**Novel genetic subgroups inform on shared pathobiology within adult and pediatric Burkitt lymphoma**

Nicole Thomas1, Kostiantyn Dreval1, Daniela S. Gerhard2, Laura K. Hilton3, Jeremy S. Abramson4, Nancy L. Bartlett5, Jeffrey Bethony6, Jay Bowen7, Anthony C. Bryan7, Corey Casper8, Manuela Cruz1, Maureen Dyer9, Julie M. Gastier-Foster7,10, Alina Gerrie3, Bruno M. Grande11, Timothy Greiner12, Nicholas B. Griner2, Thomas G. Gross13, Nancy Lee Harris14, John D. Irvin15, Elaine S. Jaffe16, Fabio Leal17, Jean Paul Martin15, Marie-Reine Martin15, Sam M. Mbulaiteye18, Charles G. Mullighan19, Andrew J. Mungall20, Karen Mungall20, Constance Namirembe21, Ariela Noy22, Martin D. Ogwang23, Jackson Orem21, German Ott24, Hilary Petrello7, Steven J. Reynolds25, Shaghayegh Soudi1, Steven H. Swerdlow26, Alexandra Traverse-Glehen27, Wyndham H. Wilson28, Jasper Wong3, Marco A. Marra20,29, Louis M. Staudt28, David W. Scott3 & Ryan D. Morin1,20

1Department of Molecular Biology and Biochemistry, Simon Fraser University, Burnaby, BC, Canada.

2Office of Cancer Genomics, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA.

3Centre for Lymphoid Cancer, BC Cancer, Vancouver, BC, Canada.

4Center for Lymphoma, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA.

5Washington University School of Medicine, St. Louis, MO, USA.

6George Washington University, Washington, DC, USA.

7Biopathology Center, Nationwide Children's Hospital, Columbus, OH, USA.

8Infectious Disease Research Institute, Seattle, WA, USA.

9Clinical Research Directorate, Frederick National Laboratory for Cancer Research sponsored by the National Cancer Institute, Frederick, MD, USA.

10Departments of Pathology and Pediatrics, The Ohio State University, Columbus, OH, USA.

11Sage Bionetworks, Seattle, WA, USA.

12Department of Pathology and Microbiology, University of Nebraska Medical Center, Omaha, NE, USA.

13Center for Global Health, National Cancer Institute, National Institutes of Health, Rockville, MD, USA.

14Department of Pathology, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA.

15Foundation for Burkitt Lymphoma Research, Geneva, Switzerland.

16Laboratory of Pathology, Clinical Center, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA.

17Programa de Oncovirologia, Instituto Nacional de Cancer Jose de Alencar, Rio de Janeiro, Brazil.

18Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Rockville, MD, USA.

19Department of Pathology, St. Jude Children's Research Hospital, Memphis, TN, USA.

20Canada's Michael Smith Genome Sciences Centre at BC Cancer, Vancouver, BC, Canada.

21Uganda Cancer Institute, Kampala, Uganda.

22Memorial Sloan Kettering Cancer Center and Weill Cornell Medical College, New York, NY, USA.

23EMBLEM Study, St. Mary's Hospital Lacor, Gulu, Uganda.

24Department of Clinical Pathology, Robert-Bosch-Krankenhaus and Dr. Margarete Fischer-Bosch Institute of Clinical Pharmacology, Department of Clinical Pathology, Stuttgart, Germany.

25Division of Intramural Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD, USA.

26Division of Hematopathology, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania, USA.

27Hospices Civils de Lyon, Université Lyon 1, Service d'Anatomie Pathologique, Hopital Lyon Sud France.

28Lymphoid Malignancies Branch, Center for Cancer Research, National Cancer Institute, Bethesda, MD, USA.

29Department of Medical Genetics, University of British Columbia, Vancouver, BC, Canada.

