blood advances

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

Comparison of MHG and DZsig reveals shared biology and a core overlap group with inferior prognosis in DLBCL

Tracking no: ADV-2023-010673R2

John Davies (University of Leeds, United Kingdom) Laura Hilton (Molecular Biology and Biochemistry, Simon Fraser University, Canada) Aixiang Jiang (BCCRC, Canada) Sharon Barrans (St James's Institute of Oncology, United Kingdom) Catherine Burton (Haematological Malignancy Diagnostic Service, St James University Hospital, United Kingdom) Peter Johnson (University of Southampton, United Kingdom) Andrew Davies (University of Southampton, United Kingdom) Ming-Qing Du (University of Cambridge, United Kingdom) Reuben Tooze (University of Leeds, United Kingdom) Francesco Cucco (Institute of Clinical Physiology (IFC), CNR, Pisa, Italy, Italy) Matthew Care (University of York, United Kingdom) Ryan Morin (Canada's Michael Smith Genome Sciences Centre, BC Cancer, Vancouver, BC, Canada, Canada) Christian Steidl (Department of Pathology and Laboratory Medicine, University of British Columbia, Vancouver, BC, Canada, Canada) Chulin Sha (Institute of Basic Medicine and Cancer, Chinese Academy of Sciences, China) David Westhead (Division of Medical Oncology, Department of Medicine, University of British Columbia, Vancouver, BC, Canada, Canada) David Scott (Division of Medical Oncology, Department of Medicine, University of British Columbia, Vancouver, BC, Canada, Canada)

Abstract:

Conflict of interest: COI declared - see note

COI notes: DWS, AJ, and RDM are named inventors on patents for the use of gene expression to subtypes aggressive B-cell lymphomas, one of which is licensed to NanoString Technologies. CSteidl reports consultancy for AbbVie, Bayer, and Seattle Genetics; research funds from Trillium Therapeutics, BMS, and Epizyme. AJD reports consultancy or honoraria for Celgene, Roche, Kite, Janssen, Acerta Pharma/AstraZeneca, Abbvie, Incyte, and Genmab; research support from Celgene, Roche, Kite, Janssen, Acerta Pharma/AstraZeneca, and Abbvie.

Preprint server: No;

Author contributions and disclosures: Experimental design: JRD, LKH, AJ, DRW, DWS. Data analysis: JRD, LKH, AJ, SB, CSha, MD, RT, MAC. Figure preparation: LKH, JRD. Clinical data collection and management: CB, PWMJ, AJD, FC, DWS Manuscript writing: LKH, JRD, DRW, DWS. All authors contributed to data interpretation and manuscript review and approved the manuscript for submission.

Non-author contributions and disclosures: No;

Agreement to Share Publication-Related Data and Data Sharing Statement: No new sequencing or gene expression data were generated for this study. Data used here may be obtained from the Gene Expression Ombibus database (GSE117556, GSE181063) or at the European Genome-Phenome Archive (EGAS00001005953, EGAS00001002657).

Clinical trial registration information (if any):

Comparison of MHG and DZsig reveals shared biology and a core overlap group with inferior prognosis in DLBCL

John R. Davies[1*], Laura K. Hilton[2, 3*], Aixiang Jiang[2], Sharon Barrans[4], Catherine Burton[4], Peter W.M. Johnson[5], Andrew J. Davies[5], Ming-Qing Du[6], Reuben Tooze[7], Francesco Cucco[6,8], Matthew A. Care[7], Ryan D. Morin[2,3,9], Christian Steidl[2,10], Chulin Sha[11], David R. Westhead[1†], and David W. Scott[2,12†]

