

LYMPHOID NEOPLASIA

CME Article

Genomic landscape of TCR $\alpha\beta$ and TCR $\gamma\delta$ T-large granular lymphocyte leukemia

HeeJin Cheon,¹⁻³ Jeffrey C. Xing,^{2,3} Katharine B. Moosic,^{2,3} Johnson Ung,^{2,3} Vivian W. Chan,^{2,3} David S. Chung,^{2,3} Mariella F. Toro,^{2,3} Omar Elghawy,^{2,3} John S. Wang,^{2,3} Cait E. Hamele,^{2,3} Ross C. Hardison,⁴ Thomas L. Olson,^{2,3} Su-Fern Tan,^{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; ²Department of Medicine, University of Virginia Cancer Center, Charlottesville, VA; 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

KEY POINTS

- **STAT3-mutated patients have increased mutational burden and neutropenia, whereas select STAT3-activating mutations are associated with anemia.**
- **Recurrent mutations in epigenetic and chromatin modifiers, especially KMT2D and SETD1B, co-occur with STAT3 mutations in LGL leukemia.**

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 T-cell receptor $\alpha\beta$ [TCR $\alpha\beta$] T-LGL and 12 TCR $\gamma\delta$ T-LGL). Seventy-six mutations were observed in 3 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 using an unbiased driver analysis approach leveraging our whole-exome cohort. 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 ($P < .01$ and $P < .05$, respectively). STAT3 was mutated in 50.5% of the patients. Most common Y640F STAT3 mutation was associated with lower absolute neutrophil count values, and N647I mutation was associated with lower hemoglobin values. Somatic activating mutations (Q160P, D170Y, L287F) in the STAT3 coiled-coil domain were characterized. STAT3-mutant patients exhibited increased mutational burden and enrichment of a mutational signature associated with increased spontaneous deamination of 5-methylcytosine. Finally, gene expression analysis revealed enrichment of interferon- γ signaling and decreased phosphatidylinositol 3-kinase-Akt signaling for STAT3-mutant patients. These findings highlight the clinical and molecular heterogeneity of this rare disorder.



Medscape Continuing Medical Education online

In support of improving patient care, this activity has been planned and implemented by Medscape, LLC and the American Society of Hematology. Medscape, LLC is jointly accredited by the Accreditation Council for Continuing Medical Education (ACCME), the Accreditation Council for Pharmacy Education (ACPE), and the American Nurses Credentialing Center (ANCC), to provide continuing education for the healthcare team.

JOINTLY ACCREDITED PROVIDER™
INTERPROFESSIONAL CONTINUING EDUCATION

Medscape, LLC designates this Journal-based CME activity for a maximum of 1.0 AMA PRA Category 1 Credit(s)[™]. Physicians should claim only the credit commensurate with the extent of their participation in the activity.

Successful completion of this CME activity, which includes participation in the evaluation component, enables the participant to earn up to 1.0 MOC points in the American Board of Internal Medicine's (ABIM) Maintenance of Certification (MOC) program. Participants will earn MOC points equivalent to the amount of CME credits claimed for the activity. It is the CME activity provider's responsibility to submit participant completion information to ACCME for the purpose of granting ABIM MOC credit.

All other clinicians completing this activity will be issued a certificate of participation. To participate in this journal CME activity: (1) review the learning objectives and author disclosures; (2) study the education content; (3) take the post-test with a 75% minimum passing score and complete the evaluation at <http://www.medscape.org/journal/blood>; and (4) view/print certificate. For CME questions, see page 3100.

Disclosures

Associate Editor Freda K. Stevenson declares no competing financial interests. CME questions author Laurie Barclay owns stock, stock options, or bonds from the following ineligible company: AbbVie Inc (former).

Learning objectives

Upon completion of this activity, participants will:

1. Describe nonsynonymous somatic variants in large granular lymphocyte (LGL) leukemia exomes, potential differences of these mutational profiles across LGL leukemia subtypes, and potential associations of mutational groups with clinical features, according to a paired exome and transcriptome study of LGL leukemia
2. Determine shared and unique molecular characteristics of LGL leukemia subtypes (T-LGL, where cells express $\alpha\beta$ T-cell receptor [TCR], and GD-LGL, where cells express TCR $\gamma\delta$) based on putative driver genes; and the transcriptomic and somatic mutational landscape of *STAT3*-mutated patients compared with *STAT3* wild type to identify their distinct molecular signatures and *STAT3*-mutation-specific clinical associations, according to a paired exome and transcriptome study of LGL leukemia
3. Identify clinical implications of mutations and clinical features associated with LGL leukemia, according to a paired exome and transcriptome study of LGL leukemia

Release date: May 19, 2022; Expiration date: May 19, 2023

Introduction

Large granular lymphocyte (LGL) leukemia is a rare lymphoproliferative disorder of primarily CD8⁺ T- or natural killer (NK)-cell lineage usually manifested in elderly patients.^{1,2} T-LGL accounts for 85% of all LGL cases.³ NK-LGL, or chronic lymphoproliferative disorder of NK cells,⁴ affects ~10% of the patients. The majority of T-LGL cells express T-cell receptor $\alpha\beta$ (TCR $\alpha\beta$; hereafter referred to as T-LGL), and few express TCR $\gamma\delta$ (hereafter referred to as GD-LGL).^{5,6}

One-third of LGL patients are asymptomatic at the time of presentation, and the remaining two-thirds display symptoms related to neutropenia, such as oral ulcers and recurrent infections.² Asymptomatic anemia is also frequently observed during routine blood exams, and transfusion dependence is seen in 10% to 30% of cases. Patients are diagnosed based on the presence of expanded clonal T or NK cells. Treatment is noncurative and includes immunosuppressives such as methotrexate, cyclophosphamide, and cyclosporine.⁷

Somatic activating *STAT3* mutations are observed in 28% to 75% of T-LGL and 30% to 48% NK-LGL leukemia patients.⁸⁻¹¹ *STAT5B* mutations are infrequently observed in LGL in general but commonly detected in rare CD4⁺ variants.^{12,13} Patients with *STAT3* mutations exhibit lower absolute neutrophil counts (ANC), whereas the association with anemia is variable in the literature.^{8,14-17} Since the discovery of *STAT3* mutations, 3 small-scale whole-exome sequencing (WES) studies have been reported for 16 T-LGL and 3 NK-LGL, 4 NK-LGL, and 13 NK-LGL cases,¹⁸⁻²⁰ along with our report of 7 whole genomes in NK-LGL with a main emphasis on *TET2* biology.²¹ Exome sequencing studies have not been reported for the GD-LGL subtype. Mutations in genes that are frequently altered in clonal hematopoiesis of indeterminate potential (CHIP) such as *TET2* and *DNMT3A* have been reported in LGL leukemia.¹⁹⁻²³

However, their prevalence and effect on the rest of the genomic landscape remain unclear.

In the present study, we completed the most extensive paired exome and transcriptome study of LGL leukemia to date with 105 patients, composed of T-LGL (n = 93) and GD-LGL (n = 12) subtypes. We defined the landscape of nonsynonymous somatic variants in LGL leukemia exomes, asked whether these mutational profiles differed across LGL leukemia subtypes, and tested whether mutational groups were associated with clinical features. We identified putative driver genes and used them to determine the shared and unique molecular characteristics of T- and GD-LGL leukemia subtypes. Finally, we compare the transcriptomic and somatic mutational landscape between *STAT3*-mutated and -wild-type (WT) patients to reveal their distinct molecular signatures and *STAT3* mutation-specific clinical associations.

Methods

Patient selection and sample processing

The study was conducted under institutional review board-approved protocols for the LGL Leukemia Registry at the University of Virginia School of Medicine. One hundred five T-LGL leukemia patients (CD8⁺ TCR $\alpha\beta$ and TCR $\gamma\delta$; no CD4⁺) meeting diagnosis criteria by the World Health Organization⁴ (evidence of clonality, persistence of chronic expansion >6 months) were recruited for this study. See supplemental Methods for additional details (available on the *Blood* Web site).

WES and data processing

WES samples were sequenced by Tempus (Chicago, IL) to a target depth of 100× and 200× for the normal (saliva) and the leukemic (peripheral blood mononuclear cell [PBMC]) genome, respectively, using an Illumina HiSeq.²⁴ The paired-end sequences were aligned with Novoalign (Novocraft, Selangor, Malaysia) to an in-house g1k_v37_novoindex of human genome version GRCh37.

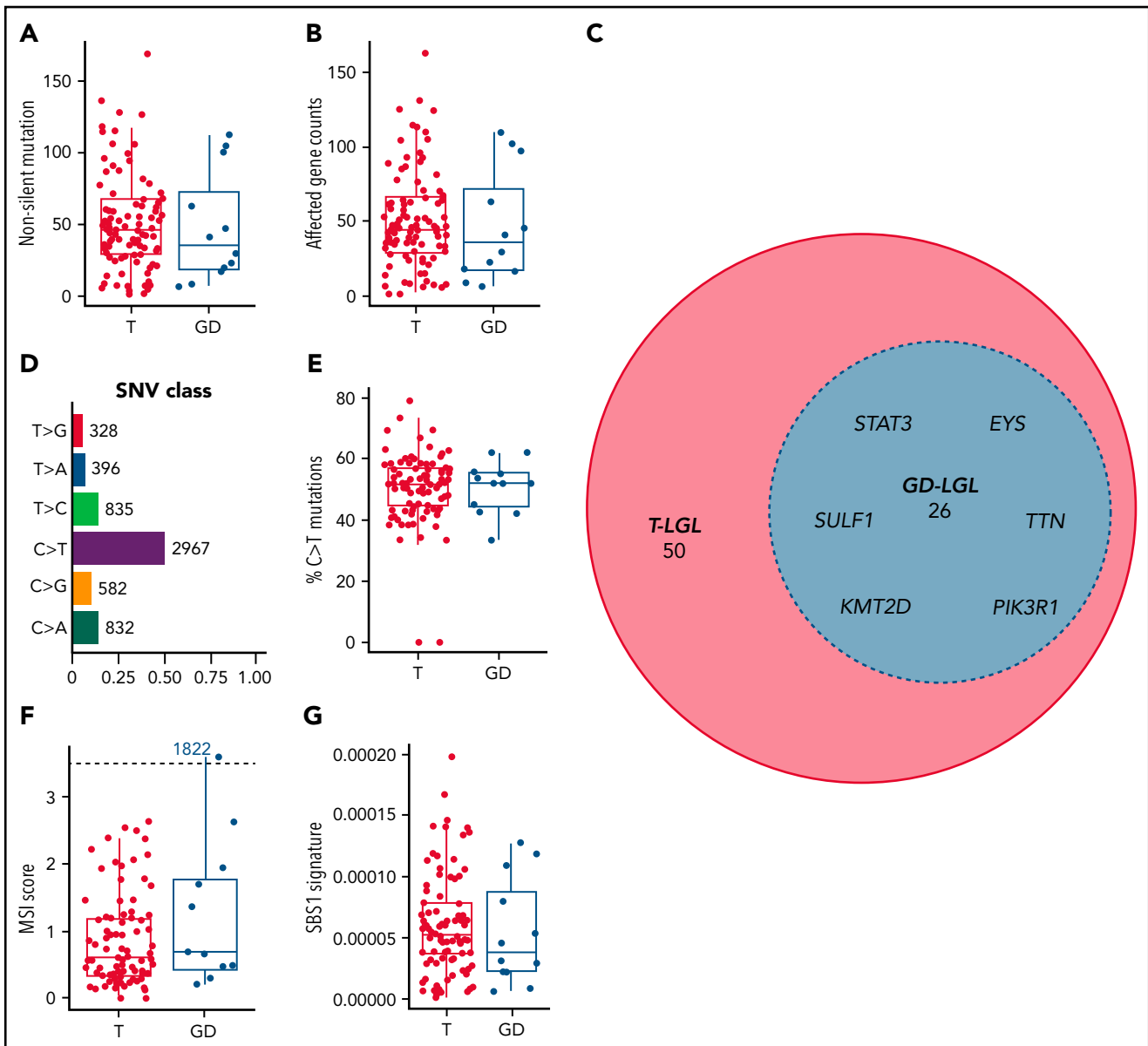


Figure 1. Mutational characteristics of T and GD LGL leukemia subtypes. (A-B) Nonsilent mutation burden and affected gene counts in T and GD subtypes. (C) Euler diagram of genes that are mutated in at least 3 patients, sectioned by LGL subtypes. Genes that are mutated in 4 or more patients and are shared between T and GD-LGL (blue area) are labeled. (D-E) Single nucleotide variant class in the cohort and percent C>T mutation comparison in T and GD subtypes. (F-G) Microsatellite instability scores and SBS1 signature contributions in T and GD subtypes.

