimes more) of the latter may be at a pre–GC stage of differentiation.

CD30⁺ GC and EF B cells have a unique gene expression pattern, characterizing them as distinctive B cell subsets with high MYC activity. We generated gene expression profiles of CD30⁺ GC and EF B cells ($n = 5$ each) and compared them with profiles of tonsillar naive B cells ($n = 5$), memory B cells ($n = 5$), plasma cells ($n = 5$), and conventional GC B cells ($n = 10$) (15, 16). By unsupervised hierarchical clustering, CD30⁺ B cells formed distinct branches, with CD30⁺ GC and EF B cells separated in independent subbranches, confirming their distinctness (Figure 2A). Separate clustering of CD30⁺ and conventional GC B cells suggested a specific dif-

Table 1. V_H gene mutation analysis of CD30⁺ GC and EF B cells

Cell type/Donor	V_H- family	Mutated sequences/ total sequences (%)	Average mutation frequency in %[A] (range)	R/S ratio in FR1-3[B]	Crippled sequences[C]	No. of clones	Cells per clone
CD30⁺ GC B cells							
1	3	23/23 (100)	5.6 (0.7–11.8)	1.7	1/23	4[D]	2–3
2	3	23/23 (100)	8 (3.4–14.2)	1.9	1/23	0	0
3	3	28/28 (100)	4.6 (1.4–9.8)	2.4	0/28	2	2 and 3 (+1 CD30⁺ EF B cell)[E]
4	1, 3, 4	39/40 (98)	4.9 (0.7–17)	1.4	0/40	1	2
CD30⁺ EF B cells							
3	3	15/19 (79)	4.3 (1.0–9.2)	1.7	0/16	1	1 (+3 CD30⁺ GC B cells)[E]
4	1, 3, 4	36/39 (92)	5.2 (0.7–12.9)	1.4	0/39	2	2
5	3	18/22 (82)	8.2 (1.7–20.3)	1.3	0/18	0	0
6	3	6/18 (33)	2.9 (0.7–9.3)	1.8	1/11	0	0

[A]For calculation of the average mutation frequency, sequences with only 1 mutation were not considered as mutated, as polymerase errors cannot be excluded. Identical sequences were counted once, as they might derive from 1 cell. In case of intraclonal diversity, each unique sequence was counted once. Both in-frame and out-of-frame rearrangements and only mutated sequences were considered. Insertions and deletions were calculated as 1 mutation. [B]For calculation of the R/S ratio, only in-frame rearrangements were considered. [C]We analyzed the sequences for crippling mutations that render originally functional IgV region genes nonfunctional, as such crippling mutations are seen in 25% of cHL cases (33). However, only 2 CD30⁺ GC B cells and 1 CD30⁺ EF B cell carried such mutations. Therefore, we cannot exclude that these represent PCR errors. [D]This donor might contain 2 additional clones with 2 cells per clone each. As they carry only 1 distinct mutation, we cannot exclude a PCR error and therefore they are not counted as distinct clones. [E]One clone of donor 3 consists of 3 cells among the CD30⁺ GC B cells and 1 cell among the CD30⁺ EF B cells.

ferentiation stage of CD30⁺ GC B cells, characterized by further upregulation of genes already higher expressed in conventional GC B cells than in naive B, memory B, and plasma cells. Similarly, CD30⁺ EF B cells clearly separated from memory B cells and plasma cells, driven by the same set of genes highly expressed in CD30⁺ GC B cells, reflecting a common gene expression signature in CD30⁺ B cells (Figure 2A). In an unsupervised principal component analysis (PCA), CD30⁺ B cell subsets were distinct from each other, and were even more distinct from the other B cell subsets (Figure 2B). Using genes best distinguishing conventional GC B cells from either naive or memory B cells or plasma cells, CD30⁺ B cells consistently located closer to conventional GC B cells than to any of the other B cell subsets in supervised PCAs (Figures 2, C–E). Hence, CD30⁺ B cells are characterized by a gene expression pattern more similar to the one of proliferative GC B cells than to patterns of resting naive, memory, and plasma cells.

We then sought to identify genes differentially expressed between the 2 CD30⁺ B cell subsets. Many of the 19 genes with at least 3-fold higher expression in CD30⁺ GC B cells encode well-known GC B cell markers (e.g., BCL6, MYBL1, LCK, RGS13, S1PR2, and KLHL6) ($P < 0.05$, false discovery rate [FDR] < 0.05, Supplemental Table 3), validating that CD30⁺CD38⁺-sorted B cells are indeed GC B cells. In contrast, many of the 33 genes higher expressed in CD30⁺ EF B cells are typically higher expressed in naive or post-GC B cells than in GC B cells, e.g., *FCRL5*, *ENTPD1 (CD39)*, *CD44*, *GPR183 (EBI2)*, and *SELL* (selectin L). EBI2 is a G protein–coupled receptor guiding B cells to the T-B border and outer follicle zones (17), and its higher expression in CD30⁺ EF B cells might contribute to their relocation outside of B cell follicles. By gene set enrichment analysis (GSEA), several transcription factor motif–based target gene sets, including NF-κB, AP-1, and MEF2, were significantly enriched in CD30⁺ EF B cells (Supplemental Table 4; $P < 0.01$, FDR < 0.1), as were gene sets of Toll-like receptor, NF-κB, MAP kinase, G protein–coupled receptor, and PI3K/AKT canonical pathways. No gene set was significantly enriched in CD30⁺ GC B cells.

Despite their GC B cell identity, CD30⁺ GC B cells clearly separated from conventional GC B cells (Figure 2A), so we defined the specific features of the former. Fifteen genes were significantly lower expressed in CD30⁺ GC B cells (fold change [FC] ≥ 3, $P < 0.01$, FDR < 0.05; Supplemental Table 5), including well-known B cell differentiation markers such as FCRL3, FCRL1, CD24, FOXP1, AICDA, and BACH2. Of note, the latter 3 genes are also downregulated during centroblast to centrocyte transition (18–20). Ninety-eight genes were significantly higher expressed in CD30⁺ GC B cells, including *IL2RB*, *CCL22*, *BATF3*, *MYC*, the MYC cooperating factor *CCDC86*, and the antiapoptotic factors LRPPRC and PRELID1. Increased expression of IL2RB by CD30⁺ GC B cells was validated at the protein level by flow cytometry (Supplemental Figure 2A). Also, mRNA expression of IL21R, another key factor for B cell–T cell interaction in the GC, was significantly upregulated by CD30⁺ GC B cells (albeit only 1.8-fold) and hence analyzed by flow cytometry. Indeed, CD30⁺ GC B cells showed significantly higher IL21R surface expression than conventional GC B cells (Supplemental Figure 2B). IL2RB and CCL22 are upregulated in centrocytes upon T helper (Th) cell interaction, suggesting a link between CD30⁺ GC B cells and centrocytes (21, 22). In agreement, centrocyte-specific genes were more highly expressed than centroblast-specific genes by GSEA in CD30⁺ GC B cells relative to conventional GC B cells (Supplemental Figure 3A). Moreover, *MYC*, a master regulator of proliferation and metabolism (23), and the MYC cooperating factor

