Evaluation of the antibody response to the EBV proteome in EBV-associated classic Hodgkin lymphoma

Zhiwei Liu [* 1], Ruth F. Jarrett [* 2], Henrik Hjalgrim [* 3,4], Carla Proietti [5], Ellen T. Chang [6,7], Karin E. Smedby [8], Kelly J. Yu [1], Annette Lake [2], Sally Troy [4], Karen A. McAulay [9], Ruth M. Pfeiffer [1], Hans-Olov Adami [10,11], Bengt Glimelius [12], Mads Melbye [3], Allan Hildesheim [1], Denise Doolan [† 4], Anna E. Coghill [† 1]

[*†]authors contributed equally to this work

1. Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD, USA

2. MRC-University of Glasgow Centre for Virus Research, Glasgow, UK

3. Statens Serum Institut, Copenhagen, DENMARK

4. Department of Haematology, Rigshospitalet, Copenhagen, DENMARK

5. Center for Molecular Therapeutics, Australian Institute of Tropical Health & Medicine, James Cook University, Cairns, QLD, AUSTRALIA

6 Stanford Cancer Institute, Stanford, CA, USA

7. Center for Health Sciences, Exponent, Inc., Menlo Park, CA, USA

8. Department of Medicine Solna, Division of Clinical Epidemiology, Karolinska Institutet, Stockholm, SWEDEN

9. REPROCELL Europe Ltd., Thomson Pavilion, Todd Campus, West of Scotland Science Park, Acre Rd, Glasgow, UK

10. Department of Medical Epidemiology and Biostatics, Karolinska Institutet, Stockholm, SWEDEN

11. Clinical Effectiveness Research Group, Institute of Health, University of Oslo, Oslo, NORWAY

12. Department of Immunology, Genetics and Pathology, Uppsala University, Uppsala, SWEDEN

Running Title: EBV antibody patterns and cHL

Key words: EBV and cancer; Hodgkin lymphoma; EBV antibody patterns

Corresponding author: Zhiwei Liu, PhD, Division of Cancer Epidemiology and Genetics, National Cancer Institute, 9609 Medical Center Drive, Bethesda, MD 20850; Phone: +1-240-276-6726, Email address: zhiwei.liu@nih.gov

Abbreviations: ANOVA, analysis of variance; AUC, area under the receiver operative curve; BL, Burkitt lymphoma; cHL, classic Hodgkin lymphoma; CI, confidence interval; CV, coefficient of variation; EA, early antigen; EBV, Epstein-Barr virus; EBERs, EBV-encoded small RNAs; EBNA, Epstein-Barr nuclear antigen; HRS, Hodgkin Reed-Sternberg; IM, infectious mononucleosis; IQR, interquartile range; LMP, latent membrane; NPC, nasopharyngeal carcinoma; OR, odds ratio; PBMC, peripheral blood mononuclear cell; SCALE, Scandinavian Lymphoma Etiology; SSI, standardized signal intensity; VCA, viral capsid antigen.

Abstract

The humoral immune response to Epstein-Barr virus (EBV) in classic Hodgkin lymphoma (cHL) stratified by EBV tumor status is unclear. We examined IgG and IgA antibody responses against 202 protein sequences representing 86 EBV proteins using a microarray and sera from 139 EBV-positive cHL cases, 70 EBV-negative cHL cases, and 141 population-based controls frequency matched to EBV-positive cHL cases on sex and age by area (UK, Denmark, and Sweden). We leveraged existing data on the proportion of circulating B-cells infected by EBV and levels of serum CCL17, a chemokine secreted by cHL tumor cells, from a subset of the cHL cases in the UK. Total IgG but not IgA response level was significantly different between EBV-positive cHL cases and controls. The distinct serological response included significant elevations in 16 IgG antibodies and 2 IgA antibodies, with odds ratios $_\text{highest vs. lowest tertile}$ >3 observed for the following EBV proteins: LMP1 (oncogene), BcLF1 (VCAp160, two variants), and BBLF1 (two variants). Our cHL IgG signature correlated with the proportion of circulating EBV-infected B-cells, but not serum CCL17 levels. We observed no differences in the anti-EBV antibody profile between EBV-negative cHL cases and controls. BdRF1(VCAp40)-IgG and BZLF1(Zta)-IgG were identified as the serological markers best able to distinguish EBV-positive from EBV-negative cHL tumors. Our results support the hypothesis that differences in the EBV antibody profile are specific to patients with EBV-positive cHL and are not universally observed as part of a systematically dysregulated immune response present in all cHL cases.

Novelty and Impact

Our data expand beyond the limited number of anti-EBV IgG antibodies evaluated in classic Hodgkin lymphoma (cHL) patients to date and provide evidence of a systemic difference in the EBV antibody profile in cHL cases that is both specific to EBV-positive tumors and includes immune aberrations reflecting exposure to multiple stages of the viral life cycle. Evidence of increased, systemic exposure to EBV lytic-cycle activity supports a role for ongoing viral activity in tumor pathogenesis.

Introduction

Epstein-Barr virus (EBV), a gamma-herpesvirus that infects lymphoid and epithelial cells and establishes lifelong latency in 90% of adults globally, is associated with a range of human diseases, including classic Hodgkin lymphoma (cHL) [1, 2]. It is estimated that EBV is causally related to 20%-50% of cHL tumors in immunocompetent people, with the virus being localized to the malignant Hodgkin Reed-Sternberg (HRS) cells [3, 4]. Serologic data provide further evidence to support the association between EBV infection and cHL. Reports evaluating anti-EBV antibodies against the viral capsid antigen [VCA], early antigen [EA], Epstein-Barr nuclear antigen 1 [EBNA1] and EBNA2, and latent membrane protein 1 [LMP1]) indicate that cHL patients harbor elevated levels of anti-EBV IgG antibodies [5-13]. However, most investigations to date have analyzed responses to complexes of proteins (e.g. VCA), rather than individual peptides, and have not investigated reactivity against the full complement of approximately 100 open reading frames translated by EBV. Whether antibodies against sequences representing additional EBV antigens can provide new etiologic insights into the nature of the association between EBV and cHL is unknown.

The EBV status of cHL tumors must be considered when characterizing the immune response to EBV in relation to cHL. Epidemiological data and molecular characteristics suggest that EBV-positive and EBV-negative cHL are likely to be etiologically distinct diseases [4, 8, 9, 14-16]. However, not all of the serological studies conducted to date have distinguished between EBV-positive and EBV-negative tumors in their analyses [5-7, 10]. Indeed, many cHL-related studies are limited by a lack of knowledge of tumor EBV status [8, 9], largely because limited availability of tumor tissue precludes EBV testing and stratified analyses. A serological tool would obviate the need for tumor tissue to conduct work accounting for cHL EBV status and could therefore prove very useful in future epidemiological research.

To address these gaps in knowledge, we applied a recently developed protein microarray technology to measure both IgG and IgA antibody responses against a comprehensive set of sequences

representing 86 EBV proteins in 350 individuals from previous cHL case-control studies conducted in European populations.

Materials and Methods

Study design

Serum samples were selected from studies of cHL in the UK [17, 18], Denmark, and Sweden [19]. Briefly, samples from the UK were derived from two population-based case-control studies and three case series (a total of 102 EBV-positive cases, 41 EBV-negative cases, and 106 controls) between 1993 and 1997 [17, 18]. Samples from Denmark and Sweden were collected as part of the Scandinavian Lymphoma Etiology (SCALE) study, a population-based case-control study conducted among adults in Denmark and Sweden between 1999 and 2002 (37 EBV-positive cHL, 37 EBV-negative cHL, and 37 controls) [19]. All study subjects were non-Hispanic Caucasians. We excluded eight EBV-negative cases and two controls without high quality serum for EBV protein microarray testing, leaving a total of 139 EBV-positive cases, 70 EBV-negative cases, and 141 controls in the present study. Tumor EBV status was determined using immunohistochemical staining of tumor biopsies for EBV latent membrane antigen (LMP)-1 and/or in situ hybridization for EBV-encoded small RNAs (EBERs). All cases were cHL and samples were collected prior to cancer treatment. Samples from controls were frequency matched to EBV-positive cHL cases on sex and age (± 5 years) by study area, and the EBV-positive and EBV-negative cHL cases were further matched on clinical stage. The EBV-negative case group was included to determine whether differences in the EBV antibody pattern were specific to EBV-positive tumors.

All contributing studies were approved by regional scientific ethics committees and data protection agencies, and all participants provided informed consent.

EBV protein microarray

We probed serum samples using an EBV protein microarray targeting IgG and IgA antibodies against 202 EBV sequences including three synthetic peptides [20] and 199 predicted EBV protein sequences representing non-redundant open reading frames in 86 EBV proteins from five prototypical EBV strains (AG876, Akata, B95-8, Mutu, and Raji), as previously described [21-25]. Details are presented in **Supplementary Materials**. Comparison with the outputs from a well-established enzyme-linked immunosorbent assay (ELISA) showed high correlations between the microarray IgA output and previously generated ELISA data for IgA antibodies against VCAp18 and EBNA1 (Spearman coefficient = 0.76 and 0.79, respectively; $P < 0.01$).[25] Each sequence, including the EBV life cycle of each probe based on updated mechanistic information from the literature, is provided in **Table S1**.

We included 25 blinded duplicate samples for quality control during testing and observed good reproducibility (*i.e.*, coefficient of variation [CV]<20%) for antibodies measured using this custom protein microarray. The average CV across the 202 EBV sequences was 12.5% (interquartile range [IQR]: 11.3-14.0 %) for IgG and 14.5% (IQR: 13.1-16.1%) for IgA. We excluded one IgG and five IgA array spots that had CVs >20% from the analysis, leaving a total of 201 IgG and 197 IgA markers.

Circulating B-cells infected with EBV and CCL17 chemokine levels

It is plausible that an elevated antibody response in cHL patients could be attributable to long-standing, uncontrolled EBV activity that pre-disposes to disease, or to increased exposure to viral antigens resulting from the presence of the tumor. To begin disentangling these two possibilities, we leveraged existing data from a subset of the cHL cases diagnosed in the UK on two metrics – the proportion of circulating B-cells infected by EBV (N=14 EBV-positive cases), and levels of the chemokine CCL17 (thymus and activation-regulated chemokine TARC; N=47 EBV-positive cases).

To determine the frequency of circulating EBV-infected B-cells, peripheral blood mononuclear cells (PBMCs) were isolated from peripheral blood (50 – 60 ml) using Lymphoprep™ (Axis-Shield, Dundee, UK), and then B-cells enriched using a CD20+ MiniMACS Separation system (Miltenyi Biotec

Ltd, Surrey, UK). Serial $\sqrt{10}$-fold (3.16-fold) dilution of the enriched B-cells was performed, and eleven

replicates of each dilution, containing 3.16×10^4 to 3.16×10^2 cells, were subjected to real-time PCR. If

insufficient cell numbers were available for the complete dilution series, the starting cell number was

reduced, as appropriate. Ten replicates of each dilution were assayed for EBV using a previously

described PCR targeting the EBV BamHI W repeat sequence, and the final replicate was assayed for β-

globin gene to check that amplifiable DNA was present [26]. PCRs contained 50 nmol/l of each primer, 200

nmol/l of probe, TaqMan Universal PCR Mastermix (ThermoFisher Scientific, Paisley, UK), and the cell

lysate in a total volume of 25 µl. PCR and analysis were performed on an Applied Biosystems 7500 Fast

Real-Time PCR System using software v2.0.6 (ThermoFisher Scientific). Each of the EBV PCRs in the

dilution series was scored as positive or negative, and results analysed using ELDA software [26]. Results

are expressed as the estimated number (and 95% confidence interval [CI]) of EBV-positive cells per 10^6

B-cells. DNA from the Namalwa cell line was used as a positive control, and nine no template controls

were included in each assay.

