Whole-slide image analysis of the tumor microenvironment identifies low B-cell content as a predictor of adverse outcome in patients with advanced-stage classical Hodgkin lymphoma treated with BEACOPP

Ron Daniel Jachimowicz,[1,2,3*] Luise Pieper,[4*] Sarah Reinke,[4,*] Artur Gontarewicz,[4,*] Annette Plütschow,[1] Heinz Haverkamp,[1] Leonie Frauenfeld,[5] Falko Fend,[5] Mathis Overkamp,[5] Franziska Jochims,[4] Christoph Thorns,[6] Martin Leo Hansmann,[7] Peter Möller,[8] Andreas Rosenwald,[9] Harald Stein,[10] Hans Christian Reinhardt,[2,3,11,12] Peter Borchmann,[1,3] Bastian von Tresckow,[1,3] Andreas Engert[1,3] and Wolfram Klapper[4]

[1]University of Cologne, Faculty of Medicine and University Hospital Cologne, Department I of Internal Medicine, Cologne; German Hodgkin Study Group, Cologne; [2]Else Kröner Forschungskolleg Clonal Evolution in Cancer, University Hospital Cologne, Cologne; [3]University of Cologne, Faculty of Medicine and University Hospital Cologne, Department I of Internal Medicine, Center for Integrated Oncology Aachen-Bonn-Cologne-Düsseldorf, Cologne; [4]Department of Pathology, Hematopathology Section, University Hospital Schleswig-Holstein, Christian-Albrechts-University, Kiel; [518]Department of Pathology, University of Tübingen, Tübingen; [6]Department of Pathology, University Hospital Schleswig-Holstein, University of Lübeck, Lübeck; [7]Department of Pathology, University Hospital Frankfurt, Frankfurt; [8]Department of Pathology, University Hospital Ulm, Ulm; [9]Institute of Pathology, University of Würzburg, and Comprehensive Cancer Center Mainfranken, Würzburg; [10]Pathodiagnostik, Berlin; [11]Cologne Excellence Cluster on Cellular Stress Response in Aging-Associated Diseases (CECAD), University of Cologne, Cologne and [12]Center for Molecular Medicine Cologne (CMMC), University of Cologne, Cologne, Germany

**RDJ, LP, SR and AG contributed equally as co-first authors.*

 doi:10.3324/haematol.2019.243287

Received: December 12, 2019.

Accepted: April 29, 2020.

Pre-published: May 7, 2020.

Correspondence: *RON D. JACHIMOWICZ* - ron.jachimowicz@uk-koeln.de

WOLFRAM KLAPPER - wklapper@path.uni-kiel.de

Supplementary methods and data

Methods

Digital image analysis

Slides were scanned centralized (Hamamatsu Nanozoomer, Hamamatsu Photonics, Ammersee, Germany) and stored as a ndp-file format in Kiel. File size per image varied between 100 MB to 2.5 GB depending on the scanned tissue size. Digital image analysis was performed using TissueStudio 64 (Definiens AG, 80636 Munich, Germany). Images of the stained slices were uploaded to the program and the magnification was defined using the metadata of the image (magnification 40x, pixel resolution (µm/pixel) ≈ 0.2270). The software offers options for the location of the immunohistochemistry (IHC) staining. For analyzing the anti-CD20 and anti-CD3 staining the parameter were set to count cells with stained membranes whereas staining of anti-CD68 and anti-CD30 was analyzed by calculation of the stained cytoplasm as area (μm^2). By using tissue-background-separation and setting regions of interest into the image, only area with tissue was analyzed. By manual adjustments in the program, areas with insufficient staining, overstaining or foldings were cut out of the tissue to be analyzed. For CD30 and CD68 analysis, the analyzed area was additionally corrected for fatty tissue, artificial holes, large vessels and necrotic areas. These areas were delineated manually and the program was instructed not to include such defined areas in the analysis. Stained cells (CD20/CD3 staining) or stained areas (CD68/CD30 staining) were defined by brown chromogen. Small subsets of the entire tissue were chosen and adjustments for cellular analysis were done in those higher magnified subsets to improve the accuracy of the analysis performance (setting the thresholds for the staining). The program offers adjustments for nucleus detection, nucleus morphology and intensities of the IHC marker. By defining the typical nucleus size as 20 μm^2 the densely laying nuclei were separated for counting by the software. For the nucleus detection the threshold for haematoxylin was set to a default of 0.01 and for the IHC Signal to a default of 0.001 for every