Capture probes from Integrated DNA Technologies' xGen Exome Research Panel were used, and samples were prepped using the KAPA HyperPrep Kit as described previously.²⁴

The somatic mutations were identified using Mutect2 in the Genome Analysis Toolkit²⁵ v4.1.2.0, following Genome Analysis Toolkit best practices. Additionally, databases of known cancer mutations from COSMIC v77 and known population polymorphisms from dbSNP v138 were used as a whitelist and blacklist, respectively, during somatic variant calling. The low-quality variants were filtered away, and the variant call format files were annotated with Ensembl's Variant Effect Predictor.²⁶ Annotated variant call format files were converted to mutation annotation format using vcf2maf.²⁷

RNA sequencing (RNA-seq) and data processing

RNA libraries were constructed and sequenced (Tempus) as described previously.²⁴ Samples were hybridized with Integrated DNA Technologies' xGen Exome Research Panel followed by amplification using a KAPA HiFi Library Amplification Kit.²⁴ The RNA libraries were sequenced to obtain at least 50 million reads on an Illumina HiSeq 4000.

Results

Somatic variants within and across LGL disease subtypes and their association with clinical phenotypes are incompletely defined by previous sequencing studies.¹⁸⁻²⁰ To better understand the underlying molecular landscape of LGL leukemia, we

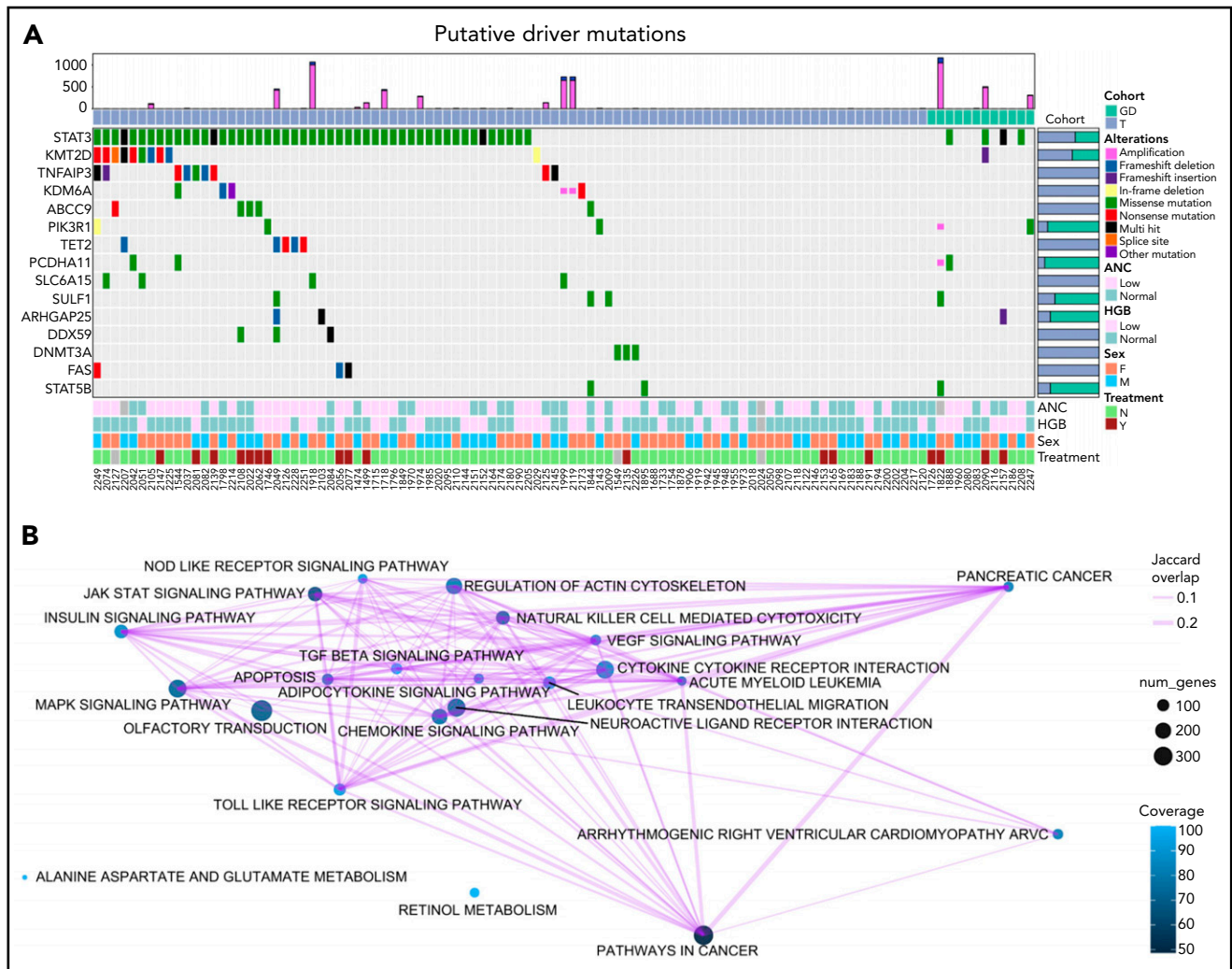


Figure 2. Oncoplot of putative drivers in LGL leukemia. (A) Seven different driver analysis tools identified 15 gene variants with putative driver roles present in at least 3 LGL leukemia samples. Each row represents 1 of the 15 putative drivers, whereas the columns represent individual patients. LGL 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 WT. 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 2 LGL cohorts. Bottom annotations represent clinical phenotypes of individual patients, with color definitions found in the rightmost annotation. ANC values above 1.5k/ μ l 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 gray. (B) Kyoto Encyclopedia of Genes and Genomes pathway analysis of somatic mutations observed in the cohort using SLAPenrich. The size of the nodes represents the number of genes in the cohort observed in the pathway. The color of the nodes represents the percent of those genes that uniquely belong in each pathway. The thickness of the edges represents the Jaccard index between the nodes.

collected saliva and peripheral blood specimens from 105 LGL patients, which included T-LGL ($n = 93$) and GD-LGL ($n = 12$). Genomic and transcriptomic analyses were performed after paired WES and leukemic bulk-RNA-seq. Patient demographics are summarized in supplemental Table 1. Detailed clinical annotations are provided in supplemental Table 2.

Landscape of somatic variants in LGL leukemia exomes

We focused our efforts on nonsilent variants to increase the likelihood of identifying changes that alter protein structure or function and contribute to disease pathogenesis. Paired saliva-leukemic WES of 105 individuals revealed a range of 1 to 169 (median: 45) nonsilent somatic variants per patient after high confidence filtering. There were no statistically significant differences between the 2 subtypes in either nonsilent mutation burden or mean affected gene count (range: 1-163; median: 45; Kruskal-Wallis test;

nonsilent somatic variants, $P = .475$; mean affected gene count, $P = .468$) (Figure 1A-B). After excluding intron, 5' flank, 3' and 5' untranslated region, intergenic region, and RNA mutation classes, 2717 genes were affected by these nonsilent mutations, with 76 genes being mutated in 3 or more patients (supplemental Table 3). Within these 76 genes, *STAT3*, *KMT2D*, *PIK3R1*, *SULF1*, *EYS*, and *TTN* (Figure 1C) were found to be mutated across T and GD subtypes. All somatic variants including silent variants are displayed in supplemental Table 4. We observed broad chromosomal arm-level somatic copy-number aberrations in 5 patients, with 22q and 3p amplifications observed in 2 patients in each category (supplemental Table 5). C>T and G>A base substitutions (seen in 50.5% of all single nucleotide variant events) dominated the somatic variant calls in this cohort (Figure 1D). Differences in the percent of C>T base substitutions were not statistically significant among the subtypes ($P = .988$) (Figure 1E). We also identified 1 patient (1822 GD-LGL) with many mutations within