CCL17 was quantified using a Human CCL17/TARC Quantikine ELISA (R&D Systems®,

Abingdon, UK), according to the manufacturer's instructions [26]. Pre-treatment serum samples were

initially tested at a 1 in 10 dilution, and further dilutions were analyzed when results fell outside the

dynamic range of the assay. All samples were analyzed in duplicate and results used only if coefficients

of variance (CVs) were <20%; mean values are reported.

Statistical analysis

To examine the association between biological groupings of antibody response (e.g., responses

directed against lytic-cycle proteins) and cHL, we used the SNP-set (Sequence) Kernel Association Test

(SKAT) [27] with a binary variable parameterization for each marker (1=positive, 0=negative). Effect

estimates for biological groupings of antibody response were considered statistically significant if their *P*-

value was ≤ 0.005, corresponding to a Bonferroni correction for 10 tests (10 different stages of the viral

life cycle evaluated).

We investigated differences in the mean standardized signal intensity (SSI) for IgG and IgA antibodies against each of the 202 array sequences across EBV-positive cHL cases, EBV-negative cHL cases, and controls using one-way analysis of variance (ANOVA). We computed p-values for all two-way comparisons between groups (e.g., EBV-positive cHL versus controls). Our original hypothesis was that case-control differences would be limited to EBV-positive cHL. Antibody associations with cHL were considered statistically significant if their P-value was $\leq 1.3 \times 10^{-4}$, corresponding to a Bonferroni correction for 400 tests (about 200 probes on IgG antibodies and IgA antibodies).

Odds ratios (ORs) and 95% CIs for the association between each 3-level categorical anti-EBV antibody variable (*i.e.*, tertiles) and cHL status were calculated using polytomous logistic regression models adjusted for sex, age, and study area. Among subjects from the SCALE study for whom information on a history of infectious mononucleosis (IM) was complete (information on IM was missing for 52% of the subjects in the UK), models were further adjusted for a history of IM. *P*-trends were calculated from a model with each 3-level antibody marker treated as an ordinal variable using Wald test. Because the etiology may differ between young-adult cHL and older-adult cHL[28-30] and the age-specific incidence decreases from age 20 to 40 years but increases after age 40 years [30], we assessed heterogeneity of associations by age group (<40 years vs. ≥40 years). A likelihood ratio test was used to compare logistic regression models with and without an interaction term between each anti-EBV antibody and age group. We also evaluated cHL in relation to the anti-EBNA1/anti-EBNA2 ratio (≤1 vs. >1), a metric that has been used in previous serological studies as a proxy for defective immunity against EBV [31, 32].

In addition to comparing anti-EBV humoral immunity between cHL cases and controls, we also conducted additional analyses to understand which set of markers was best able to distinguish EBV-positive from EBV-negative cHL. Specifically, we first computed the MeanDecreaseGini and MeanDecreaseAccuracy metrics using random forests (R package randomForest,). Second, we evaluated antibodies with differences meeting the $P \leq 1.3 \times 10^{-4}$ threshold using a stepwise logistic regression model as both continuous and categorical variables (tertiles). We set $P < 0.15$ as the model entry criterion

and $P < 0.05$ for an antibody to remain in the model (LOGISTIC procedure in the SAS statistical package). Finally, we selected markers that were selected by both the stepwise logistic regression and the two randomForest prediction metrics as our best potential predictors. We evaluated the ability of the selected anti-EBV antibody markers (as continuous levels) together with the subject's characteristics (i.e., age group, sex, and study area) to classify the EBV status of cHL tumors using the area under the receiver operative curve (AUC). We compared this AUC with that obtained based only on the subject's characteristics with a 10-fold cross-validation.

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Results

Table 1 shows the distribution of baseline demographic characteristics among the 139 EBV-positive cHL cases, 70 EBV-negative cHL cases, and 141 controls. Cases and controls had a similar age and sex distribution. As expected, nearly all adults in this population were EBV carriers, as illustrated using two of the synthetic peptides on the array. Among controls, EBV-negative cHL cases, and EBV-positive cHL cases, the sero-positive rates for IgG antibodies against VCAp18 were 92.9%, 88.6%, and 98.6%, respectively; for IgG antibodies against EBNA1, rates were 87.9%, 85.7%, and 84.2%, respectively.

Among cHL samples from the UK where information on the histopathology was available, 47 cHL cases were mixed cellularity, and 73 were nodular sclerosis subtype. We observed no significant differences in SSI between 40 EBV-positive cases with mixed cellularity and 46 with nodular sclerosis (all Bonferroni-corrected P values from t-test >0.05).

EBV-positive cHL cases vs. controls

The proteome-wide IgG repertoire (i.e., a combination of the IgG antibody responses to all array probes) was statistically significantly different between EBV-positive cHL cases and controls (P_{SKAT-C} for total IgG=0.003, P_{SKAT-C} for total IgA=0.128, **Figure 1A**). When considering anti-EBV antibody responses against proteins from specific stages of the EBV life cycle, the EBV-directed IgG repertoire to late lytic (P_{SKAT-C} = 0.001) and early lytic (P_{SKAT-C} =0.004) proteins differed significantly between EBV-positive cHL cases and controls (**Figure 1A**). We also observed a suggestive association for the IgG repertoire against sequences representing latent proteins (P_{SKAT-C} =0.020).

Comparison of the mean difference in SSI for each individual array sequence between EBV-positive cHL cases and controls revealed nominally significant differences ($P \leq 0.05$) in 56.7% of the IgG (114/201) and 4.6% of the IgA (9/197) anti-EBV antibodies. Sixteen IgG and two IgA anti-EBV antibodies remained significantly elevated in EBV-positive cHL cases after adjusting for multiple testing ($P_{t-test} \leq 1.3 \times 10^{-4}$; **Figure 1B and Table 2**). The most pronounced SSI differences between EBV-positive cHL cases and controls were observed for IgG antibodies against sequences representing three lytic-cycle enzymatic proteins (BBRF1, Thymidine Kinase [TK] and BBLF1) and one component of the viral capsid (BcLF1) (**Figure S1A-1D**). We also observed strong IgA differences between EBV-positive cases and controls for two probes on the array representing BBLF1.

After adjustment for age, sex, and residential area, ORs $_{highest\ vs.\ lowest\ tertile}$ for antibodies with $P_{t-test} \leq 1.3 \times 10^{-4}$ ranged from 1.79 to 4.99 (**Table 2**), and IgG markers representing three EBV proteins (BcLF1, BBLF1, and LMP1) had adjusted ORs > 3. The strongest OR effect was observed for antibody against sequences representing LMP-1 (adjusted OR $_{highest\ vs.\ lowest\ tertile}$=4.99, 95% CI: 2.51, 9.94, P_{trend}<0.0001), an EBV oncogene that is highly expressed by the HRS cells in EBV-positive cHL [33]. The mean array output (SSI level) representing IgG responses to these three proteins in EBV-cHL cases ranged from 1.09 for LMP1 to 2.21 for BBLF1, levels comparable to the IgG response in these cases to known disease biomarkers VCA-p18 and EBNA1 (2.00 and 1.62, respectively). No heterogeneity in the EBV-cHL

associations was observed by age group (<40 years vs. ≥40 years, all $P_{\text{heterogeneity}}$ >0.05 after Bonferroni correction).

EBV-positive cHL cases were more likely than controls to have an aberrant EBNA response pattern (i.e., anti-EBNA1:EBNA2 ratio <1), although most associations were not statistically significant after considering all possible combinations of the various EBNA1 and 2 probes on the array (**Table S2**). Among participants from the SCALE study, additional adjustment for a history of infectious mononucleosis did not materially change the observed associations (**Table S3).**

Correlation of IgG markers with the proportion of EBV-infected B-cell or chemokine CCL17 levels

To investigate potential underlying mechanisms driving the elevated anti-EBV IgG signature in EBV-positive cHL patients, we examined the correlations between our top markers and either (1) the proportion of circulating B-cells infected by EBV or (2) levels of the chemokine CC17, an indirect measure of tumor burden, using Spearman correlation coefficients. Of the 18 antibodies (16 IgG and 2 IgA) that associated with EBV-positive cHL, suggestive correlations (*P*-values<0.05) were observed between three markers (i.e., BBRF1-IgG, BBLF1-IgG, and BcLF1-IgG) and the proportion of EBV-infected B-cells in circulation, with the highest Spearman coefficient observed for BBRF1-IgG (Spearman coefficient=0.588, *P*-value=0.027, **Figure S2A**). By contrast, none of the 18 antibodies were correlated with chemokine CCL17 levels, as illustrated in **Figure S2B** for BBRF1-IgG.

EBV-negative cHL cases vs. controls

Neither the IgG nor IgA proteome-wide repertoire was significantly different between EBV-negative cHL cases and controls (**Figure 2A**). No specific IgG or IgA anti-EBV antibodies were significantly different in EBV-negative cHL cases compared to controls after adjusting for multiple testing (**Figure 2B**).

EBV-positive cHL cases vs. EBV-negative cHL cases

EBV-positive cHL patients were more likely than EBV-negative cHL patients to display an elevated anti-EBV IgG response (**Figure 3A**). A total of six IgG markers met the $P \leq 1.3 \times 10^{-4}$ significance threshold for elevation specific to EBV-positive cHL (**Table 3**). Adjusted ORs highest vs. lowest tertile for these six antibodies ranged from 2.94 to 5.27 (**Table 3**). We further assessed which of these markers were best able to distinguish EBV-positive from EBV-negative cHL based on random forests and stepwise logistic regression prediction metrics (BZLF1[Zta]-IgG and BdRF1[VCAp40]-IgG, **Figures S3A-3B**). The prediction performance for classifying EBV-positive versus EBV-negative cHL was significantly improved by including these two antibodies (continuous variable, area under the curve [AUC]=0.75, 95% CI=0.68, 0.83) versus considering only subject age, sex, and study area alone (AUC=0.64, 95% CI=0.56, 0.72, P-value=0.002, **Figure 3B**). Based on a 10-fold cross-validation, the AUC of this combination of serological markers plus patient demographics was 0.70 (95% CI: 0.63, 0.78).

Discussion

This is the first study to evaluate the antibody response to the complete EBV proteome in cHL patients. Our results demonstrate that EBV-positive cHL cases have distinct serological responses compared with both EBV-negative cHL cases and controls. The significant elevations in 16 IgG and two IgA markers in EBV-positive cHL cases are unlikely to be confounded by age, sex, and residential area. Notably, no difference in the anti-EBV antibody profile was observed between EBV-negative cHL cases and controls. Our results support the hypothesis that differences in the EBV antibody profile are specific to patients with EBV-positive cHL and are not universally observed as part of a systematically dysregulated immune response present in all cHL cases.

The EBV-positive cHL disease associations were disproportionately observed for IgG rather than IgA markers. This finding stands in contrast to the elevated levels of IgA antibody, which marks exposure to antigens along mucosal surfaces such as the oral epithelium,[11] that we and others report in

patients with nasopharyngeal carcinoma (NPC), an EBV-associated epithelial tumor.[11, 25] Previous

studies have reported significant elevations in IgA antibodies against VCA and EA among patients with

EBV-positive cHL,[6, 8] and we observed a difference between EBV-positive cHL and controls for BBLF1-

IgA. However, the lack of a broad IgA response specific to EBV-positive cHL cases supports the

hypothesis that IgG antibodies indicative of systemic exposure to EBV infection of circulating B-cells

constitute a more relevant marker for lymphoid tumors. Results from a similar study among pediatric

patients with Burkitt lymphoma (BL) also support the absence of a strong disease effect for IgA in

lymphoid tumors [34].