image. Small blue artifacts, which don´t represent a nucleus were sorted out by removing all stained areas with a size ≤7 μm^2 from the analysis. The thresholds for the antibody staining were defined individually to overcome the problem of irregular staining intensity and consequently optimize the immunohistochemical detection. For the membrane detection in the CD20 and CD3 staining the threshold for the IHC marker intensity was set to values between 0.3 and 0.75. The thresholds for the marker area detection in the anti-CD68 and anti-CD30 stained slices were set to values of 0.15-0.36. To ensure that only entire and intact macrophages were counted, areas with less than 30 μm^2 were excluded from the analysis.

After calculation the software generates Excel sheets with data of the analyzed total area in μm^2, number of nuclei, cells with and without staining as number and as percentage for the CD3 and CD20 analysis and number of stained areas and stained area as μm^2 for the CD30 and CD68 analysis. In single tiles of the entire image the original staining and the overlay with the colored and counted nuclei and staining-areas are documented as image files in jpg format. Those data collections had a size between 1-11 GB depending on the analyzed tissue size and the analysis method (analysis of area contains less data size than cell count analysis).

Calculation of cell counts for CD68 and CD30

Since B- and T cells on the one hand and macrophages and HRSC on the other hand differ strikingly in their cellular features, we used two different approaches to identify their cellular content. B- and T cells are small and virtually all stained cells in a particular slide/tissue level contain a nucleus which is detectable by hematoxilin staining. Thus, overall B- (CD20) and T cell (CD3) content was assessed using software settings that identify nuclei and count individual cells with positively stained membranes. Obtained results reflect the percentage of cells with positively stained membranes among all nuclei/cells identified. Macrophages and HRSC often show an irregular and outward bulging cytoplasm and frequently lack a detectable nucleus in the cutting level (macrophages) or display multiple nuclei (HRSC). Additionally, macrophages are often located close to each other and show variable staining intensity. Thus, automated counting of such large cells by the identification of nuclei is not accurate as it has been previously shown for macrophages [22]. We

therefore quantified macrophage and HRSC content by identification of area positively stained for CD68 and CD30, respectively, in relation to whole analyzed area independent of the localization of nuclei [23]. The results of the software analysis of macrophages and HRSC are described as the percentage of immunohistochemically positive stained area related to the whole analysed area analyzed. No numbers of cells as well as positive or negative stained cells are available in this analysis solution like for CD3 and CD20 stainings. We generated values for the average size of macrophages and HRSC by manually counting number of cells by visual inspection and divided by the positively stained area as identified by the software Definiens Tissue Studio in 10 HPF images from different cases with a size of 0.28 mm^2 and an average of 3525 nuclei. Our average calculated area for a macrophage was 127,3 μm^2 and 167,7 μm^2 for a HRSC representing a 12,7 µm and 14,6 µm diameter of an ideal circular area, respectively. The cell counts for macrophage or HRSC were calculated by dividing the positively stained area by WSI by 127,3 μm^2 and 167,6 μm^2, respectively. Since staining for CD3, CD20, CD68 and CD30 was performed in subsequent slides that reflect areas of immediate proximity, we predicted the overall cell counts for CD68 and CD30 slides based on the slides stained for CD3 and CD20. The percentage of macrophages and HRSC cells were calculated in relation to overall cell counts in analogy to CD3 and CD20 cell count analysis.