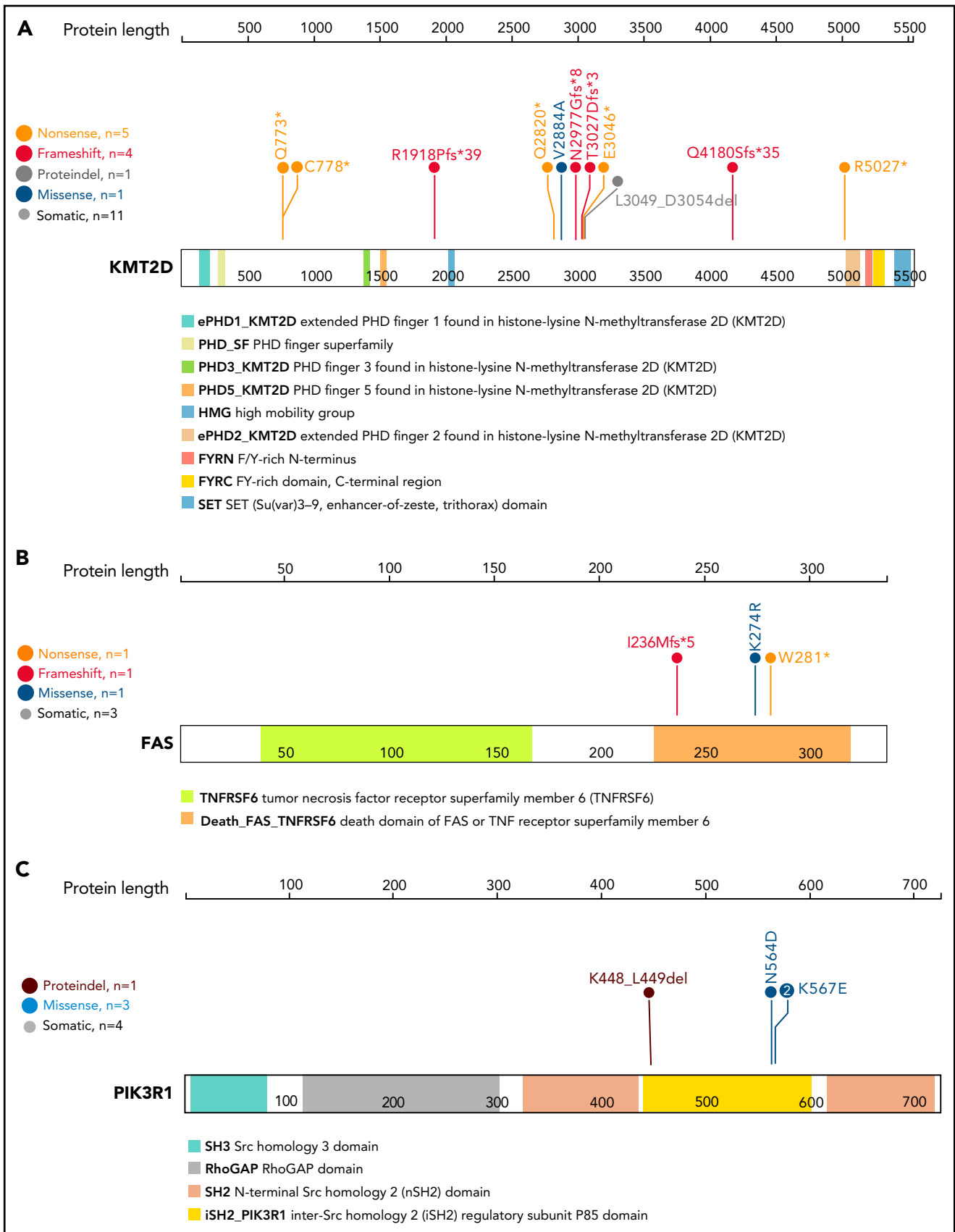


Figure 3. Lollipop plot of *KMT2D*, *FAS*, and *PIK3R1* genes. Lollipop plot of (A) *KMT2D*, (B) *FAS*, and (C) *PIK3R1* mutations detected in the 105-patient LGL leukemia cohort. Annotations for individual domains are described at the bottom. Domain information was curated from simple modular architecture research tool and conserved domain database.

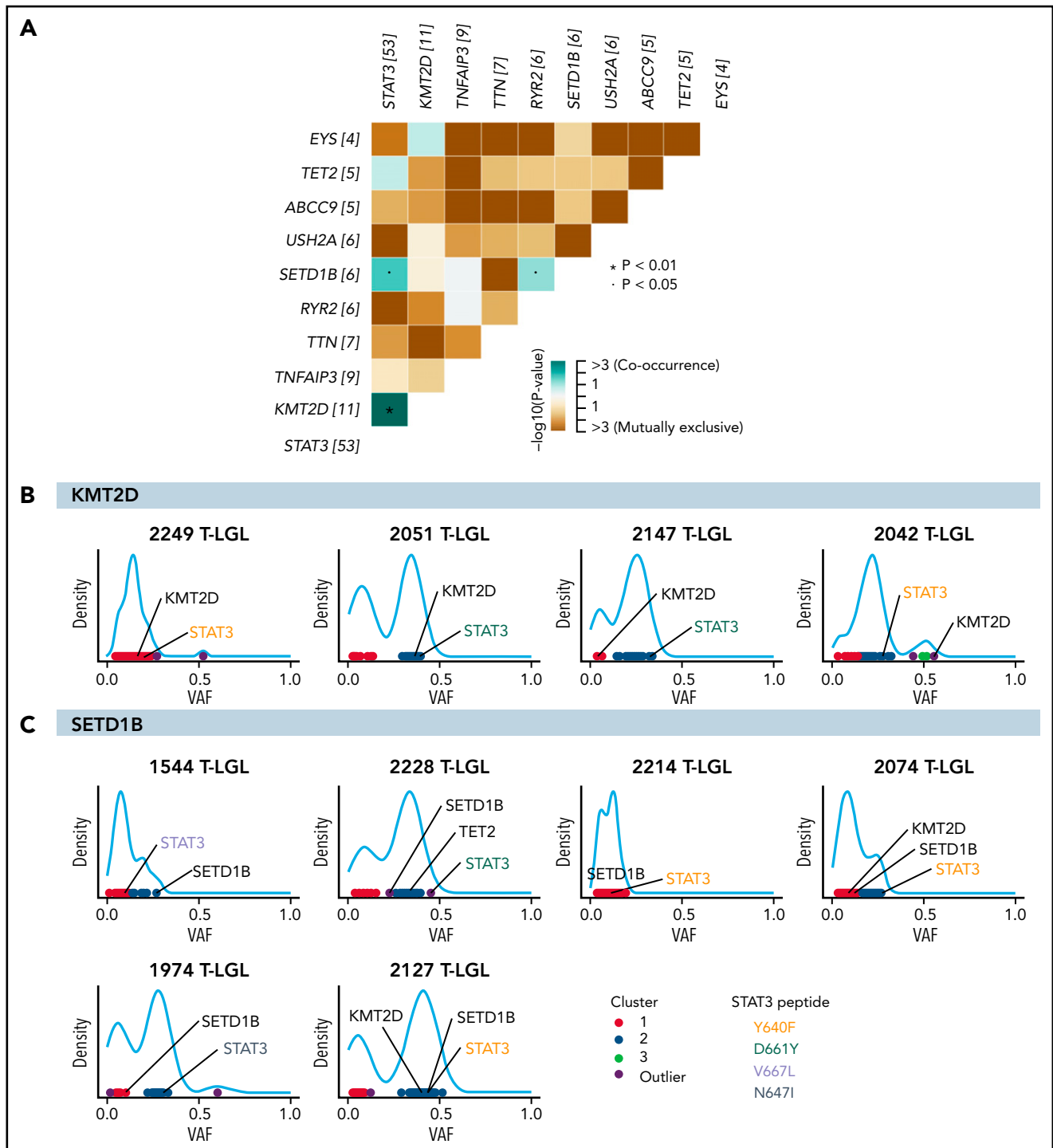


Figure 4. Significant co-occurrence of STAT3 with KMT2D or SETD1B mutations. (A) Computation plot of all nonsilent mutations observed in the 105-patient LGL leukemia cohort. Numbers in brackets after gene names represent the number of mutations in this gene in the cohort. Significant pairs of genes were detected by pairwise Fisher's exact test. (B-C) VAF density plots were used to infer clones or "clusters." Each dot represents individual somatic variants observed in a given patient. The color of the dots represents inferred clusters. Colored annotations indicate different types of STAT3 mutations. (B) Four representative KMT2D VAF density plots. The first 2 demonstrate inclusion of both STAT3 and KMT2D in the same predicted clone. The following 2 plots demonstrate 2 mutations associating with 2 different clones. (C) VAF density plot of all patients harboring SETD1B mutations in the cohort.

microsatellites (microsatellite instability [MSI]-high; MSIsensor score >3.5 ; Figure 1F), though MSI scores were not significantly different among the 2 LGL subtypes ($P = .339$). Additionally, we used signer to determine that the cohort was enriched for the SBS1 signature.²⁸ The C>T mutations dominate the SBS1 signature at CpG dinucleotides. This likely represents deamination of

5-methylcytosine at CpG dinucleotides, leading to T:G mismatches that are not repaired before DNA replication.²⁹ The relative contribution of the SBS1 signature was not significantly different between the 2 LGL subtypes ($P = .427$) (Figure 1G). Together, these results show that multiple genomic level changes and somatic variants were shared between T and GD subtypes.

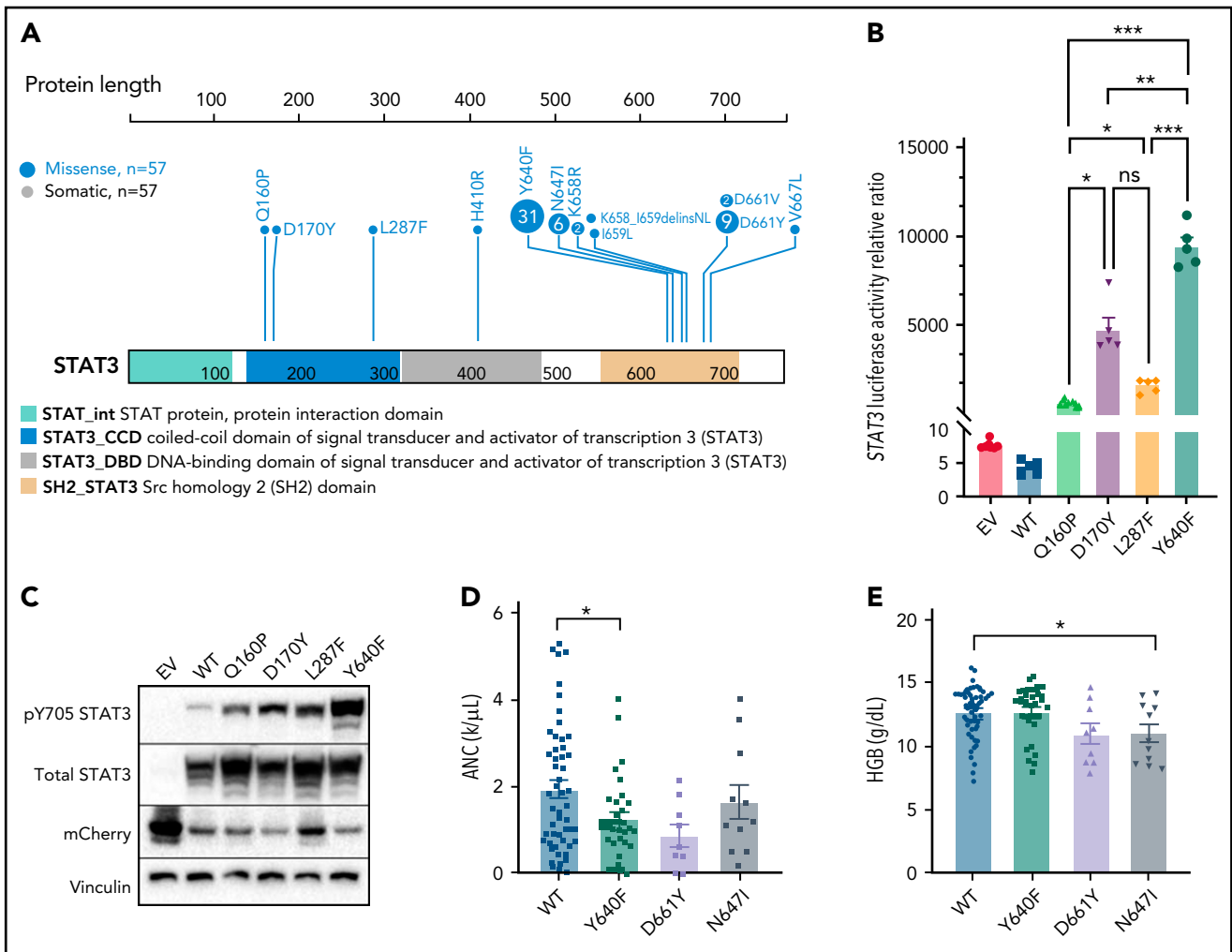


Figure 5. Novel activating *STAT3* mutations in the coiled-coil domain and clinical phenotype of patients harboring specific *STAT3* somatic mutations. (A) Lollipop plot of *STAT3* somatic mutations detected in the 105-patient LGL leukemia cohort. Annotations for individual domains are described at the bottom. (B) *STAT3* expression vectors, Q160P, L287F, D170Y, and Y640F as well as empty vector and WT *STAT3* were transfected into HEK293 cells along with *STAT3* luciferase reporter. Shown is a quantification of the relative ratio between firefly *STAT3* responsive element against cytomegalovirus-controlled renilla luciferase. Two independent experiments were performed. Shown is 1 representative experiment. (C) Western blot of phospho-Y705 *STAT3*, total *STAT3*, mCherry, and vinculin from parallel whole-cell lysates obtained from transfection of empty vector, WT, and mutant *STAT3* constructs as in panel B. Two independent experiments were performed. Shown is 1 representative experiment. (D-E) Plots of hematologic parameters ANC and HGB segregated by *STAT3* WT, Y640F, D661Y, and N6471 mutations. This includes 5 additional patients with N6471 mutations identified using Sanger sequencing outside of the original 105-patient cohort. For panel D, Dunnett's test with WT as a comparison control, $P = .045$, $P = .059$, and $P = .853$ for Y640F, D661Y, and N6471 mutations, respectively. For panel E, Dunnett's test with WT as a comparison control, $P = .999$, $P = .074$, and $P = .047$ for Y640F, D661Y, and N6471 mutations, respectively. P values are indicated as follows: * $P < .05$; ** $P < .01$; *** $P < .001$.