Historically, studies largely focused on antibodies against sequences representing four EBV

protein complexes or antigens: VCA, EA, EBNA1, and EBNA2. Here, we extend these findings to

include antibodies against a broad panel of viral proteins involved in various stages of the EBV life cycle.

Notably, we observed strong associations for immune responses representing EBV proteins involved in

replication, including the early (TK, BBLF1, and BALF2 [EA(D)_p138]) and late lytic cycle (several

components of viral capsid [BcLF1 and BdRF1] and BBRF1) antigens. We could not determine in this

retrospective study design whether this lytic activity was due to the presence of neoplastic tissue, or

whether it is reflecting an ability to control EBV in the years preceding cHL diagnosis. A prospective

study could potentially elucidate whether individuals exhibit the unique anti-EBV antibody pattern prior

to disease onset, but such a study design is challenging due to the low incidence of cHL.

To begin to understand the biology underlying the EBV-positive cHL antibody signature, we

leveraged existing data relevant to two alternative hypotheses that the elevated anti-EBV IgG responses

reflect a history of uncontrolled EBV infection that predisposes to disease, or that this antibody signature

is a reflection of EBV antigen production in the tumor. Although our data are not definitive, our cHL IgG

signature correlated with the proportion of EBV-infected circulating B-cells, rather than CCL17 level (an

indirect indication of tumor burden) [35]. This supports the assertion that the ability to control EBV

replication prior to disease is associated with cHL risk, as suggested previously [36].

In addition to elevated lytic-cycle EBV activity in cHL, we observed significant associations between EBV-positive cHL and IgG antibody responses to LMP-1, BHRF1, and BARF1, three proteins that are implicated in oncogenesis [33, 37, 38]. These antibodies do not necessarily directly reflect EBV gene expression in the tumor, but elevated expression of these molecules may be relevant for B-cell survival and eventual risk of tumor development [12, 13]. LMP-1 functions as a constitutively active CD40 molecule leading to NF-kappaB signaling and up-regulation of pro-apoptotic proteins including Bmi-1 (Bcl-2-interacting mediator of cell death) [37]. In EBV-associated cHL, it is thought to play a key role in rescuing HRS cells, or their precursors, from apoptotic death in germinal centers. BHRF1 is an EBV homolog of Bcl-2, and BARF1 also has anti-apoptotic function [39], suggesting that these proteins may also contribute to cell survival at some stage in disease pathogenesis. Despite eliciting a weaker overall immune response than some traditional disease biomarkers (e.g., VCA), reactivity to LMP-1 was significantly higher in EBV-positive cHL patients relative to disease-free controls.

Importantly, our present findings support the body of epidemiological evidence suggesting that EBV-positive and EBV-negative cHL are distinct diseases [4, 8, 9, 14-16]. Chang *et al.* reported that EBV-positive tumors were more likely than EBV-negative tumors to affect adults with less education and a history of cigarette smoking.[8] In addition, people with a history of infectious mononucleosis, which is caused by primary EBV infection in adolescence or young adulthood, are at a higher risk of developing EBV-positive cHL but not EBV-negative cHL.[14, 15, 40, 41] Despite this evidence supporting two distinct diseases, a lack of cHL tumor tissue samples for EBV status testing often precludes researchers from conducting risk factor analyses stratified by EBV status. A serological tool that could distinguish EBV status could therefore be immensely useful. Classic serological markers (e.g., EBNA1-IgG) did not yield a significant discrimination between EBV-positive and EBV-negative cHL cases in our study, which is in line with previous findings [11, 12]. Examining the antibody profile against the full viral proteome allowed us to identify a two-marker combination that classified tumor EBV status in our samples better than patient demographics alone: BdRF1(VCAp40)-IgG and BZLF1(Zta)-IgG. Whether these two markers

can achieve high accuracy for distinguishing EBV-positive from EBV-negative cHL using only blood samples merits further investigation. Such a blood-based tool for determining tumor EBV status would represent an opportunity to utilize large and racially diverse biobanks for cHL research without necessitating the presence of archived tumor for EBER staining.

Our results should be interpreted in light of certain methodologic limitations. First, although there is strong biologic plausibility for the association between antibody responses to EBV and EBV-positive cHL, our observations of case-control differences for specific markers need independent validation. In addition to identifying case-control differences in the EBV-directed antibody repertoire, our case-case comparison identified a two-marker combination that classified tumor EBV status among cases. Although this two-marker combination was internally validated using 10-fold cross-sampling in our study population, our approach is still considered as exploratory since we lacked an independent, external dataset for replication. Second, we lacked information for all subjects on education level, smoking, a history of infectious mononucleosis (IM), and other potential confounding factors. However, although IM is a risk factor for EBV-positive cHL, the prevalence of this disease in our study population overall appeared to be relatively low (i.e., ~9% among 108 subjects from the SCALE study). Among subjects for whom information on a history of IM was available, additional adjustment for IM did not materially change the observed antibody associations. Likewise, in a previous study, adjustment for potential confounding factors did not appreciably change the associations between anti-EBV antibodies (*i.e.*, VCA, EA, EBNA1, and EBNA2) and EBV-positive cHL.[8] Another limitation to note was that this array was not designed to detect antibodies to conformational epitopes, which precluded us from examining cHL associations for select transcripts that require glycosylation, including surface glycoproteins involved in virus neutralization.

In conclusion, we characterized EBV-directed antibody responses to 202 protein sequences representing 86 EBV proteins in cHL patients. Compared with patients with EBV-negative cHL and controls, patients with EBV-positive cHL displayed a distinct EBV-directed serological profile, with

15

significant elevation in several IgG antibodies. This disease-associated antibody pattern included differences in antibody responses to proteins involved in EBV replication and anti-apoptotic signaling, providing clues for future EBV-positive cHL pathogenesis research. Future studies are needed to better understand why some individuals cannot control EBV infection appropriately and how antibody responses reflect this process and can be used for risk stratification. Additional studies to elucidate etiological factors for EBV-negative cHL are also needed, as the etiology of this subset of cHL remains elusive.

Acknowledgements

Authorship Contributions

Conception and design: Zhiwei Liu, Ruth Jarrett, Henrik Hjalgrim, Allan Hildesheim, and Anna E. Coghill

Provision of study materials or patients: Ruth Jarrett, Henrik Hjalgrim, Karin E. Smedby, Annette Lake, Hans-Olov Adami, Bengt Glimelius, Mads Melbye

Collection and assembly of data: Ruth Jarrett, Henrik Hjalgrim, Carla Proietti, Ellen T. Chang, Karin E. Smedby, Kelly J. Yu, Annette Lake, Sally Troy, Hans-Olov Adami, Bengt Glimelius, Mads Melbye, Denise Doolan

Data analysis and interpretation: Zhiwei Liu, Ruth Jarrett, Henrik Hjalgrim, Carla Proietti, Ruth M. Pfeiffer, Allan Hildesheim, Denise Doolan, Anna E. Coghill

Manuscript writing: Zhiwei Liu and Anna E. Coghill

Final approval of manuscript: Zhiwei Liu, Ruth Jarrett, Henrik Hjalgrim, Carla Proietti, Ellen T. Chang, Karin E. Smedby, Kelly J. Yu, Annette Lake, Karen A. McAulay, Sally Troy, Ruth M. Pfeiffer, Hans-Olov Adami, Bengt Glimelius, Mads Melbye, Allan Hildesheim, Denise Doolan, and Anna E. Coghill

Disclosure of Conflicts of Interest. No authors report any conflicts of interest.

Reference

1. Young LS, Rickinson AB. Epstein-Barr virus: 40 years on. *Nature reviews Cancer* 2004;**4**: 757-68.
2. Ambinder RF. Epstein-barr virus and hodgkin lymphoma. *Hematology Am Soc Hematol Educ Program* 2007: 204-9.
3. Weiss LM, Movahed LA, Warnke RA, Sklar J. Detection of Epstein-Barr viral genomes in Reed-Sternberg cells of Hodgkin's disease. *The New England journal of medicine* 1989;**320**: 502-6.
4. Glaser SL, Lin RJ, Stewart SL, Ambinder RF, Jarrett RF, Brousset P, Pallesen G, Gulley ML, Khan G, O'Grady J, Hummel M, Preciado MV, et al. Epstein-Barr virus-associated Hodgkin's disease: epidemiologic characteristics in international data. *International journal of cancer Journal international du cancer* 1997;**70**: 375-82.
5. Evans AS, Gutensohn NM. A population-based case-control study of EBV and other viral antibodies among persons with Hodgkin's disease and their siblings. *Int J Cancer* 1984;**34**: 149-57.
6. Mueller N, Evans A, Harris NL, Comstock GW, Jellum E, Magnus K, Orentreich N, Polk BF, Vogelman J. Hodgkin's disease and Epstein-Barr virus. Altered antibody pattern before diagnosis. *The New England journal of medicine* 1989;**320**: 689-95.
7. Lehtinen T, Lumio J, Dillner J, Hakama M, Knekt P, Lehtinen M, Teppo L, Leinikki P. Increased risk of malignant lymphoma indicated by elevated Epstein-Barr virus antibodies--a prospective study. *Cancer Causes Control* 1993;**4**: 187-93.
8. Chang ET, Zheng T, Lennette ET, Weir EG, Borowitz M, Mann RB, Spiegelman D, Mueller NE. Heterogeneity of risk factors and antibody profiles in epstein-barr virus genome-positive and -negative hodgkin lymphoma. *The Journal of infectious diseases* 2004;**189**: 2271-81.
9. Levin LI, Chang ET, Ambinder RF, Lennette ET, Rubertone MV, Mann RB, Borowitz M, Weir EG, Abbondanzo SL, Mueller NE. Atypical prediagnosis Epstein-Barr virus serology restricted to EBV-positive Hodgkin lymphoma. *Blood* 2012;**120**: 3750-5.
10. Mueller NE, Lennette ET, Dupnik K, Birmann BM. Antibody titers against EBNA1 and EBNA2 in relation to Hodgkin lymphoma and history of infectious mononucleosis. *International journal of cancer Journal international du cancer* 2012;**130**: 2886-91.
11. Coghill AE, Hildesheim A. Epstein-Barr virus antibodies and the risk of associated malignancies: review of the literature. *American journal of epidemiology* 2014;**180**: 687-95.
12. Meij P, Vervoort MB, Bloemena E, Schouten TE, Schwartz C, Grufferman S, Ambinder RF, Middeldorp JM. Antibody responses to Epstein-Barr virus-encoded latent membrane protein-1 (LMP1) and expression of LMP1 in juvenile Hodgkin's disease. *J Med Virol* 2002;**68**: 370-7.
13. de Sanjose S, Bosch R, Schouten T, Verkuijlen S, Nieters A, Foretova L, Maynadie M, Cocco PL, Staines A, Becker N, Brennan P, Benavente Y, et al. Epstein-Barr virus infection and risk of lymphoma: immunoblot analysis of antibody responses against EBV-related proteins in a large series of lymphoma subjects and matched controls. *Int J Cancer* 2007;**121**: 1806-12.

14. Hjalgrim H, Askling J, Rostgaard K, Hamilton-Dutoit S, Frisch M, Zhang JS, Madsen M, Rosdahl N, Konradsen HB, Storm HH, Melbye M. Characteristics of Hodgkin's lymphoma after infectious mononucleosis. *The New England journal of medicine* 2003;**349**: 1324-32.