Reproducibility of image analysis

Randomly selected images (n=15) analyzed by observer were reanalyzed by a second observer blinded for the results of observer 1 using the identical software. The results of two observers of the same image shows a high correlation (Supplementary Figure 5A, Pearson's r= 0.9959, $p<0.0001$). A similar high correlation was found when a second observer analyzed newly generated scans using the same immunohistochemical slides (Supplementary Figure 5B, Pearson's r= 0.9331, $p<0.0001$).

Analysis of CD20 content by WSI by independent observers

20 cases of cHL from the cohort from the main analyzing lab (Kiel) were analyzed by a second center. The subgroup was selected to reflect different sizes of lymph node biopsies, different subgroups of cHL and different B-cell content based on the analysis in the first lab. Freshly cut, unstained slides were sent to the second center for complete work-up including staining and WSI of digitalized slides using the identical software (Definiens TissueStudio). Specimen-dependent adjustment such as exclusion of cutting and staining artefacts and manually adjusted thresholds were conducted by the second center blinded for the results of the first analysis.

B-cell counts in close proximity to HRSC

A subset of 41 cases from the cohort were selected reflecting the spectrum of high and low B-cell counts by whole-slide-image analysis. In each case 3 field of views were selected that reflect a 1000x magnification (high power field, HPF, approximately 200 nm in diameter). The fields were chosen by searching for at least one HRSC and placing it into the center of the HPF to ensure that B-cell counts were analyzed in close proximity to HRSC only. The number of CD20 positive cells per HPF were counted manually and average values for three HPF were calculated and correlated with the results for whole slides (Supplementary Figure 4).