Identification and function of putative driver mutations

We used a driver analysis approach³⁰ to identify 15 genes with putative roles ("putative drivers") in the etiology of LGL leukemia out of the 76 genes that were mutated in 3 or more patients (Figure 2A; supplemental Figure 1). Thus, only a select few genes with nonsilent mutations were identified by our analysis pipeline as putative drivers. *STAT3* was by far the most frequently mutated gene across all LGL subtypes, with nonsilent mutations in 53 out of 105 patients (50.5%). Putative drivers that passed the loss-of-function transcript effect estimator filter to identify loss-of-function mutations are annotated in supplemental Table 6. These included several genes encoding epigenetic modifiers (*KMT2D*, *KDM6A*, *TET2*, and *DNMT3A*). Putative drivers that interact with the *STAT3*-signaling pathway

(*TNFAIP3*, *PIK3R1*, and *FAS*) were also identified. Of note, *TNFAIP3* mutations in our cohort did not overlap with previously reported mutations in LGL leukemia.³¹ These 9 frame-shift and nonsense mutations spanned across the gene and were identified as high-confidence loss-of-function mutations (supplemental Table 6). Another putative driver, *ABCC9*, has a known role as a potassium channel and is a *STAT3* target gene.³² *STAT5B* mutations were detected in far fewer samples than *STAT3* as previously described.¹² Kyoto Encyclopedia of Genes and Genomes pathways that are recurrently altered by nonsilent mutations in the cohort (Figure 2B) include the pathways involved in cancer, including pancreatic cancer and acute myeloid leukemia. Additional pathways include genes involved in JAK-*STAT* signaling, chemokine signaling, and cytokine-cytokine receptor interaction.

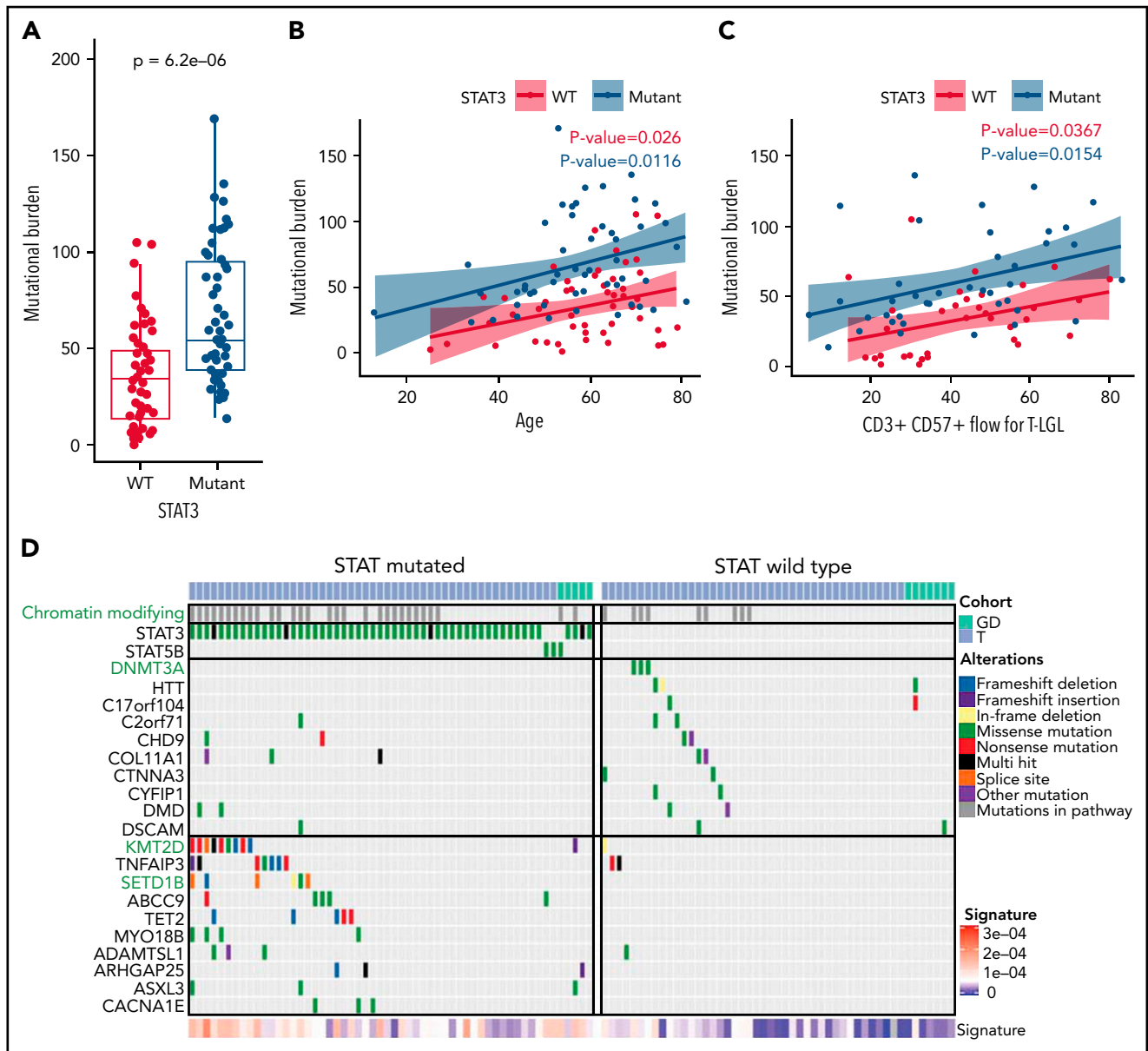


Figure 6. Increased mutational burden in *STAT3*-mutated patients. (A) Boxplot of nonsilent mutational burden comparing *STAT3* WT against mutant patients. P value from Wilcoxon t test. (B) Plot of the mutational burden against age for WT and mutant *STAT3* patients with linear regression and colored 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. (D) Top recurrent somatic mutations observed in *STAT3*-mutated and WT groups. The left oncoplot shows patients with *STAT3* or *STAT5B* mutations ($n = 56$), and the right oncoplot shows *STAT3* WT ($n = 49$) patients. The top annotation indicates the LGL subtype. Patients with mutations in the chromatin-modifying enzyme gene list (supplemental Table 9) are annotated in the second row in gray. Three chromatin-modifying genes are highlighted in green font. The 10 genes that are most frequently mutated in WT patients are shown in the top half, followed by the 10 genes most frequently mutated in *STAT3*- and *STAT5B*-mutated patients in the bottom half. Annotation of "Alterations" indicates the type of mutation affecting the gene. SBS1 signature explains the degree of intensity of the COSMIC SBS1 signature present for each patient.

Out of 15 putative driver genes, select candidates demonstrated high potential impact on LGL pathogenesis based on previous literature and the occurrence of hotspot mutations (Figure 3). We observed high-confidence loss-of-function mutations across the coding region of *KMT2D*, the second most frequently mutated driver gene (Figures 2 and 3A). *KMT2D* encodes histone H3 lysine 4 mono-methyltransferase and is thought to be a tumor suppressor as loss of *Kmt2d* is shown to promote lung tumorigenesis³³ and increase incidence of germinal center-derived lymphomas in mice.³⁴ Additionally, we detected nonsense, frameshift, and missense mutations in the death domain of *FAS* in 3 T-LGL patients (Figure 3B). Somatic

mutations in this domain are associated with a related disorder, autoimmune lymphoproliferative syndrome, characterized by lymphoid precursor resistance to Fas-mediated apoptosis.^{35,36} Finally, we observed 3 distinct mutations affecting the inter-SH2 domain of *PIK3R1* (Figure 3C). Specifically, the mutations were detected in 3 T-LGL patients (1 in-frame deletion at K448_L449del, 2 K567E mutations) and 1 GD-LGL patient (N564D mutation). Additional phosphatidylinositol 3-kinase (PI3K) family mutations were observed in *PIK3CD*, *PIK3CA*, *PIK3AP1*, *PIK3R3*, *PIK3R4*, and *PIK3C2A* (supplemental Table 7). These observations indicate that the somatic mutations in T-LGL patients frequently lead to aberrant PI3K signaling.