15. Hjalgrim H, Smedby KE, Rostgaard K, Molin D, Hamilton-Dutoit S, Chang ET, Ralfkiaer E, Sundstrom C, Adami HO, Glimelius B, Melbye M. Infectious mononucleosis, childhood social environment, and risk of Hodgkin lymphoma. *Cancer research* 2007;**67**: 2382-8.

16. Urayama KY, Jarrett RF, Hjalgrim H, Diepstra A, Kamatani Y, Chabrier A, Gaborieau V, Boland A, Nieters A, Becker N, Foretova L, Benavente Y, et al. Genome-wide association study of classical Hodgkin lymphoma and Epstein-Barr virus status-defined subgroups. *Journal of the National Cancer Institute* 2012;**104**: 240-53.

17. Jarrett RF, Krajewski AS, Angus B, Freeland J, Taylor PR, Taylor GM, Alexander FE. The Scotland and Newcastle epidemiological study of Hodgkin's disease: impact of histopathological review and EBV status on incidence estimates. *Journal of clinical pathology* 2003;**56**: 811-6.

18. Johnson PC, McAulay KA, Montgomery D, Lake A, Shield L, Gallagher A, Little AM, Shah A, Marsh SG, Taylor GM, Jarrett RF. Modeling HLA associations with EBV-positive and -negative Hodgkin lymphoma suggests distinct mechanisms in disease pathogenesis. *Int J Cancer* 2015;**137**: 1066-75.

19. Smedby KE, Hjalgrim H, Melbye M, Torrang A, Rostgaard K, Munksgaard L, Adami J, Hansen M, Porwit-MacDonald A, Jensen BA, Roos G, Pedersen BB, et al. Ultraviolet radiation exposure and risk of malignant lymphomas. *Journal of the National Cancer Institute* 2005;**97**: 199-209.

20. Liu Z, Yu KJ, Coghill AE, Brenner N, Cao SM, Chen CJ, Chen Y, Doolan DL, Hsu WL, Labo N, Middeldorp JM, Miley W, et al. Multi-laboratory assessment of EBV serologic assays: the case for standardization. *J Clin Microbiol* 2019.

21. Vigil A, Davies DH, Felgner PL. Defining the humoral immune response to infectious agents using high-density protein microarrays. *Future Microbiol* 2010;**5**: 241-51.

22. Doolan DL. Plasmodium immunomics. *Int J Parasitol* 2011;**41**: 3-20.

23. Davies DH, Duffy P, Bodmer JL, Felgner PL, Doolan DL. Large screen approaches to identify novel malaria vaccine candidates. *Vaccine* 2015;**33**: 7496-505.

24. Liu Z, Coghill AE, Pfeiffer RM, Proietti C, Hsu WL, Chien YC, Lekieffre L, Krause L, Yu KJ, Lou PJ, Wang CP, Mulvenna J, et al. Patterns of Interindividual Variability in the Antibody Repertoire Targeting Proteins Across the Epstein-Barr Virus Proteome. *The Journal of infectious diseases* 2018;**217**: 1923-31.

25. Coghill AE, Pfeiffer RM, Proietti C, Hsu WL, Chien YC, Lekieffre L, Krause L, Teng A, Pablo J, Yu KJ, Lou PJ, Wang CP, et al. Identification of a Novel, EBV-Based Antibody Risk Stratification Signature for Early Detection of Nasopharyngeal Carcinoma in Taiwan. *Clin Cancer Res* 2018;**24**: 1305-14.

26. Gallagher A, Armstrong AA, MacKenzie J, Shield L, Khan G, Lake A, Proctor S, Taylor P, Clements GB, Jarrett RF. Detection of Epstein-Barr virus (EBV) genomes in the serum of patients with EBV-associated Hodgkin's disease. *International journal of cancer Journal international du cancer* 1999;**84**: 442-8.

27. Ionita-Laza I, Lee S, Makarov V, Buxbaum JD, Lin X. Sequence kernel association tests for the combined effect of rare and common variants. *American journal of human genetics* 2013;**92**: 841-53.

28. Chang ET, Montgomery SM, Richiardi L, Ehlin A, Ekbom A, Lambe M. Number of siblings and risk of Hodgkin's lymphoma. *Cancer Epidemiol Biomarkers Prev* 2004;**13**: 1236-43.

29. Hjalgrim H, Askling J, Sorensen P, Madsen M, Rosdahl N, Storm HH, Hamilton-Dutoit S, Eriksen LS, Frisch M, Ekbom A, Melbye M. Risk of Hodgkin's disease and other cancers after infectious mononucleosis. *Journal of the National Cancer Institute* 2000;**92**: 1522-8.

30. Hjalgrim H, Askling J, Pukkala E, Hansen S, Munksgaard L, Frisch M. Incidence of Hodgkin's disease in Nordic countries. *Lancet* 2001;**358**: 297-8.

31. Henle W, Henle G, Andersson J, Ernberg I, Klein G, Horwitz CA, Marklund G, Rymo L, Wellinder C, Straus SE. Antibody responses to Epstein-Barr virus-determined nuclear antigen (EBNA)-1 and EBNA-2 in acute and chronic Epstein-Barr virus infection. *Proc Natl Acad Sci U S A* 1987;**84**: 570-4.

32. Lennette ET, Rymo L, Yadav M, Masucci G, Merk K, Timar L, Klein G. Disease-related differences in antibody patterns against EBV-encoded nuclear antigens EBNA 1, EBNA 2 and EBNA 6. *Eur J Cancer* 1993;**29A**: 1584-9.

33. Cader FZ, Vockerodt M, Bose S, Nagy E, Brundler MA, Kearns P, Murray PG. The EBV oncogene LMP1 protects lymphoma cells from cell death through the collagen-mediated activation of DDR1. *Blood* 2013;**122**: 4237-45.

34. Coghill AE, Proietti C, Liu Z, Krausse L, Troy S, Bethony J, Nkrumah F, Bhatia K, Hildesheim A, Doolan D, Mbulaiteye S. Characterizing the association between the Epstein-Barr virus (EBV) serological profile and risk of endemic Burkitt lymphoma (BL) in Ghanaian children (submitted) 2018.

35. Plattel WJ, van den Berg A, Visser L, van der Graaf AM, Pruim J, Vos H, Hepkema B, Diepstra A, van Imhoff GW. Plasma thymus and activation-regulated chemokine as an early response marker in classical Hodgkin's lymphoma. *Haematologica* 2012;**97**: 410-5.

36. Drouet E, Brousset P, Fares F, Icart J, Verniol C, Meggetto F, Schlaifer D, Desmorat-Coat H, Rigal-Huguet F, Niveleau A, Delsol G. High Epstein-Barr virus serum load and elevated titers of anti-ZEBRA antibodies in patients with EBV-harboring tumor cells of Hodgkin's disease. *J Med Virol* 1999;**57**: 383-9.

37. Dutton A, Woodman CB, Chukwuma MB, Last JI, Wei W, Vockerodt M, Baumforth KR, Flavell JR, Rowe M, Taylor AM, Young LS, Murray PG. Bmi-1 is induced by the Epstein-Barr virus oncogene LMP1 and regulates the expression of viral target genes in Hodgkin lymphoma cells. *Blood* 2007;**109**: 2597-603.

38. Tarbouriech N, Ruggiero F, de Turenne-Tessier M, Ooka T, Burmeister WP. Structure of the Epstein-Barr virus oncogene BARF1. *J Mol Biol* 2006;**359**: 667-78.

39. Wang Q, Tsao SW, Ooka T, Nicholls JM, Cheung HW, Fu S, Wong YC, Wang X. Anti-apoptotic role of BARF1 in gastric cancer cells. *Cancer Lett* 2006;**238**: 90-103.

40. Alexander FE, Lawrence DJ, Freeland J, Krajewski AS, Angus B, Taylor GM, Jarrett RF. An epidemiologic study of index and family infectious mononucleosis and adult Hodgkin's disease (HD): evidence for a specific association with EBV+ve HD in young adults. *Int J Cancer* 2003;**107**: 298-302.

41. Alexander FE, Jarrett RF, Lawrence D, Armstrong AA, Freeland J, Gokhale DA, Kane E, Taylor GM, Wright DH, Cartwright RA. Risk factors for Hodgkin's disease by Epstein-Barr virus (EBV) status: prior infection by EBV and other agents. *Br J Cancer* 2000;**82**: 1117-21.

Table 1. Baseline characteristics by case-control status [a]

Characteristics	EBV-positive cHL cases (n=139)	EBV-negative cHL cases (n=70)	Controls (n=141)
Study Area			
Scotland, UK	102 (73.4)	36 (51.4)	104 (73.8)
Demark	15 (10.8)	15 (21.4)	15 (10.6)
Sweden	22 (15.8)	19 (27.1)	22 (15.6)
Sex			
Female	45 (32.4)	25 (35.7)	45 (31.9)
Male	94 (67.6)	45 (64.3)	96 (68.1)
Age at diagnosis (years) [b]			
<30	41 (29.5)	22 (31.4)	44 (31.2)
30 - 39	24 (17.4)	18 (25.7)	25 (17.7)
40 - 49	17 (12.3)	6 (8.6)	17 (12.1)
50 - 59	27 (19.6)	10 (14.3)	29 (20.6)
60+	29 (21.0)	14 (20.0)	26 (18.4)
Histological subtype			
Mixed cellularity	40 (28.8)	7 (10.0)	--
Nodular sclerosis	46 (33.1)	27 (38.6)	--
Other/Unknown [c]	53 (38.1)	36 (51.4)	--
Clinical Stage			
I	18 (20.0)	9 (12.8)	--
II	40 (43.5)	34 (48.6)	--
III	19 (20.6)	16 (22.9)	--
IV	15 (13.9)	11 (15.7)	--
Unknown	47	0	--

Abbreviation: cHL, classic Hodgkin lymphoma

a All data are given as n (%) except where indicated

b Age information was missing for one EBV-positive cHL case

c Other includes lymphocyte-depleted, lymphocyte-rich classic, and unclassified types.

