

HeeJin Cheon,^{1,3} Jeffrey C. Xing,^{2,3} Vivian W. Chan,^{2,3} David S. Chung,^{2,3} Mariella F. Toro,^{2,3} Cait E. Hamele,^{2,3} Ross C. Hardison,⁴ Thomas L. Olson,^{2,3} David J. Feith,^{2,3} Aakrosh Ratan,^{5,6} and Thomas P. Loughran, Jr.^{2,3}

¹Department of Biochemistry and Molecular Genetics, University of Virginia School of Medicine, Charlottesville, VA; ²University of Virginia Cancer Center, Charlottesville, VA, Department of Medicine; ³Division of Hematology/Oncology, University of Virginia School of Medicine, Charlottesville, VA; ⁴Department of Biochemistry and Molecular Biology, Center for Computational Biology & Bioinformatics, The Pennsylvania State University, State College, PA; ⁵Center for Public Health Genomics, University of Virginia, Charlottesville, VA; and ⁶Department of Public Health Sciences, University of Virginia School of Medicine, Charlottesville VA

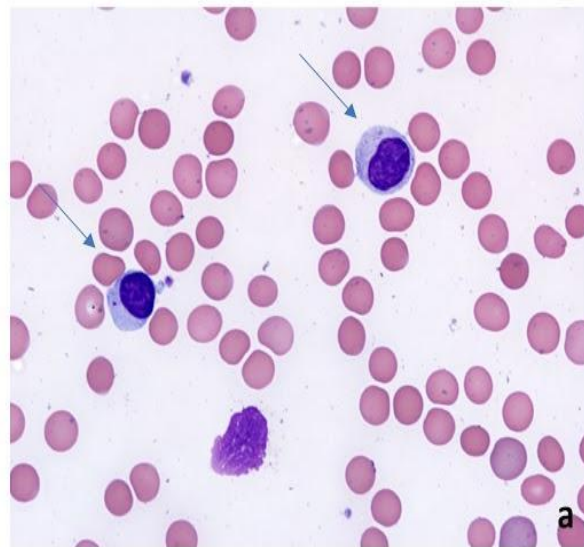
Abstract

Large granular lymphocyte (LGL) leukemia comprises a group of rare lymphoproliferative disorders whose molecular landscape is incompletely defined. We leveraged paired whole exome and transcriptome sequencing in the largest LGL leukemia cohort to date, which included 105 patients (93 TCRαβ T-LGL and 12 TCRγδ T-LGL). 76 mutations were observed in three or more patients in the cohort, and out of those, *STAT3*, *KMT2D*, *PIK3R1*, *TTN*, *EYS*, and *SULF1* mutations were shared between both subtypes. We identified *ARHGAP25*, *ABCC9*, *PCDHA11*, *SULF1*, *SLC6A15*, *DDX59*, *DNMT3A*, *FAS*, *KDM6A*, *KMT2D*, *PIK3R1*, *STAT3*, *STAT5B*, *TET2*, and *TNFAIP3* as recurrently mutated putative drivers. Pathways recurrently altered by non-silent mutations in the cohort included JAK-STAT signaling, chemokine signaling, and cytokine-cytokine receptor interaction. Hotspot mutations in *STAT3*, *PIK3R1*, and *FAS* were detected, whereas truncating mutations in epigenetic modifying enzymes such as *KMT2D* and *TET2* were observed. Moreover, *STAT3* mutations co-occurred with mutations in chromatin and epigenetic modifying genes, especially *KMT2D* and *SETD1B* ($p < 0.01$, $p < 0.05$ respectively). *STAT3* was mutated in 50.5% of the patients. Most common Y640F *STAT3* mutation was associated with lower ANC values and N647I mutation was associated with lower hemoglobin values. *STAT3* mutant patients exhibited increased non-silent mutational burden and enrichment of a mutational signature associated with increased spontaneous deamination of 5-methylcytosine. Finally, gene expression analysis revealed enrichment of interferon-gamma signaling for *STAT3* mutant patients, while increased PI3K-Akt signaling was observed for *STAT3* wild-type patients. These findings highlight the clinical and molecular heterogeneity of this rare disorder.

