



American Society of Hematology
2021 L Street NW, Suite 900,
Washington, DC 20036
Phone: 202-776-0544 | Fax 202-776-0545
editorial@hematology.org

Insertion of atypical glycans into the tumor antigen-binding site identifies DLBCLs with distinct origin and behavior

Tracking no: BLD-2021-012052R2

Georgia Chiodin (University of Southampton, United Kingdom) Joel Allen (University of Southampton, United Kingdom) Dean Bryant (University of Southampton, United Kingdom) Philip Rock (University of Rochester Medical Center, United States) Enrica Martino (UOC Ematologia Cosenza, Italy) Beatriz Valle-Argos (University of Southampton, United Kingdom) Patrick Duriez (University of Southampton, United Kingdom) Yasunori Watanabe (University of Oxford, United Kingdom) Isla Henderson (University of Southampton, United Kingdom) James Blachly (The Ohio State University, United States) Katy McCann (University of Southampton & SUHT, United Kingdom) Jonathan Strefford (University of Southampton, United Kingdom) Graham Packham (University of Southampton, United Kingdom) Teunis Geijtenbeek (Amsterdam UMC, Netherlands) Carl Figdor (Radboud Institute for Molecular Life Sciences, Radboud university medical center,) George Wright (NIH/NCI, United States) Louis Staudt (National Cancer Institute, United States) Richard Burack (University of Rochester, United States) Thomas Bowden (University of Oxford, United Kingdom) Max Crispin (University of Southampton, United Kingdom) Freda Stevenson (University of Southampton, United Kingdom) Francesco Forconi (University of Southampton, United Kingdom)

Abstract:

Glycosylation of the surface immunoglobulin variable region is a remarkable follicular lymphoma-associated feature rarely seen in normal B cells. Here, we define a subset of diffuse large B-cell lymphomas (DLBCL) which acquire N-glycosylation sites selectively in the immunoglobulin (Ig) complementary-determining-regions (CDR) of the antigen-binding sites. Mass-spectrometry and X-ray crystallography demonstrate how the inserted glycans are stalled at oligomannose-type structures due to burial in the CDR loops. Acquisition of sites occurs in ~50% of germinal center B-cell-like DLBCL, mainly of the genetic EZB subtype, irrespective of *IGHV-D-J* use. This markedly contrasts with the activated B-cell-like DLBCL Ig, which rarely has sites in the CDR, and appears not to acquire oligomannose-type structures. Acquisition of CDR-located acceptor sites associates with mutations of epigenetic regulators and *BCL2* translocations, indicating an origin shared with follicular lymphoma. Within the EZB subtype, these sites associate with more rapid disease progression and with significant gene-set enrichment of the B-cell receptor, PI3K/AKT/MTORC1, glucose metabolism, and MYC signaling pathways, particularly in the fraction devoid of MYC translocations. The oligomannose-type glycans on the lymphoma cells interact with the candidate lectin DC-SIGN, mediating low-level signals, and lectin-expressing cells form clusters with lymphoma cells. Both clustering and signaling are inhibited by antibodies specifically targeting the DC-SIGN carbohydrate-recognition-domain. Oligomannosylation of the tumor immunoglobulin is a post-translational modification that readily identifies a distinct GCB-DLBCL category with more aggressive clinical behavior, and could be a potential precise therapeutic target via antibody-mediated inhibition of the tumor Ig interaction with DC-SIGN-expressing M2-polarized macrophages.

Conflict of interest: No COI declared

COI notes:

Preprint server: No;

Author contributions and disclosures: G.C. performed research, analyzed and interpreted data, and wrote the manuscript. J.D.A. performed and interpreted glycan analysis. D.J.B. designed the bioinformatic pipeline for the immunoglobulin gene analysis and analyzed genetic variant data. P.R., E.A.M., and B.V.A. contributed to the characterization of primary samples. P.J.D. generated the lymphoma-derived Fab. I.H. and K.M. performed the sequencing and characterization of primary samples. Y.W. and T.A.B. performed and interpreted X-ray crystallography analyses. J.S.B. contributed to the immunoglobulin gene analysis pipeline. C.F. contributed anti-DC-SIGN antibodies. T.G. contributed DC-SIGN expressing and parental cell lines. D.J.B. and G.W. performed gene expression data analysis. J.C.S., G.P., T.G., C.F., L.M.S., and R.B. contributed to the interpretation and discussion of the data. M.C. and T.B. led the structural biology research and wrote the manuscript. F.K.S. and F.F. designed the study, supervised research, interpreted data and wrote the manuscript.

Non-author contributions and disclosures: No;

Agreement to Share Publication-Related Data and Data Sharing Statement: For original data, please contact f.forconi@soton.ac.uk. The tumor IG rearranged transcript sequences, genetic variants and RNAseq expression were identified from the National Cancer Institute Genomic Data Commons for Genotypes and Phenotypes, available under accession phs001444.v1.p1.

Clinical trial registration information (if any):

Insertion of atypical glycans into the tumor antigen-binding site identifies DLBCLs with distinct origin and behavior

Giorgia Chiodin,¹ Joel D. Allen,² Dean J. Bryant,¹ Philip Rock,³ Enrica Martino,^{1,4} Beatriz Valle-Argos,¹ Patrick J. Duriez,¹ Yasunori Watanabe,⁵ Isla Henderson,¹ James S. Blachly,⁶ Katy McCann,¹ Jonathan C. Strefford,¹ Graham Packham,¹ Teunis B.H. Geijtenbeek,⁷ Carl Figdor,⁸ George W. Wright,⁹ Louis M. Staudt,¹⁰ Richard Burack,³ Thomas A. Bowden,⁵ Max Crispin,² Freda K. Stevenson,¹ Francesco Forconi^{1,11}

¹School of Cancer Sciences, Cancer Research UK Southampton Centre, Faculty of Medicine, University of Southampton, Southampton, UK. ²School of Biological Sciences, University of Southampton, Southampton, UK. ³Department of Pathology & Laboratory Medicine/Hematopathology, University of Rochester Medical Center, Rochester, NY, USA. ⁴Division of Hematology, Azienda Policlinico-OVE, University of Catania, Italy. ⁵Oxford Glycobiology Institute, Department of Biochemistry, University of Oxford, South Parks Road, Oxford, OX1 3QU, UK. ⁶The Ohio State University, Columbus, Ohio, USA. ⁷Department of Experimental Immunology, Amsterdam Infection and Immunity Institute, Amsterdam University Medical Centers, University of Amsterdam, Amsterdam, The Netherlands. ⁸Department of Tumor Immunology, Radboud Institute for Molecular Life Sciences, Radboud University Medical Center, Nijmegen, The Netherlands. ⁹Biometric Research Branch, Division of Cancer Diagnosis and Treatment, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA. ¹⁰Lymphoid Malignancies Branch, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA. ¹¹Haematology Department, Cancer Care Directorate, University Hospital Southampton NHS Trust, Southampton, UK.

Running title: Antigen receptor mannosylation in DLBCL

Keywords: DLBCL, FL, B-cell receptor, glycosylation, DC-SIGN.

Correspondence: Francesco Forconi, School of Cancer Sciences, Cancer Research UK Centre, Somers Building, MP824, Tremona Road, Southampton, SO16 6YD, UK.
Email: f.forconi@soton.ac.uk. **Tel:** +44 (0)23 81205780

Word Counts: Text: 4048; Abstract: 246

Figures: 7

Tables: 10 supplemental tables

References: 48

Scientific category: lymphoid neoplasia

Key points:

- Oligomannose-type glycans are tumor-specific insertions into the antigen-binding sites of aggressive GCB-DLBCL.
- These glycans empower cells to interact with microenvironmental lectins that can be precisely targeted by novel therapeutic antibodies.

ABSTRACT

Glycosylation of the surface immunoglobulin variable region is a remarkable follicular lymphoma-associated feature rarely seen in normal B cells. Here, we define a subset of diffuse large B-cell lymphomas (DLBCL) which acquire N-glycosylation sites selectively in the immunoglobulin (Ig) complementary-determining-regions (CDR) of the antigen-binding sites. Mass-spectrometry and X-ray crystallography demonstrate how the inserted glycans are stalled at oligomannose-type structures due to burial in the CDR loops. Acquisition of sites occurs in ~50% of germinal center B-cell-like DLBCL, mainly of the genetic EZB subtype, irrespective of *IGHV-D-J* use. This markedly contrasts with the activated B-cell-like DLBCL Ig, which rarely has sites in the CDR, and appears not to acquire oligomannose-type structures. Acquisition of CDR-located acceptor sites associates with mutations of epigenetic regulators and *BCL2* translocations, indicating an origin shared with follicular lymphoma. Within the EZB subtype, these sites associate with more rapid disease progression and with significant gene-set enrichment of the B-cell receptor, PI3K/AKT/MTORC1, glucose metabolism, and MYC signaling pathways, particularly in the fraction devoid of *MYC* translocations. The oligomannose-type glycans on the lymphoma cells interact with the candidate lectin DC-SIGN, mediating low-level signals, and lectin-expressing cells form clusters with lymphoma cells. Both clustering and signaling are inhibited by antibodies specifically targeting the DC-SIGN carbohydrate-recognition-domain. Oligomannosylation of the tumor immunoglobulin is a post-translational modification that readily identifies a distinct GCB-DLBCL category with more aggressive clinical behavior, and could be a potential precise therapeutic target via antibody-mediated inhibition of the tumor Ig interaction with DC-SIGN-expressing M2-polarized macrophages.

INTRODUCTION

Diffuse large B-cell lymphomas (DLBCL) are clinically variable aggressive tumors with heterogeneous phenotypic and genetic profiles and differential responses to conventional therapies. They are commonly classified into two broad subsets and an intermediate “unclassifiable” subset based on gene expression profiling.^{1,2} The germinal center (GC) B-cell-like DLBCL (GCB-DLBCL) subset has a profile similar to normal GC B cells, while the activated B-cell-like DLBCL (ABC-DLBCL) has a profile similar to B cells activated by anti-IgM *in vitro*. Recent genomic studies have further subclassified DLBCL into 7 subtypes, with GCB-DLBCL including an EZB subtype defined by the acquisition of a *BCL2* translocation and mutations in epigenetic regulator genes, including *EZH2*, *KMT2D*, *CREBBP*, *EP300*, *ARID1A*, *IRF8*, *MEF2B*, *EBF1*.^{3,4} These lesions are early tumor events, which appear to associate with worse clinical behavior, particularly when *MYC* translocations occur.^{3,4} The *BCL2* and epigenetic regulator lesions are also early hallmarks of follicular lymphomas (FL),⁵⁻⁷ suggesting a relationship between FL and EZB subtype.³

In vivo ablation of surface immunoglobulin (slg) on mature B cells results in rapid cell death,⁸ and deletion of the tumor slg impairs the growth of lymphomas in mouse models.^{9,10} The slg is typically expressed at high levels on DLBCL cells, but the influences of B-cell receptor (BCR) signaling in ABC-DLBCL and GCB-DLBCL appear distinct. In ABC-DLBCL there is constitutive signaling, via either mutations of genes in the BCR-associated pathway or a range of interactions with putative autoantigens,^{11,12} and therapeutic inhibitors of this pathway can be effective in these patients.¹³ In contrast, there is less evidence for constitutive signaling in GCB-DLBCL and BCR-inhibitors are not as effective, raising the question of the role of a BCR-mediated drive in this subset.