Table 2. Odds ratios (OR) and 95% confidence intervals (CI) for the association between anti-EBV antibody level and EBV-positive classic Hodgkin lymphoma (cHL) vs. controls [a]

EBV Protein and Array sequence	Antibody Type	t-test P	EBV-positive cHL Mean	Control Mean	EBV-positive cHL Positivity	Control Positivity	OR tertile 2 (95% CI)	OR tertile 3 (95% CI)	P-trend
BBRF1 (Late lytic) YP_001129476.1-102746-104587	IgG	1.4×10^{-7}	1.34	1.09	68.3%	47.5%	1.33 (0.69, 2.54)	2.66 (1.45, 4.90)	1.0×10^{-3}
BcLF1 (VCA_p160) AFY97965.1-125044-120899-1	IgG	1.8×10^{-7}	1.27	1.04	63.3%	41.8%	1.73 (0.88, 3.41)	3.94 (2.08, 7.45)	1.5×10^{-5}
Thymidine kinase (Early lytic) YP_001129497.1-133399-131576	IgG	5.1×10^{-7}	1.51	1.28	90.6%	83.0%	0.96 (0.49, 1.85)	2.71 (1.50, 4.89)	3.8×10^{-4}
BBLF1 (Tegument protein) AFY97956.1-108555-108328	IgG	1.2×10^{-6}	2.09	1.83	98.6%	90.1%	1.67 (0.84, 3.32)	3.68 (1.95, 6.96)	3.2×10^{-5}
BBLF1 (Tegument protein) YP_001129480.1-109516-109289	IgA	4.6×10^{-6}	1.23	1.07	73.4%	51.8%	0.93 (0.47, 1.83)	3.03 (1.67, 5.47)	5.8×10^{-5}
BBLF1 (Tegument protein) YP_001129480.1-109516-109289	IgG	4.6×10^{-6}	2.21	1.96	99.3%	92.2%	2.23 (1.13, 4.40)	3.79 (1.98, 7.27)	5.7×10^{-5}
BBLF1 (Tegument protein) AFY97956.1-108555-108328	IgA	6.6×10^{-6}	1.23	1.07	76.3%	51.1%	1.37 (0.69, 2.72)	3.75 (2.02, 6.98)	7.0×10^{-6}
BcLF1 (VCA_p160) YP_001129493.1-126005-121860-1	IgG	1.2×10^{-5}	1.30	1.12	66.9%	53.9%	2.10 (1.09, 4.08)	3.27 (1.72, 6.20)	3.0×10^{-4}
BcLF1 (VCA_p160) CAA24794.1-137466-133321-1	IgG	1.4×10^{-5}	1.33	1.15	78.4%	60.3%	1.51 (0.80, 2.88)	2.50 (1.35, 4.62)	3.0×10^{-3}
BdRF1 (VCA_p40) AFY97974.1-136284-137321	IgG	4.0×10^{-5}	1.69	1.46	90.6%	83.0%	1.09 (0.56, 2.09)	2.54 (1.39, 4.63)	1.1×10^{-3}
BDLF3 (glycoprotein 150) AFY97964.1-118644-117940	IgG	5.1×10^{-5}	1.87	1.65	98.6%	91.5%	1.91 (1.00, 3.65)	2.82 (1.52, 5.24)	1.2×10^{-3}
BBRF3 (glycoprotein M) YP_001129479.1-107679-108896	IgG	6.4×10^{-5}	1.69	1.50	96.4%	85.8%	2.25 (1.19, 4.27)	2.56 (1.37, 4.78)	4.8×10^{-3}

BDLF3 (glycoprotein 150) YP_001129490.1-119605-118901	IgG	9.2×10^{-5}	1.88	1.66	97.1%	87.9%	1.71 (0.89, 3.26)	2.80 (1.52, 5.18)	1.0×10^{-3}
BHRF1 (Bcl-2 homolog) YP_001129442.1-42204-42779	IgG	1.1×10^{-4}	1.43	1.29	91.4%	73.8%	2.87 (1.5, 5.48)	2.80 (1.46, 5.36)	4.2×10^{-3}
BFLF2 (Late lytic) YP_001129443.1-44763-43807	IgG	1.1×10^{-4}	1.10	0.99	50.4%	35.5%	1.74 (0.92, 3.29)	2.61 (1.41, 4.83)	2.6×10^{-3}
LMP-1 (Oncogene) YP_401722.1-168507-167702	IgG	1.2×10^{-4}	1.09	0.96	43.9%	19.9%	1.90 (1.01, 3.57)	4.99 (2.51, 9.94)	2.7×10^{-6}
BALF2 (EA(D)_p138) single-stranded DNA binding protein YP_001129510.1-165796-162410-1	IgG	1.3×10^{-4}	1.26	1.13	74.8%	62.4%	1.48 (0.77, 2.83)	2.69 (1.45, 4.98)	1.3×10^{-3}
BARF1 (Oncogene) YP_001129453.1-66746-67654	IgG	1.3×10^{-4}	1.45	1.27	82.0%	74.5%	0.72 (0.38, 1.37)	1.79 (1.02, 3.16)	2.9×10^{-2}
VCA-p18 (synthetic peptide)	IgG	0.0029	2.00	1.86	98.6%	92.9%	1.91 (1.02, 3.60)	2.50 (1.33, 4.71)	5.2×10^{-3}
EBNA1 (synthetic peptide)	IgG	0.46	1.62	1.65	84.2%	87.9%	1.13 (0.64, 2.00)	0.67 (0.37, 1.22)	0.203
EAd (synthetic peptide)	IgG	0.05	0.97	0.92	10.8%	10.6%	1.49 (0.81, 2.74)	1.69 (0.92, 3.11)	0.101

a Table is ordered by t-test p-value (lowest to highest). ORs were adjusted for age group (<30, 30-39, 40-49, 50-59, and 60+ years), sex, and study area (UK, Denmark, and Sweden). The tertiles were calculated using the underlying antibody distribution among controls. All odds ratios are expressed relative to the referent group of tertile 1 (lowest third of antibody distribution). Three synthetic peptides printed in the array were shown for comparison.

Table 3. Odds ratios (OR) and 95% confidence intervals (CI) for the association between anti-EBV antibody level and EBV-positive classic Hodgkin lymphoma (cHL) vs. EBV-negative cHL[a]

EBV Protein and Array sequence	Antibody Type	t-test P	EBV-positive cHL Mean	EBV-negative cHL Mean	EBV-positive cHL Positivity	EBV-negative cHL Positivity	OR tertile 2 (95% CI)	OR tertile 3 (95% CI)	P-trend
BBRF1 (Late lytic) YP_001129476.1-102746-104587	IgG	7.6×10^{-6}	1.34	1.08	68.3%	51.4%	1.54 (0.69, 3.43)	2.94 (1.32, 6.52)	7.2×10^{-3}
BZLF1 (Zta) CAA24861.1-102338-102210	IgG	1.4×10^{-5}	1.79	1.51	97.1%	90.0%	2.53 (1.17, 5.48)	4.94 (2.13, 11.5)	2.1×10^{-4}
BARF1 (Oncogene) YP_001129453.1-166746-167654	IgG	1.9×10^{-5}	1.45	1.20	82.0%	68.6%	1.09 (0.51, 2.35)	3.29 (1.58, 6.83)	1.4×10^{-3}
BdRF1 (VCA_p40) AFY97974.1-136284-137321	IgG	3.5×10^{-5}	1.69	1.40	90.6%	74.3%	1.50 (0.67, 3.39)	4.16 (1.89, 9.14)	2.5×10^{-4}
BVRF2 (viral capsid) YP_001129501.1-136465-138282 redesigned	IgG	1.0×10^{-4}	1.02	0.91	43.2%	24.3%	2.54 (1.16, 5.57)	5.27 (2.15, 12.9)	3.1×10^{-4}
BKRF4 (Late lytic) YP_001129474.1-99676-100329	IgG	1.1×10^{-4}	1.14	0.99	49.6%	25.7%	1.50 (0.68, 3.33)	3.36 (1.58, 7.14)	1.6×10^{-3}
VCA-p18 (synthetic peptide)	IgG	0.0028	2.00	1.83	98.6%	88.6%	1.84 (0.84, 4.03)	2.90 (1.25, 6.68)	0.013
EBNA1 (synthetic peptide)	IgG	0.57	1.62.	1.65	84.2%	85.7%	0.72 (0.35, 1.47)	0.74 (0.33, 1.64)	0.44
EAd (synthetic peptide)	IgG	0.09	0.97	0.91	10.8%	11.4%	2.29 (1.03, 5.08)	1.51 (0.71, 3.22)	0.33

a Table is ordered by t-test p-value (lowest to highest). ORs were adjusted for age group (<30, 30-39, 40-49, 50-59, and 60+ years), sex, and study area (UK, Denmark, and Sweden). The tertiles were calculated using the underlying antibody distribution among controls. All odds ratios are expressed relative to the referent group of tertile 1 (lowest third of antibody distribution). Three synthetic peptides printed in the array were shown for comparison.

Figure legends

Figure 1. Case-control differences in the mean antibody response for EBV-positive cHL cases vs. controls. A) Association between anti-EBV antibodies against proteins from different EBV life cycles and EBV-positive cHL (red, IgA; blue, IgG). The dashed line represents the statistically significant *P* value threshold after Bonferroni correction. *P* values were obtained from SKAT-C tests. **B)** The x-axis displays the fold change (case vs. control ratio of standardized signal intensity) for all antibodies with CV<20%. The y-axis illustrates the *P* value corresponding to the t-test for a difference in SI between cases and controls. Sixteen IgG antibodies and two IgA antibodies were significantly elevated in EBV-positive cHL cases compared to controls at the P≤0.00013 (Bonferroni P ≤0.05) threshold. The four antibodies with the smallest *P* values are highlighted.

Figure 2. Case-control differences in the mean antibody response for EBV-negative cHL cases vs. controls. A) Association between anti-EBV antibodies against proteins from different EBV life cycles and EBV-negative cHL (red, IgA; blue, IgG). The dashed line represents the statistically significant *P* value threshold after Bonferroni correction. *P* values were obtained from SKAT-C tests. **B)** The x-axis displays the fold change (case vs. control ratio of standardized signal intensity) for all antibodies with CV≤20%. The y-axis illustrates the *P* value corresponding to the t-test for a difference in standardized signal intensity between cases and controls. No anti-EBV antibodies were significantly elevated in EBV-negative cHL cases compared to controls

Figure 3. Case-case comparison. A) Case-case differences in standardized signal intensity (SSI) for cases with EBV-positive cHL vs. EBV-negative cHL. The x-axis displays the fold change (ratio of SSI) for all antibodies with CV≤20%. The y-axis illustrates the *P* value corresponding to the t-test for a difference in SI between EBV-positive vs. EBV-negative cHL. A total of six IgG antibodies were significantly elevated in EBV-positive vs. EBV-negative cHL at the *P*≤0.00013 (Bonferroni *P* ≤0.05) threshold. The four antibodies with the smallest *P* values are highlighted. **B)** Receiver operating curve

(ROC) for classifying cHL tumors as either EBV positive or negative using the two selected serological markers BZLF1-IgG and BdRF-IgG. Dotted line: patient demographics (age, sex, and residential area) alone. Solid line: a combination of BZLF1-IgG and BdRF-IgG and patient demographics.

Figure S1. **Average standardized signal intensity for the four anti-EBV IgG antibodies with the lowest *P* values for the comparison between EBV-positive cHL cases and controls. A)** BBRF1; **B)** BcLF1; **C)** Thymidine Kinase; **D)** BBLF1. The dashed line represents the cutoff for positivity. *P* values from the global ANOVA test and each two-way t-test (e.g., EBV-positive cHL vs. EBV-negative cHL) are listed. Boxes show interquartile range (IQR).

Figure S2. Correlation between BBRF-IgG and the A) number of circulating B-cells infected with **EBV and B)** chemokine CCL17 (thymus and activation-regulated chemokine TARC) levels. The x-axis displays the antibody levels of BBRF-IgG. The y-axis is presented in logarithmic scale. The regression line is based on linear regression and the 95% confidence interval (shadows) is illustrated. Correlation (cor) and *P* value are obtained from Spearman correlation.