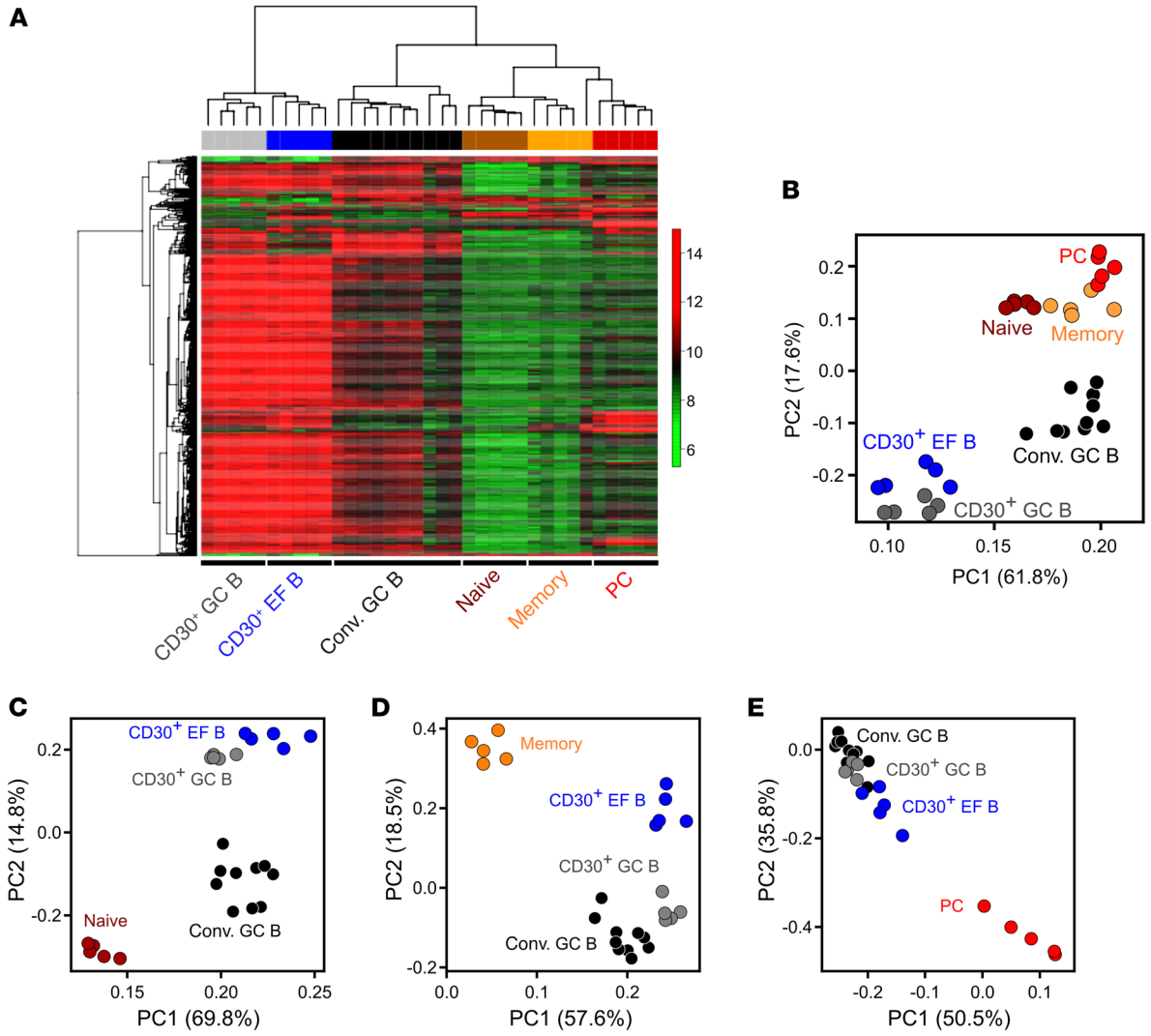


Figure 2. Unsupervised hierarchical clustering and PCA of normal human B cell subsets. (**A**) Unsupervised hierarchical clustering was performed on 683 probe sets with SD ≥ 1. For calculating the distance matrix we used the Manhattan distance method. The dendrogram was generated with the average linkage method from the R package geneplotter. (**B**) PCA was conducted on 683 probe sets with SD ≥ 1. The figure shows the first 2 principal components. The first principal component covers 61.8% of the variance, the second one 17.6%. (**C**) Supervised PCA was performed using 229 differentially expressed probe sets with FC ≥ 4 or ≤ −4 and FDR ≤ 0.05 between GC and naive B cells. (**D**) Supervised PCA was performed using 128 differentially expressed probe sets with FC ≥ 4 or ≤ −4 and FDR ≤ 0.05 between conventional GC and memory B cells. (**E**) Supervised PCA was performed using 130 differentially expressed probe sets with FC ≥ 4 or ≤ −4 and FDR ≤ 0.05 between GC B cells and plasma cells. In **B**–**E**, the first 2 principal components are shown. Conv. GC, conventional GC B cells; PC, plasma cells; PC1/PC2, principal component 1/2.

CCDC86 were upregulated in CD30$^+$ GC B cells along with many genes involved in nucleotide metabolism, cell proliferation, protein biosynthesis, or mitochondrial function. A GSEA comparing CD30$^+$ and conventional GC B cells identified several E2F and MYC DNA-binding motif–based target gene sets, as well as gene sets of cell cycle and replication pathways, DNA repair, antigen processing, and transcription as enriched in CD30$^+$ GC B cells ($P < 0.01$, FDR < 0.3; Supplemental Table 4). No gene set was significantly upregulated in conventional GC B cells. As it is well-known that GC B cells quickly undergo apoptosis when cultured in vitro (24), presumably reflecting their strong dependency on microenvironmental survival signals, we wondered whether the CD30$^+$ GC B cells show the same feature. Indeed, CD30$^+$ GC B cells massively died after several hours of culturing in vitro (Supplemental Figure 4A). Collectively, CD30$^+$ GC B cells display late centrocyte, but simultaneously proliferative or proliferation-prone, features in comparison with conventional GC B cells.

IgV gene and flow cytometric analyses revealed a post-GC B cell differentiation stage of most CD30$^+$ EF B cells, prompting their comparison with memory B cells and post-GC plasma cells. Many of the 253 genes with at least 5-fold higher expression in CD30$^+$ EF B cells than in memory B cells play roles in proliferation, mitosis, metabolism, mitochondrial activity, and proteasome function (FDR < 0.05; Supplemental Table 6). *MYC* was significantly upregulated in CD30$^+$ EF B cells by approximately 2-fold. By GSEA, several E2F, MYC, and AP-1 motif–based target gene sets, as well as several metabolic signaling pathways and functionally determined STAT3 and STAT6 gene sets, were significantly

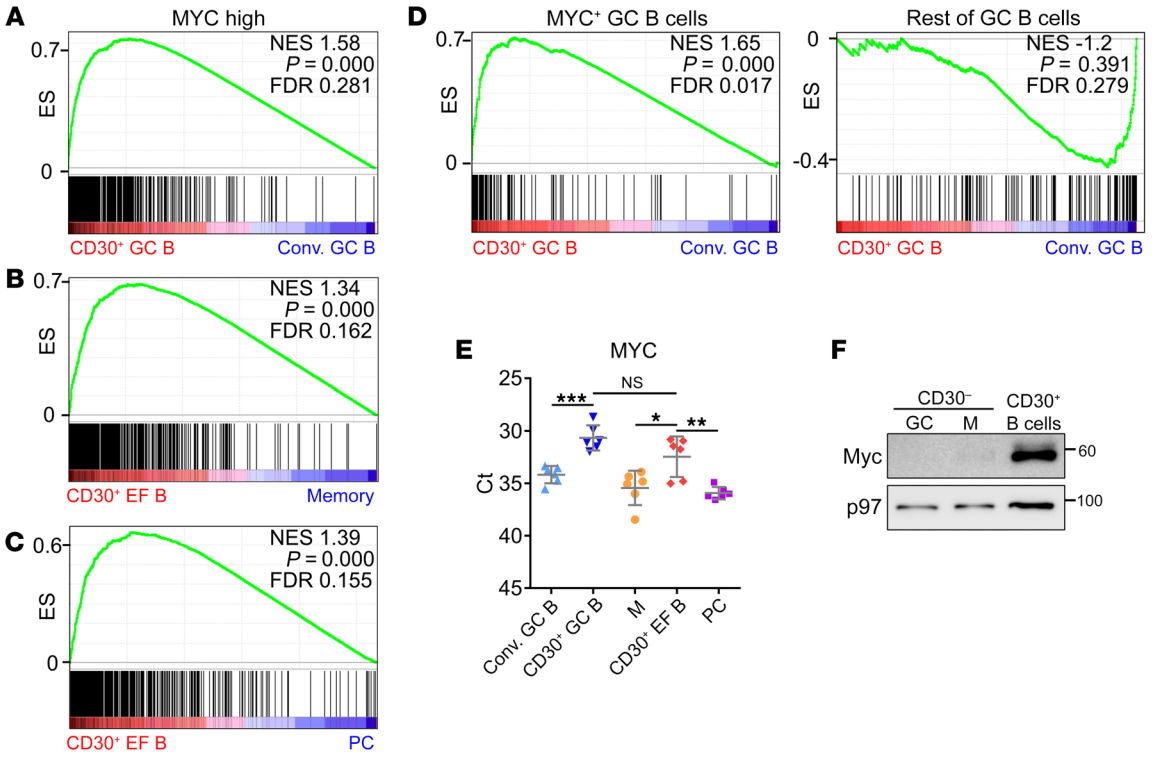


Figure 3. MYC expression and activity in CD30⁺ B cells. Expression of MYC and MYC target genes in CD30⁺ and conventional B cells. Enrichment plots of functionally validated MYC target genes (482 Myc high) (69) for comparisons of (**A**) CD30⁺ versus conventional (Conv.) GC B cells, (**B**) CD30⁺ EF B cells versus memory B cells, and (**C**) CD30⁺ EF B cells versus plasma cells (PC). (**D**) A comparison of CD30⁺ GC B cells versus conventional GC B cells is shown for a gene set specifically enriched in murine MYC⁺ GC B cells, and a gene set enriched in murine MYC⁻ GC B cells (42). (**E**) MYC expression in tonsillar CD30⁻ cells (M, memory B cells; PC, plasma cells) and CD30⁺ GC and EF B cell subsets by RT-qPCR (n = 6 donors). Ct values are shown. Graphs indicate mean ± SD. (**F**) MYC protein expression in CD30⁺ B cells, CD30⁻ GC B cells, and memory (M) B cells. A representative image of 2 independent experiments is shown. $^*P < 0.05$; $^{**}P < 0.01$, $^{***}P < 0.001$ (unpaired, 2-tailed t test).