[1] School of Molecular and Cellular Biology, Faculty of Biological Sciences, University of Leeds, Leeds, United Kingdom
[2] BC Cancer Centre for Lymphoid Cancer, Vancouver, British Columbia, Canada
[3] Department of Molecular Biology and Biochemistry, Simon Fraser University, Burnaby, British Columbia, Canada
[4] Haematological Malignancy Diagnostic Service, Leeds Cancer Centre, Leeds Teaching Hospitals, Leeds, United Kingdom
[5] School of Cancer Sciences, Faculty of Medicine, University of Southampton, Southampton, United Kingdom
[6] Department of Pathology, University of Cambridge, Cambridge, United Kingdom
[7] Section of Experimental Haematology, University of Leeds, Leeds, United Kingdom
[8] Institute of Clinical Physiology (IFC), CNR, Pisa, Italy
[9] Canada's Michael Smith Genome Sciences Centre, BC Cancer Research Centre, Vancouver, BC, Canada
[10] Department of Pathology and Laboratory Medicine, University of British Columbia, Vancouver, BC, Canada
[11] Institute of Basic Medicine and Cancer, Chinese Academy of Science, Hangzhou, Zhejiang Province, China
[12] Division of Medical Oncology, Department of Medicine, University of British Columbia, Vancouver, BC, Canada

* Equal Contributions
† Co-Corresponding Authors
David R. Westhead, University of Leeds, D.R.Westhead@leeds.ac.uk
David W. Scott, BC Cancer Centre for Lymphoid Cancer, dscott8@bccancer.bc.ca

Running title: Comparing DZsig and MHG classifications in DLBCL

Word Count: 1296
Figures: 2
References: 18

Scientific Category: Lymphoid Neoplasia

No new sequencing or gene expression data were generated for this study. Data used here may be obtained from the Gene Expression Ombibus database (GSE117556, GSE181063) or at the European Genome-Phenome Archive (EGAS00001005953, EGAS00001002657). For other forms of data sharing, contact the corresponding authors: D.R.Westhead@leeds.ac.uk and dscott8@bccancer.bc.ca.

Diffuse large B-cell lymphoma (DLBCL) is a heterogeneous disease identified by morphology, immunophenotype, and a typically aggressive clinical course.[1] DLBCL has long been stratified based on gene expression profiling (GEP) into activated B-cell-like (ABC) and germinal center B-cell-like (GCB) cell-of-origin (COO) subtypes.[2] Recently, several studies have stratified DLBCL into genetic subgroups based on co-occurrence of mutational features, with strong associations with COO.[3-6] Previously, our two groups independently reported gene expression signatures associated with dark zone-like biology in DLBCL. The molecular high-grade signature (MHG) identifies DLBCLs expressing a Burkitt lymphoma (BL)-like GEP signature,[7] while the double-hit signature (since renamed dark zone signature [DZsig][8]) identifies DLBCLs with a GEP signature like high-grade B-cell lymphoma with *MYC* and *BCL2* rearrangement (HGBCL-DH-*BCL2)* (whether the tumors harbor *MYC* and *BCL2* rearrangements or not).[9,10] Remarkably, despite the small overlap in genes that make up each signature, both classifiers identify a subset of DLBCL tumors enriched for certain genetic aberrations including concomitant *MYC* and *BCL2* rearrangements.[7,9]

Here we present analyses that directly compare MHG and DZsig classifications applied to the same data sets, demonstrating that most tumors positive for one signature are positive for both. We evaluate agreement between the two scores in several cohorts and investigate the association of the group of tumors positive for both signatures, 'DZSig&MHG', with outcome. Finally, we compare mutation frequencies in tumors positive for one or both signatures and the association with two DLBCL genetic subgroup classifications. Our results demonstrate clear biological similarity between the DZsig and MHG classifications, and strong associations with the EZB genetic subgroup. All data used here were generated as part of studies reviewed and approved by institutional review boards at each site in accordance with the Declaration of Helsinki.

GEP matrices were obtained for three cohorts: REMoDL-B (N=928; Illumina DASL),[7,11] HMRN (N=1024; Illumina DASL),[12] and DLC (N=304; PolyA-selected RNAseq) (Table S1-2).[9] MHG classifications were generated in each dataset using the BDC classifier.[13] DZsig classification was performed using the PRPS-ST classifier.[14] We limited positive calls for either signature to GCB COO to improve biological coherence, but analyses including the small number of ABC/unclassified tumors positive for these signatures did not produce significantly different conclusions. Mutation data were available for the HMRN (N=515)[5] and DLC (N=302) [9,15] cohorts. Genes covered by targeted capture panels and evaluable in both cohorts are provided in Table S3. The hotspot *EZH2*^{Y646} and *CREBBP* lysine acetyltransferase (KAT) domain missense mutations were treated separately from other mutations in these genes. Fisher's exact test was used to identify associations between binarized mutation status per gene and signature classifications. Multiple tests were corrected with Benjamini-Hochberg using a false discovery rate (FDR) threshold of 0.1 for significance.