Figure S3. Average standardized signal intensity for the two anti-EBV IgG antibodies selected to distinguish EBV tumor status. A) BZLF1 and **B)** BdRF1. The dashed line represents the cutoff for positivity. *P* values from the global ANOVA test and each two-way t-test are listed.
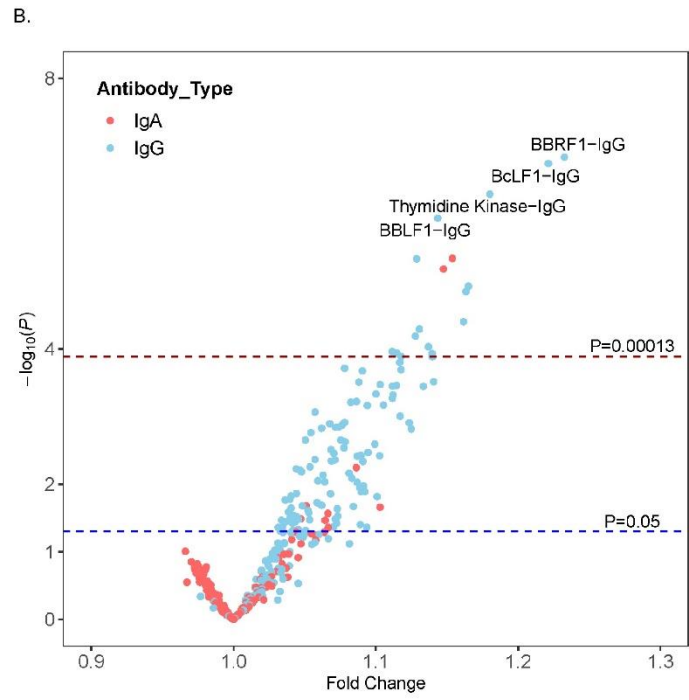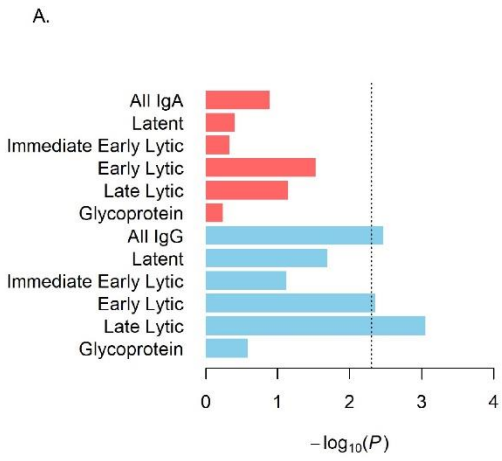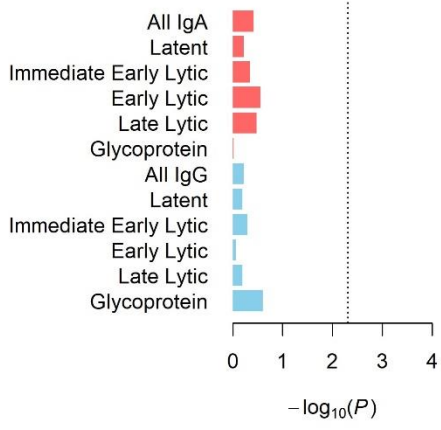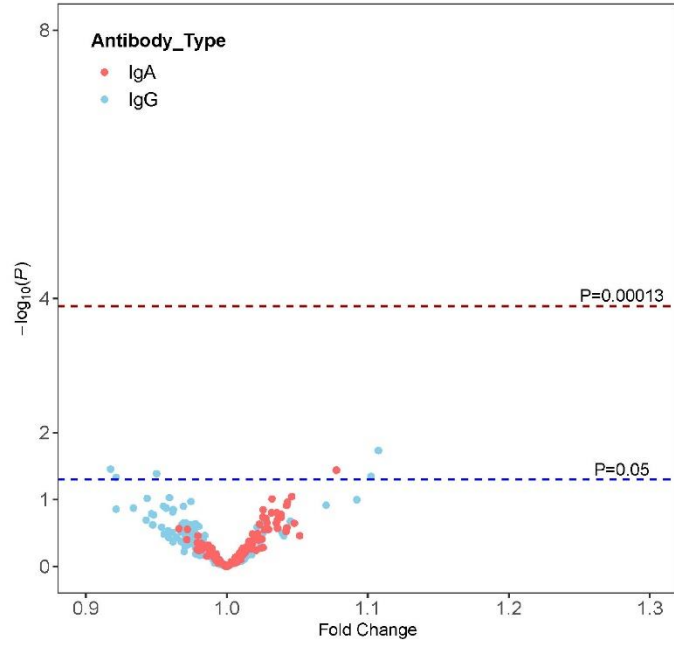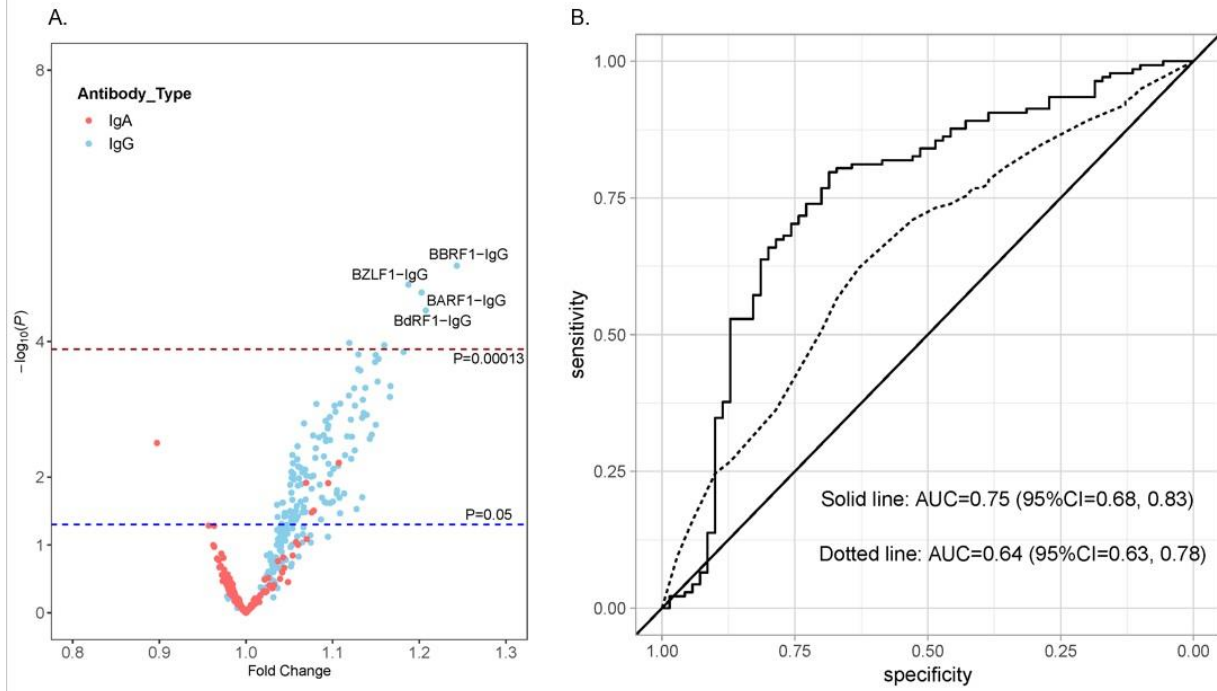
Figure 1

A.

B.

Figure 2

A.

B.

Figure 3

Supplementary Materials

We developed an EBV protein microarray targeting IgG and IgA antibodies against 199 predicted EBV protein sequences representing non-redundant open reading frames (86 EBV proteins) from five prototypical EBV strains (AG876, Akata, B95-8, Mutu, and Raji). Each of the protein sequences was cloned into the pXT7 expression vector, expressed using the *E. coli* cell-free protein system, and printed onto the microarray. Sequences included N-terminal 10x histidine (His) and C-terminal hemagglutinin (HA) tags to confirm expression on the microarray. High coverage was achieved across the five prototypical EBV strains and ten Chinese strains, with $\geq$97% of the predicted sequences from each strain represented on the microarray at $\geq$99% homology. We also included three synthetic EBV peptides that are putative cancer biomarkers (VCAp18, EBNA1, and early antigen [EA] p47), bringing the total number of anti-EBV probes on the array to 202.

After testing, raw signal intensities were corrected for spot-specific background using the Axon GenePix Pro 7 software and were variant log-transformed using variance stabilizing normalization transformation in Gmine (http://cgenome.net/wiki/index.php/Genomics_Data_Miner). In addition to the 202 EBV sequences, we included four "no DNA" (no translated protein) spots to assess person-specific background (e.g., *E.coli* reactivity). We defined a lower limit of detection corresponding to the upper bound of the lowest quartile (Q1) of this "no DNA control" distribution and assigned that level to all spots that fell into Q1. To remove potential differences in this background between cases and controls,, the array output for each participant was divided by the person-specific background (mean +1.5 standard deviations of four "no DNA" spots) prior to analysis, referred to as the standardized signal intensity (SSI). Positivity was defined as a SSI >1.0, and the SSI for each marker was further grouped into three categories, with cutoffs for the categories defined using tertiles of the distribution of a given marker among the 141 controls.

Table S1. Array Description

Spot.ID	Marker	Life_cycle
synthetic antigen	VCA_p18	Late lytic
synthetic antigen	EBNA1	Latent
synthetic antigen	EAD_p47	Early lytic
YP_001129436.1-1026-1196	LMP2A	Latent
YP_001129465.1-86654-87013	EBNA3C	Latent
YP_001129474.1-99676-100329	BKRF4	Late lytic
YP_001129477.1-104490-105326	BBRF2	Late lytic
YP_001129485.1-117754-118890	BGRF1/BDRF1	Late lytic
YP_001129470.1-94844-96457	BRRF2	Late lytic
YP_001129467.1-90855-90724	BZLF1 (ZEBRA)	Immediate early lytic
YP_001129464.1-83074-83430	EBNA3B	Latent
YP_001129507.1-157772-154725-2	BALF5 (EARLY GENE; DNA POL)	Early lytic
YP_001129459.1-76320-75484	DUTPASE	Early lytic
YP_001129451.1-63084-64178	CAPSID	Late lytic
YP_001129444.1-46353-44776	BFLF1	Late lytic
YP_001129494.1-126004-128256	HYPOTHETICAL	Early lytic
YP_001129463.1-80447-82888	EBNA3A	Latent
YP_001129440.1-20824-20955	EBNA-LP	Latent
YP_001129436.1-167587-167942	LMP2A	Latent
YP_001129464.1-83509-86532-2	EBNA3B	Latent
YP_001129500.1-136454-135636	BVLF1	Late lytic
YP_001129455.1-68964-70037	BMRF2	Glycoprotein
YP_001129438.1-1736-5692-2	FGAM	Other/Unknown
YP_001129496.1-131574-129454	BXLF2 (GP85/GH)	Glycoprotein
YP_001129467.1-91045-90941	BZLF1 (ZEBRA)	Immediate early lytic
YP_001129463.1-80026-80361	EBNA3A	Latent
YP_001129478.1-105928-105323	BBLF2/3	Early lytic
YP_001129473.1-98895-99662	GLYCOSYLASE	Other/Unknown
YP_001129446.1-46719-47729	BFRF1 (UL34 homolog)	Late lytic
YP_001129481.1-110883-109471	BGLF5 (EARLY GENE; ALK EXO)	Early lytic
YP_001129509.1-162392-160335	BALF3	Late lytic
YP_001129507.1-157772-154725-1	BALF5 (EARLY GENE; DNA POL)	Early lytic
YP_001129440.1-35558-35662	EBNA-LP	Latent
YP_001129460.1-76393-76701	BLRF1	Late lytic
YP_001129442.1-42204-42779	BHRF1 (BCL2 ANALOGUE)	Early lytic
YP_001129515.1-169948-169188	LMP1	Latent
YP_001129486.1-115415-114405	BGLF2	Early lytic
YP_001129505.1-153178-151769	LF1	Other/Unknown
YP_001129493.1-126005-121860-1	BCLF1 (MAJOR CAPSID PROTEIN)	Late lytic
YP_001129436.1-1574-1680	LMP2A	Latent