enriched in CD30⁺ EF B cells (Supplemental Table 4), reflecting increased metabolic demands of the highly active and proliferating CD30⁺ EF B cells. In addition, genes engaged in T cell interaction(s) were upregulated in CD30⁺ EF B cells: CCL22 and CXCL10 attract CD4⁺ Th cells (25, 26), and IL2RB and IL21R are receptors for interleukins produced by Th cells. By flow cytometry, we confirmed higher surface expression of IL2RB and IL21R on CD30⁺ EF B cells than on memory B cells (Supplemental Figure 2). Of note, some antiapoptotic factors were highly expressed by CD30⁺ EF B cells, such as BIRC5, which is upregulated in plasmablasts (27), and the chemokine macrophage migration inhibitory factor (MIF). MIF binds to the CD44/CD74 complex on B cells and promotes their survival (28). CD44 is significantly higher expressed in CD30⁺ EF B cells than in memory B cells by approximately 2-fold, hence MIF may promote survival of CD30⁺ EF B cells in an autocrine manner. To assess the intrinsic survival capacity of CD30⁺ EF B cells in comparison with memory B cells, we isolated CD27⁺ EF B cells and determined their survival upon culturing in vitro. The fraction of CD30⁺ B cells among surviving cells increased over time, indicating a slightly enhanced survival capacity of these cells compared with conventional memory B cells (Supplemental Figure 4B). Among 44 genes significantly downregulated in CD30⁺ EF B cells relative to memory B cells (FC ≥ 3, FDR ≤ 0.05; Supplemental Table 6), we identified known B cell differentiation factors, including CD24, which is downregulated upon plasma cell differentiation (29).

Also in comparison to plasma cells, CD30⁺ EF B cells have many more genes upregulated than downregulated (Supplemental Table 7). Again, numerous genes involved in proliferation are higher expressed by CD30⁺ EF B cells, as are E2F and MYC motif–based gene sets (Supplemental Table 4 and Supplemental Figure 3, B–D). Plasma cells have significantly higher levels of Ig heavy and light chain transcripts and the typical plasma cell markers *XBP1* and *CD38*. Collectively, CD30⁺ EF B cells display characteristics of T cell interactions, cell survival, and proliferative or proliferation-prone features.

Recently, a distinct subset of human CD21^{lo} post-GC B cells was described that are recent GC graduates primed for plasma cell differentiation (30). The CD30⁺ EF B cells also show lower CD21 surface expression than conventional memory B cells (Supplemental Figure 2C). However, for other key features, the 2 types of B cells clearly differ from each other (Supplemental Figure 2D). We therefore conclude that the CD30⁺ EF B cells are distinct from the CD21^{lo} recent GC emigrants.

We confirmed high MYC activity in CD30⁺ versus conventional GC B cells, and in CD30⁺ EF versus memory B cells or plasma cells (Figure 3, A–D), using functionally validated sets of MYC target genes in human B cells and a set of genes upregulated

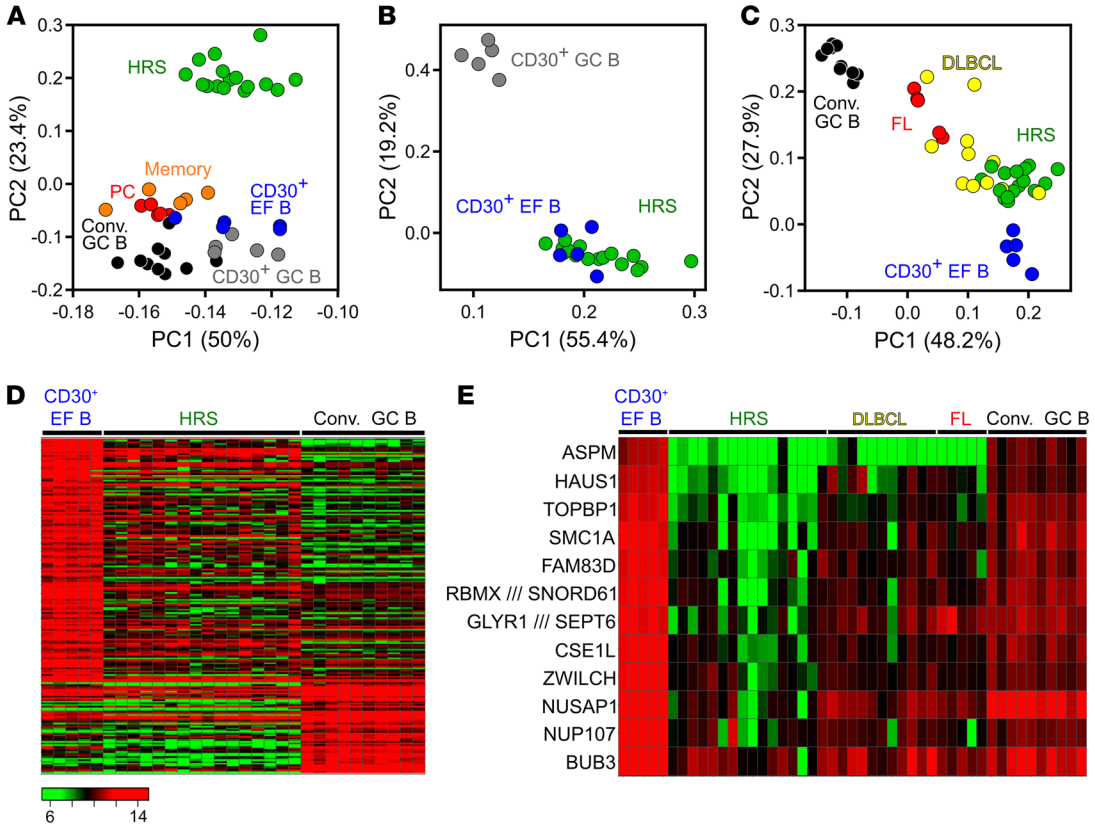


Figure 4. PCA and heat maps of the comparison of HRS cells to CD30⁺ and CD30⁻ B cells. (**A**) Unsupervised 2D PCA of probe sets with a SD > 1 (765 probe sets) from the comparison of HRS cells to normal GC and post-GC B cell subsets. (**B**) Supervised PCA of CD30⁺ B cells versus HRS cells, based on genes differentially expressed between CD30⁺ GC and CD30⁺ EF B cells (254 probe sets; FC ≥ 2, FDR ≤ 0.05). (**C**) Two-dimensional PCA based on probe sets at least 4-fold differentially expressed between CD30⁺ EF and conventional (Conv.) GC B cells (70 probe sets). Besides these 2 normal B cell subsets, HRS cell samples as well as FL and DLBCL are displayed. (**D**) Heat map of a supervised analysis of genes differentially expressed between conventional GC B cells and CD30⁺ EF B cells (199 probe sets, FC ≥ 3). HRS cell samples are shown for comparison. (**E**) Heat map shows the expression of 12 of 169 genes down-regulated at least 5-fold in HRS cells in comparison with CD30⁺ EF B cells that are involved in the regulation of the spindle apparatus, cytokinesis, and polyploidy. Samples of conventional GC B cells as well as of FL and DLBCL are included for comparison. Scale as in **D**.

in murine MYC⁺ GC B cells (not included in the Broad Institute collections initially used). Consistently, both CD30⁺ B cell subsets expressed higher MYC mRNA and protein levels than conventional GC, memory B, or plasma cells, thus establishing high expression of MYC and MYC signatures as a distinctive feature of human CD30⁺ B cells (Figure 3, E and F).

Finally, we validated lower BACH2 and FOXP1 expression in CD30⁺ than in CD30⁻ GC B cells, high CCL22 expression in both CD30⁺ B cell subsets, a tendency for higher EBI2 transcript levels in CD30⁺ EF than CD30⁺ GC B cells, and higher expression of PRDM1 in CD30⁺ EF B cells than in GC or memory B cells, but lower expression than in plasma cells (Supplemental Figure 5).

Taken together, CD30⁺ GC B cells show centrocytic characteristics paired with high activation and proliferation features and a strong MYC signature. The MYC-associated, proproliferative gene expression signature is shared by CD30⁺ EF B cells.