One of the key features of FL is the almost universal acquisition of N-linked glycosylation sites in the Ig variable region.¹⁴ The acquired N-glycosylation sites (AGS) are tumor-specific and are introduced during somatic hypermutation.¹⁵ They persist through the entire disease history, despite ongoing somatic hypermutation, and those lymphoma cells that lose the sites (which is a rare event) reacquire a different site or are lost from subsequent samples, indicating that they are not selected for expansion or long-term survival.¹⁶

In FL, the glycans acquired on the slg variable region are unusual in that their biosynthetic processing is stalled at the initial oligomannose-type state.¹⁷ Termination of any cell surface glycoprotein at oligomannose is rare in normal human cells, but a common feature of infective agents, where it facilitates diverse roles, including immune evasion by glycan shielding and enhancement of immune cell colonization.¹⁸⁻²⁰ Mannosylated slg (slg-Mann) interacts with the C-type lectin dendritic cell-specific intercellular adhesion molecule 3-grabbing non-integrin (DC-SIGN),^{21,22} which is naturally expressed by dendritic cells and macrophages with M2 polarization.^{23,24} This interaction induces low-level intracellular signaling and Myc activation in FL.^{21,22} This antigen-independent signal is relatively weak as compared to anti-Ig, but is sustained,²² and appears to protect slg-Mann+ve FL cells against apoptosis through direct interaction with DC-SIGN-expressing macrophages.^{25,26}

In an early study, we analyzed the tumor *Ig-heavy-chain-variable* region (*IGHV*) sequences from a small group of DLBCL and found that 13/32 (41%) acquired N-glycosylation sites.^{14,27} However, the characteristics of the subsets of DLBCL have not been defined.

Here we have investigated the nature and three-dimensional structure of the glycans which appear affected by location in the variable region. We have mapped incidence, distribution, and functional consequences of the acquired sites in DLBCL, stratifying the GCB category into two distinct subsets and pointing to a potential opportunity to target this feature by blocking antibodies when sites are located in the antigen-binding region.

MATERIALS AND METHODS

Primary samples, cell lines, phenotypic and intracellular signaling analyses

Details about primary samples, cell lines, phenotypic and intracellular signaling analyses are outlined in the Supplemental Methods. The study was approved by the Universities of Southampton (H228/02/t) and Rochester (ULAB03012) Institutional Review Boards. All patients and healthy individuals provided informed consent.

DC-SIGN binding assay

Cells were incubated with 20 μ g/ml DC-SIGN-Fc (R&D_Systems) or left untreated for 30 minutes on ice in RPMI-1640 with 10% FBS and 1mM calcium, washed, and stained with FITC-conjugated or APC-Cy7-conjugated anti-human-Fc (Biolegend) for 30 minutes on ice. Binding was analyzed by flow cytometry. In primary samples, DC-SIGN binding was measured in the CD20+ve or CD20+ve/CD10+ve cells, depending on tumor phenotype.

Immunoglobulin gene analysis

The tumor *IG* transcript sequences were identified from public RNAseq data,^{3,28} or viable DLBCL primary samples and cell lines, as described in the Supplemental Methods.

Genetic variant and RNAseq data analysis

Genetic data were obtained via the National Cancer Institute Genomic Data Commons for Genotypes and Phenotypes (accession phs001444.v1.p1) and analyzed for genetic variants and RNAseq expression as described in the Supplemental Methods.

Glycopeptide mass spectrometry of lymphoma-derived Fabs

Lymphoma Fabs generation and purification procedures are described in the Supplemental Methods. Fabs were digested separately with trypsin and chymotrypsin (Promega) at a 1:30 (w/w) ratio in 50mM Tris/HCl, pH 8.0. Glycopeptides were eluted using C18 Zip-tip (MerckMilipore), dried, resuspended in 0.1% formic acid and analyzed by nanoLC-electrospray ionization mass spectrometry with an Easy-nLC 1200 system and a Fusion mass spectrometer (Thermo_Fisher_Scientific) using higher-energy collisional dissociation fragmentation. Peptides were separated using an Easy-Spray PepMap RSLC C18 column (75 μ m x 75cm) with a 0-32% acetonitrile linear gradient in 0.1% formic acid for 240 minutes followed by 80% acetonitrile in 0.1% formic acid for 35 minutes (flow rate 200nl/minute; spray voltage 2.8 kV; capillary temperature 275°C; collisional energy

50%). Glycopeptide fragmentation data were extracted using Byonic™ and Byologic™ softwares (Protein_Metrics). The peptide was scored as true-positive when the correct b and y fragment ions were observed along with oxonium ions corresponding to the glycan identified (Table S1). The chromatographic areas for each true-positive peptide with the same amino acid sequence were compared to determine the relative amounts of each glycoform at individual sites.

Determination of L14 Fab crystal structure

Determination of the purified L14 Fab crystal structure following X-ray diffraction data recording, indexing, integration, and scaling was performed as per Supplemental Methods. Fab atomic coordinates and structure factors have been deposited in the Protein Data Bank (access code 6ZEC, Table S2 and PDB report).

Heterotypic cell clustering

Heterotypic cell clustering was determined by flow cytometry and by inverted fluorescence microscopy.

For flow cytometry, Raji, Raji/DC-SIGN and monocyte-derived dendritic cells (MoDCs) (Supplemental Methods) were labeled with CFSE (10 minutes at 37°C in PBS), while lymphoma cells were stained with DiD (Invitrogen) (30 minutes at 37°C in PBS). Cells were washed 3 times before co-culture in 1mM calcium medium. Lymphoma cells were co-cultured with Raji or Raji/DC-SIGN (1:1 ratio, 30 minutes, 37°C), or with MoDCs (1:5 ratio, 2 hours, 37°C). Anti-DC-SIGN antibody hlgG1-D1 (humanized AZN-D1 05E03) was added (10nM) to Raji/DC-SIGN and MoDCs for 20 minutes before or for 2 hours during co-culture at 37°C. Cells were fixed in 0.5% paraformaldehyde and clustering was measured in the live cell population using a FACS Canto-II and FlowJo software.

For microscopy, Raji and Raji/DC-SIGN were labeled with CFSE (10 minutes), while lymphoma cells were stained with PKH26 (Sigma-Aldrich) (5 minutes). Clustering was analyzed using an Olympus IX81 microscope.

RESULTS

Incidence and location of acquired N-glycosylation sites in the IGHV of DLBCL subsets

Incidence of AGS was analyzed in 321 primary DLBCL. These included 309 samples from the GDC portal,^{3,28} and 12 from our tissue banks, and were 92 GCB-DLBCL and 180 ABC-DLBCL (Table S3, Figure 1). In this cohort, selected for DLBCL with IGHV homology to germline <98%, the mean somatic hypermutation frequency was higher in GCB than in ABC-DLBCL (Figure 1A), in line with previous studies.^{12,29} AGS incidence was strikingly different between GCB (60%) and ABC (13%) subsets ($p < 0.0001$), while the unclassified subset had an intermediate level (Figure 1B).

In GCB-DLBCL there was no skewed *IGHV* use and AGS accumulated in multiple *IGHV*, irrespective of IgM or IgG isotype (Table S3), and 85% AGS were located in the complementarity-determining regions (CDR), accounting for 51% of all GCB-DLBCL, Figure 1C). This differed significantly from the ABC-DLBCL, where only 6% AGS were located in the CDR. This difference was even more striking in the ABC-

DLBCL using *IGHV4-34* (95% IgM), known to be overused in this subset (25% of all ABC-DLBCL in this analysis), which acquired sites infrequently (16%) and exceptionally in the CDR (2%, Figure 1D).

These data indicated that the differential frequency of AGS between subsets was particularly evident in the CDR of GCB-DLBCL. The frequency in the framework regions (FR) was instead low and similar between GCB-DLBCL and the other subtypes (Figure 1B).

Genetic features of the GCB-DLBCL with N-glycosylation sites acquired in the CDR

Three-hundred-seven samples were available for analysis of genetic subtype and mutational profile (Figure S1A, Tables S3-S4).³ The GCB-DLBCL with AGS in the CDR (CDR+ve) were mainly the EZB subtype (74%), while 26% were either assigned to no specific subtype (18%) or, less frequently, to A53 (5%) or BN2 (3%) (Figure 1E). In contrast, the CDR-ve GCB-DLBCL were more heterogeneous with fewer assigned to EZB (36%, $p < 0.001$), and the remaining 64% to no specific subtype (33%), BN2 (17%), ST2 (9%), or A53 (5%) (Figure S1B). Concerning location, the majority (97%) of AGS in EZB were in the CDR (Figure 1F).

We focused on CDR+ve EZB (65% of all EZB, Figure 1F) and compared their genetic features to CDR-ve EZB (Tables S4-S5, Figure S2). The epigenetic modifier lesions were similarly frequent in the CDR+ve and in the CDR-ve EZB. In contrast, lesions of *BCL2* were more frequent in the CDR+ve (78% translocations and 62.5% mutations) than in the CDR-ve EZB (33% and 15%, $p = 0.008$ and $p = 0.013$). Another striking difference was the significant enrichment of either single copy deletion or mutation of *PTEN* in the CDR+ve (45.4%) compared to the CDR-ve EZB (7.7%). *PTEN* regulates the PI3K/AKT pathway, which may be preferentially induced by lectin-engaged sig-Mann.²² Mutations of *PTEN*, leading to its downregulation, may amplify the antigen-independent signal of CDR+ve lymphomas.^{22,30}

These data documented that the majority of CDR+ve GCB-DLBCL were of the EZB subtype (74%). This subtype could now be divided into CDR+ve EZB, enriched with *BCL2* translocations/mutations reminiscent of the cell of origin (COO) of FL, and CDR-ve EZB with a potentially distinct COO.

Structure and composition of the acquired glycans in the Ig variable region

To characterize the nature of the glycans acquired at CDR and FR sites, we analyzed the structural features of the variable region oligosaccharides of two lymphomas, L14 and L29 (Figure 2A and Figure S3). L14 was chosen also because it used *IGHV4-34*, known to contain the germline N57 site in the heavy chain CDR2 (HCDR2) which would be expected not to be glycosylated.³¹ The single acquired site in L14 was located in the HCDR1 (N38). The 3 acquired sites in L29 were in the HCDR2 (N55), HFR3 (N84), and kappa light chain CDR3 (KCDR3, N107).

The two CDR glycan sites of L29 terminated at mannoses, with $\text{Man}_8\text{GlcNAc}_2$ dominating the HCDR2 site and $\text{Man}_5\text{GlcNAc}_2$ the KCDR3 site (Figure 2A-B). The HFR3 site was occupied by complex glycans, indicating that it was fully processed by the endoplasmic reticulum and Golgi α -mannosidases and was not under the same processing restraints as the CDR sites (Figure 2A). The L14 Fab glycans detected at the HCDR1 site were also oligomannose-type (Figure 2A-B). The predominant

glycan detected was Man₆GlcNAc₂. In line with our previous observations, no glycans could be detected at the N57 germline site (Figure 2A).

To understand the local steric environment surrounding the N-glycans, we resolved the crystal structure of the L14 Fab to 1.65 Å (Figure 2C). This structure revealed that the N38 site was buried amongst the CDR loops. Due to the intrinsically flexible nature of N-glycans, we could not accurately resolve the N-glycan beyond the first two GlcNAc residues. However, the presence of oligomannose-type glycans at N38 can be rationalized as the glycan-processing enzymes are likely to be prevented from making extensive contacts with the target glycan due to its interaction with residues within the HCDR2, HCDR3, and KCDR3 loops. Prominent interactions with GlcNAc residues of the mannosylated glycan were mediated by S108 and E110 in the HCDR3. The final key contact observed was between the Q44 side chain in HFR2. Q44 side chain is oriented internally and inaccessible to solvent. Hence, its interaction with the glycan appears possible only because the glycan itself is buried in the CDR loops. Lack of electron density for any glycan at N57 in HCDR2 provided further evidence that this germline site was unoccupied. This position contacts the GlcNAc acetyl group, suggesting that N57 cannot be occupied by a glycan without disrupting the Fab structure.

These data suggest that the buried nature of the glycan region likely inhibits the trimming activity of the mannosidases in the endoplasmic reticulum and early Golgi apparatus, resulting in the presentation of oligomannose-type glycans on the mature protein.

Mannosylated Ig is expressed at the cell surface

The mannosylation status of slg was analyzed enzymatically in CDR+ve primary samples and cell lines (Tables S6-S7). In all cases, SDS-PAGE separation showed a single band of slg_μ or slg_γ chain, that increased mobility after mannose cleavage by Endo H (Figures 3 and S4). Removal of all glycan by PNGase F increased mobility further. A CDR-ve cell line and normal B cells were unaffected by Endo H, but were susceptible to PNGase F. These results confirm that the slgs (either IgM or IgG) of CDR+ve DLBCL, but not CDR-ve DLBCL and normal B cells, are mannosylated.