Background

Large granular lymphocyte (LGL) leukemia is a rare hematologic malignancy

- Often diagnosed in elderly population, with more than 75% of patients over the age of 50
- LGL leukemia subtypes
 - T-LGL (major) CD3+ TCR αβ
 - Less frequent CD3+ TCR γδ (GD)
 - Chronic NK-LGL
- Aggressive NK cell disease
- Neutropenia and anemia are major sympto
- Associated with autoimmune diseases
 - Rheumatoid arthritis
- Treatment
 - Immunosuppressive agents
 - No cure



Somatic *STAT3* mutations are found in ~30% of LGL leukemia patients

- Since the discovery of *STAT3* mutations, three small-scale whole exome sequencing (WES) studies have been reported for 16 T-LGL and three NK-LGL, four NK-LGL, and 13 NK-LGL cases along with our report of seven whole genomes in NK-LGL with a main emphasis on *TET2* biology.

Objectives

We completed the most extensive paired exome and transcriptome study of LGL leukemia to date with 105 patients, composed of 93 T-LGL and 12 GD-LGL patients.

Results

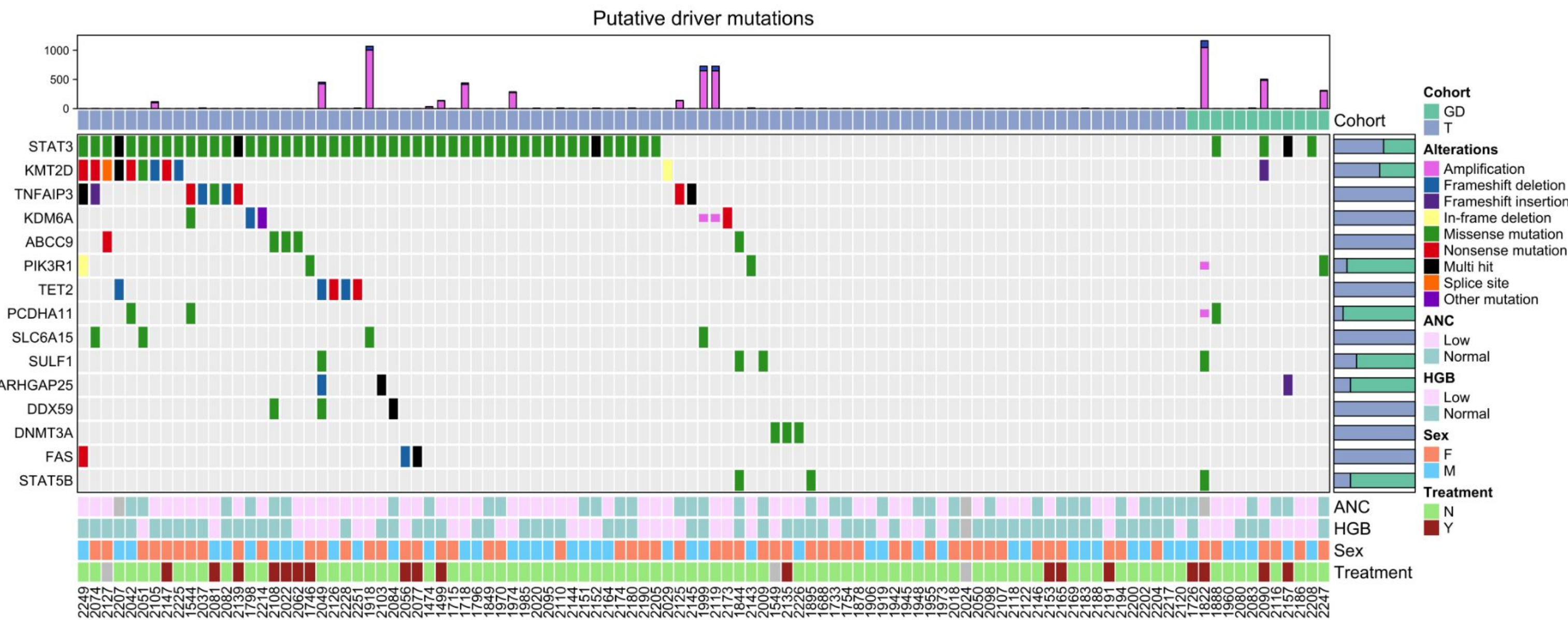


Figure 1. Onco-plot of putative drivers in LGL leukemia. Seven different driver analysis tools identified 15 gene variants with putative driver roles present in at least three LGL leukemia samples. Each row represents one of the 15 putative drivers, whereas the columns represent individual patients. LGL disease cohorts are color coded at the top. Type of alteration in putative driver genes is indicated by the colored boxes, which are defined under “Alterations,” and gray indicates the wild type. The top bar graph shows the number of amplification and deletion events. Right bar graph shows the distribution of variants for a given gene across the two LGL cohorts. Bottom annotations represent clinical phenotypes of individual patients, with color definitions found in the rightmost annotation. ANC values above 1.5k/μl and HGB values above 12 are annotated as ‘Normal’ and anything below is ‘Low.’ Treatment indicates if patients were on LGL treatment (methotrexate, cyclophosphamide, or cyclosporine) at the time of sample acquisition. Unknown values are colored in grey.