Normal CD30⁺ B cells share many gene expression patterns with HRS cells, but also show critical differences linked to cHL pathogenesis. The transcriptome of primary HRS cells is very different from that of conventional GC and post-GC B cells (16), but it has not been compared with gene expression profiles of normal CD30⁺ B cells. In the first component of an unsupervised PCA, HRS cells were more similar to CD30⁺ B cells than to conventional GC B cells, memory B cells, or plasma cells (Figure 4A). HRS cells were widely distributed in the first component, likely reflecting their known heterogeneity in terms of gene expression (31, 32). In a supervised PCA (FC ≥ 2, FDR ≤ 0.05) focusing on genes differentially expressed between CD30⁺ GC and EF B cells, HRS cells were closely positioned to and partly intermingled with CD30⁺ EF B cells, and clearly separated from CD30⁺ GC B cells (Figure 4B). The relatedness of CD30⁺ EF B cells to HRS cells was confirmed in a further supervised PCA using probe sets with at least 4-fold differential expression between conventional GC B cells and CD30⁺ EF B cells (Figure 4C). For comparison, samples of follicular lymphoma (FL) as prototypical GC B cell–derived lymphoma and DLBCL were included (Figure 4C). FL samples clustered closer to conventional GC B cells, as expected, while DLBCL samples showed a more heterogeneous distribution, in line with the known heterogeneity of this type of lymphoma. Hence, HRS cells display a gene expression pattern more similar to CD30⁺, and in particular CD30⁺ EF B cells, than to other B cell subsets, which is also visible in a heatmap based on genes differentially expressed between CD30⁺ EF B cells and conventional GC B cells (Figure 4D).

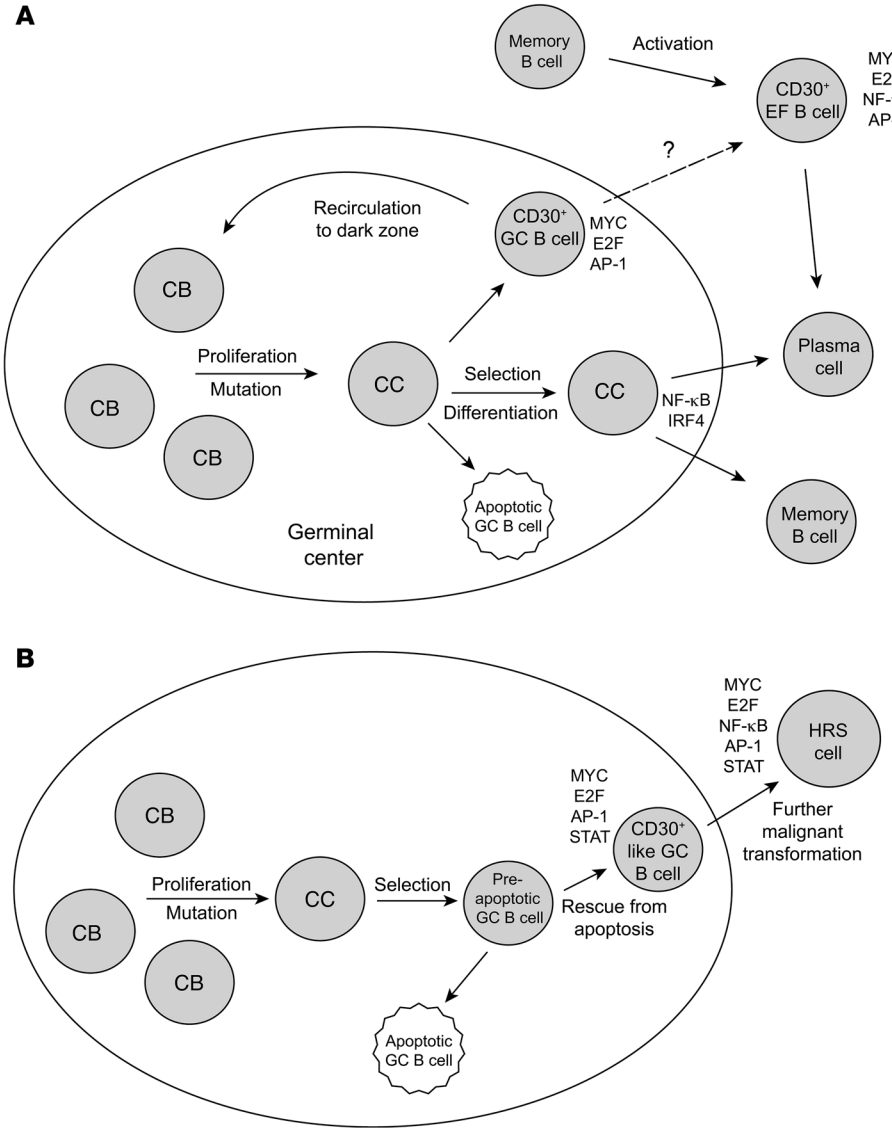


Figure 5. Scenarios for the generation of normal CD30⁺ B cells and HRS cells. (**A**) The key features of CD30⁺ GC B cells as outlined in the text indicate that CD30⁺ GC B cells represent the positively selected centrocytes (CC) that are preparing to return to the dark zone to become centroblasts (CB) again and undergo an additional round of proliferation and selection. High activity of MYC and E2F and signatures for AP-1 activity are hallmarks of these cells. CD30⁺ EF B cells likely represent reactivated memory B cells that undergo proliferation before they presumably differentiate into plasma cells. It is unclear whether CD30⁺ GC B cells may also directly differentiate into CD30⁺ EF B cells. (**B**) Scenario for the generation of HRS cells. Genetic features of HRS cells (crippled IgV genes in at least a quarter of cases) strongly indicate that these cells derive from preapoptotic GC B cells that were rescued from apoptosis by some transforming events (e.g., EBV infection in a fraction of cases). The similarities in their gene expression between HRS and CD30⁺ B cells indicate that HRS precursor cells differentiated into the direction of CD30⁺ B cells in the course of their further malignant transformation. Mainly because of the downregulation of GC B cell–specific genes and the strong constitutive NF-κB activity, HRS cells more closely resemble CD30⁺ EF B cells than CD30⁺ GC B cells.

Next, we sought to identify genes differentially expressed between CD30⁺ EF B cells and HRS cells. We found 169 genes with at least 5-fold higher expression in CD30⁺ EF B cells and 73 genes with at least 3-fold higher expression in HRS cells (Supplemental Table 8). These genes largely overlap with those deriving from the comparison of CD30⁺ GC B cells and HRS cells, which was expected, given the high overall similarity of the 2 normal CD30⁺ B cell subsets (Supplemental Table 9). Genes expressed at higher levels in HRS cells reflected well-known pathophysiological features of these cells, including components and regulators of the extracellular matrix (e.g., collagen, fibronectin, metallopeptidases, and cathepsins), numerous chemokines (e.g., CCL3, CCL5, CCL17), and several non–B cell markers (e.g., CD3δ, LAT, granzyme B, and the NK-cell factor ID2) (33).

The genes higher expressed in CD30⁺ EF B cells than in HRS cells included numerous B cell factors (e.g., CD20, BOB1, EBF1, BLNK) (Supplemental Table 8), in line with the downregulation of the B cell expression program in HRS cells (34). Strikingly, 30 of the 169 genes at least 5-fold lower expressed in HRS than in CD30⁺ EF B cells play roles in cell-cycle and spindle apparatus, genomic stability, cytokinesis, DNA repair, or polyploidy (Supplemental Table 8). Multinuclearity is a hallmark of the Reed/Sternberg cells, hence we compared the expression of 12 genes regulating spindle apparatus, cytokinesis, and polyploidy in CD30⁺ EF B and HRS cells with their expression in conventional GC B cells, DLBCL, and FL (Figure 4E). Downregulation of these important regulators in HRS cells is also seen in comparison to conventional GC B cells, and is more pronounced in HRS cells than in the 2 other types of lymphomas included for comparison (Supplemental Table 8). Thus, the strong downregulation of key factors regulating cytokinesis and the spindle apparatus is a specific feature of HRS cells.

We then compared CD30⁺ GC and EF B cells with HRS cells by GSEA to identify pathway and transcription factor gene sets differentially expressed between these cells. In line with the lost B cell phenotype of HRS cells, several B cell transcription factor–related gene sets (e.g., PAX5, MYB, YY1, and ERG) were significantly enriched in CD30⁺ B cells, as were E2F and MYC signatures (Supplemental Table 10). Moreover, motif-based gene sets of AP-1 family members MAF and ATF3 were enriched in CD30⁺ B cells, suggesting strong AP-1 activity in these cells (Supplemental Table

10). Similarly, individual STAT5a and STAT6 motif-based target gene sets were significantly enriched in CD30$^+$ B cells. HRS cells have strong STAT5 and STAT6 activity (35–37), but STAT signaling may be stronger in normal CD30$^+$ B cells. Taken together, HRS cells prominently downregulate genes associated with regulation and fidelity of cell division, and gene sets of transcription factors with known activities in HRS cells were also enriched in CD30$^+$ B cells.

Discussion

CD30$^+$ B cells are an enigmatic B cell population, potentially related to several types of B cell malignancies. We comprehensively characterized these cells and identified distinctive features of CD30$^+$ B cells. We distinguished CD30$^+$ GC and EF B cells and compared them to their respective conventional normal B cell subsets, as well as to HRS cells of cHL to assess their relatedness and the pathophysiological features of the latter.