C-type lectin binding of the mannosylated Ig

We assessed the ability of slg-Mann to bind to DC-SIGN, a calcium-dependent lectin known to recognize oligomannose-type glycans.³² Specificity for slg-Mann was confirmed by the failure of DC-SIGN to bind to normal B cells, including naïve, memory, and germinal center B cells (Figure S5A) and by the loss of DC-SIGN binding following Endo H digestion (Figure S5B).

Ten DLBCL primary samples were available for DC-SIGN binding assay (Table S6). DC-SIGN binding to the phenotypically defined tumor cells was observed in 7/10 samples (Figure 4A). The 7 cases which bound DC-SIGN were all CDR+ve GCB-DLBCL. Conversely, the 3 DC-SIGN non-binders were 1 CDR-ve GCB-DLBCL, and 2 ABC-DLBCL with no sites.

Analysis of 10 cell lines (Table S7) confirmed that, like in the primary samples, binding occurred in all the CDR+ve GCB-DLBCL, but not in the CDR-ve GCB-DLBCL with FR sites (OCI-Ly7 and WSU-DLCL2) or in the ABC-DLBCL (Table S7, Figures 4B and S5B-C). Binding to tumor cells appeared homogeneous, with no subpopulations unable to bind.

The conclusion is that AGS in the CDRs are occupied by oligomannose-type glycans able to bind DC-SIGN, while those in the FRs are not. Thus, location of the glycan acceptor motifs in the CDR appears critical for mannosylation and for conferring the ability to bind DC-SIGN. This is a tumor-related feature of a subset of GCB-DLBCL which is not evident in normal B cells and not found in ABC-DLBCL.

DC-SIGN induces slg-Mann-mediated low-level intracellular signals in DLBCL

To assess functional competence of slg-Mann in DLBCL, we measured SYK phosphorylation (pSYK) in slg-Mann+ve samples and cell lines following DC-SIGN or anti-Ig exposure (Figure 5A). We selected a Phosflow assay that allowed detection of signals within the tumor (CD20+/BCL2^{hi}) population and excluded non-tumor cells (Figure S6A).³³ We found pSYK induction by DC-SIGN in both the primary DLBCL samples and the GCB-DLBCL lines (Figures 5A and S6B-C). Levels of phosphorylation induced by DC-SIGN were lower than by anti-Ig, and levels and kinetics varied between cases. In the ABC cell line TMD8, phosphorylation of SYK was not induced by DC-SIGN, while cells responded to anti-Ig, as expected. By additional immunoblot analyses, we also found that pAKT and pERK levels were increased after 15 minutes of treatment with DC-SIGN (Figure 5B). Pre-incubation of DC-SIGN with the hlgG1-D1 antibody, which is specific for the DC-SIGN carbohydrate-recognition-domain,^{23,34-36} inhibited the capacity of DC-SIGN to induce signaling in slg-Mann+ve lymphoma cells of either a primary sample or a cell line (Figure 5C and S6D), confirming that DC-SIGN-induced signaling occurred by engagement with the slg oligomannose-type glycans.

Direct interaction between slg-Mann+ve lymphoma cells and DC-SIGN-expressing cells

We assessed the ability of slg-Mann+ve lymphoma cell lines to interact with the DC-SIGN-transduced Raji cell line (Raji/DC-SIGN) or with MoDCs (Figures 6 and S7), which express DC-SIGN after treatment with IL-4.²³ The flow cytometry and fluorescence microscopy assays documented that slg-Mann+ve lymphoma cells clustered with both Raji/DC-SIGN (but not with parental Raji) and MoDCs (Figure 6A-B). Conversely, slg-Mann-ve lymphoma cells failed to cluster with Raji/DC-SIGN (Figure S8A).

Incubation of Raji/DC-SIGN or MoDCs with hlgG1-D1 efficiently prevented cluster formation with slg-Mann+ve cells (Figures 6C-D and S8B). Furthermore, the addition of hlgG1-D1 following co-culture disrupted pre-formed clusters (Figure 6E), confirming the specificity of the slg-Mann:DC-SIGN interaction. These data show that cell-cell interaction occurs between lymphoma cells expressing slg-Mann and DC-SIGN expressing cells, including MoDCs.

Maintenance of slg-Mann function following transformation from FL to DLBCL

IGHV-IGHD-IGHJ rearrangement, DC-SIGN binding, and slg-Mann-mediated signaling capacity were investigated in a patient with biopsies available both at the time of FL diagnosis and following DLBCL transformation (Figure S9). *IGHV-IGHD-IGHJ* rearrangements were clonally related, and the AGS in the CDR3 of the primary FL was preserved at transformation. The AGS was occupied by an oligomannose-type glycan, as documented by DC-SIGN binding of the entire clonal population with similar pattern at both time points. DC-SIGN induced SYK phosphorylation and the anti-DC-SIGN hlgG1-D1 efficiently prevented cluster formation at both times.

The finding in three previous patients that the AGS persists through a transformation from FL to DLBCL despite ongoing somatic hypermutation suggested that the influence of slg-Mann continued.¹⁶ The results from this new case demonstrate that oligomannose-type glycans occupying the slg antigen-binding site preserve structure and function with retention of DC-SIGN binding continuing after transformation.

Association of slg-Mann with DLBCL clinical progression

The observation that 65% EZB acquired sites in the CDR, while 35% did not (Figure 1F) prompted us to investigate the association of slg-Mann with clinical progression in this subtype (Table S8). The CDR+ve EZB progressed more rapidly than the CDR-ve ($p=0.007$, Figure 7A). However, 44% CDR+ve EZB had a *MYC* translocation (EZB MYC+ve), compared to 18% CDR-ve EZB ($p=0.1$), and the more rapid progression of the CDR+ve lost statistical significance within the EZB MYC+ve ($p=0.152$), possibly due to small numbers and *MYC* overexpression mastering the transcriptional profile (Figure S10).³ To avoid this, EZB MYC-ve lymphomas were investigated separately. Here, although numbers were limited, CDR+ve EZB progressed significantly faster than CDR-ve EZB ($p=0.037$, Figure 7B). These data suggested that the acquisition of oligomannose-type glycans in the antigen-binding sites associates with more rapid progression.

Association of slg-Mann with the differential slg signaling in patients

To verify the hypothesis that the slg-Mann might determine a chronic antigen-independent BCR signal affecting clinical behavior, we compared the gene expression signature of CDR+ve against CDR-ve EZB MYC-ve lymphomas. CDR+ve EZB MYC-ve had 22 genes upregulated and 96 genes downregulated (Table S9, Figure 7C). Gene set enrichment analysis revealed a significant enrichment of the B-cell receptor signaling pathway ($p=0.024$, Figure 7D) with upregulation of genes involved with PI3K/AKT/MTORC1 pathway, glucose metabolism, mitosis, and translation (Table S10). Although these cases did not have a documentation of *MYC* translocation in the public dataset,³ there was also upregulation of *MYC*-associated genes, suggesting some association between slg-Mann and *MYC* activation.

DISCUSSION

Our data describe the acquisition of N-glycosylation motifs in the antigen-binding site of cases of DLBCL and add a stable and discriminatory feature to current gene-based disease subdivisions. The presence of unusual oligomannose-type glycans in the antigen-binding sites (CDR) has now identified a new GCB-DLBCL subset with a specific ability to interact with environmental lectins through the glycan-modified slg. Although its frequency might be slightly higher if the *IG* light chain is also investigated,³⁷ our previous analysis of heavy and light chain V regions in 44 cases of FL revealed only 3 cases with sites in the light chain only and the nature of those sites remains to be assessed. The sites in CDR identify 51% of GCB-DLBCL cases and contrast strikingly with the ABC-DLBCL subset (6%).^{11,12}

A recent probabilistic classification of GCB-DLBCL based on genetic features has delineated a subtype, EZB, identified by the presence of a *BCL2* translocation and mutations of epigenetic modifiers.³ Our study adds another dimension of heterogeneity by documenting that a large percentage (74%) of CDR+ve GCB-

DLBCL lies within the EZB subtype. The sites in EZB are located mainly in the CDR as might be expected since the COO is likely to be related to FL, where such sites are universal and almost exclusively (~90%) in the CDR.³⁷ The acquisition of sites reflects the fact that the COO of a potentially new subtype has undergone somatic hypermutation and positive selection for this feature, and that this clonal marker persists across FL and DLBCL.

However, 35% EZB have no AGS in the CDR, pointing to a different derivation and revealing heterogeneity within this category. Also 26% GCB-DLBCL have AGS in the CDR but are not EZB, perhaps suggesting a COO in common with FL, but with a distinct subsequent development. As well as revealing the COO, the presence of sites is likely to confer environmental support which then sets the scene for further genetic events.

Probing for an overall function of the slg-Mann has to account for the fact that sites are acquired by a wide range of IGV rearrangements and in different positions within the CDR. We looked in detail at the added glycans in recombinant Fabs derived from two lymphomas and found strong evidence that structural features control glycan processing. It is rare for mammalian-expressed glycoproteins to present such homogenous mannosylation. However, it is reminiscent of pathogens such as HIV where, on the densely glycosylated viral envelope, several sites have exclusively under-processed oligomannose-type glycans.³⁸ The reason for stalling further sugar addition can be seen when ER α 1,2-mannosidase-I is docked onto a particular glycosylation site where adjacent glycans clash and prevent processing.³⁹ A similar mechanism appears to be occurring here where the Fab inhibits glycan processing by steric protection of the AGS from the catalytic site buried inside the mannosidase.⁴⁰ In lymphoma, where AGS can be added to any available IGV during somatic hypermutation, this appears therefore as a tumor-specific process selecting only those sites in which mannosidase processing is stalled and the resulting oligomannose-type glycans can bind DC-SIGN.⁴¹

DC-SIGN is expressed on M2-polarised macrophages and expected to serve for heterotypic interactions with slg-Mann in lymphoma cells.^{23,24,42-44} The mechanism whereby the interaction with microenvironmental DC-SIGN supports tumor cells may involve a low-level antigen-independent signal similar to those in FL cells, inducing survival and possibly MYC-mediated growth.^{22,25} We have found that cell-cell interaction occurs between lymphoma-expressed slg-Mann and MoDC-expressed DC-SIGN, and it is possible that this interaction may fine-tune the slg for adequate survival signals in DLBCL too,^{25,45} while at the same time shielding the slg from its original antigen,⁴⁶ and protecting it from potential slg overstimulation by an alternative ligand.⁴⁷

We probed the impact of the slg-Mann on clinical behavior by investigating progression-free survival in the EZB subset. The CDR+ve EZB were enriched in *MYC* translocations mastering the transcriptional profile over the slg-Mann mediated signals. However, the concurrence of *MYC* translocations and AGS is also found in endemic Burkitt lymphoma,^{14,27} suggesting a pathogenetic link between these two events. To identify a signature associated with slg-Mann *in vivo*, we focused on those EZB that did not carry a *MYC* translocation. While in the *MYC*+ve EZB the number of cases investigated was small to reach statistical value, in the *MYC*-ve EZB the acquisition of N-glycosylation sites identified a subset of patients with significantly more aggressive lymphoma. Validation of this observation will be required in independent studies rigorously accounting for *BCL2* and *MYC* status.⁴⁸ However, the

upregulation of the BCR pathway was evident in these lymphomas, which also carried an enrichment of MYC- and glycolytic-pathway genes.

The interaction between slg-Mann and DC-SIGN depends on a non-templated post-translational modification, revealing a new opportunistic strategy exploited by tumor cells that behave more aggressively in patients. The interaction can be inhibited by antibodies blocking DC-SIGN carbohydrate-recognition-domain.³⁵ Even pre-formed clusters of interacting cells can be disrupted by antibody, raising the possibility that this specific dependence can be blocked *in vivo*, separating any slg-Mann+ve cells from their natural habitat. This focused microenvironmental disruption could lead to tumor cell stress while avoiding significant effects on normal cells. Blockade by antibody could offer a potential new specific therapeutic approach for slg-Mann+ve lymphomas.