Association of the signature classifications with progression-free (PFS) and overall survival (OS) up to five years from diagnosis was evaluated using Kaplan-Meier curves and Cox proportional hazard models adjusted for age and sex, and IPI score. Multiple imputation was used to account for missing data contributing to the IPI score for the HMRN cohort (see supplementary methods). Survival analysis was performed only on patients with *de novo* DLBCL treated with R-CHOP, and patients with transformed disease or who received any other treatment including RB-CHOP were excluded. All statistical analyses were performed in R 4.1 or Stata 14.2.

To directly compare the biology of both signatures within GCB COO, DZsig and MHG classifications were applied to three GEP cohorts (Table S2). A comparison of DZsig score and p(MHG) shows a strong linear correlation between the scores (Fig. S1). In total, 82.4% (range 79.1-84.0% across cohorts) of MHG DLBCLs were also classified as DZsig, while 57.6% (range 55.2-68%) of DZsig DLBCLs were also classified as MHG (Figure 1A-C, Table S4). Of the 28 genes used in MHG classification and 104 genes used in DZsig classification (or 30 genes for the DLBCL90 NanoString assay), only five are used in both classifiers (Table S5). Naïve clustering on the superset of MHG and DLBCL90 DZsig genes consistently identified a cluster enriched for DZsig and MHG positive tumors across all three cohorts (Fig. S2). Together, these data show that the classifiers identify a largely overlapping set of tumors with dark zone biology and consistent GEP features, albeit with the DZSig-only group generally being larger than the MHG-only group.

We next explored the relationship of each classification to mutation prevalence across the DLC and HMDS cohorts. Both cohorts were subjected to targeted sequencing, with 76 genes captured in both (Table S3). Owing to the small number of DZsig and MHG tumors in either cohort alone, mutation data from both cohorts were combined to increase power to detect differentially mutated genes. To confirm the validity of this approach, we subset the pooled data to tumors classified as either ABC or GCB and verified that all genes mutated with at least 10% frequency were not significantly differentially mutated between cohorts (Fig. S3). We then proceeded to compare the mutation frequencies across signature classifications, considering only GCB tumors, using the set of genes mutated in at least 10% of GCB tumors plus targeted genes known to be associated with DZsig, MHG, or BL (Fig. 1D). Comparing the mutation frequencies in tumors within each classification to GCB-DLBCL reveals similar levels of enrichment for many mutation features (Fig. 1E). Several genetic features were consistently significantly enriched across all DZsig, MHG, and DZsig&MHG tumors including *MYC*, *DDX3X*, *FOXO1*, and *EZH2*^{Y646}. In contrast, MHG tumors were significantly enriched for *TP53* mutations, while DZsig tumors were enriched for mutations in *KMT2D*, *MEF2B*, and *RB1*.

When comparing MHG-only or DZsig-only to DZsig&MHG tumors, few comparisons reached statistical significance owing to small group sizes and sparse features (Fig. S4, Table S6). To overcome this limitation, we used published LymphGen classifications (both DLC and HMDS data)[6,16] and HMRN classifications (HMDS only)[5], which consider the tumor's overall mutation profile. Both DZsig&MHG and DZsig-only tumors are significantly enriched for the EZB subgroup relative to GCB and MHG-only tumors (Fig. 1F). No significant associations with other LymphGen classifications were observed. The HMRN BCL2 classification was not significantly enriched in any signature group, which may be due to limited sample size in the HMDS cohort alone (Fig. 1G). Together these data suggest that the overlap DZsig&MHG and DZsig-only groups share significant biology with the EZB genetic subgroup, and the smaller MHG-only group may bring in related cases with *TP53* mutation driven biology.

Both MHG and DZsig are established prognostic biomarkers of poor outcomes, especially in comparison to the relatively favorable GCB-DLBCL group. We compared outcomes of DZsig&MHG, MHG-only and DZsig-only tumors to GCB-DLBCL across all three cohorts. In the DLC data, the DZsig&MHG and DZsig-only, but not MHG-only, were associated with inferior PFS (Fig. 2A-B, Fig. S5A-B), while only DZsig&MHG tumors were associated with inferior OS (Fig. 2C-D, Fig. S5C-D). However, in the HMDS and REMoDL-B data, only the DZsig&MHG and MHG-only were associated with inferior outcomes (Fig. 2E-J, Fig. S5E-J).