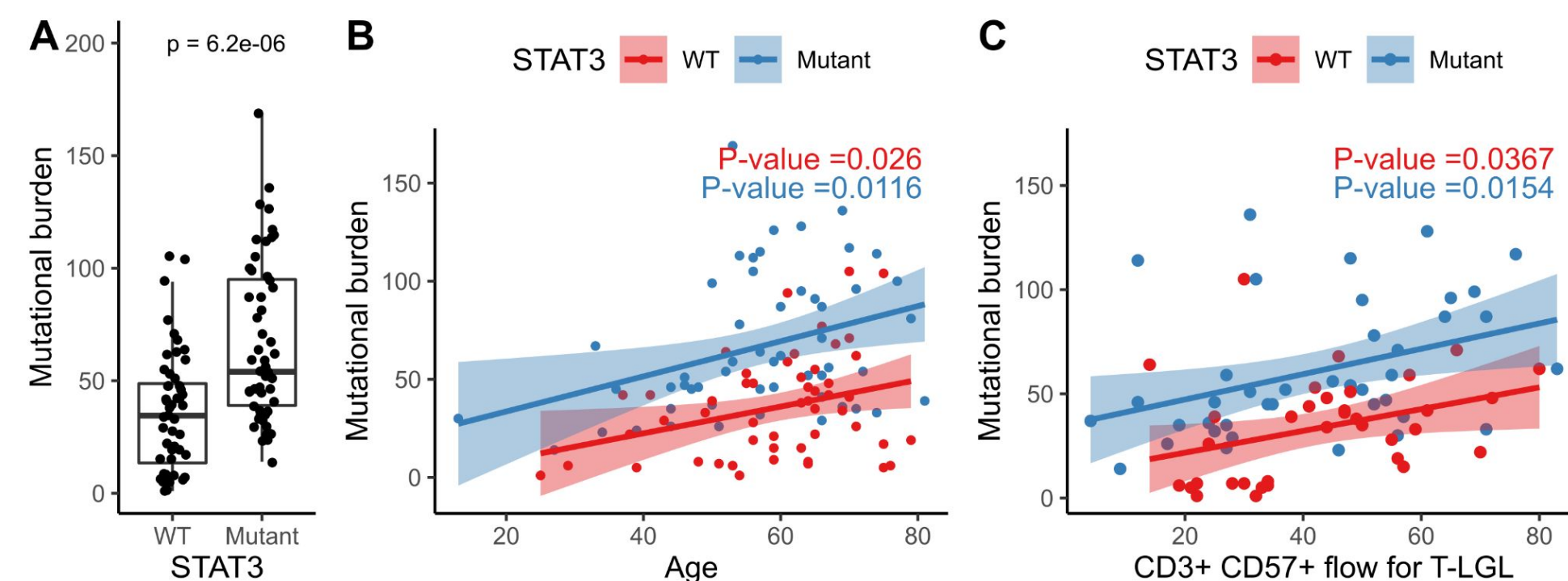


Figure 2. Increased mutational burden in *STAT3* mutated patients. (A) Boxplot of non-silent mutational burden comparing *STAT3* WT against mutant patients. P-value from Wilcoxon t-test. (B) A plot of the mutational burden against age for WT and mutant *STAT3* patients with linear regression, and colored annotation representing *STAT3* mutation status. (C) Plot of mutation burden against CD3+ CD57+ fraction for T-LGL patients with linear regression, and colored annotation representing *STAT3* mutation status.