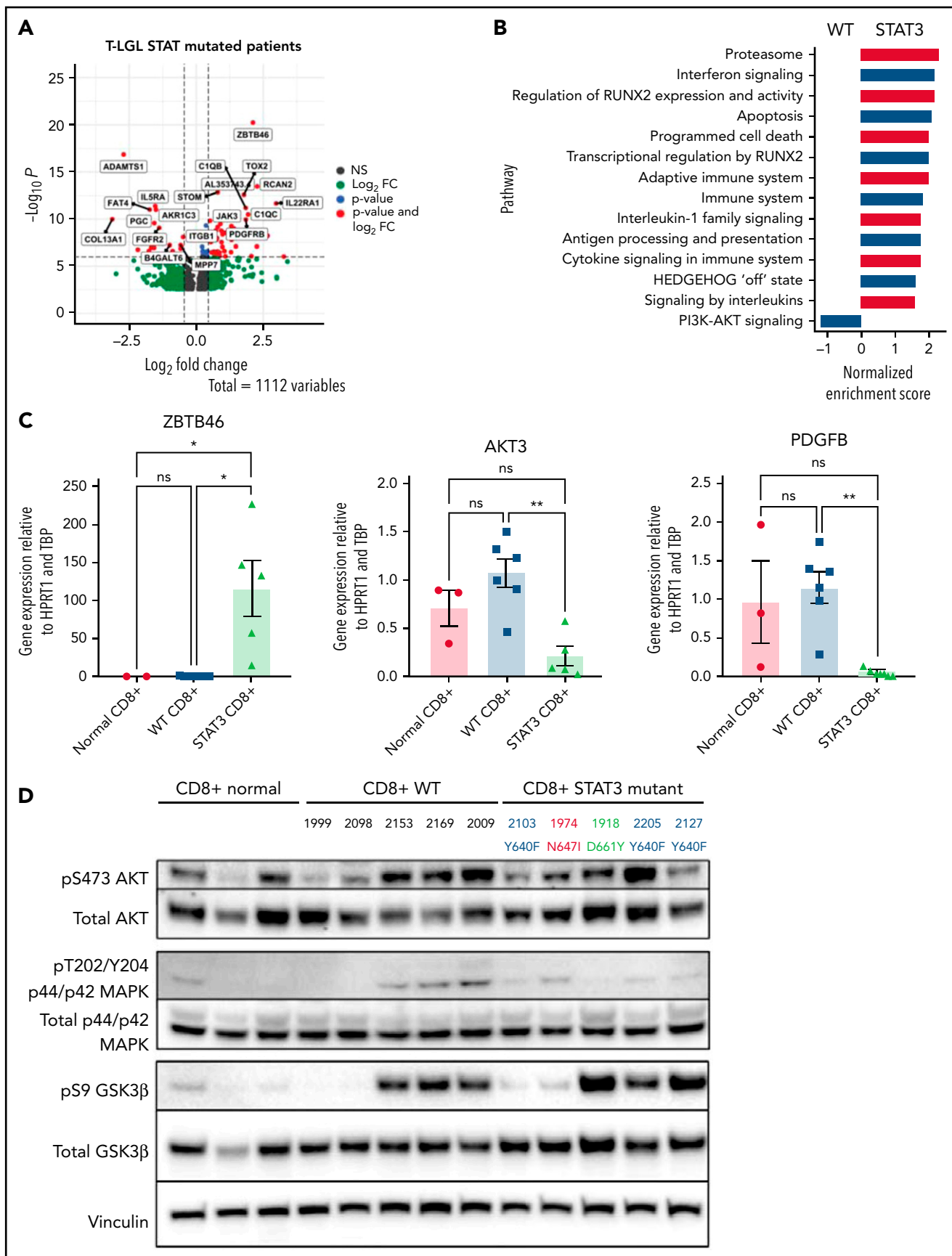


Figure 7.

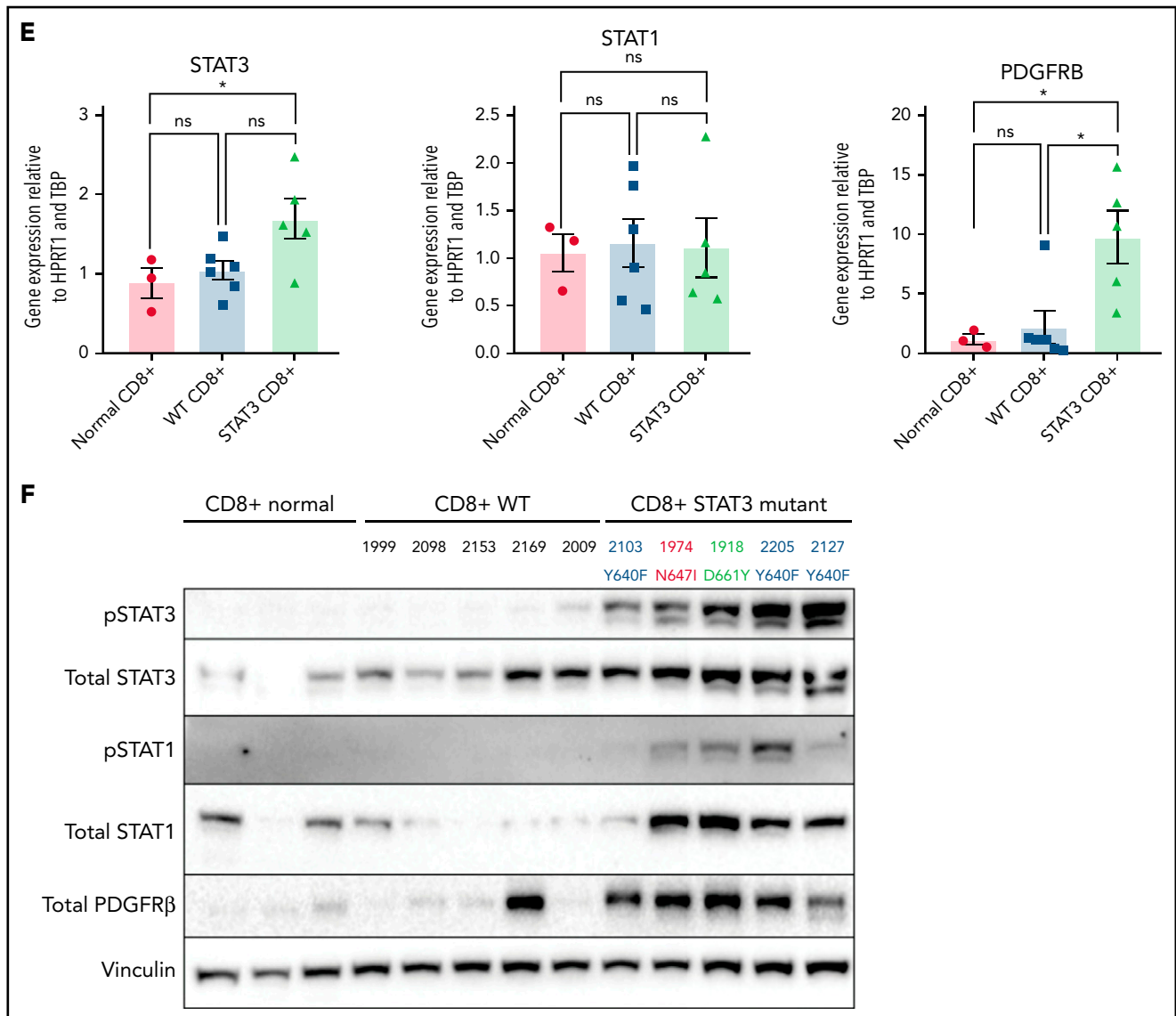


Figure 7. Transcriptomic and protein expression comparison of STAT3-mutated patients against STAT3-WT patients. (A-B) RNA-seq analysis of T-LGL leukemia patients. (A) Volcano plot of DESeq2 output of T-LGL STAT3 mutant vs STAT3 WT. Genes with positive \log_2 fold changes are more highly expressed in STAT3 mutants, and genes with negative \log_2 fold changes are more highly expressed in STAT3 WT. (B) Gene set enrichment analysis of the nodes from the functional modular network analysis. Positive enrichment scores indicate pathways that are positively enriched in STAT3-mutant patients. Negative enrichment scores indicate pathways enriched in STAT3-WT patients. (C) RT-qPCR comparison of mRNA levels of ZBTB46, AKT3, and PDGFB from CD8⁺-isolated normal controls (n = 3; ZBTB46 n = 2 due to signal below the limit of detection in 1 sample), WT (n = 6; ZBTB46 n = 5 due to signal below the limit of detection in 1 sample), and STAT3-mutant LGL patient samples (n = 5). Analysis of variance (ANOVA) $P = .028, .014, \text{ and } .018$ for ZBTB46, AKT3, and PDGFB, respectively. (D) Western blotting of pS473 AKT, total AKT, phospho-p44/42 MAPK, total p44/42 MAPK, pS9 GSK3β, and total GSK3β from CD8⁺-isolated normal controls (n = 3) compared with WT and STAT3-mutant LGL patient samples (n = 5). STAT3 mutation type and LGL registry ID are indicated. (E) RT-qPCR comparison of mRNA levels of STAT3, STAT1, and PDGFRB from CD8⁺-isolated normal controls, WT, and STAT3-mutant LGL patients. ANOVA $P = .049, .963, \text{ and } .026$ for STAT3, STAT1, and PDGFRB, respectively. (F) Western blotting of pSTAT3, total STAT3, pSTAT1, total STAT1, total PDGFRβ, and vinculin loading control from CD8⁺-isolated normal controls compared with WT and STAT3-mutant LGL patient samples. STAT3 mutation type and LGL registry ID are indicated. Welch's or Brown-Forsythe ANOVA was used for all RT-qPCR analyses depending on the distribution of the data. Unpaired t with Welch's correction as a post hoc test. P values are indicated as follows: * $P < .05$; ** $P < .01$.

Other notable drivers included genes with known tumor-suppressor function in leukemia and solid cancers. Four T-LGL patients exhibited mutations in *KDM6A*, which encodes a lysine-specific demethylase with deleterious mutations previously reported in T-cell acute lymphoblastic leukemia and other cancers (supplemental Table 4).³⁷⁻⁴¹ High-confidence loss-of-function mutations (S495Pfs*65, Y129Mfs*80, M90Hfs*24) in *ARHGAP25*, a GTPase-activating protein for Rac that is

associated with tumor-suppressive functions,^{42,43} were observed for 2 T-LGL and 1 GD-LGL patient (supplemental Table 4).

Co-occurrence and mutual exclusivity of putative driver mutations

Comutation analysis was completed to identify putative redundant or cooperating molecular events. *KMT2D* mutation exhibited significant co-occurrence with STAT3 mutation, being observed in 10

out of 11 total *KMT2D* mutant patients ($P = .0084$; Figure 4A). *SETD1B* mutations also significantly co-occurred with *STAT3* mutations ($P = .027$; Figure 4A). Furthermore, approximately half of the *KMT2D*-mutated samples showed similar variant allele frequencies (VAF) of *KMT2D* and *STAT3* mutations, compatible with coexistence in the same clone (Figure 4B; see supplemental Figure 2 for analysis on all *KMT2D*-mutated samples). Three out of 6 *SETD1B*-mutated samples also showed coexistence in the same clone with *STAT3* (Figure 4C). Of note, all 5 T-LGL patients with *TET2* and *STAT3* mutations had *TET2* mutations in the same predicted clone as *STAT3* based on VAF density clustering (supplemental Figure 3). No direct association with specific *STAT3* amino acid change and *TET2*, *KMT2D*, or *SETD1B* mutation was observed (Figure 4B-C; supplemental Figure 3). The significant co-occurrence of *STAT3* mutations with *KMT2D*, *SETD1B*, and *TET2* mutations suggests that *STAT3* activation may cooperate with epigenetic modifications in driving the disease.

Functional and clinical characterization of *STAT3* mutation

Previous studies have found recurrent, somatic activating mutations affecting the SH2 domain of *STAT3*. In agreement with those previous studies, we found nonsilent mutations in the SH2 domain of *STAT3* in 51 out of 105 (48.6%) patients (Figure 5A; supplemental Table 4). We also observed a single mutation in the DNA-binding domain (H410R) and previously unreported mutations in the coiled-coil domain (Q160P, D170Y, and L287F; Figure 5A; supplemental Figure 4). To explore the functional significance of novel mutations in the coiled-coil domain, we transfected HEK293 cells with expression vectors for WT, Q160P, D170Y, L287F, and Y640F (positive control⁸) along with a *STAT3* luciferase reporter vector containing a *STAT3*-responsive cis-inducible element. All novel mutants exhibited luminescence signal 129- to 794-fold higher than empty vector and WT *STAT3* controls (Figure 5B; supplemental Figure 5). Y640F exhibited the strongest signal and was statistically significant compared with all other examined mutations. Phospho-Y705 *STAT3* levels were also strongly increased (1.5- to 2.7-fold higher) in all novel mutants relative to WT (Figure 5C; supplemental Figure 5).