These associations were consistent after adjusting for age, sex and/or IPI (Table S7). Although it is intriguing that the association between outcomes and the non-overlap groups are stronger in the cohorts in which they were originally developed, it is important to note that both signatures were developed to distinguish biological entities and were locked prior to un-blinding of clinical outcomes to prevent overfitting the model to outcomes.

Overall, this study demonstrates broad agreement between DZsig and MHG classifications, both in terms of unifying tumor biology and prognostic significance, underscoring the importance of dark zone biology in DLBCL. This study inferred classification across technology platforms – further studies applying both signatures on the same platform(s) may be informative. The inferior outcomes of patients with dark zone DLBCL, even with chimeric antigen receptor T-cell (CAR-T) therapy[17], and the prospect of dose intensification and/or targeted therapy[18], should motivate wide adoption of clinical assays to identify these tumors. The power of genome-wide approaches such as RNAseq should form a critical part of correlative studies in clinical trials allowing exploration of the predictive power of both signatures alongside other biology.

Acknowledgments

The study was funded by Blood Cancer UK grant numbers 15002 and 19009, and the Terry Fox Research Institute (1061, 1043). The Randomized Evaluation of Molecular-Guided Therapy for DLBCL With Bortezomib (REMoDL-B) trial was endorsed by Cancer Research UK, reference number CRUKE/10/024, and Janssen-Cilag provided funding. D.R.W. acknowledges UK Medical Research Council grant MR/L01629X/1 for infrastructure support. D.W.S is supported by a Michael Smith Foundation for Health Research Health Professional Investigator award (18646).

Authorship contributions

Experimental design: JRD, LKH, AJ, DRW, DWS,
Data analysis: JRD, LKH, AJ, SB, CSha, MD, RT, MAC,
Figure preparation: LKH, JRD,
Clinical data collection and management: CB, PWMJ, AJD, FC, DWS,
Manuscript writing: LKH, JRD, DRW, DWS,
All authors contributed to data interpretation and manuscript review and approved the manuscript for submission.

Disclosure of Conflicts of Interest

DWS, AJ, and RDM are named inventors on patents for the use of gene expression to subtypes aggressive B-cell lymphomas, one of which is licensed to NanoString Technologies. CSteidl reports consultancy for AbbVie, Bayer, and Seattle Genetics; research funds from Trillium Therapeutics, BMS, and Epizyme. AJD reports consultancy or honoraria for Celgene, Roche, Kite, Janssen, Acerta Pharma/AstraZeneca, Abbvie, Incyte, and Genmab; research support from Celgene, Roche, Kite, Janssen, Acerta Pharma/AstraZeneca, and Abbvie.