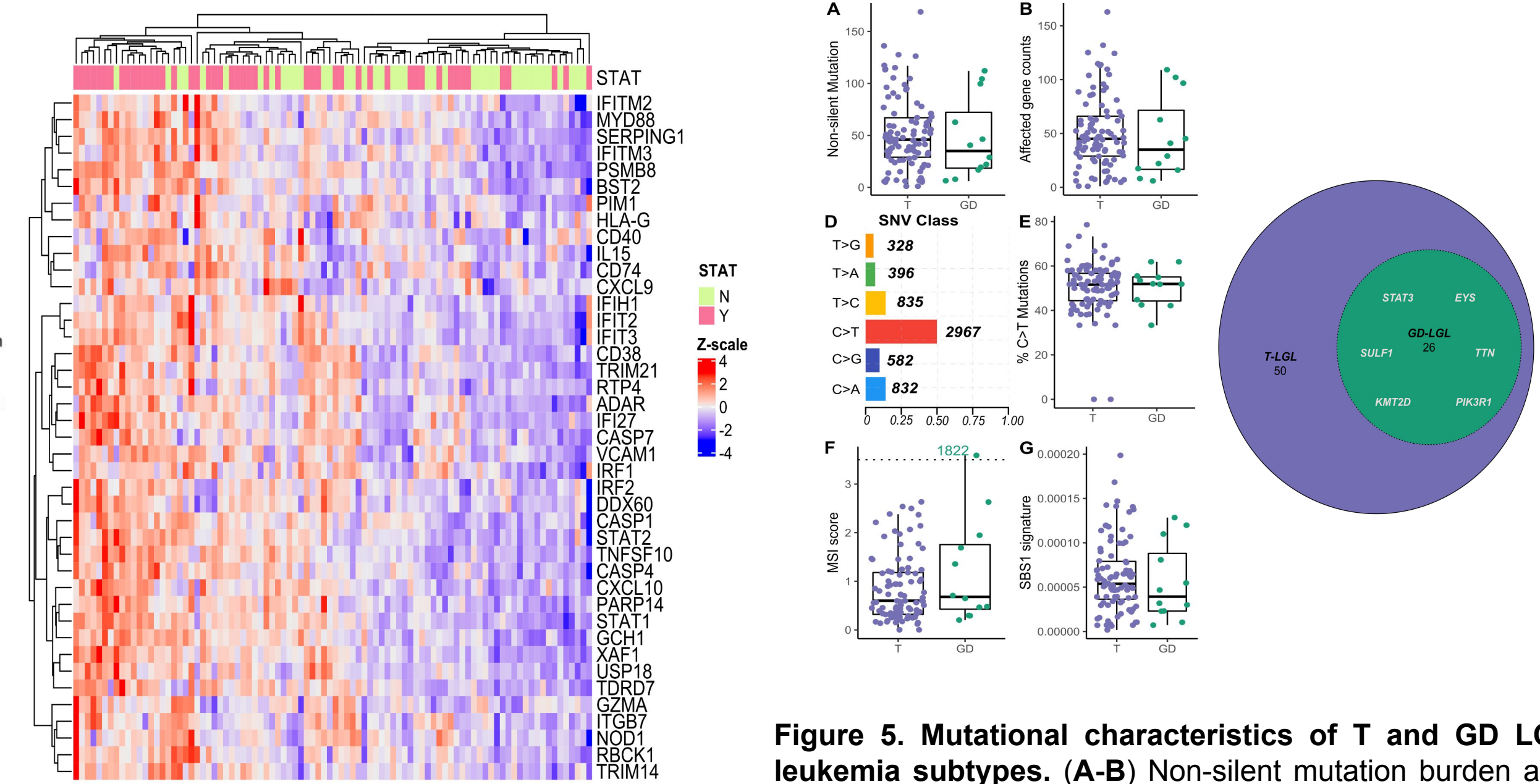


Figure 3. Transcriptomic comparison of *STAT3* mutated patients compared to *STAT3* WT. Heatmap of leading edges from Hallmark interferon-gamma response GSEA gene set. Z-scaled expression values after correcting for batch effect, sex, and tumor counts are shown. Columns are clustered using complete linkage clustering, and rows are clustered using Ward's minimum variance method.