ACKNOWLEDGMENTS

We thank Prof. Riccardo Dalla-Favera and Dr. Laura Pasqualucci (Columbia University, NY) for generously providing the DLBCL cell lines. We are grateful to Bas van der Schoot and Dr. Martijn Verdoes (Radboud University) for providing the anti-DC-SIGN antibody. We thank Prof. Peter Johnson for his contribution to the provision of lymphoma specimens in Southampton. We are grateful to Dr. Kathy Potter at the Faculty of Medicine Tissue Bank (Cancer Sciences Unit, University of Southampton) for the processing and storage of the primary DLBCL specimens. We thank Dr. Tom Walter (University of Oxford) for assistance with crystal freezing and Diamond Light Source for beamtime (proposal mx14744) and crystallographic data collection. The Genomic Variation in Diffuse Large B Cell Lymphomas study was supported by the Intramural Research Program of the National Cancer Institute, National Institutes of Health, Department of Health and Human Services. This work was funded by Blood Cancer UK/Bloodwise (grant 18009), the Keanu Eyles Hematology Fellowship, and Cancer Research UK (grant C2750/A23669, C36811/A29101, and C42023/A29370). The Southampton Experimental Cancer Medicine Centre and the Cancer Research Cancer Research UK Southampton Centre core funding (C24563/A15581, C34999/A18087) supported I.W., P.J.D., and part of the research of J.A. and M.C. The mass spectrometer was funded by the Bill and Melinda Gates Foundation through the Collaboration for AIDS Vaccine Discovery (OPP1115782 to M.C.). Crystallographic work was supported by the Medical Research Council (MR/L009528/1 and MR/S007555/1 to T.A.B.). The Wellcome Centre for Human Genetics is supported by Wellcome Centre grant 203141/Z/16/Z.

AUTHORSHIP CONTRIBUTIONS

G.C. performed research, analyzed and interpreted data, and wrote the manuscript. J.D.A. performed and interpreted glycan analysis. D.J.B. designed the bioinformatic pipeline for the immunoglobulin gene analysis and analyzed genetic variant data. P.R., E.A.M., and B.V.A. contributed to the characterization of primary samples. P.J.D. generated the lymphoma-derived Fab. I.H. and K.M. performed the sequencing and characterization of primary samples. Y.W. and T.A.B. performed and interpreted X-ray crystallography analyses. J.S.B. contributed to the immunoglobulin gene analysis pipeline. C.F. contributed anti-DC-SIGN antibodies. T.G. contributed DC-SIGN expressing and parental cell lines. D.J.B. and G.W. performed gene

expression data analysis. J.C.S., G.P., T.G., C.F., L.M.S., and R.B. contributed to the interpretation and discussion of the data. M.C. and T.B. led the structural biology research and wrote the manuscript. F.K.S. and F.F. designed the study, supervised research, interpreted data and wrote the manuscript. All authors reviewed and approved the manuscript.

DISCLOSURE OF CONFLICT OF INTEREST

The authors declare no potential conflicts of interest.

REFERENCES

1. Alizadeh AA, Eisen MB, Davis RE, et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* 2000;403:503-11.
2. Rosenwald A, Wright G, Chan WC, et al. The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma. *N Engl J Med* 2002;346:1937-47.
3. Wright GW, Huang DW, Phelan JD, et al. A Probabilistic Classification Tool for Genetic Subtypes of Diffuse Large B Cell Lymphoma with Therapeutic Implications. *Cancer Cell* 2020;37:551-68 e14.
4. Pasqualucci L, Dalla-Favera R. Genetics of diffuse large B-cell lymphoma. *Blood* 2018;131:2307-19.
5. Green MR, Kihira S, Liu CL, et al. Mutations in early follicular lymphoma progenitors are associated with suppressed antigen presentation. *Proc Natl Acad Sci U S A* 2015;112:E1116-25.
6. Pasqualucci L, Dominguez-Sola D, Chiarenza A, et al. Inactivating mutations of acetyltransferase genes in B-cell lymphoma. *Nature* 2011;471:189-95.
7. Okosun J, Bodor C, Wang J, et al. Integrated genomic analysis identifies recurrent mutations and evolution patterns driving the initiation and progression of follicular lymphoma. *Nat Genet* 2014;46:176-81.
8. Lam KP, Kuhn R, Rajewsky K. In vivo ablation of surface immunoglobulin on mature B cells by inducible gene targeting results in rapid cell death. *Cell* 1997;90:1073-83.
9. George AJ, Spellerberg MB, Stevenson FK. Idiotype vaccination leads to the emergence of a stable surface Ig-negative variant of the mouse lymphoma BCL1, with different growth characteristics. *J Immunol* 1988;140:1695-701.
10. Varano G, Raffel S, Sormani M, et al. The B-cell receptor controls fitness of MYC-driven lymphoma cells via GSK3beta inhibition. *Nature* 2017;546:302-6.
11. Phelan JD, Young RM, Webster DE, et al. A multiprotein supercomplex controlling oncogenic signalling in lymphoma. *Nature* 2018;560:387-91.
12. Young RM, Wu T, Schmitz R, et al. Survival of human lymphoma cells requires B-cell receptor engagement by self-antigens. *Proc Natl Acad Sci U S A* 2015;112:13447-54.
13. Wilson WH, Young RM, Schmitz R, et al. Targeting B cell receptor signaling with ibrutinib in diffuse large B cell lymphoma. *Nat Med* 2015;21:922-6.
14. Zhu D, McCarthy H, Ottensmeier CH, Johnson P, Hamblin TJ, Stevenson FK. Acquisition of potential N-glycosylation sites in the immunoglobulin variable region by somatic mutation is a distinctive feature of follicular lymphoma. *Blood* 2002;99:2562-8.
15. Kupperts R, Stevenson FK. Critical influences on the pathogenesis of follicular lymphoma. *Blood* 2018;131:2297-306.
16. Odabashian M, Carlotti E, Araf S, et al. IGHV sequencing reveals acquired N-glycosylation sites as a clonal and stable event during follicular lymphoma evolution. *Blood* 2020;135:834-44.
17. Radcliffe CM, Arnold JN, Suter DM, et al. Human follicular lymphoma cells contain oligomannose glycans in the antigen-binding site of the B-cell receptor. *J Biol Chem* 2007;282:7405-15.
18. Giovannone N, Liang J, Antonopoulos A, et al. Galectin-9 suppresses B cell receptor signaling and is regulated by I-branching of N-glycans. *Nat Commun* 2018;9:3287.
19. Watanabe Y, Bowden TA, Wilson IA, Crispin M. Exploitation of glycosylation in enveloped virus pathobiology. *Biochim Biophys Acta Gen Subj* 2019;1863:1480-97.
20. Loke I, Kolarich D, Packer NH, Thaysen-Andersen M. Emerging roles of protein mannosylation in inflammation and infection. *Mol Aspects Med* 2016;51:31-55.
21. Coelho V, Krysov S, Ghaemmaghami AM, et al. Glycosylation of surface Ig creates a functional bridge between human follicular lymphoma and microenvironmental lectins. *Proc Natl Acad Sci U S A* 2010;107:18587-92.
22. Linley A, Krysov S, Ponzoni M, Johnson PW, Packham G, Stevenson FK. Lectin binding to surface Ig variable regions provides a universal persistent activating signal for follicular lymphoma cells. *Blood* 2015;126:1902-10.

23. Geijtenbeek TB, Torensma R, van Vliet SJ, et al. Identification of DC-SIGN, a novel dendritic cell-specific ICAM-3 receptor that supports primary immune responses. *Cell* 2000;100:575-85.
24. Soilleux EJ, Morris LS, Leslie G, et al. Constitutive and induced expression of DC-SIGN on dendritic cell and macrophage subpopulations in situ and in vitro. *J Leukoc Biol* 2002;71:445-57.
25. Amin R, Mourcin F, Uhel F, et al. DC-SIGN-expressing macrophages trigger activation of mannosylated IgM B-cell receptor in follicular lymphoma. *Blood* 2015;126:1911-20.
26. Savelyeva N, King CA, Vitetta ES, Stevenson FK. Inhibition of a vaccine-induced anti-tumor B cell response by soluble protein antigen in the absence of continuing T cell help. *Proc Natl Acad Sci U S A* 2005;102:10987-92.
27. Forconi F, Capello D, Berra E, et al. Incidence of novel N-glycosylation sites in the B-cell receptor of lymphomas associated with immunodeficiency. *Br J Haematol* 2004;124:604-9.
28. Schmitz R, Wright GW, Huang DW, et al. Genetics and Pathogenesis of Diffuse Large B-Cell Lymphoma. *N Engl J Med* 2018;378:1396-407.
29. Xu-Monette ZY, Li J, Xia Y, et al. Immunoglobulin somatic hypermutation has clinical impact in DLBCL and potential implications for immune checkpoint blockade and neoantigen-based immunotherapies. *J Immunother Cancer* 2019;7:272.
30. Young RM, Phelan JD, Wilson WH, Staudt LM. Pathogenic B-cell receptor signaling in lymphoid malignancies: New insights to improve treatment. *Immunol Rev* 2019;291:190-213.
31. McCann KJ, Ottensmeier CH, Callard A, et al. Remarkable selective glycosylation of the immunoglobulin variable region in follicular lymphoma. *Mol Immunol* 2008;45:1567-72.
32. Guo Y, Feinberg H, Conroy E, et al. Structural basis for distinct ligand-binding and targeting properties of the receptors DC-SIGN and DC-SIGNR. *Nat Struct Mol Biol* 2004;11:591-8.
33. Myklebust JH, Brody J, Kohrt HE, et al. Distinct patterns of B-cell receptor signaling in non-Hodgkin lymphomas identified by single-cell profiling. *Blood* 2017;129:759-70.
34. Tacke PJ, de Vries IJ, Gijzen K, et al. Effective induction of naive and recall T-cell responses by targeting antigen to human dendritic cells via a humanized anti-DC-SIGN antibody. *Blood* 2005;106:1278-85.
35. Tacke PJ, Ginter W, Berod L, et al. Targeting DC-SIGN via its neck region leads to prolonged antigen residence in early endosomes, delayed lysosomal degradation, and cross-presentation. *Blood* 2011;118:4111-9.
36. Geijtenbeek TB, Kwon DS, Torensma R, et al. DC-SIGN, a dendritic cell-specific HIV-1-binding protein that enhances trans-infection of T cells. *Cell* 2000;100:587-97.
37. McCann KJ, Johnson PW, Stevenson FK, Ottensmeier CH. Universal N-glycosylation sites introduced into the B-cell receptor of follicular lymphoma by somatic mutation: a second tumorigenic event? *Leukemia* 2006;20:530-4.
38. Struwe WB, Chertova E, Allen JD, et al. Site-Specific Glycosylation of Virion-Derived HIV-1 Env Is Mimicked by a Soluble Trimeric Immunogen. *Cell Rep* 2018;24:1958-66 e5.
39. Behrens AJ, Crispin M. Structural principles controlling HIV envelope glycosylation. *Curr Opin Struct Biol* 2017;44:125-33.
40. Xiang Y, Karaveg K, Moremen KW. Substrate recognition and catalysis by GH47 alpha-mannosidases involved in Asn-linked glycan maturation in the mammalian secretory pathway. *Proc Natl Acad Sci U S A* 2016;113:E7890-E9.
41. van Liempt E, Bank CM, Mehta P, et al. Specificity of DC-SIGN for mannose- and fucose-containing glycans. *FEBS Lett* 2006;580:6123-31.
42. Lenz G, Wright GW, Emre NC, et al. Molecular subtypes of diffuse large B-cell lymphoma arise by distinct genetic pathways. *Proc Natl Acad Sci U S A* 2008;105:13520-5.
43. Scott DW, Gascoyne RD. The tumour microenvironment in B cell lymphomas. *Nat Rev Cancer* 2014;14:517-34.
44. Lamaison C, Tarte K. B cell/stromal cell crosstalk in health, disease, and treatment: Follicular lymphoma as a paradigm. *Immunol Rev* 2021;302:273-85.
45. Akkaya M, Traba J, Roesler AS, et al. Second signals rescue B cells from activation-induced mitochondrial dysfunction and death. *Nat Immunol* 2018;19:871-84.