STAT3 mutation at any site was significantly associated with lower neutrophil (ANC, $P = .0042$) counts (supplemental Figure 6) as previously observed.⁸ We also observed lower hemoglobin (HGB, $P = .071$) and hematocrit (HCT, $P = .063$) values in *STAT3*-mutated patients (supplemental Figure 6). Lastly, we retrospectively collected bone marrow pathology on 42 patients to detect differences between *STAT3*-WT ($n = 18$) vs -mutant patients ($n = 24$) (supplemental Figures 7 and 8). No association with bone marrow cellularity, lymphocyte percent, myeloid to erythroid ratio, reticulin grade, presence of TCR clone, T-cell intracellular antigen, or granzyme staining was found with *STAT3* mutation. Other somatic mutations, such as *TET2*, were not analyzed given the very low number of mutated samples ($n = 2$) with associated bone marrow data.

Our large patient cohort allowed us to further stratify the *STAT3* mutation group into common somatic activating variants Y640F ($n = 31$), D661Y ($n = 9$), and N647I ($n = 11$) to determine their association with the key clinical features,

neutropenia, and anemia, which are indicators for treatment in LGL leukemia. Previous studies and supplemental Figure 6 analyzed all *STAT3* mutants together in comparisons with *STAT3* WT.^{8,14-17} We found that Y640F mutation exhibited significantly lower ANC values ($P = .045$; Figure 5D). ANC values for D661Y ($P = .059$) and N647I ($P = .853$) mutation also trended toward lower levels than WT. Notably, N647I mutation was associated with significantly lower HGB values ($P = .047$), with D661Y mutation trending similarly ($P = .074$). However, the range of HGB values was not significantly different for the most frequent somatic activating mutation, Y640F, compared with WT ($P = .9999$) (Figure 5E). HCT and red blood cell values showed similar changes (supplemental Figure 8). There was no association with white blood cell, absolute lymphocyte count, mean corpuscular volume, red cell distribution width, platelets, or age for *STAT3* mutation groups.

Mutational landscape of *STAT3*-WT and -mutant LGL leukemia

As defined by the counts of nonsilent somatic variants in exomes, the total mutational burden has been reported to be a function of age in other hematological malignancies.⁴⁴ We investigated if there were a relationship between the mutational burden and clinical features in LGL leukemia, which includes mutation status for *STAT3* and flow-cytometric profiling. We used least absolute shrinkage and selection operator regularization and determined that the *STAT3* mutation status, the CD3⁺CD57⁺ fraction, and the patient's age at the time of sequencing were all significant predictors of the mutational burden (Figure 6A-C). Additional analysis including silent variants also showed increased mutational burden in *STAT3*-mutated samples (supplemental Figure 9). Together, these variables explain only ~35% of the variance in the mutation burden, highlighting the heterogeneity of the disease. Additionally, overall mutational burden was not significantly different among the Y640F, D661Y, and N647I *STAT3* mutation groups ($P = .7512$, Kruskal-Wallis).

Next, we compared the mutational landscapes of *STAT3*-mutant and -WT LGL leukemia. *STAT3*-mutant samples exhibit increased mutation prevalence for genes such as *KMT2D* and *SETD1B* involved in chromatin modification (Figure 6D; see supplemental Table 9 for entire list) as well as epigenetic modifier *TET2*. Conversely, *DNMT3A* mutations were found exclusively in *STAT3*-WT patients in our cohort. Further analysis revealed that the relative contribution of the SBS1 signature, which is enriched in LGL leukemia (Figure 1E), was significantly higher in *STAT3*-mutated patients compared with *STAT3* WT (Figure 6D; supplemental Figure 10). Taken together, this suggests that *STAT3*-mutated patients harbor a distinct genomic profile characterized by epigenetic and mutational aberrations.

Transcriptomic analysis of *STAT3*-WT and -mutant LGL leukemia

We next examined the transcriptome differences of patients with somatic mutations in *STAT3* compared with *STAT3* WT. Six hundred fifty four genes were significantly upregulated and 458 genes were significantly downregulated in T-LGL patients with *STAT3* mutations after accounting for batch effects, sex, and purity of the samples (supplemental Table 10). The top 10 differentially expressed genes with higher expression in the

STAT3-mutated group (Figure 7A) included *PDGFRB*, which was previously reported as highly expressed in LGL leukemic cells,⁴⁵ and *JAK3*, which is activated in various hematologic malignancies.^{46,47} Aberrant expression of myeloid lineage-specific genes was also observed, including *ZBTB46*, *C1QB*, and *C1QC* for *STAT3* mutants, and *PGC* and *FGFR2* for *STAT3* WT. Functional enrichment analysis using gProfiler followed by EnrichmentMap visualization revealed key module differences in cytokine signaling, antigen processing, apoptosis, PI3K-Akt signaling, Runx2 activity, and hedgehog signaling between *STAT3*-mutant and -WT patients (supplemental Figure 11A). Gene set enrichment analysis of these network modules demonstrated overrepresentation in *STAT3* mutants, except for enrichment of PI3K-Akt signaling in *STAT3*-WT samples (Figure 7B).

RNA-seq analysis was confirmed by sorting CD8⁺ cells from normal controls as well as the original date-matched sample used for whole-exome and RNA-seq. RNA and protein expression were determined using RT-qPCR (Figure 7C,E) and western blotting (Figure 7D,F). *STAT3*-mutated CD8⁺ samples strongly overexpress the myeloid lineage-specific *ZBTB46* compared with WT samples and CD8⁺ normal controls (Figure 7C). We also demonstrated decreased expression of *AKT3* and *PDGFB*, both genes in the PI3K-Akt signaling axis, in *STAT3* mutants compared with *STAT3*-WT samples and normal controls (Figure 7C). We then probed for PI3K-Akt signaling cascades in CD8⁺ isolated normal, WT, and *STAT3*-mutated samples. Three patients in both groups exhibited increased levels of phospho-AKT, phospho-p44/42 MAPK, and phospho-GSK3 β compared with normal CD8⁺. Conversely, 2 patients in both groups showed little activation of these readouts (Figure 7D). There was no statistical difference in PI3K-Akt signaling, primarily due to this patient heterogeneity (supplemental Figure 12A-C).

Finally, unbiased clustering of patients with leading edges from the hallmark interferon- γ gene set revealed increased expression of interferon- γ pathway genes in *STAT3*-mutant patients (supplemental Figure 11B). We examined *STAT3* and *STAT1* gene expression in CD8⁺ isolated samples (Figure 7E) but did not observe statistically significant differences in messenger RNA (mRNA) levels. However, immunoblots showed high abundance of phospho-*STAT3* and phospho-*STAT1* in *STAT3* mutant relative to WT patient samples and normal controls (Figure 7F; supplemental Figure 12D-E). We also detected high *PDGFR β* protein expression levels in *STAT3* mutants (Figure 7F; supplemental Figure 12F), consistent with RNA-seq analysis (Figure 7A,E). In sum, these analyses show a distinct transcriptomic profile of *STAT3*-mutant patients with highlights in interferon and interleukin signaling.

Discussion

Here, we define the genomic landscape of LGL leukemia using an integrated and comprehensive genomic approach in 105 samples. This is the first WES report for GD-LGL subtype. We identified shared somatic mutations and putative drivers between T-LGL and GD-LGL leukemia. We detected nonsilent mutations in 2717 genes, including novel somatic mutations that are minor in frequency in LGL leukemia. Moreover, 76 genes were mutated in 3 or more patients, thus more clearly defining the prevalence and gene targets of recurrent mutational events as well as high-confidence loss-of-function variants. Overall, the

numbers of nonsilent mutations, affected genes, and mutational signatures were similar across both subtypes. Multiple algorithms were used to predict driver mutations in 15 genes, which were generally detected across subtypes. The 12 GD-LGL samples did not reveal any unique mutational events from 93 T-LGL samples. Most studies that rely on *STAT3* molecular profiling for mutational subgroup analysis in LGL leukemia target the SH2 domain. We found mutations in the coiled-coil and DNA-binding domains in 3.8% of the cohort. Our functional assays determined that the novel mutations in the coiled-coil domain of *STAT3* are also activating and should be included in any subgroup analysis.

We observed multiple mutational and transcriptomic landscape differences between *STAT3*-mutated and -WT groups and clinical differences between different *STAT3* point mutations that previous studies were underpowered to detect. Thus, we have refined the molecular characterization of this disease and identified genes that may contribute to lymphoproliferation, such as recurrent mutations in the death domain of *FAS* and hotspot *PIK3R1* mutations. *PDGFRB* was strongly overexpressed at the mRNA and protein levels in *STAT3*-mutated samples, providing additional insight into our previous observation.⁴⁵

Our mutational signature analysis showed that a SBS1 replicative signature has a higher relative contribution toward the somatic mutation profile in *STAT3*-mutated samples, indicating increased cell division in those samples relative to *STAT3* WT. Unlike acute myeloid leukemia, where age is the most significant predictor of mutation burden, both *STAT3* status and the CD3⁺CD57⁺ fraction are better predictors of mutational burden in LGL. CD57 expression in T lymphocytes is a well-recognized marker of replicative senescence (clonal exhaustion).⁴⁸ In summary, our analysis indicates that the increased mutational burden in *STAT3*-mutant patients may be associated with lymphocyte replication stress.

The frequent co-occurrence of *STAT3* mutations with mutations in epigenetic modifiers in LGL leukemia is an important theme emerging from this study. Others recently reported that *STAT3*-mutated patients have high protein levels of epigenetic regulator *DNMT1* and are characterized by global hypermethylation.⁴⁹ In our study, all *TET2* mutations were found in the same predicted clone as *STAT3* if *STAT3* mutations were present. *KMT2D* and *SETD1B* mutations were associated with a *STAT3*-mutated clone approximately half the time. Since our original discoveries in LGL leukemia,^{8,9,12,50} *STAT3* or *STAT5B* activation/mutation has been a prominent pathogenic mechanism in T-cell leukemia/lymphomas.⁵¹⁻⁵⁵ However, to our knowledge, concurrent epigenetic modifier mutations in patients with predominant *STAT3* mutation have only been observed in a rare T-cell lymphoma, breast implant-associated anaplastic large cell lymphoma.⁵⁶ The significant co-occurrence of loss-of-function mutations in *KMT2D* and *TET2* with *STAT3* suggests a unique cooperative relationship in LGL leukemia.

Previously, we reported that dysregulation of Fas-mediated apoptosis is a fundamental pathogenic mechanism in LGL leukemia.⁵⁷ Genetic lesions in the death domain of Fas were previously not detected in studies that included 5 and 7 patient PBMC samples, respectively.^{57,58} Somatic *FAS* mutations, clustered mainly in the death domain, are observed in autoimmune lymphoproliferative syndrome type 3, which is characterized by

elevated numbers of double-negative TCR $\alpha\beta$ CD3⁺ CD4⁻ CD8⁻ T cells.^{35,36} These mutations impair normal FAS signaling, leading to accumulation of autoreactive lymphocytes.⁵⁹ Here, we discovered that T-LGL leukemia cells also harbor mutations in the death domain of FAS. FAS mutation was present in a patient harboring *STAT3*, *KMT2D*, and *TNFAIP3* mutations, and 2 patients with a concurrent *STAT3* mutation. Somatic FAS mutations in T-LGL leukemia provide another etiologic mechanism for LGL leukemic cells to resist apoptosis, persist in the circulation, and accumulate to elevated numbers.