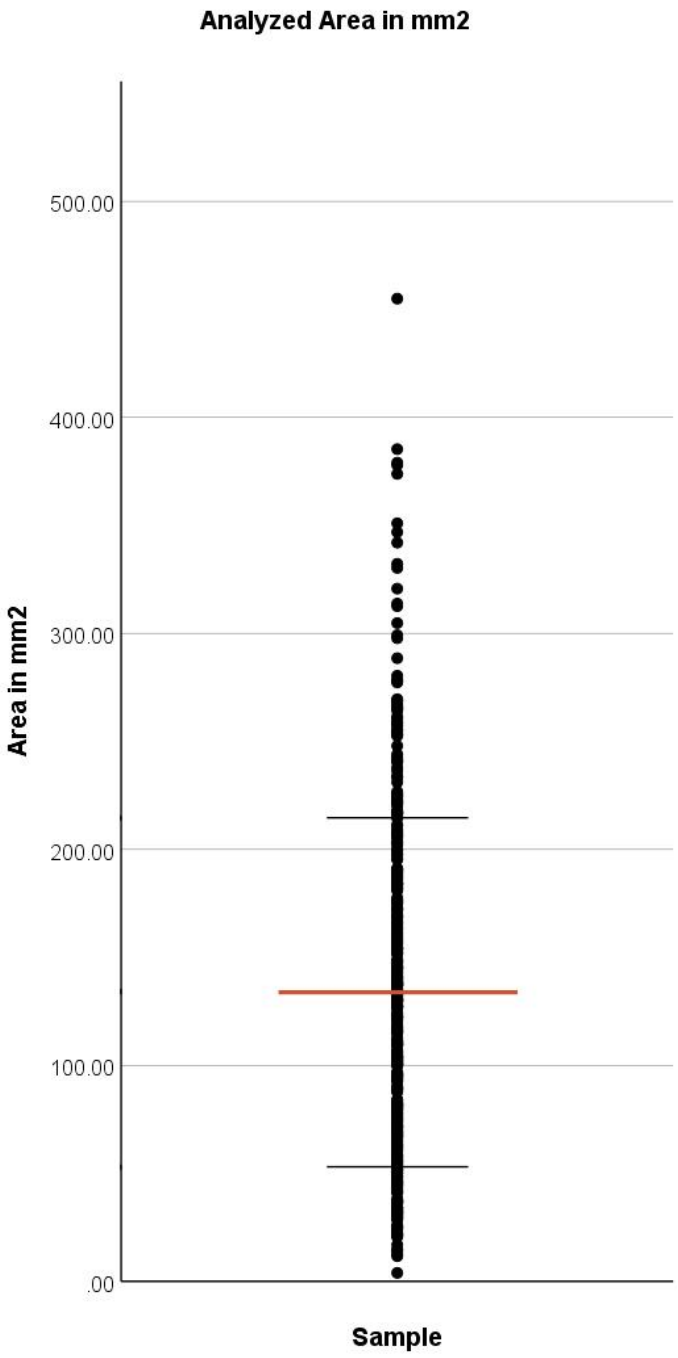
Statistical Analysis

Univariate statistics were used to describe and summarize analysis variables. Log2-transformation was applied to computer counts in order to normalize the data. We used Pearson product-moment correlation as parametric measure of linear relationship between two variables and t-test to compare CD3, CD20, CD30 and CD68 content between patients with or without disease progression or relapse. Survival was estimated by the Kaplan-Meier method and compared using a log-rank test. The hazard ratio and corresponding confidence interval were calculated using a univariate Cox proportional hazards regression model. Progression-free survival (PFS) was measured from the date of diagnosis to

either disease progression or disease- or treatment-related death or censoring date. Receiver operating characteristic (ROC) curves were analyzed to select an optimum biomarker cut-off value for prediction of PFS-failure as a binary event. Multivariate logistic regression with backward selection was used to assess the prognostic value of binary B cell content for PFS-failure, while adjusting for other individual parameters with a significance level below 0.10 at a univariate level. All statistical analyses were done using SAS software version 9.4.