46. Schneider D, Duhren-von Minden M, Alkhatib A, et al. Lectins from opportunistic bacteria interact with acquired variable-region glycans of surface immunoglobulin in follicular lymphoma. *Blood* 2015;125:3287-96.
47. Valle-Argos B, Chiodin G, Bryant DJ, et al. DC-SIGN binding to mannosylated B-cell receptors in follicular lymphoma down-modulates receptor signaling capacity. *Scientific Reports* 2021;11:11676.
48. Ennishi D, Jiang A, Boyle M, et al. Double-Hit Gene Expression Signature Defines a Distinct Subgroup of Germinal Center B-Cell-Like Diffuse Large B-Cell Lymphoma. *J Clin Oncol* 2019;37:190-201.

FIGURE LEGENDS

Figure 1. Frequency and location of AGS in the Ig heavy chain variable region of DLBCL subsets.

Tumor *IGHV-IGHD-IGHJ-IGHC* rearrangements from primary DLBCL cases were analyzed according to the IMGT/V-QUEST numbering system and N_xS/T motifs acquired by somatic hypermutation (AGS) were identified. The pie-charts represent the frequency of rearrangements with at least 1 AGS (red slice) or no AGS (white slice). Bar-charts identify the percentage of rearrangements with sites, containing at least 1 AGS in the CDR (blue bars) or only in the FR (white bars). **(A)** Homology (%) of the tumor *IGHV* transcript sequence to the closest germline sequence in GCB-, ABC- and Unclassified DLBCL. **(B)** Frequency of rearrangements with AGS (pie-charts) and their distribution in the CDR or FR only (bar-charts), divided by COO subset (GCB-, ABC- and Unclassified DLBCL). **(C)** Frequency of rearrangements with AGS in the CDR (blue bars) or the FR only (white bars) relative to the total number of GCB-, ABC- and Unclassified DLBCL. **(D)** Frequency of rearrangements with AGS (pie-chart) and their distribution in the CDR or FR only (bar-chart) in the ABC-DLBCL using *IGHV4-34*. **(E)** Distribution of genetic subtypes (according to LymphGen algorithm) within GCB-DLBCL having acquired sites in the CDR. **(F)** Frequency of rearrangements with AGS (pie-chart) and their percent distribution in the CDR or FR only (bar-chart) in GCB-DLBCL assigned to the EZB genetic subtype according to the LymphGen algorithm.

Figure 2. Glycan composition and structure of the lymphoma-derived Fabs.

Determination of the crystal structure of L14 and glycan analysis of Fabs L14 and L29 that have N-glycan sites in both the CDR loops and the FR. **(A)** Schematic representation of the Ig variable heavy (H) and kappa light (K) chain pairs of L29 and L14 displaying the location of the AGSs and of the natural N57 glycosylation site. Sites are numbered according to the IMGT/V-QUEST numbering system. Distribution of each site is represented relative to CDR or FR and colored by predominant glycan composition. Pie charts are colored according to the proportion of oligomannose-type glycans (green), processed complex-type glycans (magenta) and no glycans (grey) at each site. At least one CDR-AGS is always occupied by glycans terminating at oligomannose-type in both L14 and L29, while the site in FR3 of L29 is predominantly but not always occupied by complex glycans. **(B)** Site-specific glycan compositions detected for L29 and L14 as determined by liquid chromatography-mass spectrometry (LC-MS). Bars represent the relative abundance of each category of glycan. Oligomannose-type (green) glycans are categorized according to the number of mannose residues (M9-M5), hybrids by the presence/absence of fucose, and complex-type glycans (magenta) are categorized according to the number of N-acetyl hexosamine structures detected (HexNAc) and the presence/absence of fucose (F). The proportion of AGS without a glycan attached are shown as a grey bar. **(C)** Structure of the L14 Fab at 1.65 Å resolution. The Ig heavy chain (HC) is colored light grey and the Ig kappa light chain (KC) is colored dark grey. CDR loops are colored magenta (HCDR1), purple (HCDR2), blue (HCDR3), red (KCDR1), yellow (KCDR2) and teal (KCDR3). The amino acids Q44, N57, E110 which interact with the glycan at position N38 are represented as sticks. Their carbon atoms are colored according to the CDR which they are present on, while their oxygen atoms are colored in red and

their nitrogen atoms in blue. The two resolved GlcNAc residues at N38 are colored by atom with carbon in green, oxygen in red and nitrogen in blue. Atoms likely to form electrostatic interactions with the N38 glycan are shown as dashed yellow lines. The representation in the left panel is rotated by 100° on the Y-axis and 20° on the X-axis in the right panel to visualize the contacts of the glycan to itself and to the protein. Maximum likelihood-weighted $2Fo-Fc$ electron density obtained for the glycan at N38 is shown (green mesh).

Figure 3. N-linked mannoses in the antigen-binding site of GCB-DLBCL are susceptible to Endo H treatment.

The glycosylation pattern of GCB-DLBCL primary samples, known to have AGS in the CDR, or PBMCs from healthy donors, was analyzed by digestion with Endo H (which cleaves mannose only) or PNGase F (which removes all glycan) following biotinylation and isolation of the cell surface proteins. Primary anti- μ or anti- γ antibodies were used to detect the surface Ig μ (16-TB0006 and 16-TB0084 primary samples) or Ig γ heavy chains (12-TB0153 and 16-TB0014) by immunoblotting, respectively. Numbers on the left to each gel indicate the molecular weight in kiloDaltons of the reference ladder. The characteristics of the primary samples are described in Table S6. The glycosylation pattern of DLBCL cell lines with or without AGS in the CDR is shown in Figure S4.

Figure 4. DC-SIGN binds specifically to the slg-Mann GCB-DLBCL cells.

Binding of recombinant DC-SIGN was analyzed in primary samples and DLBCL cell lines by flow cytometry after incubation with FITC-conjugated anti-Fc antibody. Red lines represent binding of DC-SIGN, black lines represent binding of the secondary antibody in the non-treated control. **(A)** DC-SIGN binding to the clonal CD20+/CD10+ or CD20+/CD10- tumor population of GCB-DLBCL or ABC-DLBCL primary samples, respectively. **(B)** DC-SIGN binding in GCB-DLBCL (with AGS in the CDR or only the FR) and ABC-DLBCL cell lines (not having AGS). Also, refer to Tables S6-S7 for the intensity of DC-SIGN binding to the primary samples and cell lines.

Figure 5. DC-SIGN mediates low-level antigen-independent signaling in slg-Mann+ve cells.

(A) The histograms show SYK phosphorylation at Y525/526 in a representative slg-Mann+ve primary sample (16-TB0014) and in the WSU-FSCCL cell line after stimulation with DC-SIGN (red) or anti-Ig (blue), or no treatment (NT, grey). Geometric mean fluorescence intensity (MFI) levels of pSYK are shown for each condition. The lower panel shows a heat-map of DC-SIGN mediated and anti-Ig-mediated signaling in all cell lines and primary samples analyzed at 1, 5, and 15 minutes. MFI of pSYK after stimulation was normalized with pSYK MFI of the NT at the respective time point (NT was normalized to 1). The statistical difference between DC-SIGN and anti-Ig stimulation was significant at each time point ($p=0.03$, Wilcoxon signed-rank test). **(B)** Immunoblotting of AKT phosphorylation at S473 and ERK phosphorylation at T202/Y204 in the primary DLBCL sample 17-TB0084 after exposure to DC-SIGN or anti-IgM or left untreated (NT) for 15 minutes. GAPDH was used as the loading control. **(C)** Soluble recombinant DC-SIGN was incubated with 500nM hlgG1-D1 before treatment of NU-DHL1 cells. Phosphorylation of SYK at Y525/526 was measured by flow cytometry. Data are represented as mean \pm SEM of

three independent experiments. Also, refer to Tables S6-S7 for slg characteristics and DC-SIGN binding of the primary samples and cell lines.

Figure 6. The specific interaction of slg-Mann+ve lymphoma cells and DC-SIGN expressing cells.

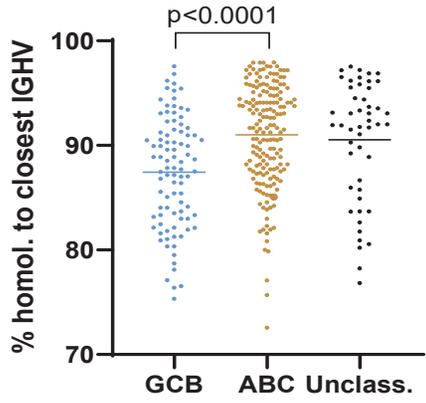
Panels A-B: slg-Mann+ve lymphoma cells form clusters round DC-SIGN expressing cells. **(A)** Flow cytometry analysis of clustering between slg-Mann+ve DLBCL lines (WSU-FSCCL or NU-DHL1) and either Raji/DC-SIGN cells (upper panels) or parental Raji cells (middle panels) or MoDCs (lower panels). **(B)** Inverted fluorescence microscopy images of clustering between slg-Mann+ve DLBCL lines (WSU-FSCCL or NU-DHL1) and either Raji/DC-SIGN (upper panels), or Rajicells (lower panels). **Panels C-D-E:** the interaction of slg-Mann+ve lymphoma cells and DC-SIGN expressing cells is specifically inhibited or interrupted by the anti-DC-SIGN antibody hlgG1-D1. **(C-D)** Raji/DC-SIGN or MoDCs were treated with 10nM hlgG1-D1, or left untreated (NT), before co-culture with WSU-FSCCL or NU-DHL1. Clustering of WSU-FSCCL with Raji/DC-SIGN was determined by flow cytometry (C, left) and inverted fluorescence microscopy (C, right). Percent (%) of clustering, as determined by flow cytometry, in the presence or absence of hlgG1-D1 was calculated as $(\text{double-positive population}) \times 2 / (2 \times \text{double-positive population} + \text{single positive CFSE} + \text{single positive DiD})$ (D). **(E)** Raji/DC-SIGN were cultured with WSU-FSCCL or NU-DHL1 for 30 minutes. HlgG1-D1 (10nM) or medium (NT) was subsequently added to the co-culture and clustering was measured after 2 hours by flow cytometry. Clustering with WSU-FSCCL is shown. Data are represented as mean \pm SEM of at least two independent experiments.

Figure 7. Progression-free survival and gene expression profile of EZB lymphomas acquiring N-glycosylation sites in the CDR.