YP_001129515.1-170457-170190	LMP1	Latent
YP_001129448.1-49335-49865	BFRF3	Late lytic
YP_001129503.1-139063-138317	BILF2 (GP55-78)	Glycoprotein
YP_001129484.1-113481-112483	BGLF3	Late lytic
YP_001129456.1-71967-70589	BSLF2/BMLF1	Immediate early lytic
YP_001129476.1-102746-104587	BBRF1	Late lytic
YP_001129436.1-360-458	LMP2A	Latent
YP_001129436.1-540-788	LMP2A	Latent
YP_001129439.1-9659-10171	BCRF1	Late lytic
YP_001129498.1-133398-134144	BXRF1	Late lytic
YP_001129510.1-165796-162410-2	DNA BINDING	Early lytic
YP_001129497.1-133399-131576	THY.KINASE	Early lytic
YP_001129449.1-59370-49906-3	BPFL1	Late lytic
YP_001129515.1-170111-170025	LMP1	Latent
YP_001129489.1-117772-117539	BDLF3 (GP100-150)	Glycoprotein
YP_001129467.1-91697-91197	BZLF1 (ZEBRA)	Immediate early lytic
YP_001129490.1-119605-118901	BDLF3 (GP100-150)	Glycoprotein
YP_001129443.1-44763-43807	BFLF2	Late lytic
YP_001129468.1-93725-91908	BRLF1 (IMMEDIATE EARLY)	Immediate early lytic
YP_001129449.1-59370-49906-2	BPFL1	Late lytic
YP_001129436.1-871-951	LMP2A	Latent
YP_001129480.1-109516-109289	BBLF1	Early lytic
YP_001129461.1-76771-77259	BLRF2	Late lytic
YP_001129488.1-117560-116883	BDLF4	Early lytic
YP_001129506.1-154125-153187	BILF1	Glycoprotein
YP_001129504.1-151808-150519	HYPOTHETICAL	Early lytic
YP_001129501.1-136465-138282	BVRF2 (VCAP40)	Late lytic
YP_001129440.1-29887-29952	EBNA-LP	Latent
YP_001129436.1-1280-1495	LMP2A	Latent
YP_001129483.1-112496-112035	BGLF3	Late lytic
YP_001129466.1-90630-89959	BZLF2 (GP42)	Glycoprotein
YP_001129469.1-93724-94656	BRRF1	Immediate early lytic
YP_001129491.1-120928-119666	BDLF2	Glycoprotein
YP_001129464.1-83509-86532-1	EBNA3B	Latent
YP_001129440.1-35441-35473	EBNA-LP	Latent
YP_001129436.1-58-272	LMP2A	Latent
YP_001129472.1-98500-98913	BKRF2 (GP25/GL)	Glycoprotein
YP_001129512.1-166530-167195	BARF1	Early lytic
YP_001129453.1-166746-167654	BARF1	Early lytic
YP_001129479.1-107679-108896	BBRF3	Glycoprotein
YP_001129493.1-126005-121860-2	BCLF1 (MAJOR CAPSID PROTEIN)	Late lytic

YP_001129438.1-1736-5692-1	FGAM	Other/Unknown
YP_001129445.1-46352-46759	BFRF1A	Other/Unknown
YP_001129458.1-74770-75426	BSRF1	Late lytic
YP_001129492.1-121844-120939	CAPSID	Late lytic
YP_001129454.1-67745-68959	BMRF1 (EAD)	Early lytic
YP_001129499.1-133954-135666	BVRF1	Late lytic
YP_001129510.1-165796-162410-1	DNA BINDING	Early lytic
AFY97829.1-82733-83089	EBNA3B	Latent
AFY97840.1-62772-59044-2	BOLF1	Late lytic
CAA24859.1-98371-98730	EBNA3C	Latent
CAA24827.1-122341-120929	BGLF5 (EARLY GENE; ALK EXO)	Early lytic
AFY97909.1-540-788	LMP2A	Latent
AFY97894.1-154809-155094	RPMS1	Latent
CAA24811.1-167303-166998	BNLF2B	Early lytic
CAA24828.1-123692-122328	BGLF4	Early lytic
CAA24858.1-95788-98247	EBNA3B	Latent
AFY97958.1-116811-116578	BDLF3 (GP100-150)	Glycoprotein
AFY97916.1-36198-37658	EBNA2	Latent
AFY97906.1-168513-168246	LMP1	Latent
AFY97882.1-118329-117625	BDLF3 (GP100-150)	Glycoprotein
AFY97841.1-59051-49692-1	BPFL1	Late lytic
CAB56339.1-48385-48552	UNCHARACTERIZED	Other/Unknown
CAA24796.1-139642-140916	BTRF1	Late lytic
CAA24805.1-156746-153699-1	BALF5 (EARLY GENE; DNA POL)	Early lytic
AFY97956.1-108555-108328	BBLF1	Early lytic
AFY97955.1-109922-108510	BGLF5 (EARLY GENE; ALK EXO)	Early lytic
AFY97832.1-35494-35598	EBNA-LP	Latent
AFY97877.1-116284-115607	BDLF4	Early lytic
CAA24861.1-102338-102210	BZLF1 (ZEBRA)	Immediate early lytic
CAA24873.1-40189-41340	BWRF1	Other/Unknown
CAA24794.1-137466-133321-1	BCLF1 (MAJOR CAPSID PROTEIN)	Late lytic
AFY97910.1-59-272	LMP2B	Latent
AFY97929.1-67486-68700	BMRF1 (EAD)	Early lytic
AFY97906.1-168167-168081	LMP1	Latent
AFY97901.1-160450-158393	BALF3	Late lytic
CAB56340.1-84288-84169	BSLF2/BMLF1	Immediate early lytic
CAA24829.1-124938-125915	BGRF1/BDRF1	Late lytic
CAA24839.1-71527-62078-3	BPFL1	Late lytic
AFY97988.1-166888-166706	BNLF2A	Late lytic
AFY97987.1-168367-167562	LMP1	Latent
AFY97832.1-35377-35409	EBNA-LP	Latent
AFY97842.1-95349-97142	EBNA1	Latent

CAA24861.1-102530-102423	BZIP	Immediate early lytic
CAA24798.1-144860-145606	BXRF1	Late lytic
CAA24839.1-71527-62078-2	BPFL1	Late lytic
AFY97910.1-1026-1196	LMP2B	Latent
AFY97946.1-98716-99369	BKRF4	Late lytic
AFY97868.1-104653-104048	BBLF2/3	Early lytic
AFY97838.1-47422-49197	BFRF2	Late lytic
CAB56341.1-92663-92767	BLRF3	Late lytic
CAA24860.1-102116-101445	BZLF2 (GP42)	Glycoprotein
AFY97941.1-90112-90008	BZLF1 (ZEBRA)	Immediate early lytic
AFY97980.1-156149-153102-2	BALF5 (EARLY GENE; DNA POL)	Early lytic
AFY97883.1-124729-120584-2	BCLF1 (MAJOR CAPSID PROTEIN)	Late lytic
AFY97856.1-86125-88794	EBNA3C	Latent
CAA24810.1-165517-164855	BALF1	Early lytic
CAA24807.1-161678-159312	BALF3	Late lytic
AFY97917.1-35572-35676	EBNA-LP	Latent
AFY97950.1-104968-104363	BBLF2/3	Early lytic
YP_001129457.1-74727-72103	BSLF1 (EARLY GENE)	Early lytic
AFY97861.1-93701-95314	BBRF2	Late lytic
CAA24838.1-61507-62037	BFRF3	Late lytic
AFY97981.1-159642-159726	A73	Other/Unknown
AFY97924.1-49199-49729	BFRF3	Late lytic
AFY97830.1-80050-82545	EBNA3A	Latent
CAA24861.1-103155-102655	BZLF1 (ZEBRA)	Immediate early lytic
CAA24824.1-119080-117515	BBLF2	Early lytic
AFY97981.1-156513-156598	A73	Other/Unknown
AFY97909.1-165963-166318	LMP2A	Latent
YP_001129452.1-64253-66733	BORF2 (EARLY GENE; RIBO UNIT)	Early lytic
AFY97836.1-46138-46545	BRFR1A	Early lytic
CAA24817.1-109958-110371	BKRF2 (GP25/GL)	Glycoprotein
CAA24832.1-128374-126851	BGLF1	Late lytic
AFY97917.1-35455-35487	EBNA-LP	Latent
AFY97989.1-166696-166400	BNLF2B	Early lytic
AFY97856.1-85691-86050	EBNA3C	Latent
AFY97897.1-151239-149830	LF1	Other/Unknown
AFY97883.1-124729-120584-1	BCLF1 (MAJOR CAPSID PROTEIN)	Late lytic
CAA24856.1-92243-92602	EBNA3A	Latent
AFY97976.1-155128-155413	RPMS1	Latent
YP_401636.1-20698-20763	EBNA-LP	Latent
YP_401707.1-143344-140570	LF3	Other/Unknown

AFY97915.1-80252-82747	EBNA3A	Latent
YP_001129475.1-102801-100372 redesigned	BBLF4 (EARLY GENE)	Early lytic
YP_001129441.1-36201-37565 redesigned	EBNA2	Latent
CAA24816.1-107950-109875 redesigned	EBNA1	Latent
AFY97965.1-125044-120899-1	BCLF1 (MAJOR CAPSID PROTEIN)	Late lytic
YP_401715.1-160908-158851	BALF3	Late lytic
CAA24858.1-95353-95709	EBNA3B	Latent
CAA24839.1-71527-62078-4	BPFL1	Late lytic
AFY97987.1-168876-168609	LMP1	Latent
CAA24877.1-48504-49967 redesigned	EBNA2	Latent
AFY97966.1-125043-127295	HYPOTHETICAL	Early lytic
YP_401645.1-40269-38287	BHLF1	Early lytic
AFY97990.1-85953-86312	EBNA3C	Latent
YP_001129449.1-59370-49906-1 redesigned	BPFL1	Late lytic
AFY97974.1-136284-137321	BDRF1	Late lytic
AFY97964.1-118644-117940	BDLF3 (GP100-150)	Glycoprotein
CAA24806.1-159322-156749	BALF4 (GP110/GB)	Glycoprotein
AFY97913.1-95532-97457	EBNA1	Latent
YP_401722.1-168507-167702	LMP1	Latent
AFY97978.1-151556-150147	LF1	Other/Unknown
YP_001129449.1-59370-49906-4 redesigned	BPFL1	Late lytic
CAA24854.1-92153-89430 redesigned	BLLF1 (GP350)	Glycoprotein
AFY97959.1-116599-115922	BDLF4	Early lytic
YP_001129465.1-87088-89937 redesigned	EBNA3C	Latent
YP_401669.1-80382-82877	EBNA3A	Latent
AFY97943.1-93884-95497	BRRF2	Late lytic
YP_401636.1-35590-35694	EBNA-LP	Latent
YP_401722.1-168670-168584	LMP1	Latent
YP_001129501.1-136465-138282 redesigned	BVRF2 (VCAP40)	Late lytic
CAA24859.1-98805-101423 redesigned	EBNA3C	Latent
AFY97921.1-46216-46623	BFRF1A	Other/Unknown
YP_001129462.1-79936-77276 redesigned	BLLF1 (GP350)	Glycoprotein
CAA24821.1-114259-111830	BBLF4 (EARLY GENE)	Early lytic
YP_401722.1-168670-168584	LMP1	Latent
YP_001129447.1-47636-49411 redesigned	BFRF2	Late lytic
CAA24839.1-71527-62078-1 redesigned	BPFL1	Late lytic
CAA24880.1-59808-61583	NA	Immediate early lytic
YP_001129508.1-160348-157775 redesigned	BALF4 (GP110/GB)	Glycoprotein
CAA24841.1-75239-71520-1	BOLF1	Late lytic