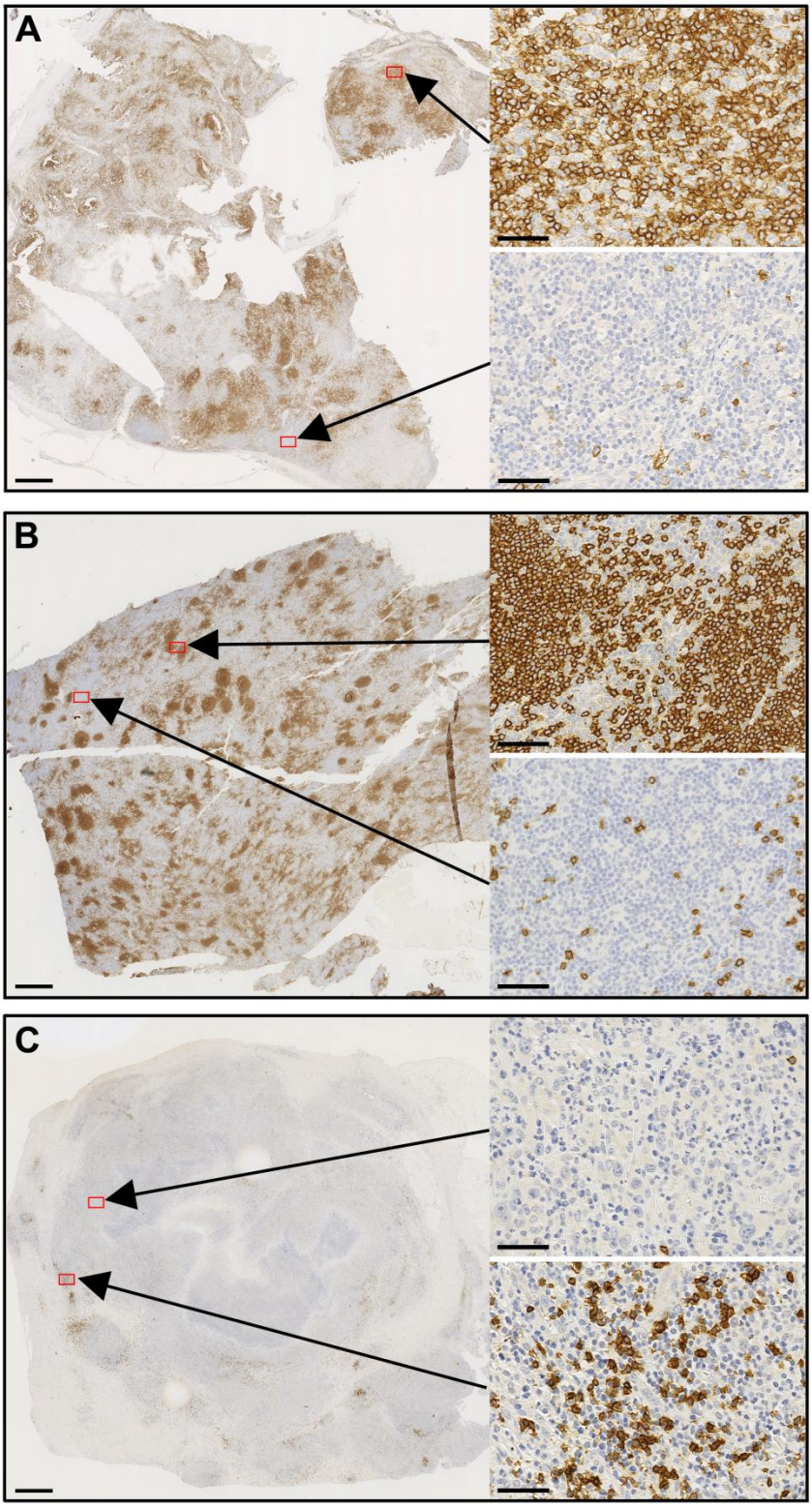
Supplementary data

Supplementary Figure 1



Supplementary Figure 1: Size of analyzed area. The size of the CD30 marked sections is shown exemplarily for the size of the analyzed areas in mm^2. The analyzed size of the individual samples varied between with a mean of 133.81 mm^2 (red bar) and a SD of ±81.84 mm^2 (black bars).