CD30$^+$ GC B cells carry mutated IgV genes and are positively selected for BCR expression during the GC reaction. Similar to conventional GC B cells, CD30$^+$ GC B cells are CD27$^+$ and mostly class-switched to IgG or IgA. Transcriptome analysis of CD30$^+$ GC B cells revealed that they express key GC B cell factors, validating their GC B cell identity. Their striking dissimilarity to conventional GC B cells was mainly due to their upregulation of hundreds of genes. A major reason for this is likely their strong expression and activity of MYC, which is an amplifier of gene expression (38, 39). Expression of MYC by CD30$^+$ GC B cells has been reported recently; however, that study was restricted to immunohistochemical analysis and did not investigate MYC activity or the overall gene expression pattern of these cells (40). Here we show that MYC activity in CD30$^+$ GC B cells links them to a specific transitional state of GC B cell differentiation. Transgenic mouse studies characterized MYC$^+$ centrocytes, a small subset of positively selected light zone cells in established GCs that is preparing for a return to the dark zone to become proliferating centroblasts again (41, 42). Our findings support the idea that human CD30$^+$ GC B cells are positively selected centrocytes preparing to become proliferating centroblasts again (Figure 5A). First, genes involved in B cell–T cell interaction, which is essential for B cell selection in the light zone, were higher expressed in CD30$^+$ GC than conventional GC B cells. Second, centrocyte-specific genes were higher expressed in CD30$^+$ GC B cells than in conventional GC B cells (remarkably despite the many proliferative genes expressed by CD30$^+$ GC B cells). Third, factors that are downregulated in late GC B cells (AID, BACH2, MYBL1) (18, 19) were also expressed at lower levels in CD30$^+$ than in conventional GC B cells. Fourth, prosurvival factors such as LRPPRC and PRELID1 (43, 44) were expressed at elevated levels in CD30$^+$ GC B cells, potentially reflecting positive selection of these cells. Fifth, CD30$^+$ GC B cells upregulate many physiological genes, preparing them for high proliferation and metabolic demands. In addition, immunohistochemical studies for the location of CD30$^+$ and/or MYC$^+$ human GC cells revealed that these cells are mostly positioned in the apical light zone and at the periphery of the GC (1, 4, 40, 45, 46), in line with their identity as centrocytes initiating their return to the dark zone. We conclude that human CD30$^+$ GC B cells represent MYC$^+$ centrocytes, positively selected and undergoing a cyclic reentry into the centroblast pool, thereby sustaining the GC reaction. Of note, a subset of murine centrocytes expressing MYC and the transcription factor AP-4 has higher transcript levels of CD30 than MYC$^-$ GC B cells (47) (not mentioned in the text of that study, but retrievable from the RNA-Seq data), suggesting that CD30 may also serve as a surface marker for these cells in the mouse.

An additional feature of CD30$^+$ GC B cells is upregulated expression of the transcription factor BATF3, which is a member of the AP-1 transcription factor family, as is JUNB, which is also expressed by CD30$^+$ B cells (4). Thus, AP-1 activity appears to be a specific feature of CD30$^+$ GC B cells. Notably, AP-1 is an activator of CD30 transcription (48, 49). Therefore, elevated AP-1 activity in CD30$^+$ GC B cells likely is a main factor for their CD30 expression.

The phenotypic and IgV gene analysis of CD30$^+$ EF B cells showed that they have a post-GC identity, because most of them carry mutated IgV genes, are CD27$^+$, and are class-switched to IgG or IgA. Transcriptome analysis revealed, however, that CD30$^+$ EF B cells are very distinct from memory B cells and plasma cells. A main factor for the specific gene expression pattern of CD30$^+$ EF B cells was high MYC activity and a high proliferative and metabolic signature. A number of the genes highly expressed by CD30$^+$ EF B cells indicate an interaction with CD4$^+$ Th cells, i.e. production of CCL22 and CXCL10, which attract CD4$^+$ Th cells (25, 26), and the upregulation of IL2RB and IL21R may render the CD30$^+$ B cells more responsive to IL2 and IL21, which are produced by Th cells. A T cell–dependent stimulation of AID induction could also explain the high expression of AID in CD30$^+$ EF B cells, which may promote class-switching in these cells. As strong MYC activity often goes along with increased propensity for apoptosis, this feature of MYC presumably needs to be counteracted. The upregulation of the antiapoptotic factors BIRC5 and MIF in CD30$^+$ EF B cells may represent such antiapoptotic mechanisms. In the GSEA of CD30$^+$ EF B cells, a top signature in comparison with memory B cells was the AP-1 transcription factor signature, indicating that AP-1 activity is a hallmark also for this CD30$^+$ B cell subset, and might hence be a main factor for its CD30 expression.

We conclude that CD30$^+$ EF B cells are mostly post-GC memory B cells and are highly activated and proliferative. Their activation likely involves interaction with Th cells. Thus, they appear to represent reactivated memory B cells undergoing extra-GC proliferation, perhaps before they then undergo plasma cell differentiation, as IRF4 and BLIMP1 transcripts are upregulated, but BLIMP1 protein is not (yet) expressed (50). Considering the extensive differences of both subsets of CD30$^+$ B cells to the other main mature B cell subsets, and their unique features, it can be concluded that these cells represent distinct B cell subsets. However, as opposed to most other mature B cell subpopulations, the existence of CD30$^+$ B cells at specific differentiation stages is very dynamic and transitory, with CD30$^+$ EF B cells presumably existing a few days and CD30$^+$ GC B cells only a few hours before they further differentiate into CD30$^-$ B cells.

cHL HRS cells are derived from mature B cells, but show a transcriptome very different from all major mature B cell subsets (16). Our analysis revealed that the gene expression pattern of HRS cells is much more similar to the expression pattern of CD30$^+$ B cells than to the patterns of other main B cell subsets (Figure 4). Notably, although genetic analyses revealing crippling IgV gene mutations in HRS cells of 25% of cHL cases clearly indicate that HRS cells are

derived from apoptosis-prone GC B cells (33), in terms of their gene expression, HRS cells showed a much higher similarity to CD30$^+$ EF than GC B cells. This is mainly due to downregulation of GC B cell–specific genes both in CD30$^+$ EF B cells and HRS cells, in comparison with CD30$^+$ GC B cells (Supplemental Table 3 and Supplemental Table 8). Moreover, CD30$^+$ EF B cells and HRS cells share a strong NF-κB signature, which is less pronounced in CD30$^+$ GC B cells (Supplemental Table 4) (51). NF-κB and JAK/STAT activities in HRS cells are not only mediated by stimuli from other cells in the microenvironment but are also due to numerous genetic lesions and/or EBV infection in a fraction of cases (32). Moreover, MYC activity in HRS cells is at least partly mediated through STAT3/5 activation via induction of the AP-1 family member BATF3, which directly induces MYC transcription in HRS cells (52). In the course of their malignant transformation, GC B cell–specific features were lost by HRS cells and NF-κB activity was increased due to genetic lesions, so that they resemble post-GC EF CD30$^+$ B cells more than CD30$^+$ GC B cells. Thus, we propose a scenario in which HRS cells derive from crippled GC B cells that manage to escape apoptosis by differentiating into the direction of CD30$^+$ B cells and acquiring key features of these cells (Figure 5B). In this sense, CD30$^+$ B cells can be considered the origin of HRS cells.

Assuming that CD30$^+$ B cells are the specific cellular origin of HRS cells, a comparison of these cells can reveal which features of HRS cells reflect physiological characteristics of their cellular origin and which hallmarks of HRS cells can be considered as aberrant and disease-specific. MYC activity, as known for HRS cells (16, 53), is also seen in normal CD30$^+$ B cells, and at even higher levels. Similarly, STAT activities, which are hallmarks of HRS cells (36, 37, 54), now turn out not to be aberrant, as normal CD30$^+$ B cells can also show STAT activity. A further similarity between CD30$^+$ normal B cells and HRS cells is expression and presumably activity of AP-1 family members. JUN, JUNB, and ATF3 expression and their role for survival of HRS cells are well-known (55, 56). Moreover, we recently identified high expression and activity of BATF3 as being essential for survival of HRS cells (52). As CD30$^+$ B cells also express BATF3 as well as JUNB (4), presumed activity of the AP-1 pathway is an additional similarity between CD30$^+$ B cells and HRS cells. Although AP-1, STAT, and MYC activities are shared by CD30$^+$ B cells and HRS cells, they are nevertheless of major pathophysiological impact for cHL and are deregulated in HRS cells, as normal CD30$^+$ B cells represent a transitory differentiation stage, whereas HRS cells show these activities constitutively.