*Introduction:* Burkitt lymphoma (BL) accounts for approximately 50% of all pediatric non-Hodgkin lymphomas compared to 1-2% in adults. Adult BL (aBL) remains a poorly understood entity and its relationship to pediatric BL (pBL) and to DLBCL has not been fully elucidated. The variable treatment outcomes between these entities necessitate a more thorough understanding of the genetic and molecular features underlying their biology to enable better prognostication and more effective treatments. We sought to comprehensively determine genetic features shared with DLBCL and those that are unique to BL, to further delineate genetic subgroupings within each entity.

*Methods:* Samples for this study were collected through the Burkitt Lymphoma Genome Sequencing Project (BLGSP). We sequenced the tumor genomes of 139 pBL and 92 aBL, consisting of both EBV-positive (EBV+) and EBV-negative (EBV-) BLs, and compared these to the genomes of 252 DLBCL patients. All cases were analyzed for simple somatic mutations (SSM), recurrent copy number variations (CNV), structural variations (SV), aberrant somatic hypermutation (aSHM), and SSM hotspots. Mutations were used as features for the identification of genetic subgroups using non-negative matrix factorization (NMF) clustering.

*Results:* Clustering of BL and DLBCL revealed six distinct genetic subgroups (Figure 1) with three primarily representing DLBCLs (DLBCL-1, DLBCL-2, and DLBCL-3) and three predominantly comprising BLs (M53-BL, IC-BL, and DGG-BL). The DLBCL-predominant subgroups partially overlapped with those previously described and resembled features of EZB and ST2-like subgroups. The frequency of aBLs within these subgroups was higher than that of pBL patients (*p*=0.0005). The new cluster M53-BL consists of both pBL (9/27) and aBL (13/27) samples and is characterized by the highest prevalence of mutations in *TP53* accompanied by the paucity of other driver mutations but without the aneuploidy associated with the A53 subgroup described in DLBCL. Enrichment of EBV- samples in this cluster further corroborate our previous findings of *TP53* mutations being associated with EBV- BL. IC-BL is characterized by mutations in *ID3*, *CCND3*, and *SMARCA4*. In contrast, DGG-BL, where 65% of the cluster consisted of EBV+ BL samples, had mutations in *DDX3X*, *GNA13*, and *GNAI2*. Using a linear model, we compared the rates of aSHM in BL genomes from all clusters and identified the DLBCL-3 cluster to harbor the highest aSHM rates at common sites while the M53-BL cluster harbored the lowest rates. To further establish the biological basis of unique clusters within BL, we conducted differential gene expression analyses between the two major BL genetic subgroups, DGG-BL and IC-BL. We identified a total of 86 differentially expressed genes between the two clusters (p.adj < 0.01 and |log2foldChange| > 1). Among the genes with the strongest differential expression were *IRF4*, *SERPINA9*, and *TNFRSF13B.* Each of these are notable as their expression is a component of the DLBCL cell-of-origin and double-hit signature classifiers. Further, we found *IRF4* expression to be one of the strongest predictors of cluster membership, with high *IRF4* expression associated with IC-BL membership. Using TP53 and ID3 mutations as a proxy for M53-BL and IC-BL clusters in aBL, we found mutations in TP53 to be associated with significantly inferior progression free survival (PFS) at 2 year follow up, while mutations in ID3 were associated with overall better PFS at 2 year follow up.

*Conclusion*: This work identifies novel genetic subgroups within BL with characteristic genetic and gene expression differences and some bearing relationship to DLBCL subgroups. The three subgroups with predominantly BL samples (DGG-BL, IC-BL, and M53-BL) each comprised a mixture of aBL and pBL samples, confirming similar molecular features in these entities. The IC-BL cluster is associated with mutations in *ID3* and *CCND3*, high *IRF4* expression, and ID3 mutated cases exhibited significantly better outcomes. M53-BL is associated with *TP53* mutations and inferior PFS in aBL, representing a subset of patients to be considered for novel treatment approaches. These findings highlight shared pathogenesis between aBL and pBL and establish genetic subtypes within BL that delineate cases with distinct molecular and clinical features. This provides a new framework for new diagnostic and therapeutic strategies.



Figure 1. Overview of genetic subgroups of Burkitt lymphoma.