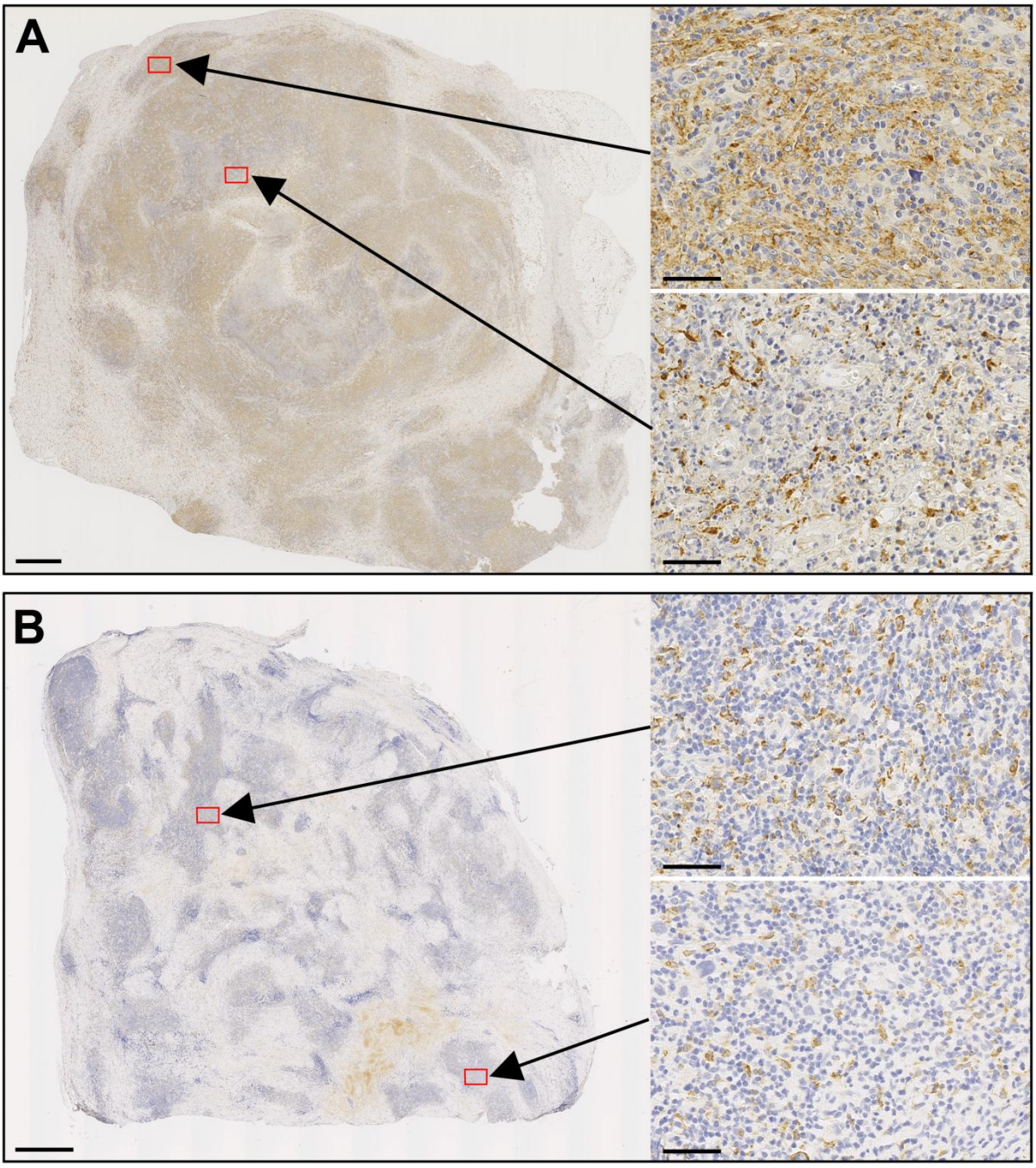
Supplementary Figure 2



Supplementary Figure 2. Distribution of B cells in lymph node tissue of cHL.

(A-C) Representative examples of lymph node tissue specimen stained for CD20 reflecting a high (A), moderate (B) and low B cell content (C). The left panel shows an overview picture and the right panel high magnifications of areas with either high and low B cell content illustrating the uneven distribution of cells. Scale bar left panel corresponds to 1000 µm, and in right panel to 50 µm.