Downregulation of numerous B cell differentiation factors is a specific feature of HRS cells in comparison with normal CD30$^+$ B cells, confirming that the "lost B cell identity" of HRS cells (34) is indeed an aberrant feature of those cells. Also, expression of numerous markers of non–B cells by HRS cells is confirmed here as HRS cell–specific, as normal CD30$^+$ B cells lacked expression of such genes. In addition, the strong expression of numerous chemokines, immune suppressive factors, and regulators of the extracellular matrix is seen only in HRS but not normal CD30$^+$ B cells. Hence, the capacity of HRS cells to remodel their microenvironment by attracting other immune cells and by altering the extracellular matrix are characteristic and specific features of HRS cells.

As a further feature of HRS cells, we identified strong downregulation of numerous key regulators of the cell cycle, spindle apparatus, and cytokinesis (Figure 4E). HRS cells show high genomic instability, and the generation of bi- or multinucleated Reed/Sternberg cells by a process of incomplete cytokinesis and refusion of daughter cells is a hallmark of cHL (57, 58). Hence, suppression of these factors may, perhaps cooperatively, account for this key pathobiological feature of cHL.

Our finding that CD30$^+$ GC B cells represent the MYC$^+$ subset of centrocytes is also of clinical relevance, as numerous patients affected by B or T cell lymphomas with CD30$^+$ tumor cells are currently treated with the anti–CD30 antibody brentuximab vedotin, which eradicates CD30-expressing cells (9–11, 59). Considering that elimination of MYC$^+$ centrocytes in the mouse leads to a breakdown of the GC reaction (41, 42), treatment of patients with lymphoma with brentuximab vedotin may impair T-dependent humoral immune responses, which should be closely monitored in clinical studies.

Methods

Isolation and phenotypic characterization of CD30$^+$ B cells from tonsils. CD30$^+$ GC and EF B cells were isolated from human tonsils obtained from children and adolescents undergoing routine tonsillectomy. Mononuclear cells were obtained by Ficoll density centrifugation. CD3$^+$ T cells were depleted and CD30$^+$ cells were enriched afterwards by magnetic activated cell separation (MACS) using anti-CD3 and anti-CD30 magnetically labeled antibodies (catalog 130-050-101 and 130-051-401, respectively, Miltenyi Biotec). Enriched CD30$^+$ B cell suspensions were stained with anti–CD20-FITC, anti–CD38-APC (catalog 555622 and 555462, respectively, BD Biosciences) and anti–CD30-PE (catalog 550041, Miltenyi Biotec) antibodies. Approximately 2,000–3,000 CD30$^+$ GC B cells (CD20^{hi}CD30$^+$CD38$^+$) and CD30$^+$ EF B cells (CD20$^+$CD30$^+$CD38$^{lo/-}$) were sorted with a FACS Diva cell sorter (Becton Dickinson) from different donors (n = 5 for each CD30$^+$ B cell subset).

For phenotypic analysis, from MACS-enriched CD30$^+$ B cell suspensions from 10 tonsils, CD30$^+$ GC B cells (CD20^{hi}CD30$^+$CD38$^+$), CD30$^+$ EF B cells (CD20$^+$CD30$^+$CD38$^{lo/-}$), and CD30$^-$ GC B cells (CD20^{hi}CD30$^-$CD38$^+$) were distinguished. Analytical stainings were performed with one of the following additional antibodies: anti–CD27-FITC (catalog 555440), anti–IgM-FITC (catalog 555782), anti–IgD-PE-Cy7 (catalog 561314), anti–IgG-FITC (catalog 555786), anti–IgA-FITC (catalog 130-093-071), and anti–IgE-FITC (catalog 130-099-376), with IgG1κ-FITC (catalog 555748) as isotype control; anti–IL2RB-PE (catalog 554525) with IgG1κ-PE (catalog 554680) as isotype control; and anti–IL21R-PE (catalog 560264) and anti–CD21-PE (catalog 555422), with IgG1κ-PE (catalog 555749) as isotype control. In further stainings, antibodies against CD30, CD38, IgM, and IgD were combined. All antibodies were purchased from Becton Dickinson, except the anti–IgA-FITC and anti–IgE-FITC antibodies (Miltenyi Biotec).

For in vitro survival analysis, tonsillar GC B cells (CD20^{hi}CD38$^+$) and EF B cells (CD20$^+$CD38$^{-/lo}$CD27$^+$) were sorted from 5 donors and maintained ex vivo in RPMI 1640 supplemented with 20% fetal calf serum, penicillin (100 U/ml), and streptomycin (100 μg/ml) at 37°C and 5% CO_2. After 0, 6, 16, and 30 hours, cells were stained with propidium iodide (PI) and anti-human CD30 to determine proportions of PI$^+$ cells as well as PI$^-$CD30$^+$ cells.

IgV gene analysis of CD30$^+$ B cells. Aliquots of 2,000 or 3,000 CD30$^+$ GC B cells and CD30$^+$ EF B cells were sort-purified and their

V_H gene rearrangements of the large V_H3 family (donors 1–3) or V_H1, V_H3, and V_H4 families (donor 4) were amplified by seminested PCR. DNA was isolated using the Gentra Puregene Cell Kit (Qiagen) or cell samples were digested with 0.5 mg/ml proteinase K at 55°C for 2 hours, followed by an 8-minute incubation at 95°C. DNA samples were processed in 2 to 3 replicates. Family-specific primers for the V_H leader regions were combined with 2 sets of nested primers for the *IGHJ* segments as described (60, 61), using Expand High Fidelity DNA polymerase mixture (Roche). The first round of PCR was run for 31 cycles, the second for 33–36 cycles. PCR products were gel-purified and cloned into pCR 4-TOPO vector (Invitrogen, Life Technologies) for donor 1, or pGEM-T-easy vector (Promega) for donors 2–4. Plasmids were isolated from XL1-Blue competent cells (Stratagene, Agilent Technologies), and sequenced with an ABI 3130 sequencer (Applied Biosystems). Sequence analysis was performed with MegAlign (DNASTAR) and the IMGT/V-Quest database (http://www.imgt.org/IMGT_vquest/vquest).

Generation of gene expression profiles. RNA isolation, generation of cRNA by 2 rounds of in vitro transcription, cRNA fragmentation and hybridization to Affymetrix HG-U133-Plus2.0 arrays, and microarray washing and scanning were previously described (15, 16). The gene expression data set is available through http://www.ncbi.nlm.nih.gov/geo/ (accession numbers GSE 12453, 14879, and 40160). The $CD30^+$ B cell samples are available under accession no. GSE 83441.

Analysis of differential gene expression by reverse transcription real-time PCR. Differential gene expression between $CD30^+$ and $CD30^-$ B cells and plasma cells and/or between the 2 $CD30^+$ B cell subsets was analyzed for the genes *PRDM1*, *BACH2*, *FOXP1*, *CCL22*, *MYC*, and *EBI2* by reverse transcription real-time PCR. Total RNA was extracted using peqGOLD Micro Spin Columns (Peqlab) following the manufacturer's instructions, including on-column DNase digestion (Qiagen). Synthesis of cDNA was performed with Sensiscript (Qiagen) using anchored oligo(dT) primer (ThermoFisher Scientific). Target gene expression was analyzed in triplicate with cDNA templates equivalent to 500 cells using TaqMan Universal PCR Master Mix No AmpErase UNG (Applied Biosystems). As typical housekeeping genes showed varying expression in $CD30^+$ versus $CD30^-$ B cells, presumably due to the effect of MYC activity in $CD30^+$ B cells on overall mRNA expression levels (38, 39), cycle threshold (Ct) values of the real-time PCR were normalized to the same cell equivalents in each reaction. The following TaqMan assays (Applied Biosystems) were used for target gene detection: MYC (Hs00153408_m1), PRDM1 (Hs00153357_m1), BACH2 (Hs00222364_m1), FOXP1 (Hs00908900_m1), CCL22 (Hs01574247_m1), and GPR183 (EBI2) (Hs00270639_s1). An ABI PRISM 7900HT (Applied Biosystems) was used to run 384-well plates, and Ct values were calculated in SDS software (v2.2, Applied Biosystems). B cells isolated from 6 tonsils were analyzed.

Immunoblot. Equal numbers of sorted B cells from distinct subsets were lysed in RIPA buffer (Santa Cruz Biotechnology) supplemented with 1% NP-40 and complete protease and phosph-STOP phosphatase inhibitors (Roche). Samples were separated using 8%–16% precast TGX gels (BioRad) and transferred to nitrocellulose MiniStacks (Life Technologies) using an iBlot transfer device (Invitrogen). The following primary and secondary antibodies were used: anti-MYC antibody (ab32072, Abcam), anti-p97 antibody (catalog 2648, Cell Signaling), and HRP-coupled secondary anti-mouse/anti-rabbit antibodies (catalog 115-036-003 and catalog 711-036-152; Jackson ImmunoResearch).

Statistics. The statistical analysis was performed as described elsewhere (15, 62). The $CD30^+$ B cell samples were isolated and processed later than the samples of GSE14879, but a coanalysis of newly processed conventional GC B cells and $CD30^+$ B cells revealed that there was no strong batch effect after normalization, so that the new $CD30^+$ samples could be included in the prior set without having to apply a batch correction.

