

FROM ADULT PATIENTS ANALYZED IN THE ICGC MMML-SEQ CONSORTIUM

D. Huebschmann^{1*} | K. Kleinheinz¹ | R. Wagener² |
H. Kretzmer³ | U.H. Toprak¹ | S.H. Bernhart³ |
C. Lopez Gonzales² | M. Kreuz⁴ | R. Eils¹ | M. Hansmann⁵ |
S. Hoffmann³ | M. Hummel⁶ | W. Klapper⁷ | C. Lawerenz¹ |
M. Loeffler⁴ | P. Möller⁸ | J. Richter⁷ | P. Rosenstiel⁹ |
A. Rosenwald¹⁰ | S. Stilgenbauer¹¹ | M. Weniger¹² |
L. Trümper¹³ | R. Küppers¹² | M. Schlesner¹ | R. Siebert²

¹Theoretical Bioinformatics, German Cancer Research Center (DKFZ), Heidelberg, Germany; ²Medicine, Institute of Human Genetics, Ulm, Germany; ³University Leipzig, Leipzig Research Center for Civilization Diseases, Leipzig, Germany; ⁴Leipzig University, Institute for Medical Informatics Statistics and Epidemiology, Leipzig, Germany; ⁵University of Frankfurt Medical School, Senckenberg Institute of Pathology, Frankfurt am Main, Germany; ⁶University Medicine Berlin, Institute of Pathology, Charité, Berlin, Germany; ⁷University Hospital Schleswig-Holstein Campus Kiel, Section of Hematopathology Institute of Pathology, Kiel, Germany; ⁸Medical Faculty of the Ulm University, Institute of Pathology, Ulm, Germany; ⁹University Hospital Schleswig-Holstein Campus Kiel, Institute of Clinical Molecular Biology, Kiel, Germany; ¹⁰University of Wuerzburg, institute of Pathology, Würzburg, Germany; ¹¹University Medical Center Ulm, Department for Internal Medicine III, Hematology, Oncology and Rheumatology and infectious diseases, Ulm, Germany; ¹²University of Duisburg-Essen, Medical School, Institute of Cell Biology (Cancer Research), Essen, Germany; ¹³Georg August University of Göttingen, Department of Hematology and Oncology, Göttingen, Germany

Introduction: The ICGC MMML-seq consortium aims at a precise characterization of germinal center derived B-cell lymphomas (gcBCL). Mutational signatures are patterns of single nucleotide variants (SNVs) taking into account the motif context. 30 mutational signatures had previously been extracted from a cross-entity data set, half of which could be attributed to mutational mechanisms. The goal of this work was to identify mutational mechanisms active in gcBCL and link these to B-cell biology.

Methods: Matched tumor normal control pairs of gcBCL (76 diffuse large B-cell lymphomas (DLBCL), 85 follicular lymphomas (FL), 16 FL/DLBCLs, two double hit lymphomas, and one B-cell lymphoma not otherwise specified (B-NOS)) from adult patients were analyzed by whole genome sequencing. SNVs were called with the DKFZ inhouse pipeline. An unsupervised analysis of mutational signatures was performed with non-negative matrix factorization. This analysis was complemented by a supervised analysis of mutational signatures using non-negative least squares and thereby enabling the extraction of enrichment and depletion patterns of the mutational signatures. Clusters at different mutation density were extracted based on intermutation distance.

Results: Clusters of mutations at different levels of mutation density were extracted: Kataegis (rainfalls, high mutation density at single sample level) and Psichales (intermediate mutation density at single sample level). Genomic regions affected by the respective processes recurrently across the cohort were called regions of interest (ROIs).

253 Psichales-ROIs were identified and were enriched in late replicating regions of the genome. 166 Kataegis-ROIs were identified, 42/64/17/4 of which were known targets of aberrant somatic hypermutation (SHM) / were located within the immunoglobulin (IG) loci / overlapped with lymphoma-associated genes / overlapped with cancer genes known from other entities, respectively. Kataegis-ROIs were enriched in early replicating regions of the genome. In an analysis of mutational signatures, 11 known signatures including clocklike signatures (e.g. spontaneous deamination), DNA repair defect signatures, an APOBEC signature and the B-cell specific signature AC9 (attributed to AID and polymerase η) were found. Furthermore, we discovered three new signatures (L1 – L3). L1 was enriched at the IG loci. L2 was specifically enriched in the constant domains of the IGH locus.

Conclusions: L1 and L2 may be interpreted as an imprint of the action of AID on the genome, L2 with a high amount of modulation by altered repair pathways, L1 with a lower amount of modulation. Both L1 and L2 contribute to SHM, whereas class switch recombination (CSR) may be explained with only L1. Kataegis clusters may be classified into two groups, one of which is attributable to aberrant SHM, the other to dysregulated CSR.

Keywords: B-cell lymphoma; germinal center (GC); immunoglobulins (Ig).