Our discovery of *PIK3R1* mutations in T-LGL (K567E) and GD-LGL (N564D), combined with a previously reported mutation in NK-LGL (S565IinsEYREIDKRMNS),¹⁸ identifies a hotspot of mutations in *PIK3R1* in all 3 LGL subtypes. This hotspot region resides in the binding domain of the p110 α subunit of PI3K, and mutations in this region lead to its activation.^{60,61} *PIK3CA* encodes p110 α , and activating mutations of *PIK3CA* are found in other cancers and reported to be oncogenic.^{60,62} Notably, heterozygous germline mutations affecting the inter-SH2 domain of *PIK3R1*, where LGL leukemia mutations were observed, led to activated PI3 kinase δ syndrome 2 that is characterized by prominent expansion of peripheral blood CD8⁺ T cells.⁶³⁻⁶⁶ *PIK3R1* mutation is also reported in other cancers, with the highest frequency in endometrial and pancreatic cancer.⁶⁷ Additionally, activation of the PI3K-Akt pathway promotes resistance to apoptosis in LGL leukemia.⁶⁸ Furthermore, additional PI3K family mutations in our cohort (supplemental Table 7), including *PIK3CA* and *PIK3CD*, provide further support for the potential role of this pathway in T-LGL leukemia, consistent with our previous report of *PIK3CD* and *PIK3AP1* mutations in NK-LGL.²¹ The PI3K-Akt pathway is a central hub for many signaling pathways and is known to be activated in LGLs relative to normal controls.⁶⁹ Our data suggest decreased transcriptional activation of the pathway in *STAT3*-mutant leukemic samples that exhibit other mechanisms of PI3K-Akt activation at the protein level. Heterogeneity of PI3K-Akt signaling in patient samples (Figure 7) limits the ability to draw firm conclusions and highlights the need for additional studies in a larger cohort. In sum, observance of somatic mutations in PI3K family genes and increased phospho-protein levels in select patients demonstrate multiple mechanisms of PI3K pathway activation in both WT and *STAT3*-mutant patients.

Previous studies have described an association between *STAT3* mutation status and lower ANC values.^{8,14,15,17} However, the association of *STAT3* mutation with indicators of anemia has not been reported consistently.^{8,14-17} Here, we found that the most common Y640F mutation was associated with lower ANC values, and N647I mutation was associated with lower HGB values. Unexpectedly, patients with the most common somatic activating mutation, Y640F, exhibited similar HGB and HCT values to *STAT3* WT in our large cohort with 52 *STAT3*-WT, 31 Y640F, 9 D661Y, and 11 N647I patients. This observation may explain why general *STAT3* mutation status is inconsistently associated with anemia in previous studies that did not define or separate specific *STAT3* mutation groups. Consistent with this result, male patients harboring D661Y, but not Y640F, mutations were previously associated with higher mean corpuscular volume and lower HGB values.¹⁷ Exact mechanisms to explain how

different *STAT3* mutations lead to such striking and selective differences in blood parameters remain to be elucidated. Further, bone marrow pathology was similar between *STAT3*-mutated patients and WT patients, including when *STAT3* mutation types were stratified.

Our use of nonsorted PBMC to generate whole-exome and RNA-seq data are a potential limitation of this study. Even though the median leukemic fraction as assessed by clinical flow cytometry in the study is 66%, and the samples were sequenced to an average coverage of 200-fold, this approach has limitations in its ability to characterize the various minor clones. In a sample with a leukemic fraction of 66% and an average coverage of 200-fold, somatic mutations with a mean expected allelic fraction of 0.03 with power 0.8 can be detected. The approach of using bulk PBMC samples may also detect somatic variants within nonlymphocyte compartments. For example, 2 putative driver genes in this study, *KMT2D* and *TET2*, have been proposed to reflect early stages of hematopoiesis, resulting in mutation detection within myeloid as well as NK-LGL leukemia compartments.²⁰ However, the precise definition of clonal compartments and evolution in LGL leukemia, and potential interplay with clonal hematopoiesis of indeterminate potential mutations, awaits more detailed single-cell and longitudinal analyses. Finally, an additional potential concern is that gene expression profiles may represent signals that are not autonomous to the leukemic LGLs. However, we modeled the leukemic fraction along with gender and batch differences so that the gene expression reflected sample purity in our analyses. Importantly, we directly confirmed overexpression of representative genes identified by RNA-seq in purified CD8⁺ cells from LGL patients.

In 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. These findings suggest the possibility that mutations in nonexonic regions of the genome, epigenetic mechanisms, or germline involvement could contribute to pathogenesis and are worthy of future research. 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.

Acknowledgments

The authors thank Bryna Shemo, Andrea Hines, Rachel Stidham, Matthew Schmachtenberg, Wendy Dunton, and Shubha Dighe for LGL Leukemia Registry and technical support. They also thank Kristine C. Olson, Jun Yang, and Mary Poss for their critical reading of the manuscript and feedback.

LGL leukemia research in the Loughran laboratory is supported by the National Cancer Institute, National Institutes of Health grants R01CA178393 and P30CA044579 (T.P.L.); T32LM012416, T32GM007267, and T32CA009109 (H.C., J.C.X., and J.U.); and F30CA225046 (J.C.X.). Additional support was provided by the Bess Family Charitable Fund, the LGL Leukemia Foundation, Charles and Katharine Hutton Tweedy, William J. Branch, Szabolcs Szentpetery, Robert R. Wagner Fellowship (H.C.), and a generous anonymous donor (T.P.L.).

Authorship

Contribution: H.C., J.C.X., R.C.H., T.L.O., S.-F.T., D.J.F., A.R., and T.P.L. conceived the study; H.C., D.J.F., and A.R. wrote the manuscript; H.C. and A.R. made the figures; H.C., J.C.X., K.B.M., J.U., V.W.C., and A.R. analyzed results; H.C., J.C.X., K.B.M., J.U., V.W.C., D.S.C., M.F.T., O.E., J.S.W., and C.E.H. designed experiments and collected data; and all authors have edited and approved the manuscript.

Conflict-of-interest disclosure: T.P.L. is on the scientific advisory board and has stock options for Keystone Nano, Bioniz Therapeutics, and Dren Bio. T.P.L. and D.J.F. received honoraria from Kymera Therapeutics. D.J.F. has research funding from AstraZeneca. There are no conflicts of interest with the work presented in this manuscript. The remaining authors declare no competing financial interests.

ORCID profiles: H.C., 0000-0003-0149-9651; J.C.X., 0000-0003-0696-1390; K.B.M., 0000-0001-6736-4392; J.U., 0000-0002-3275-0530; V.W.C., 0000-0002-6834-3013; D.S.C., 0000-0002-3861-7693; C.E.H., 0000-0003-0077-7840; R.C.H., 0000-0003-4084-7516; T.L.O., 0000-0001-7906-205X; D.J.F., 0000-0003-4981-1691; A.R., 0000-0002-0782-3056.

Correspondence: Thomas P. Loughran Jr, University of Virginia Cancer Center, PO Box 800334, Charlottesville, VA 22908-0334; e-mail: tl7cs@virginia.edu.

Footnotes

Submitted 2 July 2021; accepted 18 December 2021; prepublished online on *Blood* First Edition 11 January 2022. DOI 10.1182/blood.2021013164.

Presented in abstract form at the 62nd annual meeting of the American Society of Hematology, 5 December 2020.

The whole exome and RNA sequencing data can be accessed in the NCBI dbGaP database under accession number phs002359.v1.p1. Additional patient selection, whole-exome and RNA-seq analysis, STAT3 reporter assay, western blotting, quantitative reverse transcription polymerase chain reaction (RT-qPCR), and statistical analysis details are presented in supplemental Methods. Requests for data sharing may be submitted to Thomas P. Loughran Jr at tl7cs@virginia.edu.

The online version of this article contains a data supplement.

There is a *Blood* Commentary on this article in this issue.

The publication costs of this article were defrayed in part by page charge payment. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 USC section 1734.

REFERENCES

- Loughran TP Jr, Kadin ME, Starkebaum G, et al. Leukemia of large granular lymphocytes: association with clonal chromosomal abnormalities and autoimmune neutropenia, thrombocytopenia, and hemolytic anemia. *Ann Intern Med*. 1985; 102(2):169-175.
- Lamy T, Moignet A, Loughran TP Jr. LGL leukemia: from pathogenesis to treatment. *Blood*. 2017;129(9):1082-1094.
- Cheon H, Dzieluska KH, Moosic KB, et al. Advances in the diagnosis and treatment of large granular lymphocytic leukemia. *Curr Hematol Malig Rep*. 2020;15(2):103-112.
- Swerdlow SH, Campo E, Pileri SA, et al. The 2016 revision of the World Health Organization classification of lymphoid neoplasms. *Blood*. 2016;127(20):2375-2390.
- Chen Y-H, Chadburn A, Evens AM, et al. Clinical, morphologic, immunophenotypic, and molecular cytogenetic assessment of CD4-/CD8- $\gamma\delta$ T-cell large granular lymphocytic leukemia. *Am J Clin Pathol*. 2011; 136(2):289-299.
- Sandberg Y, Almeida J, Gonzalez M, et al. TCR $\gamma\delta$ large granular lymphocyte leukemias reflect the spectrum of normal antigen-selected TCR $\gamma\delta$ T-cells. *Leukemia*. 2006;20(3):505-513.
- Lamy T, Loughran TP Jr. How I treat LGL leukemia. *Blood*. 2011;117(10):2764-2774.
- Koskela HLM, Eldfors S, Ellonen P, et al. Somatic STAT3 mutations in large granular lymphocytic leukemia. *N Engl J Med*. 2012; 366(20):1905-1913.
- Jerez A, Clemente MJ, Makishima H, et al. STAT3 mutations unify the pathogenesis of chronic lymphoproliferative disorders of NK cells and T-cell large granular lymphocyte leukemia. *Blood*. 2012;120(15):3048-3057.
- Fasan A, Kern W, Grossmann V, Haferlach C, Haferlach T, Schnittger S. STAT3 mutations are highly specific for large granular lymphocytic leukemia. *Leukemia*. 2013;27(7):1598-1600.
- Moosic KB, Paila U, Olson KC, et al. Genomics of LGL leukemia and select other rare leukemia/lymphomas. *Best Pract Res Clin Haematol*. 2019;32(3):196-206.
- Rajala HLM, Eldfors S, Kuusanmäki H, et al. Discovery of somatic STAT5b mutations in large granular lymphocytic leukemia. *Blood*. 2013;121(22):4541-4550.
- Andersson EI, Tanahashi T, Sekiguchi N, et al. High incidence of activating STAT5B mutations in CD4-positive T-cell large granular lymphocyte leukemia. *Blood*. 2016; 128(20):2465-2468.
- Barilà G, Teramo A, Calabretto G, et al. Stat3 mutations impact on overall survival in large granular lymphocyte leukemia: a single-center experience of 205 patients. *Leukemia*. 2012;34(4):1116-1124.
- Kawakami T, Sekiguchi N, Kobayashi J, et al. STAT3 mutations in natural killer cells are associated with cytopenia in patients with chronic lymphoproliferative disorder of natural killer cells. *Int J Hematol*. 2019; 109(5):563-571.
- Teramo A, Barilà G, Calabretto G, et al. STAT3 mutation impacts biological and clinical features of T-LGL leukemia. *Oncotarget*. 2017;8(37):61876-61889.
- Olson KC, Moosic KB, Jones MK, et al. Large granular lymphocyte leukemia serum and corresponding hematological parameters reveal unique cytokine and sphingolipid biomarkers and associations with STAT3 mutations. *Cancer Med*. 2020; 9(18):6533-6549.
- Gasparini VR, Binatti A, Coppe A, et al. A high definition picture of somatic mutations in chronic lymphoproliferative disorder of natural killer cells. *Blood Cancer J*. 2020; 10(4):42.
- Coppe A, Andersson EI, Binatti A, et al. Genomic landscape characterization of large granular lymphocyte leukemia with a systems genetics approach. *Leukemia*. 2017; 31(5):1243-1246.
- Pastoret C, Desmots F, Drillet G, et al. Linking the KIR phenotype with STAT3 and TET2 mutations to identify chronic lymphoproliferative disorders of NK cells. *Blood*. 2021;137(23):3237-3250.
- Olson TL, Cheon H, Xing JC, et al. Frequent somatic TET2 mutations in chronic NK-LGL leukemia with distinct patterns of cytopenias. *Blood*. 2021;138(8):662-673.
- Teramo A, Barilà G, Calabretto G, et al. Insights into genetic landscape of large granular lymphocyte leukemia. *Front Oncol*. 2020;10:152.
- Raess PW, Cascio MJ, Fan G, et al. Concurrent STAT3, DNMT3A, and TET2 mutations in T-LGL leukemia with molecularly distinct clonal hematopoiesis of indeterminate potential. *Am J Hematol*. 2017;92(1): E6-E8.
- Beaubier N, Tell R, Huether R, et al. Clinical validation of the Tempus xO assay. *Oncotarget*. 2018;9(40):25826-25832.
- McKenna A, Hanna M, Banks E, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*. 2010; 20(9):1297-1303.
- McLaren W, Gil L, Hunt SE, et al. The ensembl variant effect predictor. *Genome Biol*. 2016;17(1):122.
- Kandath C, Gao J, Qwangmsk, et al. Mskcc/Vcf2Maf: Vcf2Maf V1.6.16. *Zenodo*. 2018.
- Rosales RA, Drummond RD, Valieris R, Dias-Neto E, da Silva IT. signeR: an empirical