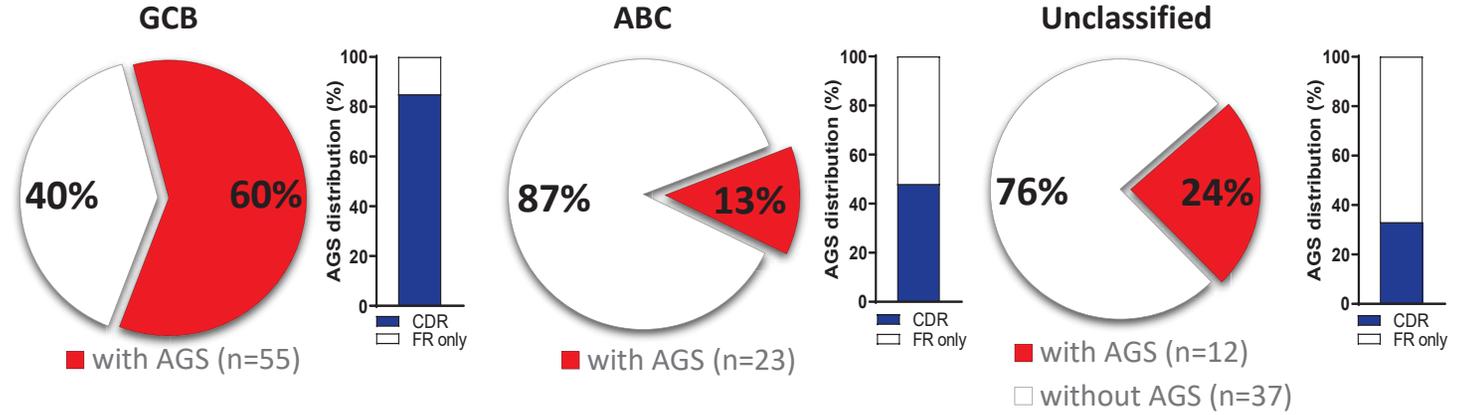
A-B: Progression-free survival (PFS) was determined by the Kaplan-Meier method using log-rank statistics. The number of patients at risk are indicated in blue (CDR-ve) or in red (CDR+ve) at each time point (years). **(A)** Progression-free survival in EZB GCB-DLBCL with N-glycosylation sites acquired in the CDR of the tumor Ig (CDR+ve) or not (CDR-ve). **(B)** Progression-free survival in CDR+ve and CDR-ve EZB GCB-DLBCL without MYC translocation (EZB MYC-ve). **C-D:** Differential gene expression and B-cell receptor (BCR) gene set enrichment in CDR+ve compared to CDR-ve EZB MYC-ve DLBCL. **(C)** Each point represents a gene and the fold change (FC) and P-value for differential expression in CDR+ve vs CDR-ve EZB MYC-ve DLBCL. Positive \log_2 fold changes were expressed higher in CDR+ve while negative \log_2 fold changes were expressed higher in CDR-ve. Grey points were not differentially expressed genes. Light blue/pink points were differentially expressed at $P=0.05$, whilst dark blue/red points were differentially expressed at false discovery rate (FDR)=0.05 after controlling for multiple testing (Benjamini-Hochberg procedure). **(D)** Gene set enrichment plot showing the enrichment of BCR signaling pathway genes in the \log_2 fold change ranked genes for CDR+ve vs CDR-ve. Enrichment of the BCR pathway towards the beginning of the ranked list indicated that the BCR pathway was enriched in genes more highly expressed in CDR+ve than CDR-ve EZB MYC-ve. Leading edge genes (i.e. those observed in the left edge of the green curve) in the BCR pathway are listed in the order that they appear on the x-axis.

Figure 1

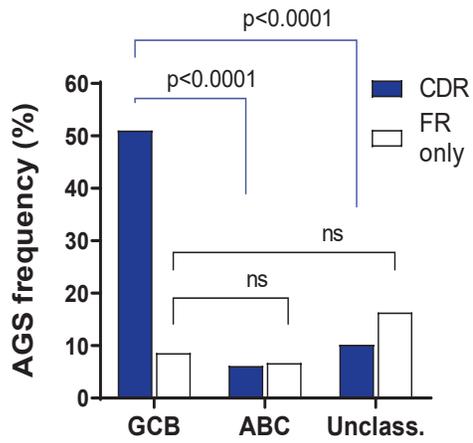
A



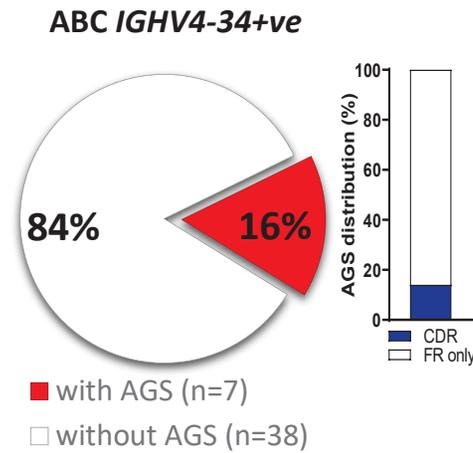
B



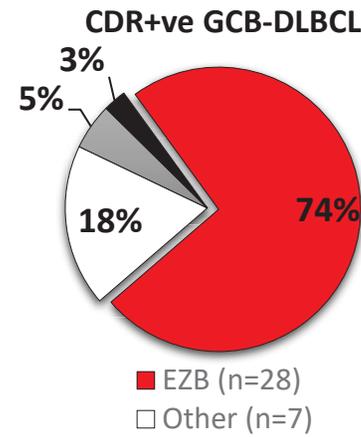
C



D



E



F

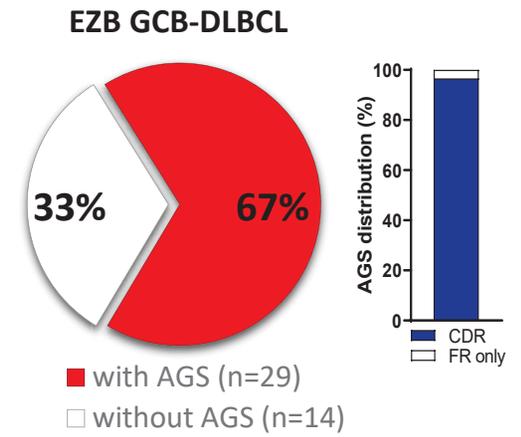
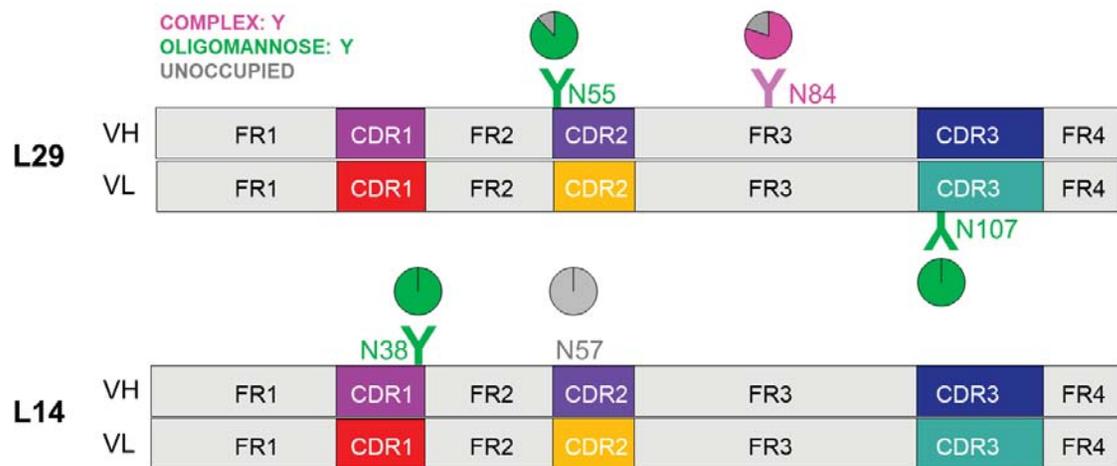
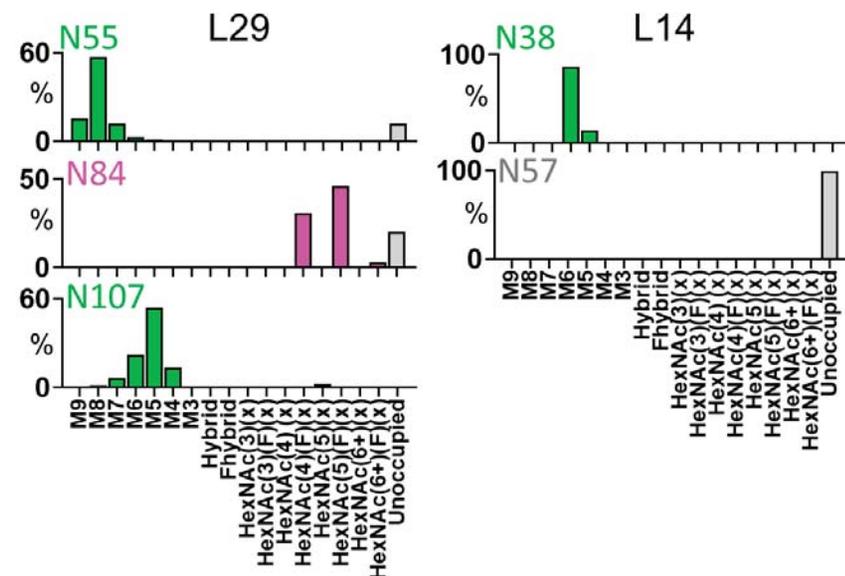