Table S2. Odds ratio and 95% confidence interval for the association between EBNA-1:EBNA-2 ratio ≤1 vs. >1 and classic Hodgkin lymphoma (cHL) *

EBNA1	EBNA2	EBV-negative cases vs. controls		EBV-positive cases vs. controls	
		OR (95% CI)	*P* value	OR (95% CI)	*P* value
Synthetic peptide	AFY97916.1-36198-37658	1.45 (0.60, 3.51)	0.409	1.63 (0.81, 3.29)	0.169
Synthetic peptide	YP_001129441.1-36201-37565	1.03 (0.55, 1.94)	0.927	1.22 (0.74, 2.01)	0.433
Synthetic peptide	CAA24877.1-48504-49967	1.67 (0.67, 4.11)	0.269	1.81 (0.87, 3.77)	0.110
CAA24816.1-107950-109875	AFY97916.1-36198-37658	1.18 (0.62, 2.24)	0.611	1.09 (0.65, 1.83)	0.738
CAA24816.1-107950-109875	YP_001129441.1-36201-37565	0.81 (0.45, 1.47)	0.489	1.59 (0.97, 2.62)	0.068
CAA24816.1-107950-109875	CAA24877.1-48504-49967	1.63 (0.71, 3.70)	0.247	1.81 (0.94, 3.49)	0.077
AFY97913.1-95532-97457	AFY97916.1-36198-37658	0.98 (0.31, 3.06)	0.973	2.30 (0.77, 6.87)	0.136
AFY97842.1-95349-97142	YP_001129441.1-36201-37565	0.95 (0.30, 3.03)	0.929	3.42 (0.92, 12.7)	0.067
AFY97842.1-95349-97142	CAA24877.1-48504-49967	1.03 (0.48, 2.21)	0.933	1.95 (0.98, 3.88)	0.056
AFY97913.1-95532-97457	AFY97916.1-36198-37658	0.46 (0.17, 1.22)	0.119	0.34 (0.15, 0.77)	0.010
AFY97913.1-95532-97457	YP_001129441.1-36201-37565	0.65 (0.14, 3.10)	0.588	0.37 (0.11, 1.22)	0.103
AFY97913.1-95532-97457	CAA24877.1-48504-49967	0.40 (0.20, 0.82)	0.012	0.42 (0.23, 0.77)	0.005

* ORs were adjusted for age group (<30, 30-39, 40-49, 50-59, and 60+ years), sex, and study area (UK, Denmark, and Sweden)

Table S3. Odds ratios (OR) and 95% confidence intervals (CI) for the association between anti-EBV antibody level and EBV-positive classic Hodgkin lymphoma (cHL) vs. controls among participants from the Scandinavian Lymphoma Etiology study, additionally adjusted for a history of infectious mononucleosis. *

EBV Protein and Array sequence	Antibody Type	OR tertile 2 (95% CI)	OR tertile 3 (95% CI)	*P*-trend
BBRF1 (Late lytic) YP_001129476.1-102746-104587	IgG	1.58 (0.44, 5.63)	2.01 (0.55, 7.32)	0.292
BcLF1 (VCA_p160) AFY97965.1-125044-120899-1	IgG	1.17 (0.27, 5.03)	4.50 (1.18, 17.2)	0.016
Thymidine kinase (Early lytic) YP_001129497.1-133399-131576	IgG	0.64 (0.15, 2.70)	2.96 (0.85, 10.4)	0.044
BBLF1 (Tegument protein) AFY97956.1-108555-108328	IgG	1.61 (0.37, 6.94)	4.87 (1.17, 20.2)	0.014
BBLF1 (Tegument protein) YP_001129480.1-109516-109289	IgA	0.98 (0.28, 3.52)	2.51 (0.77, 8.18)	0.114
BBLF1 (Tegument protein) YP_001129480.1-109516-109289	IgG	2.10 (0.50, 8.88)	5.05 (1.24, 20.5)	0.016
BBLF1 (Tegument protein) AFY97956.1-108555-108328	IgA	1.37 (0.36, 5.21)	3.30 (0.99, 11.0)	0.042
BcLF1 (VCA_p160) YP_001129493.1-126005-121860-1	IgG	1.87 (0.49, 7.13)	2.44 (0.64, 9.23)	0.201
BcLF1 (VCA_p160) CAA24794.1-137466-133321-1	IgG	3.09 (0.73, 13.1)	4.48 (1.03, 19.4)	0.052
BdRF1 (VCA_p40) AFY97974.1-136284-137321	IgG	0.76 (0.19, 3.03)	4.41 (1.20, 16.2)	0.008
BDLF3 (glycoprotein 150) AFY97964.1-118644-117940	IgG	0.74 (0.19, 2.82)	3.42 (0.94, 12.4)	0.038
BBRF3 (glycoprotein M) YP_001129479.1-107679-108896	IgG	1.05 (0.29, 3.84)	1.51 (0.42, 5.39)	0.480
BDLF3 (glycoprotein 150) YP_001129490.1-119605-118901	IgG	1.82 (0.42, 7.95)	8.91 (2.04, 39.0)	0.001
BHRF1 (Bcl-2 homolog) YP_001129442.1-42204-42779	IgG	3.03 (0.73, 12.5)	2.87 (0.70, 11.7)	0.207
BFLF2 (Late lytic) YP_001129443.1-44763-43807	IgG	1.27 (0.33, 4.88)	8.52 (2.12, 34.2)	0.001
LMP-1 (Oncogene) YP_401722.1-168507-167702	IgG	2.84 (0.68, 11.9)	4.74 (1.13, 19.9)	0.035
BALF2 (EA(D)_p138) single-stranded DNA binding protein YP_001129510.1-165796-162410-1	IgG	1.55 (0.39, 6.24)	2.75 (0.77, 9.81)	0.111

BARF1 (Oncogene) YP_001129453.1-166746-167654	IgG	0.62 (0.17, 2.27)	2.72 (0.81, 9.06)	0.080

* Table is ordered by t-test p-value (lowest to highest) from the **Table 2**. ORs were adjusted for age group (<30, 30-39, 40-49, 50-59, and 60+ years), sex, study area (Denmark and Sweden), and a history of infectious mononucleosis. The tertiles were calculated using the underlying antibody distribution among controls. All odds ratios are expressed relative to the referent group of tertile 1 (lowest third of antibody distribution).
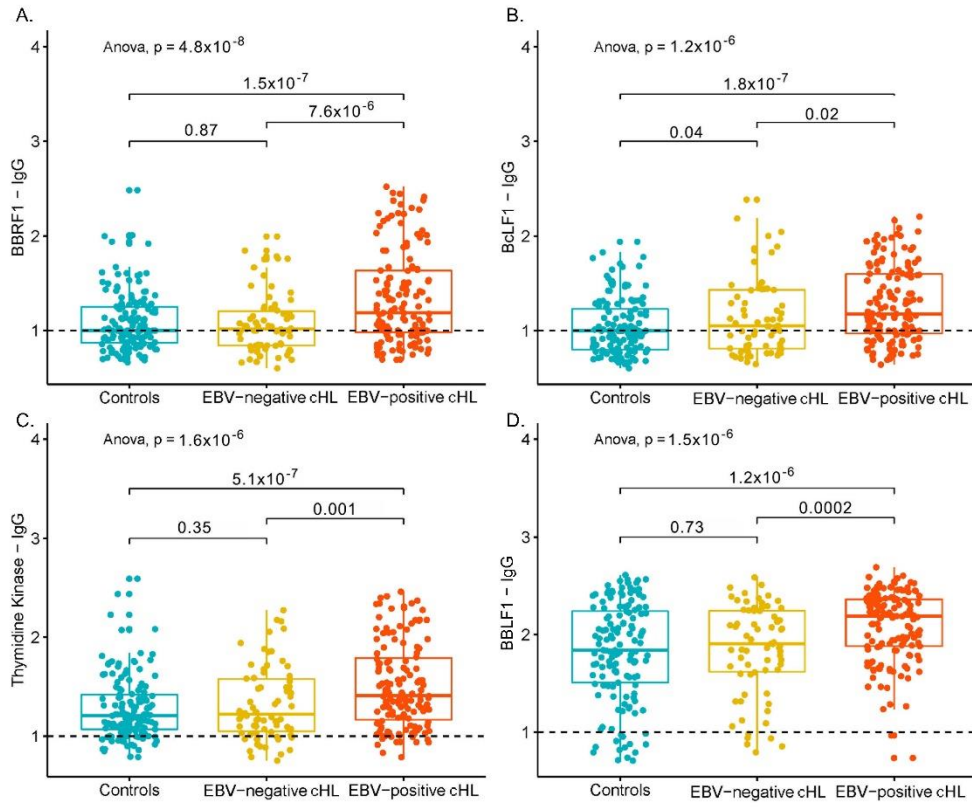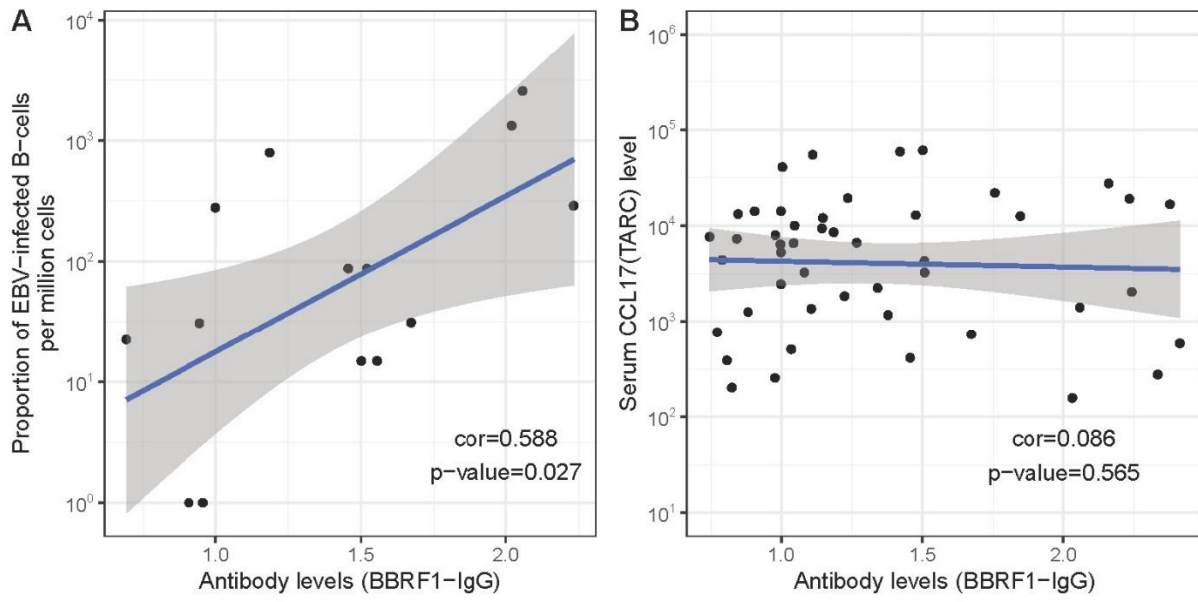
Figure S1

A.

Anova, p = 3.9×10^{-5}

0.003

1.4×10^{-5}

0.046

BZLF1 – IgG

Controls   EBV−negative cHL   EBV−positive cHL

B.

Anova, p = 6.4×10^{-6}

4.0×10^{-5}

3.5×10^{-5}

0.42

BdRF1 – IgG

Controls   EBV−negative cHL   EBV−positive cHL