- Bayesian approach to mutational signature discovery. *Bioinformatics*. 2017;33(1):8-16.
29. Alexandrov LB, Jones PH, Wedge DC, et al. Clock-like mutational processes in human somatic cells. *Nat Genet*. 2015;47(12):1402-1407.
30. Martínez-Jiménez F, Muñíos F, Sentís I, et al. A compendium of mutational cancer driver genes. *Nat Rev Cancer*. 2020;20(10):555-572.
31. Johansson P, Bergmann A, Rahmann S, et al. Recurrent alterations of TNFAIP3 (A20) in T-cell large granular lymphocytic leukemia. *Int J Cancer*. 2016;138(1):121-124.
32. Mathelier A, Zhao X, Zhang AW, et al. JASPAR 2014: an extensively expanded and updated open-access database of transcription factor binding profiles. *Nucleic Acids Res*. 2014;42(D1):D142-D147.
33. Alam H, Tang M, Maitiuheti M, et al. KMT2D deficiency impairs super-enhancers to confer a glycolytic vulnerability in lung cancer. *Cancer Cell*. 2020;37(4):599-617.e7.
34. Zhang J, Dominguez-Sola D, Hussein S, et al. Disruption of KMT2D perturbs germinal center B cell development and promotes lymphomagenesis. *Nat Med*. 2015;21(10):1190-1198.
35. Holzelova E, Vonarbourg C, Stolzenberg M-C, et al. Autoimmune lymphoproliferative syndrome with somatic Fas mutations. *N Engl J Med*. 2004;351(14):1409-1418.
36. Dowdell KC, Niemela JE, Price S, et al. Somatic FAS mutations are common in patients with genetically undefined autoimmune lymphoproliferative syndrome. *Blood*. 2010;115(25):5164-5169.
37. Ler LD, Ghosh S, Chai X, et al. Loss of tumor suppressor KDM6A amplifies PRC2-regulated transcriptional repression in bladder cancer and can be targeted through inhibition of EZH2. *Sci Transl Med*. 2017;9(378):eaai8312.
38. Kobatake K, Ikeda K-I, Nakata Y, et al. *Kdm6a* deficiency activates inflammatory pathways, promotes M2 macrophage polarization, and causes bladder cancer in cooperation with p53 dysfunction. *Clin Cancer Res*. 2020;26(8):2065-2079.
39. van Haaften G, Dalgliesh GL, Davies H, et al. Somatic mutations of the histone H3K27 demethylase gene UTX in human cancer. *Nat Genet*. 2009;41(5):521-523.
40. Lawrence MS, Stojanov P, Mermel CH, et al. Discovery and saturation analysis of cancer genes across 21 tumour types. *Nature*. 2014;505(7484):495-501.
41. Van der Meulen J, Sanghvi V, Mavrakis K, et al. The H3K27me3 demethylase UTX is a gender-specific tumor suppressor in T-cell acute lymphoblastic leukemia. *Blood*. 2015;125(1):13-21.
42. Xu K, Liu B, Ma Y. The tumor suppressive roles of ARHGAP25 in lung cancer cells. *OncoTargets Ther*. 2019;12:6699-6710.
43. Thuault S, Comunale F, Hasna J, et al. The RhoE/ROCK/ARHGAP25 signaling pathway controls cell invasion by inhibition of Rac activity. *Mol Biol Cell*. 2016;27(17):2653-2661.
44. Shaver AC, Seegmiller AC, Strickland SA, et al. Mutational burden in acute myeloid leukemia is largely age dependent. *Blood*. 2015;126(23):2605.
45. Yang J, Liu X, Nyland SB, et al. Platelet-derived growth factor mediates survival of leukemic large granular lymphocytes via an autocrine regulatory pathway. *Blood*. 2010;115(1):51-60.
46. Cornejo MG, Kharas MG, Werneck MB, et al. Constitutive JAK3 activation induces lymphoproliferative syndromes in murine bone marrow transplantation models. *Blood*. 2009;113(12):2746-2754.
47. Koo GC, Tan SY, Tang T, et al. Janus kinase 3-activating mutations identified in natural killer/T-cell lymphoma. *Cancer Discov*. 2012;2(7):591-597.
48. Wood KL, Twigg HL III, Doseff AI. Dysregulation of CD8+ lymphocyte apoptosis, chronic disease, and immune regulation. *Front Biosci*. 2009;14:3771-3781.
49. Kim D, Park G, Huuhtanen J, et al. STAT3 activation in large granular lymphocyte leukemia is associated with cytokine signaling and DNA hypermethylation. *Leukemia*. 2021;35(12):3430-3443.
50. Epling-Burnette PK, Liu JH, Catlett-Falcone R, et al. Inhibition of STAT3 signaling leads to apoptosis of leukemic large granular lymphocytes and decreased Mcl-1 expression. *J Clin Invest*. 2001;107(3):351-362.
51. Takemoto S, Mulloy JC, Cereseto A, et al. Proliferation of adult T cell leukemia/lymphoma cells is associated with the constitutive activation of JAK/STAT proteins. *Proc Natl Acad Sci USA*. 1997;94(25):13897-13902.
52. Chen J, Zhang Y, Petrus MN, et al. Cytokine receptor signaling is required for the survival of ALK- anaplastic large cell lymphoma, even in the presence of JAK1/STAT3 mutations. *Proc Natl Acad Sci USA*. 2017;114(15):3975-3980.
53. Ohgami RS, Ma L, Merker JD, Martinez B, Zehnder JL, Arber DA. STAT3 mutations are frequent in CD30+ T-cell lymphomas and T-cell large granular lymphocytic leukemia. *Leukemia*. 2013;27(11):2244-2247.
54. Kiel MJ, Velusamy T, Rolland D, et al. Integrated genomic sequencing reveals mutational landscape of T-cell lymphocytic leukemia. *Blood*. 2014;124(9):1460-1472.
55. Netchiporouk E, Litvinov IV, Moreau L, Gilbert M, Sasseville D, Duvic M. Deregulation in STAT signaling is important for cutaneous T-cell lymphoma (CTCL) pathogenesis and cancer progression. *Cell Cycle*. 2014;13(21):3331-3335.
56. Laurent C, Nicolae A, Laurent C, et al. Gene alterations in epigenetic modifiers and JAK-STAT signaling are frequent in breast implant-associated ALCL. *Blood*. 2020;135(5):360-370.
57. Lamy T, Liu JH, Landowski TH, Dalton WS, Loughran TP Jr. Dysregulation of CD95/CD95 ligand-apoptotic pathway in CD3(+) large granular lymphocyte leukemia. *Blood*. 1998;92(12):4771-4777.
58. Perzova R, Loughran TP Jr. Constitutive expression of Fas ligand in large granular lymphocyte leukaemia. *Br J Haematol*. 1997;97(1):123-126.
59. Sneller MC, Straus SE, Jaffe ES, et al. A novel lymphoproliferative/autoimmune syndrome resembling murine *lpr/gld* disease. *J Clin Invest*. 1992;90(2):334-341.
60. Gymnopoulos M, Elsliger M-A, Vogt PK. Rare cancer-specific mutations in PIK3CA show gain of function. *Proc Natl Acad Sci USA*. 2007;104(13):5569-5574.
61. Jaiswal BS, Janakiraman V, Kljavin NM, et al. Somatic mutations in p85alpha promote tumorigenesis through class IA PI3K activation. *Cancer Cell*. 2009;16(6):463-474.
62. Cancer Genome Atlas Research Network. Comprehensive genomic characterization defines human glioblastoma genes and core pathways [published correction appears in *Nature*. 2013;494(7438):506]. *Nature*. 2008;455(7216):1061-1068.
63. Lucas CL, Zhang Y, Venida A, et al. Heterozygous splice mutation in PIK3R1 causes human immunodeficiency with lymphoproliferation due to dominant activation of PI3K. *J Exp Med*. 2014;211(13):2537-2547.
64. Deau M-C, Heurtier L, Frange P, et al. A human immunodeficiency caused by mutations in the PIK3R1 gene. *J Clin Invest*. 2014;124(9):3923-3928.
65. Michalovich D, Nejentsev S. Activated PI3 kinase delta syndrome: from genetics to therapy. *Front Immunol*. 2018;9:369.
66. Wentink M, Dalm V, Lankester AC, et al. Genetic defects in PI3K δ affect B-cell differentiation and maturation leading to hypogammaglobulinemia and recurrent infections. *Clin Immunol*. 2017;176:77-86.
67. Thorpe LM, Yuzugullu H, Zhao JJ. PI3K in cancer: divergent roles of isoforms, modes of activation and therapeutic targeting. *Nat Rev Cancer*. 2015;15(1):7-24.
68. Schade AE, Powers JJ, Wlodarski MW, Maciejewski JP. Phosphatidylinositol-3-phosphate kinase pathway activation protects leukemic large granular lymphocytes from undergoing homeostatic apoptosis. *Blood*. 2006;107(12):4834-4840.
69. Hoxhaj G, Manning BD. The PI3K-AKT network at the interface of oncogenic signalling and cancer metabolism. *Nat Rev Cancer*. 2020;20(2):74-88.