Figure 2

A



B



C

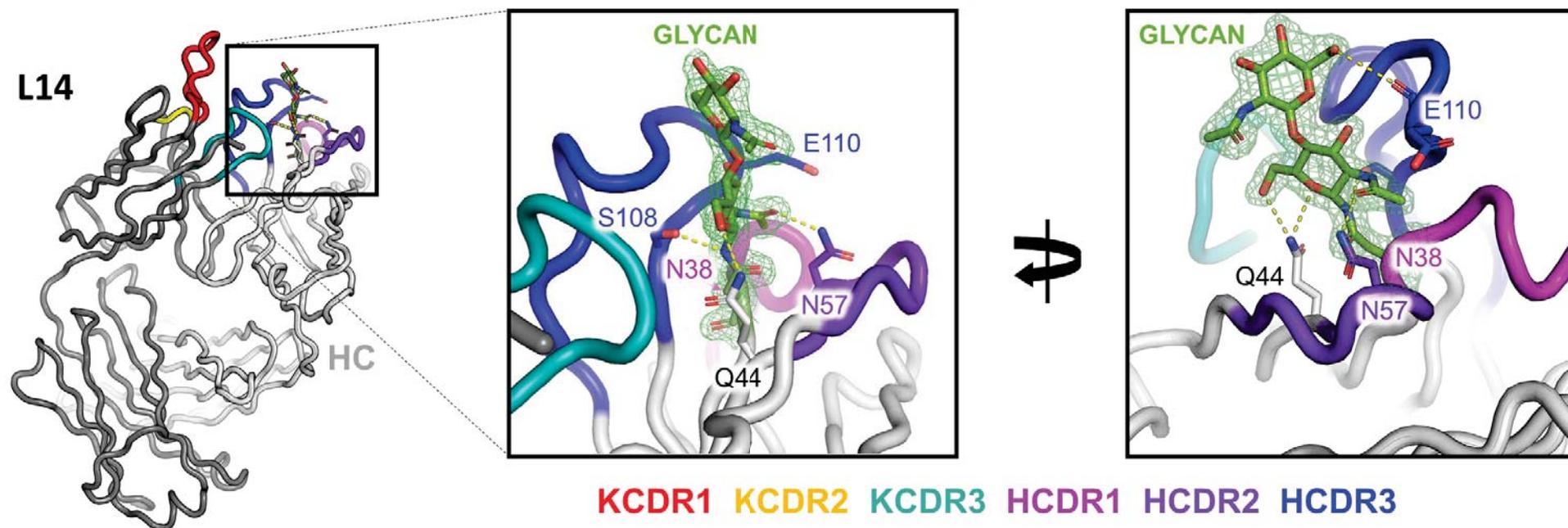


Figure 3

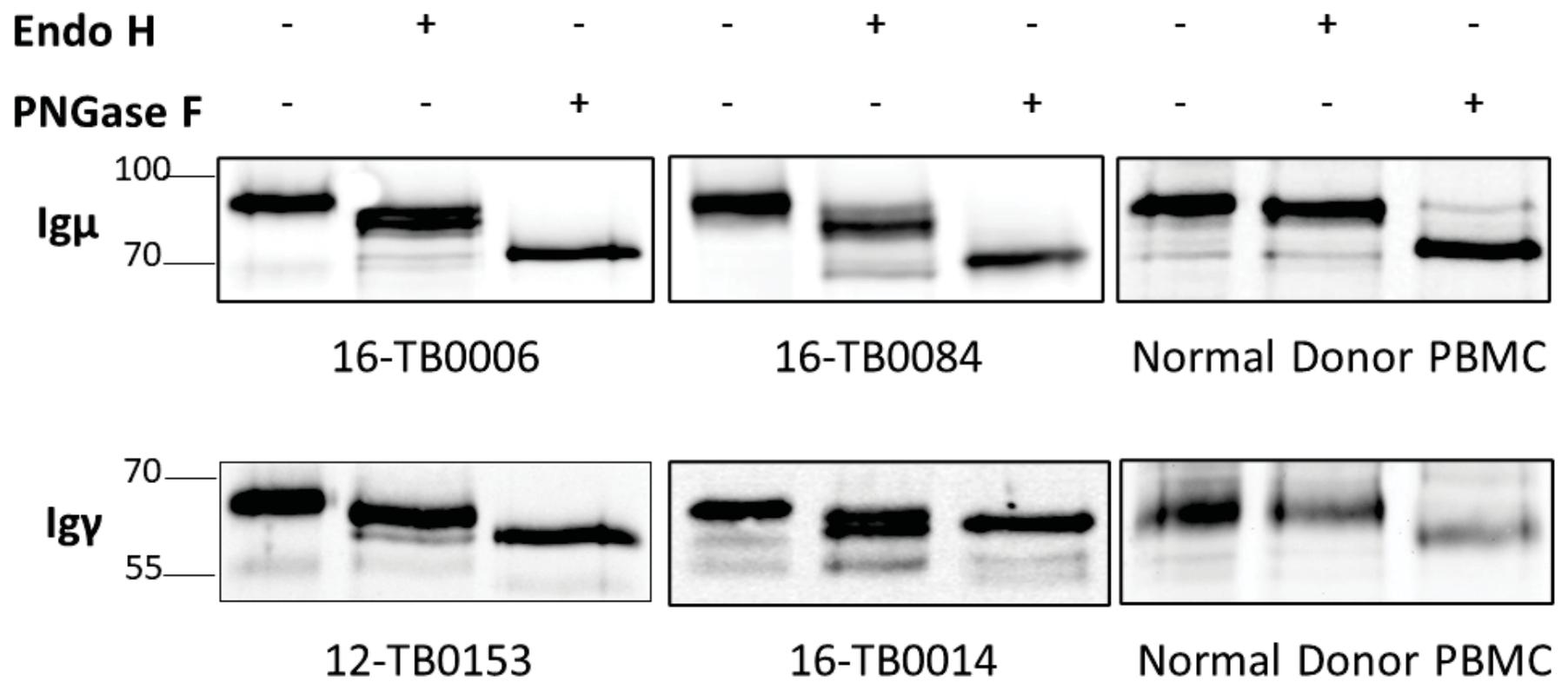
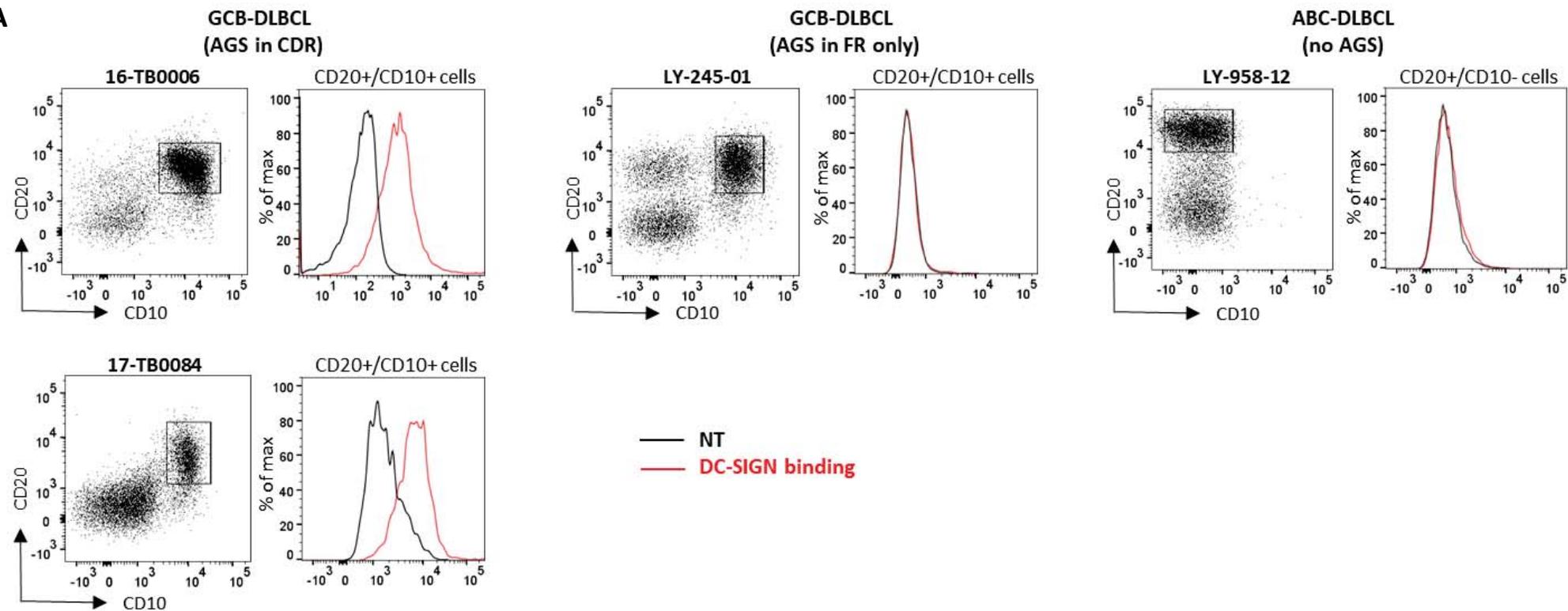


Figure 4

A



B

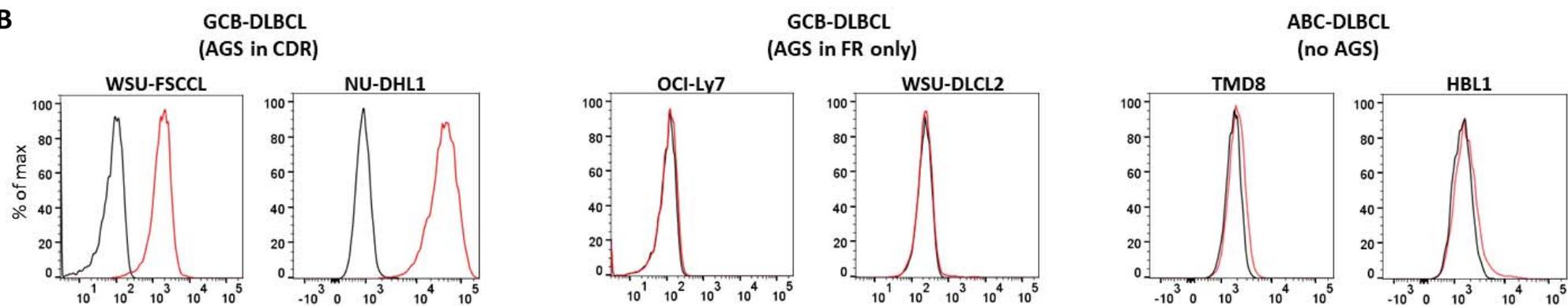


Figure 5

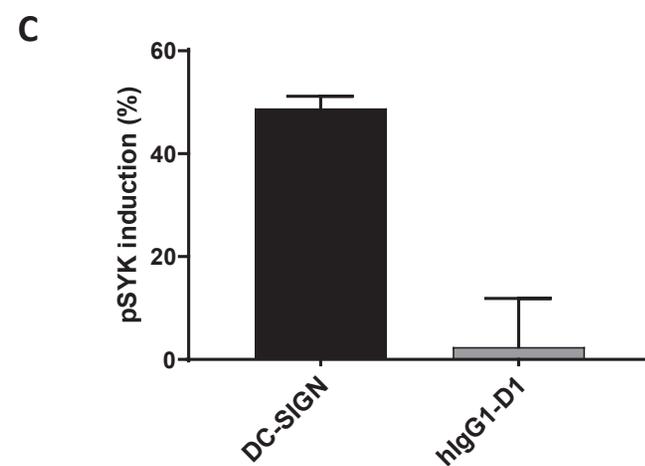
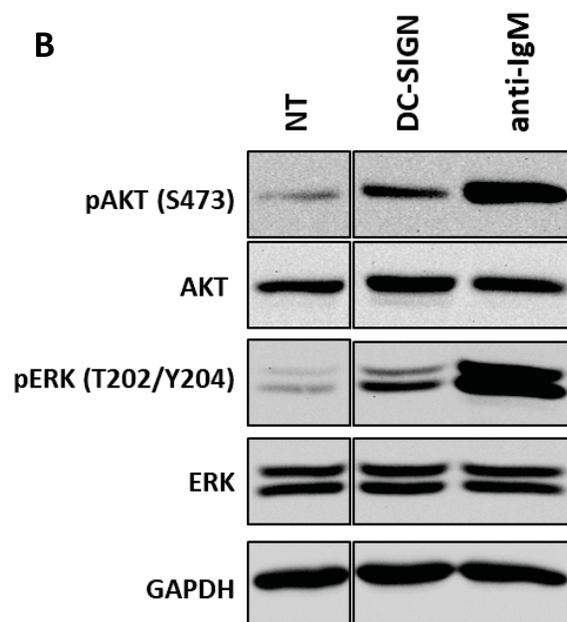
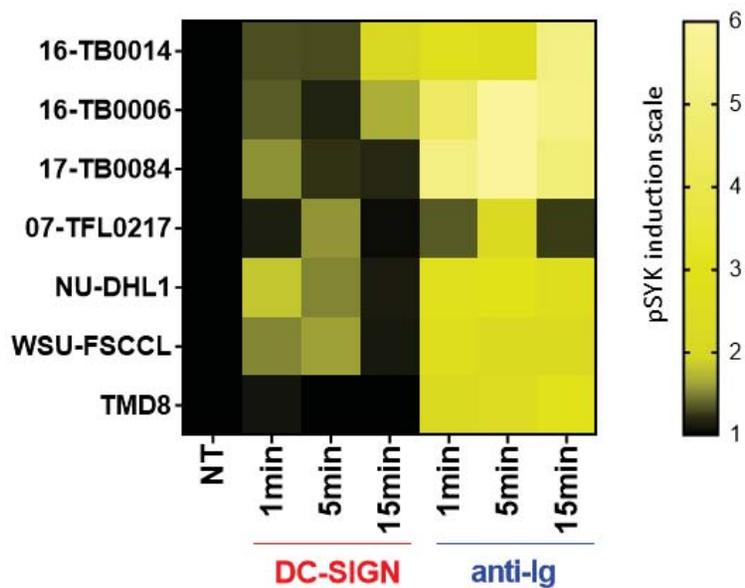
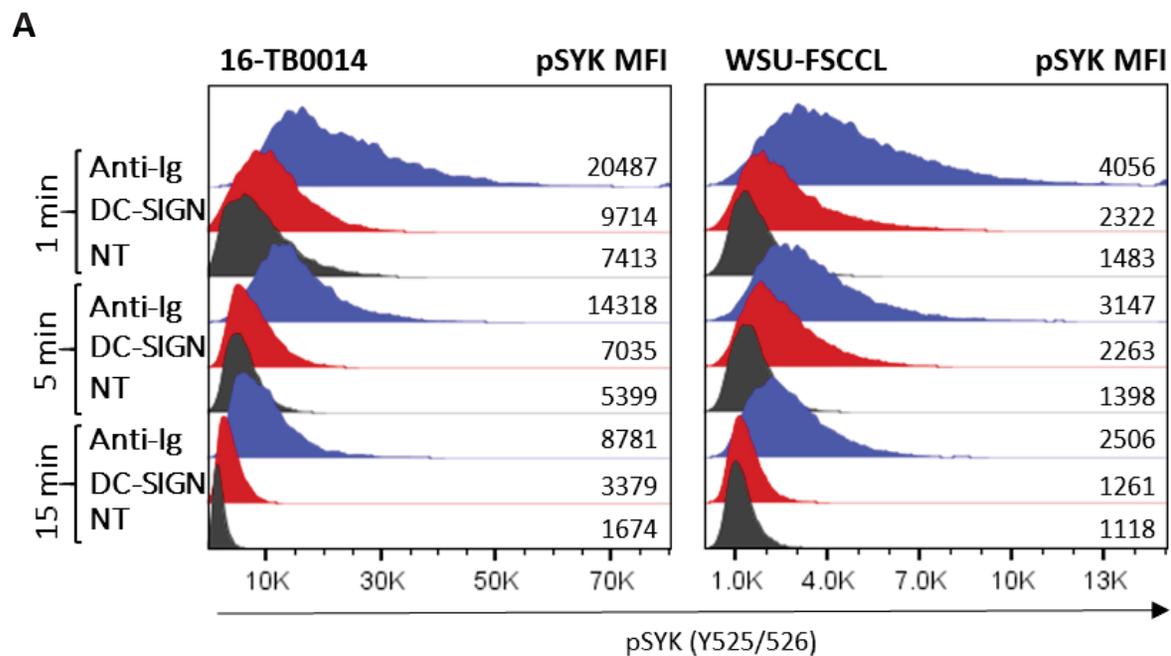