Statistical analysis was performed using the computing environment R (R Development Core Team, 2005). Additional software packages (affy, geneplotter, multtest, vsn) were taken from the Bioconductor project (63). For microarray preprocessing, the variance stabilization method of Huber et al. was applied for probe level normalization (64). The variance of probe intensities is rendered independent of their expected expression levels by this method. Assuming that the majority of genes are not differentially expressed across the samples, parameters (offset and a scaling factor) were estimated for each microarray. With regard to the computational complexity of the algorithm, parameters are estimated on a random subset of probes and are then used to transform the complete arrays. By application of the robust median polish method on the normalized data probe set, summarization was calculated. Considering the different probe affinities via the probe effect, an additive robust model on the logarithmic scale (base 2) was fitted across the arrays for each probe set (65, 66).

Unsupervised hierarchical clustering was performed for the probe sets with a standard deviation greater than 1 across all samples using the Manhattan distance and the average linkage method. The stability of the resulting dendrogram was tested with Pvclust (67).

For differential gene expression analysis, to reduce the dimension of the microarray data before conducting pairwise comparisons, an intensity filter (the signal intensity of a probe had to be ≥100 in at least 25% of the samples, if the group sizes were equal) and a variance filter (the interquartile range of $\log^2$ intensities should be ≥0.5, if the group sizes were equal) were applied. In case of unequal group sizes, the signal intensity of a probe set had to be above 100 in at least a fraction (smaller group size minus 1, divided by the total sample size of the 2 groups) of samples. The interquartile range of $\log^2$ intensities had to be at least 0.1 for unequal group sizes. After global filtering, a 2-sample *t* test was applied to identify genes that were differentially expressed between 2 groups. The FDR according to Benjamini and Hochberg was used to account for multiple testing (68). FC between the 2 groups of each supervised analysis was calculated for each gene. PCA was performed using preselected gene sets (unsupervised with SD filter or significant differentially expressed genes) and predecided groups of samples. The first and second principal components are visualized in a 2D plot.

For the other statistical evaluation, either a paired or unpaired 2-tailed *t* test was used, as specified in the respective figure legends. A *P* value smaller than 0.05 was considered significant.

Study approval. As approved by the local ethics committee of the Medical School of the University Duisburg-Essen, all samples were collected with the informed consent of the donors. Written informed consent was received from participants prior to inclusion in the study.

Author contributions

MAW, ET, MLH, and RK designed the study. MAW, ET, SS, SR, and JD performed the experiments. JA provided essential material. MS provided technical expertise. CD and MAW performed the

bioinformatical analysis. MAW, ET, and RK evaluated the data. MAW and RK wrote the manuscript.

Acknowledgments

We thank Klaus Lennartz for expert help with cell sorting, Ludger Klein-Hitpass for expert help to perform Affymetrix genechip analysis, Kerstin Heise for expert technical assistance (all from the Institute of Cell Biology, University of Duisburg-Essen), and Ulf Klein (Leeds University) for critically reading the manuscript. This work was supported by the Deutsche Forschungsgemeinschaft (KU1315/7-1, KU1315/10-1), the Bundesministerium für Bildung und Forschung through the International Cancer Genome Consortium for malignant lymphomas (01KU1002F), and the German José Carreras Leukemia Foundation (R08/04). ET was the recipient of fellowships granted by the German José Carreras Leukemia Foundation (F05/01) and by Ms. Livia Benedetti.

Address correspondence to: Ralf Küppers, Institute of Cell Biology (Cancer Research), University of Duisburg-Essen, Faculty of Medicine, Virchowstrasse 173, 45122 Essen, Germany. Phone: 49.2017233384; Email: ralf.kueppers@uk-essen.de.

ET's present address is: Institute of Hematology, University of Perugia, Perugia, Italy.

1. Falini B, et al. CD30 (Ki-1) molecule: a new cytokine receptor of the tumor necrosis factor receptor superfamily as a tool for diagnosis and immunotherapy. *Blood*. 1995;85(1):1–14.
2. Stein H, et al. The expression of the Hodgkin's disease associated antigen Ki-1 in reactive and neoplastic lymphoid tissue: evidence that Reed-Sternberg cells and histiocytic malignancies are derived from activated lymphoid cells. *Blood*. 1985;66(4):848–858.
3. Kurth J, et al. EBV-infected B cells in infectious mononucleosis: viral strategies for spreading in the B cell compartment and establishing latency. *Immunity*. 2000;13(4):485–495.
4. Cattoretti G, Büttner M, Shaknovich R, Kremmer E, Alobeid B, Niedobitek G. Nuclear and cytoplasmic AID in extrafollicular and germinal center B cells. *Blood*. 2006;107(10):3967–3975.
5. Cattoretti G, Shaknovich R, Smith PM, Jäck HM, Murty VV, Alobeid B. Stages of germinal center transit are defined by B cell transcription factor coexpression and relative abundance. *J Immunol*. 2006;177(10):6930–6939.
6. Swerdlow SH, et al. *Classification of Tumours of Haematopoietic and Lymphoid Tissues*. Lyon, France: IARC Press; 2008:323–325.
7. Hu S, et al. CD30 expression defines a novel subgroup of diffuse large B-cell lymphoma with favorable prognosis and distinct gene expression signature: a report from the International DLBCL Rituximab-CHOP Consortium Program Study. *Blood*. 2013;121(14):2715–2724.
8. Campuzano-Zuluaga G, Cioffi-Lavina M, Lossos IS, Chapman-Fredricks JR. Frequency and extent of CD30 expression in diffuse large B-cell lymphoma and its relation to clinical and biologic factors: a retrospective study of 167 cases. *Leuk Lymphoma*. 2013;54(11):2405–2411.
9. Duvic M, Tetzlaff MT, Gangar P, Clos AL, Sui D, Talpur R. Results of a phase II trial of brentuximab vedotin for CD30+ cutaneous T-cell lymphoma and lymphomatoid papulosis. *J Clin Oncol*. 2015;33(32):3759–3765.
10. Pro B, et al. Brentuximab vedotin (SGN-35) in patients with relapsed or refractory systemic anaplastic large-cell lymphoma: results of a phase II study. *J Clin Oncol*. 2012;30(18):2190–2196.
11. Brown MP, Staudacher AH. Could bystander killing contribute significantly to the antitumor activity of brentuximab vedotin given with standard first-line chemotherapy for Hodgkin lymphoma? *Immunotherapy*. 2014;6(4):371–375.
12. Klein U, Rajewsky K, Küppers R. 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*. 1998;188(9):1679–1689.
13. Jung J, Choe J, Li L, Choi YS. Regulation of CD27 expression in the course of germinal center B cell differentiation: the pivotal role of IL-10. *Eur J Immunol*. 2000;30(8):2437–2443.
14. Klein U, et al. Somatic hypermutation in normal and transformed human B cells. *Immunol Rev*. 1998;162:261–280.
15. Brune V, et al. Origin and pathogenesis of nodular lymphocyte-predominant Hodgkin lymphoma as revealed by global gene expression analysis. *J Exp Med*. 2008;205(10):2251–2268.
16. Tiacci E, et al. Analyzing primary Hodgkin and Reed-Sternberg cells to capture the molecular and cellular pathogenesis of classical Hodgkin lymphoma. *Blood*. 2012;120(23):4609–4620.
17. Gatto D, Paus D, Basten A, Mackay CR, Brink R. Guidance of B cells by the orphan G protein-coupled receptor EBI2 shapes humoral immune responses. *Immunity*. 2009;31(2):259–269.
18. Kjeldsen MK, et al. Multiparametric flow cytometry for identification and fluorescence activated cell sorting of five distinct B-cell subpopulations in normal tonsil tissue. *Am J Clin Pathol*. 2011;136(6):960–969.
19. Golay J, et al. The A-Myb transcription factor is a marker of centroblasts in vivo. *J Immunol*. 1998;160(6):2786–2793.
20. Victora GD, Dominguez-Sola D, Holmes AB, Deroubaix S, Dalla-Favera R, Nussenzweig MC. Identification of human germinal center light and dark zone cells and their relationship to human B-cell lymphomas. *Blood*. 2012;120(11):2240–2248.
21. Klein U, et al. Transcriptional analysis of the B cell germinal center reaction. *Proc Natl Acad Sci U S A*. 2003;100(5):2639–2644.
22. Lin L, Nonoyama S, Oshiba A, Kabasawa Y, Mizutani S. TARC and MDC are produced by CD40 activated human B cells and are elevated in the sera of infantile atopic dermatitis patients. *J Med Dent Sci*. 2003;50(1):27–33.
23. van Riggelen J, Yetil A, Felsher DW. MYC as a regulator of ribosome biogenesis and protein synthesis. *Nat Rev Cancer*. 2010;10(4):301–309.
24. Liu YJ, Joshua DE, Williams GT, Smith CA, Gordon J, MacLennan IC. Mechanism of antigen-driven selection in germinal centres. *Nature*. 1989;342(6252):929–931.
25. Andrew DP, et al. STCP-1 (MDC) CC chemokine acts specifically on chronically activated Th2 lymphocytes and is produced by monocytes on stimulation with Th2 cytokines IL-4 and IL-13. *J Immunol*. 1998;161(9):5027–5038.
26. Loetscher M, et al. Chemokine receptor specific for IP10 and mig: structure, function, and expression in activated T-lymphocytes. *J Exp Med*. 1996;184(3):963–969.
27. Jourdan M, et al. Gene expression of anti- and pro-apoptotic proteins in malignant and normal plasma cells. *Br J Haematol*. 2009;145(1):45–58.
28. Gore Y, et al. Macrophage migration inhibitory factor induces B cell survival by activation of a CD74-CD44 receptor complex. *J Biol Chem*. 2008;283(5):2784–2792.
29. Harada H, et al. Phenotypic difference of normal plasma cells from mature myeloma cells. *Blood*. 1993;81(10):2658–2663.
30. Lau D, et al. Low CD21 expression defines a population of recent germinal center graduates primed for plasma cell differentiation. *Sci Immunol*. 2017;2(7):eaai8153.
31. Drexler HG. Recent results on the biology of Hodgkin and Reed-Sternberg cells. I. Biopsy material. *Leuk Lymphoma*. 1992;8(4-5):283–313.
32. Schmitz R, Stanelle J, Hansmann ML, Küppers R. Pathogenesis of classical and lymphocyte-predominant Hodgkin lymphoma. *Annu Rev Pathol*. 2009;4:151–174.
33. Küppers R. The biology of Hodgkin's lymphoma. *Nat Rev Cancer*. 2009;9(1):15–27.
34. Schwering I, et al. Loss of the B-lineage-specific gene expression program in Hodgkin and Reed-Sternberg cells of Hodgkin lymphoma. *Blood*. 2003;101(4):1505–1512.
35. Lamprecht B, et al. Aberrant expression of the Th2 cytokine IL-21 in Hodgkin lymphoma cells regulates STAT3 signaling and attracts Treg cells via regulation of MIP-3alpha. *Blood*. 2008;112(8):3339–3347.
36. Scheeren FA, et al. IL-21 is expressed in Hodgkin lymphoma and activates STAT5: evidence that activated STAT5 is required for Hodgkin lymphomagenesis. *Blood*. 2008;111(9):4706–4715.
37. Baus D, et al. STAT6 and STAT1 are essential antagonistic regulators of cell survival in classical Hodgkin lymphoma cell line. *Leukemia*. 2009;23(10):1885–1893.
38. Nie Z, et al. c-Myc is a universal amplifier of