References

1. Sehn LH, Salles G. Diffuse Large B-Cell Lymphoma. *N Engl J Med*. 2021;384(9):842-858.

2. Alizadeh AA, Eisen MB, Davis RE, et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature*. 2000;403(6769):503-511.
3. Schmitz R, Wright GW, Huang DW, et al. Genetics and Pathogenesis of Diffuse Large B-Cell Lymphoma. *N Engl J Med*. 2018;378(15):1396-1407.
4. Chapuy B, Stewart C, Dunford AJ, et al. Molecular subtypes of diffuse large B cell lymphoma are associated with distinct pathogenic mechanisms and outcomes. *Nature medicine*. 2018;24(5):679-690.
5. Lacy SE, Barrans SL, Beer PA, et al. Targeted sequencing in DLBCL, molecular subtypes, and outcomes: a Haematological Malignancy Research Network report. *Blood*. 2020;135(20):1759-1771.
6. 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(4):551-568.e514.
7. Sha C, Barrans S, Cucco F, et al. Molecular High-Grade B-Cell Lymphoma: Defining a Poor-Risk Group That Requires Different Approaches to Therapy. *J Clin Oncol*. 2018;37(3):202-212.
8. Alduaij W, Collinge B, Ben-Neriah S, et al. Molecular determinants of clinical outcomes in a real-world diffuse large B-cell lymphoma population. *Blood*. 2023;141(20):2493-2507.
9. 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(3):190-201.
10. Hilton LK, Tang J, Ben-Neriah S, et al. The double-hit signature identifies double-hit diffuse large B-cell lymphoma with genetic events cryptic to FISH. *Blood*. 2019;134(18):1528-1532.
11. Davies A, Cummin TE, Barrans S, et al. Gene-expression profiling of bortezomib added to standard chemoimmunotherapy for diffuse large B-cell lymphoma (REMoDL-B): an open-label, randomised, phase 3 trial. *Lancet Oncol*. 2019;20(5):649-662.
12. Painter D, Barrans S, Lacy S, et al. Cell-of-origin in diffuse large B-cell lymphoma: findings from the UK's population-based Haematological Malignancy Research Network. *Br J Haematol*. 2019;185(4):781-784.
13. Sha C, Barrans S, Care MA, et al. Transferring genomics to the clinic: distinguishing Burkitt and diffuse large B cell lymphomas. *Genome Medicine*. 2015;7(1):64.
14. Jiang A, Hilton LK, Tang J, et al. PRPS-ST: A protocol-agnostic self-training method for gene expression-based classification of blood cancers. *Blood Cancer Discov*. 2020;1(3):244-257.
15. Arthur SE, Jiang A, Grande BM, et al. Genome-wide discovery of somatic regulatory variants in diffuse large B-cell lymphoma. *Nat Commun*. 2018;9(1):4001.
16. Runge HFP, Lacy S, Barrans S, et al. Application of the LymphGen classification tool to 928 clinically and genetically-characterised cases of diffuse large B cell lymphoma (DLBCL). *Br J Haematol*. 2021;192(1):216-220.
17. Olson NE, Ragan SP, Reiss DJ, et al. Exploration of Tumor Biopsy Gene Signatures to Understand the Role of the Tumor Microenvironment in Outcomes to Lisocabtagene Maraleucel. *Mol Cancer Ther*. 2023;22(3):406-418.
18. Davies AJ, Barrans S, Stanton L, et al. Differential Efficacy From the Addition of Bortezomib to R-CHOP in Diffuse Large B-Cell Lymphoma According to the Molecular Subgroup in the REMoDL-B Study With a 5-Year Follow-Up. *J Clin Oncol*. 2023;41(15):2718-2723.

Figure Legends

Figure 1. Mutational spectra of dark zone DLBCL. A-C. Venn diagrams showing the overlap in cases classified as MHG and DZsig across the DLC (**A**), HMDS (**B**), and REMoDL-B (**C**) cohorts. Percentages are relative to the total number of cases classified as GCB-DLBCL in each cohort, and the size of each circle is proportional to the number of cases it represents. Numbers outside the circles indicate GCB-DLBCL not classified as either MHG or DZsig. **D.** The mutational profile of GCB COO tumors across genes mutated in at least 10% of GCB tumors, stratified according to GEP classifications and annotated according to published LymphGen or HMRN genetic classification and cohort. **E.** Enrichment for coding or hotspot mutations among the genes shown in A. Positive values in the forest plot (left panel) indicate enrichment relative to GCB-DLBCL negative for the signatures. The percentage of tumors within each group harboring one or more coding mutation per gene is shown in the right panel. MHG includes all MHG-classified tumors (MHG and DZsig&MHG), and DZsig includes all DZsig tumors (DZsig and DZsig&MHG). * Hotspot mutations (see Methods). **F.** The proportion of tumors belonging to each GEP classification groups classified as EZB by the LymphGen classifier across the DLC and HMDS cohorts, separating MHG, DZsig, and DZsig&MHG tumors. ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$. **G.** The percentage of tumors belonging to each GEP classification group classified as BCL2 by the HMRN classifier, separating MHG, DZsig, and DZsig&MHG tumors. No comparisons reached statistical significance.

Figure 2. Outcomes in dark zone DLBCL. Left panels show Kaplan-Meier curves indicating the survival probability across the GEP classification subgroups, including tables of numbers at risk below. Right panels show forest plots indicating the unadjusted hazard ratio and 95% confidence of each dark zone group relative to GCB-DLBCL. **A, B.** PFS and **C, D.** OS in the DLC cohort. **E, F.** OS in the HMDS cohort. **G, H.** PFS and **I, J.** OS in the REMoDL-B cohort. All survival analyses were performed in patients with *de novo* DLBCL treated with R-CHOP.