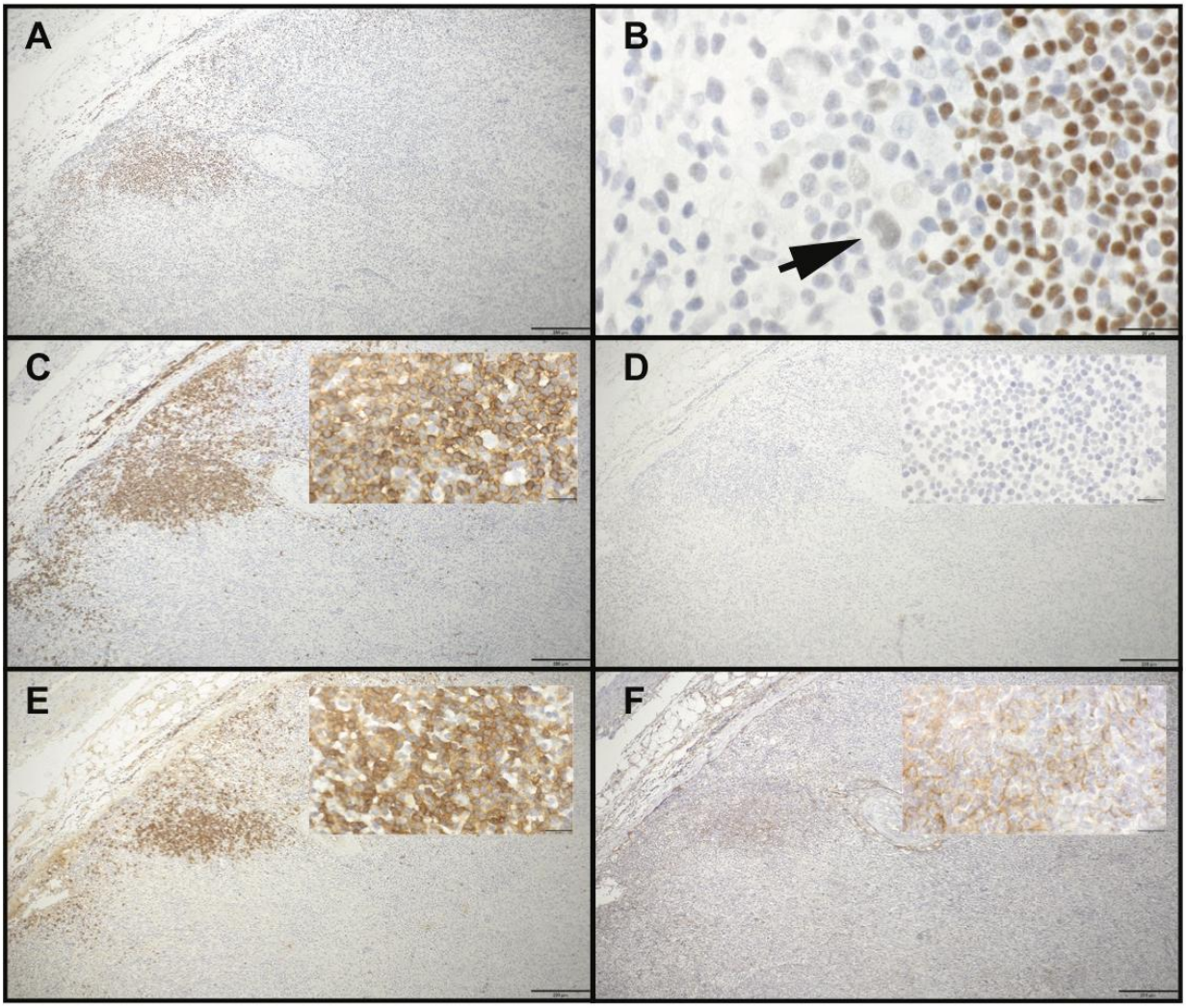
Supplementary Figure 3



Supplementary Figure 3. Distribution of macrophages in lymph node tissue of cHL.

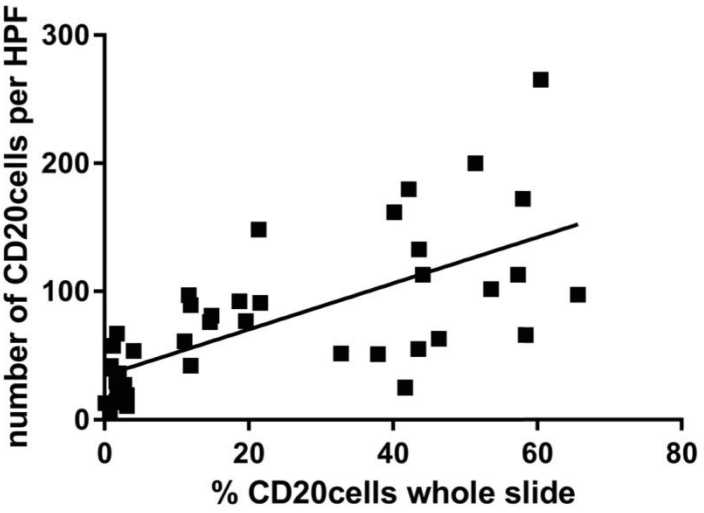
(A-B) Representative examples of lymph node tissue specimen stained for CD68 reflecting a high (A) and low macrophage content (B). The left panel shows an overview picture and the right panel high magnifications of areas with either high and low macrophage content illustrating the uneven distribution of cells. Scale bar overview left panel corresponds to 1000 µm, and in high magnification in the right panel to 50 µm.