Figure 6

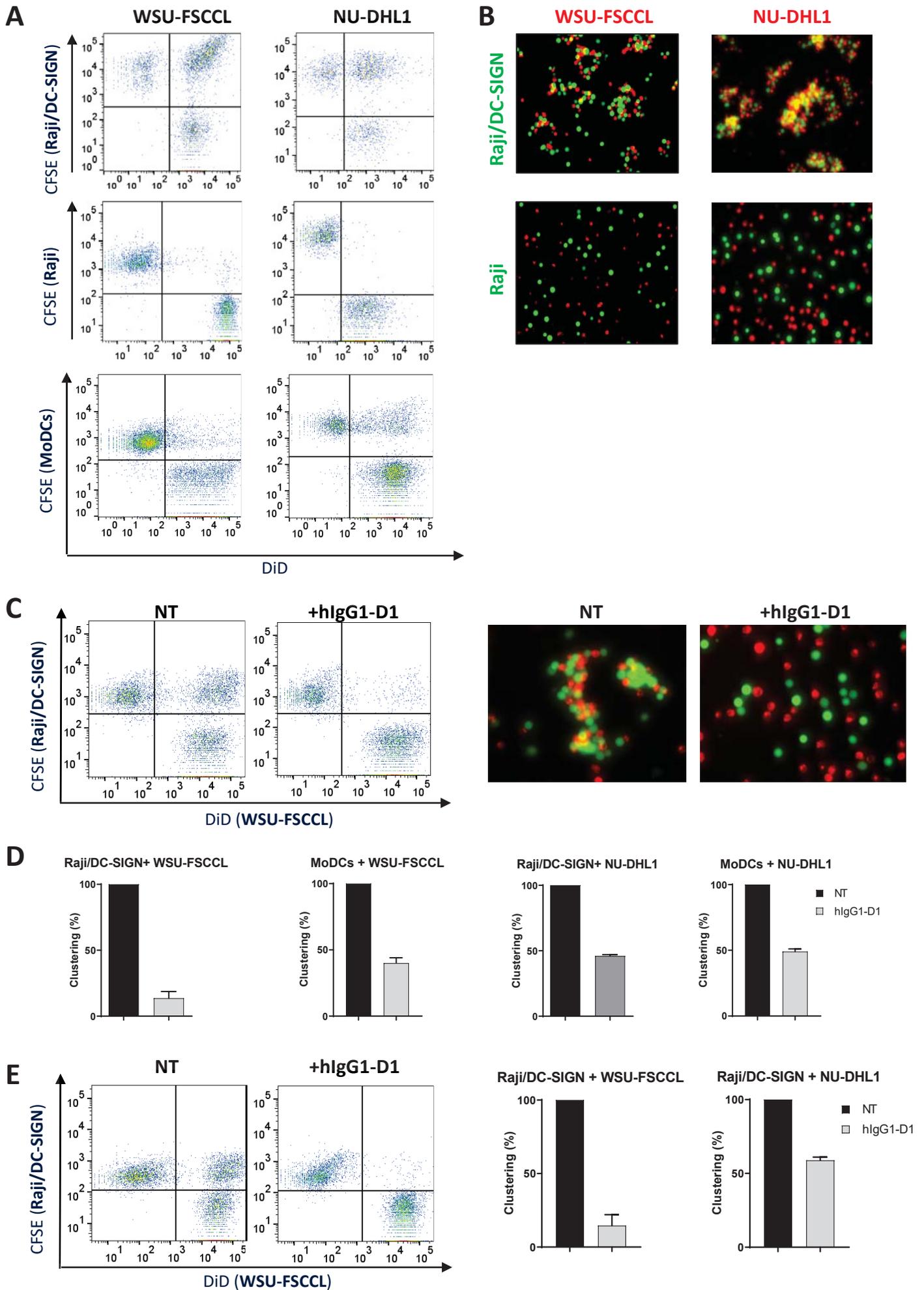
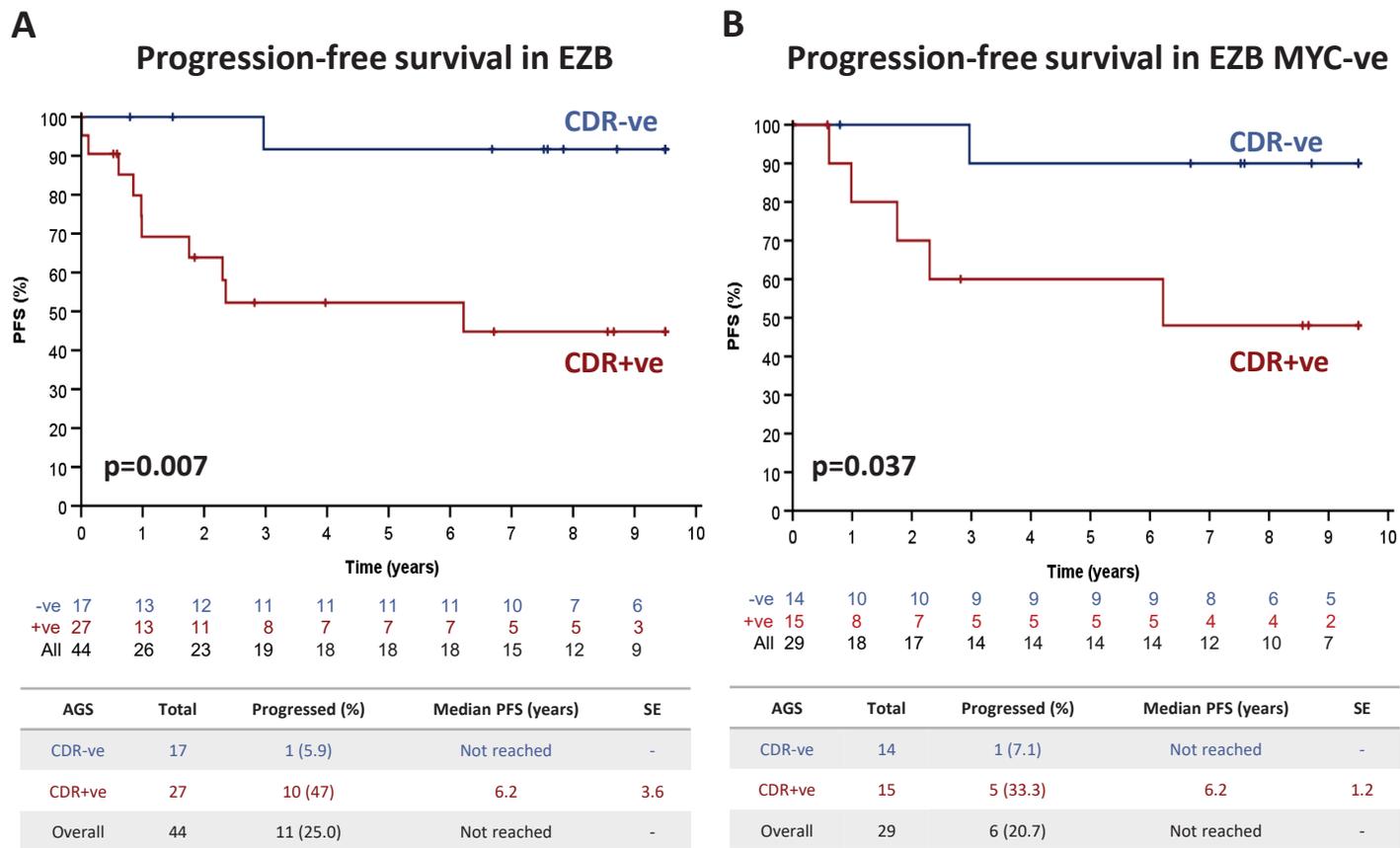
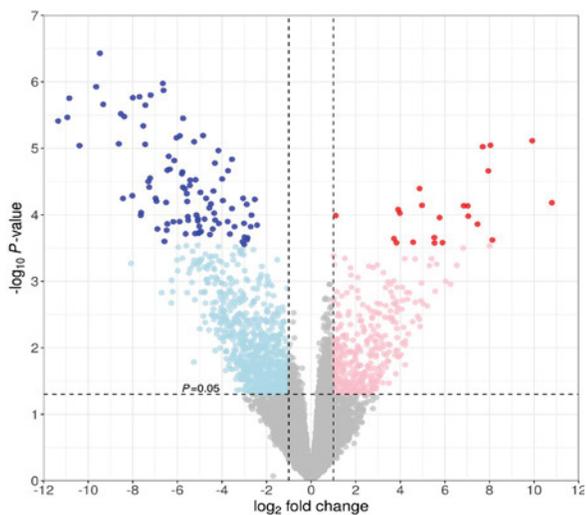


Figure 7



C

Differential gene expression in EZB MYC-ve

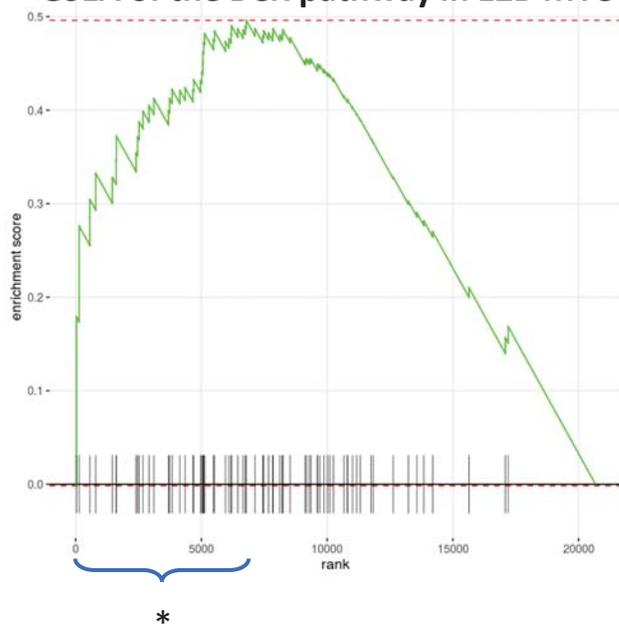


Differentially expressed genes in CDR+ve vs CDR-ve EZB MYC-ve

Upregulated	22
Downregulated	96
Not significant	20556

D

GSEA of the BCR pathway in EZB MYC-ve



*Leading edge genes (from left to right): *CHP2*, *FCGR2B*, *CR2*, *PLCG2*, *CD79A*, *CD22*, *RAC2*, *RELA*, *BTK*, *PIK3AP1*, *PTPN6*, *CD72*, *RASGRP3*, *BLNK*, *NFKBIE*, *VAV1*, *PIK3R1*, *IFITM1*, *NFAT5*, *MAP2K2*, *HRAS*, *INPP5D*, *VAV3*, *IKBKG*, *MAP2K1*, *GSK3B*, *SYK*, *NFATC3*, *CD81*, *MAPK1*, *LYN*, *LILRB3*, *PPP3CA*, *RAF1*, *SOS2* & *RAC1*