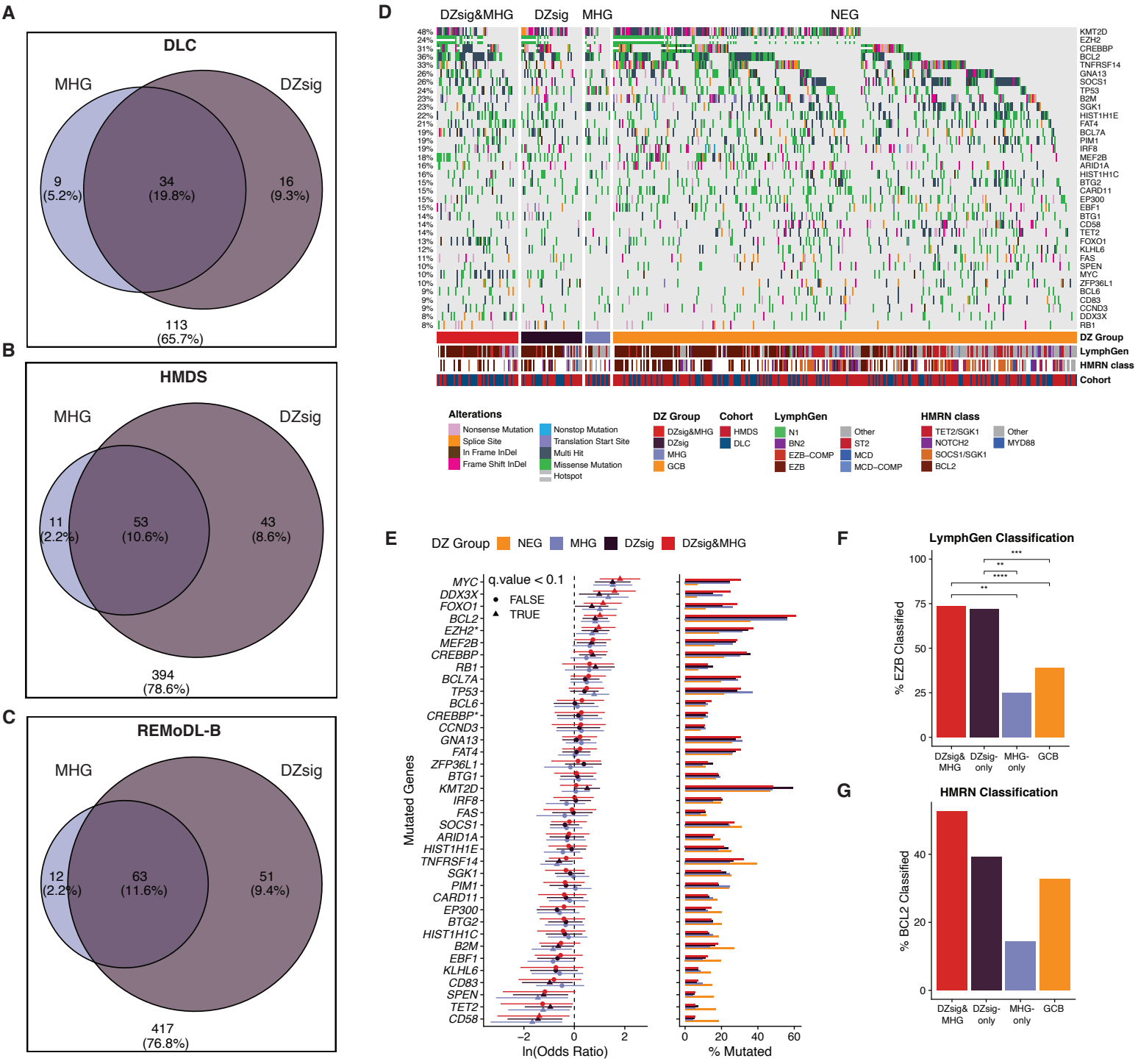

Figure 2