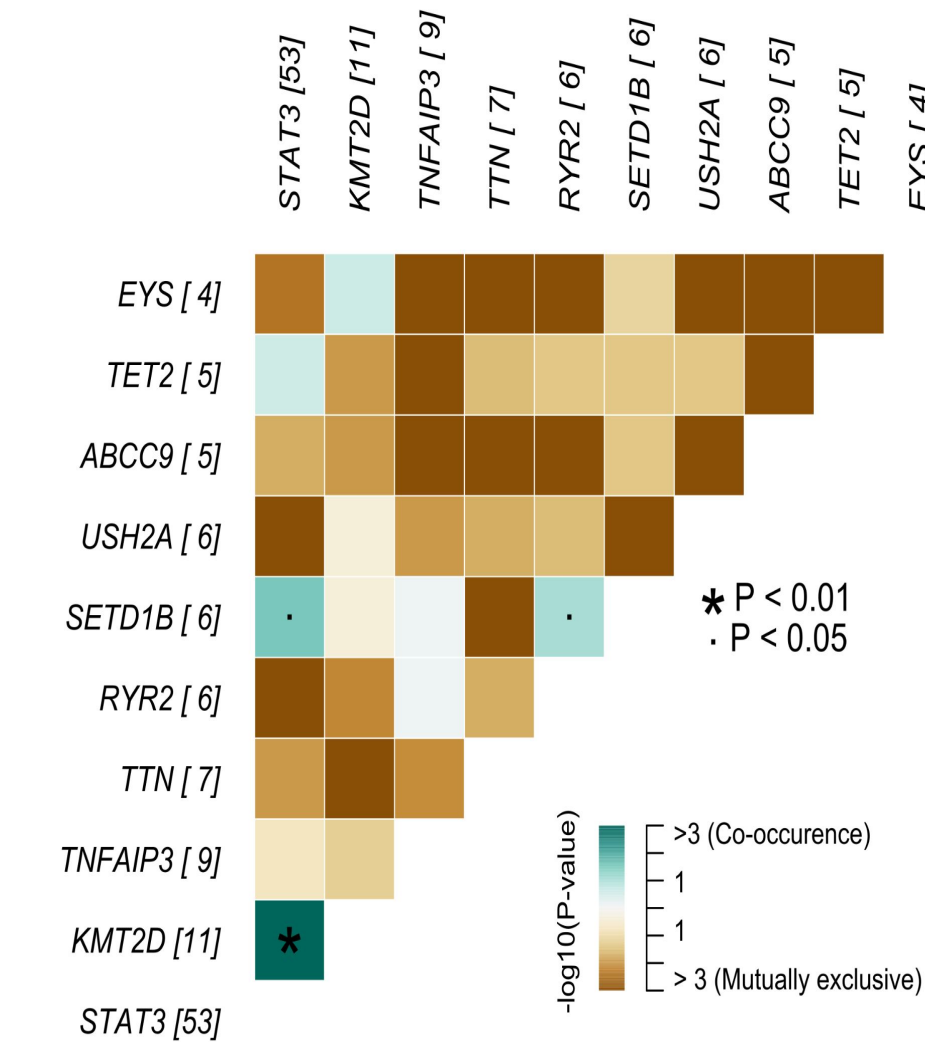


Figure 4. Significant co-occurrence of *STAT3* with *KMT2D* or *SETD1B* mutations. Co-mutation plot of all non-silent mutations observed in the 105 patient LGL leukemia cohort. Numbers in brackets after gene names represent number mutations in this gene in the cohort. Significant pairs of genes were detected by Pair-wise Fisher's Exact test.

Conclusion

We identified recurrent putative driver mutations that may underlie the pathogenesis of LGL leukemia. However, we also observed that a significant percentage of LGL patients sequenced do not harbor notable recurrent somatic mutations in the exonic region. Finally, we found unique differences between *STAT3* mutant and WT leukemic samples in mutational burden and signature, transcriptome, and clinical associations. These findings highlight the emerging etiologic insights into this rare and heterogeneous disorder, which in turn may be leveraged as potential targets to improve therapeutic outcomes in this disease.

Acknowledgements

We thank Bryna Shemo, Andrea Hines, Rachel Stidham, Matthew Schmachtenberg and Shubha Dighe for LGL Leukemia Registry and technical support. We also thank Drs. Kristine C. Olson and Mary Poss for their critical reading of the manuscript and feedback. LGL leukemia research in the Loughran lab is supported by the National Cancer Institute of the National Institutes of Health under award numbers R01CA178393 and P30CA044579 (to TPL). HC and JX are supported by T32LM012416 and T32GM007267. JX is also supported by F30CA225046. Additional support was provided by the Bess Family Charitable Fund, the LGL Leukemia Foundation, Dr. Charles and Katharine Hutton Tweedy, William J. Branch, Dr. Szabolcs Szentpetery, Robert R. Wagner Fellowship (to HC), and a generous anonymous donor (to TPL).