expressed genes in lymphocytes and embryonic stem cells. *Cell*. 2012;151(1):68–79.
39. Lin CY, et al. Transcriptional amplification in tumor cells with elevated c-Myc. *Cell*. 2012;151(1):56–67.
40. Cattoretti G. MYC expression and distribution in normal mature lymphoid cells. *J Pathol*. 2013;229(3):430–440.
41. Calado DP, et al. The cell-cycle regulator c-Myc is essential for the formation and maintenance of germinal centers. *Nat Immunol*. 2012;13(11):1092–1100.
42. Dominguez-Sola D, et al. The proto-oncogene MYC is required for selection in the germinal center and cyclic reentry. *Nat Immunol*. 2012;13(11):1083–1091.
43. Potting C, et al. TRIAP1/PRELI complexes prevent apoptosis by mediating intramitochondrial transport of phosphatidic acid. *Cell Metab*. 2013;18(2):287–295.
44. Tian T, et al. Role of leucine-rich pentatricopeptide repeat motif-containing protein (LRPPRC) for anti-apoptosis and tumourigenesis in cancers. *Eur J Cancer*. 2012;48(15):2462–2473.
45. Schwarting R, Gerdes J, Dürkop H, Falini B, Pileri S, Stein H. BER-H2: a new anti-Ki-1 (CD30) monoclonal antibody directed at a formol-resistant epitope. *Blood*. 1989;74(5):1678–1689.
46. Gerdes J, Schwarting R, Stein H. High proliferative activity of Reed Sternberg associated antigen Ki-1 positive cells in normal lymphoid tissue. *J Clin Pathol*. 1986;39(9):993–997.
47. Chou C, et al. The transcription factor AP4 mediates resolution of chronic viral infection through amplification of germinal center B cell responses. *Immunity*. 2016;45(3):570–582.
48. Watanabe M, et al. AP-1 mediated relief of repressive activity of the CD30 promoter microsatellite in Hodgkin and Reed-Sternberg cells. *Am J Pathol*. 2003;163(2):633–641.
49. Watanabe M, et al. JunB induced by constitutive CD30-extracellular signal-regulated kinase 1/2 mitogen-activated protein kinase signaling activates the CD30 promoter in anaplastic large cell lymphoma and reed-sternberg cells of Hodgkin lymphoma. *Cancer Res*. 2005;65(17):7628–7634.
50. Cattoretti G, Angelin-Duclos C, Shaknovich R, Zhou H, Wang D, Alobeid B. PRDM1/Blimp-1 is expressed in human B-lymphocytes committed to the plasma cell lineage. *J Pathol*. 2005;206(1):76–86.
51. Weniger MA, Küppers R. NF-κB deregulation in Hodgkin lymphoma. *Semin Cancer Biol*. 2016;39:32–39.
52. Lollies A, et al. An oncogenic axis of STAT-mediated BATF3 upregulation causing MYC activity in classical Hodgkin lymphoma and anaplastic large cell lymphoma. *Leukemia*. 2018;32(1):92–101.
53. Rui L, et al. Cooperative epigenetic modulation by cancer amplicon genes. *Cancer Cell*. 2010;18(6):590–605.
54. Weniger MA, et al. Mutations of the tumor suppressor gene SOCS-1 in classical Hodgkin lymphoma are frequent and associated with nuclear phospho-STAT5 accumulation. *Oncogene*. 2006;25(18):2679–2684.
55. Janz M, et al. Classical Hodgkin lymphoma is characterized by high constitutive expression of activating transcription factor 3 (ATF3), which promotes viability of Hodgkin/Reed-Sternberg cells. *Blood*. 2006;107(6):2536–2539.
56. Mathas S, et al. Aberrantly expressed c-Jun and JunB are a hallmark of Hodgkin lymphoma cells, stimulate proliferation and synergize with NF-kappa B. *EMBO J*. 2002;21(15):4104–4113.
57. Rengstl B, et al. Incomplete cytokinesis and re-fusion of small mononucleated Hodgkin cells lead to giant multinucleated Reed-Sternberg cells. *Proc Natl Acad Sci U S A*. 2013;110(51):20729–20734.
58. Xavier de Carvalho A, et al. Reed-Sternberg cells form by abscission failure in the presence of functional Aurora B kinase. *PLoS One*. 2015;10(5):e0124629.
59. Chen R, et al. Five-year survival and durability results of brentuximab vedotin in patients with relapsed or refractory Hodgkin lymphoma. *Blood*. 2016;128(12):1562–1566.
60. Bräuninger A, et al. Molecular analysis of single B cells from T-cell-rich B-cell lymphoma shows the derivation of the tumor cells from mutating germinal center B cells and exemplifies means by which immunoglobulin genes are modified in germinal center B cells. *Blood*. 1999;93(8):2679–2687.
61. Küppers R, Schneider M, Hansmann ML. Laser-based microdissection of single cells from tissue sections and PCR analysis of rearranged immunoglobulin genes from isolated normal and malignant human B cells. *Methods Mol Biol*. 2013;971:49–63.
62. Eckerle S, et al. Gene expression profiling of isolated tumour cells from anaplastic large cell lymphomas: insights into its cellular origin, pathogenesis and relation to Hodgkin lymphoma. *Leukemia*. 2009;23(11):2129–2138.
63. Gentleman RC, et al. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol*. 2004;5(10):R80.
64. Huber W, von Heydebreck A, Sültmann H, Poustka A, Vingron M. Variance stabilization applied to microarray data calibration and to the quantification of differential expression. *Bioinformatics*. 2002;18(Suppl 1):S96–104.
65. Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B, Speed TP. Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Res*. 2003;31(4):e15.
66. Tukey JW. *Exploratory data analysis*. New York, NY: Pearson;1977.
67. Suzuki R, Shimodaira H. Pvclust: an R package for assessing the uncertainty in hierarchical clustering. *Bioinformatics*. 2006;22(12):1540–1542.
68. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Series B Stat Methodol*. 1995;57(1):289–300.
69. Seitz V, et al. Deep sequencing of MYC DNA-binding sites in Burkitt lymphoma. *PLoS One*. 2011;6(11):e26837.