Supplementary Figure 4



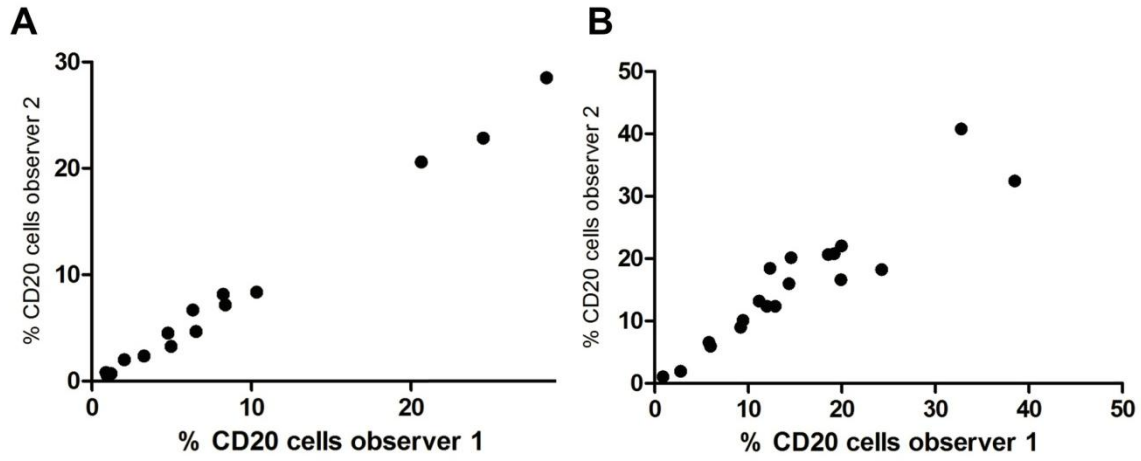
Supplementary Figure 4. B cell arrangement in follicles. B cells arranged in ill-defined follicles (A: Pax5). HRSC (arrow) reside outside of these follicles (B: Pax5). B cell follicles (C: CD20) lack germinal centers (D: BCL6) and are mainly composed of IgD-positive follicle mantle cells (E: IgD) of which some express CD73, a marker that has been used to identify memory cells (F: CD73). Bar represents 200µm in large picture and 20 µm in inserts with high magnification.

Supplementary Figure 5



Supplementary Figure 5. B cell content in whole slides and close proximity to HRSC. Correlation of CD20 cell content in whole slide (assessed by image analysis) and in close proximity to Hodgkin-Reed-Sternberg-cells (high power field, HPF by visual inspection and manual counting, n=41, Pearson r=0.676, p<0.0001); online only.

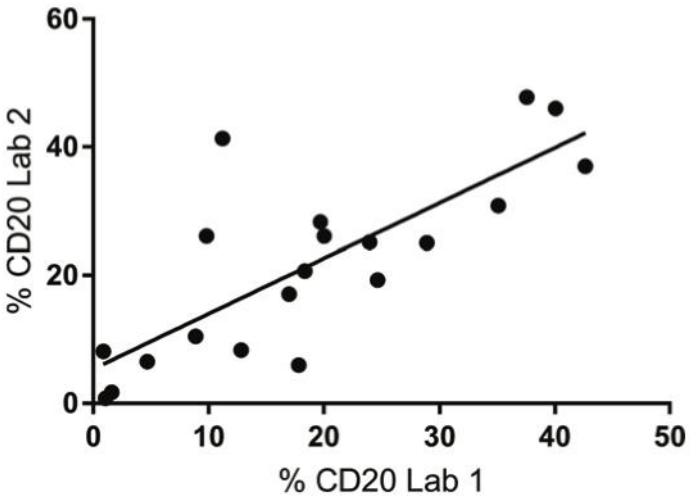
Supplementary Figure 6



Supplementary Figure 6. Inter-observer comparison of identical images. Inter-observer comparison of n=15 cases using identical images/scans (**A**, Pearson's r=0.9959, p<0.0001) or newly generated images/scan (n=20) of the original

immunohistochemical staining (**B**, Pearson's r = 0.9331, p<0.0001).

Supplementary Figure 7



Supplementary Figure 7. Inter-observer comparison of newly stained slides at different centers. CD20 staining and WSI performed in two laboratories independently. Percentage of B cell for each specimen indicated. The line represents the linear regression (r^2=0.6101).