The Genetic Basis of Hepatosplenic T Cell Lymphoma

Matthew McKinney, Duke University
Andrea B. Moffitt, Duke University
Philippe Gaulard, Hôpital Henri Mondor
Marion Travert, Hôpital Henri Mondor
Laurence De Leval, CHUV Lausanne
Alina Nicolae, National Institutes of Health
Mark Raffeld, National Institutes of Health
Elaine S. Jaffe, National Institutes of Health
Stefania Pittaluga, National Institutes of Health
Liqiang Xi, National Institutes of Health

Only first 10 authors above; see publication for full author list.

Journal Title: Cancer Discovery
Volume: Volume 7, Number 4
Publisher: American Association for Cancer Research | 2017-04-01, Pages 369-379
Type of Work: Article | Post-print: After Peer Review
Publisher DOI: 10.1158/2159-8290.CD-16-0330
Permanent URL: https://pid.emory.edu/ark:/25593/s8vs3

Final published version: http://dx.doi.org/10.1158/2159-8290.CD-16-0330

Copyright information:
© 2017 American Association for Cancer Research.
Accessed September 15, 2019 8:20 PM EDT
The Genetic Basis of Hepatosplenic T Cell Lymphoma

A full list of authors and affiliations appears at the end of the article.

Abstract

Hepatosplenic T cell lymphoma (HSTL) is a rare and lethal lymphoma; the genetic drivers of this disease are unknown. Through whole exome sequencing of 68 HSTLs, we define recurrently mutated driver genes and copy number alterations in the disease. Chromatin modifying genes including SETD2, INO80 and ARID1B were commonly mutated in HSTL, affecting 62% of cases. HSTLs manifest frequent mutations in STAT5B (31%), STAT3 (9%), and PIK3CD (9%) for which there currently exist potential targeted therapies. In addition, we noted less frequent events in EZH2, KRA5 and TP53. SETD2 was the most frequently silenced gene in HSTL. We experimentally demonstrated that SETD2 acts as a tumor suppressor gene. In addition, we found that mutations in STAT5B and PIK3CD activate critical signaling pathways important to cell survival in HSTL. Our work thus defines the genetic landscape of HSTL and implicates novel gene mutations linked to HSTL pathogenesis and potential treatment targets.

Keywords

exome sequencing; genetic mutations; copy number; lymphoma; hepatosplenic T cell lymphoma; HSTL; peripheral T cell lymphoma; angioimmunoblastic T cell lymphoma; diffuse large B cell lymphoma; mantle cell lymphoma; Burkitt lymphoma; FFPE exome sequencing; deep sequencing; high throughput sequencing; germinal center B cell; DERL2; DERL7; SETD2; survival; JAK-STAT; PIK3CD; STAT5B; STAT3; isochromosome

INTRODUCTION

Hepatosplenic T cell lymphoma (HSTL) is a rare type of peripheral T cell lymphoma (PTCL) that affects fewer than 2% of all lymphoma patients each year. Compared to other PTCL subtypes, HSTL is striking for its predilection for young individuals and a dismal prognosis; most HSTL patients succumb to their disease in less than a year (1,2).

While the vast majority of PTCLs arise from αβ T cells, HSTL arises predominantly (>80%) from γδ T cells. Immunosuppression is a major risk factor for HSTL, particularly in the setting of organ transplantation and exposure to thiopurines or TNF-α antagonists, but most HSTL cases arise sporadically in the absence of any known factors. Most HSTL patients are treated with combination chemotherapy and, when feasible, bone marrow
transplantation, but relapses are frequent and the overwhelming majority of patients succumb to HSTL.

The most frequent known genetic abnormalities in HSTL are isochromosome 7q (iso7q) (3,4) and trisomy 8 (5,6), but the role of somatic mutations and other genomic alterations in HSTL has yet to be defined. Application of next generation sequencing in other peripheral T cell lymphomas (7–18) has revealed a strikingly heterogeneous genetic landscape implicating over a dozen different genetic alterations including novel therapeutic targets. Similar systematic studies in HSTLs have been lacking.

In this study, we sought to define the genetic landscape of HSTL through whole exome sequencing of 68 HSTL cases, including 20 cases with paired germline DNA. We found that HSTLs have a distinct pattern of genetic alterations. In addition to iso7q and trisomy 8, HSTLs manifested frequent mutations in the genes SETD2, INO80, STAT5B, SMARCA2, TET3, and PIK3CD. Interestingly, mutations that occur frequently in other T cell lymphomas in genes such as RHOA, CD28 and CCR4 were notably absent or occurred at much lower frequency in HSTLs. Finally, through functional studies in HSTL cells, we demonstrate SETD2 to be a novel tumor suppressor gene that mediates increased cellular proliferation in HSTL. In addition, we found that mutations in STAT5B and PIK3CD activate critical signaling pathways important to cell survival in HSTL.

To our knowledge, this study is the largest series of HSTL patients ever described and defines in detail the genetic landscape of mutations, implicating a number of novel genes and molecular pathways in this disorder.

RESULTS

Whole exome sequencing of HSTL identifies gene coding mutations and chromosomal copy number alterations

We performed whole exome sequencing on 68 primary HSTL tumors, 20 with paired normal tissue, and 2 HSTL cell lines (i.e. total 90 exomes). The bioinformatics methods are described more comprehensively in the Supplementary Methods. The 20 cases with paired normal tissue were deemed a “discovery set” and the remaining samples, a “validation set”. Only genes found to be somatically mutated in the discovery set were initially considered as candidate driver genes in HSTL.

As expected, HSTL patients in our study had a dismal prognosis with a median survival of 11.9 months following the diagnosis. The majority (63%) of the patients died within two years of diagnosis (Figure 1a). HSTL driver genes are depicted as a heat map in Figure 1b. The distribution of these gene mutations in the different HSTL cases is shown in Figure 1b and the frequency of these events is shown in Figure 1c. Chromatin modifying genes constituted the most frequently mutated group of genes in HSTL, accounting for 62% of the cases. SETD2 was the most commonly mutated chromatin modifying gene in HSTL (24 mutations found in 17 patients) with 71% (12/17) of the cases manifesting at least one loss-of-function (nonsense or frameshift) mutation. SETD2 showed strong selection for protein altering mutations, with only one synonymous mutation observed. Other frequently mutated
chromatin modifier genes included \textit{INO80} (21\%), \textit{TET3} (15\%), and \textit{SMARCA2} (10\%). Mutations in signaling pathways comprised the next most common group of mutations and manifested as predominantly missense mutations in \textit{STAT5B}, \textit{STAT3}, and \textit{PIK3CD}. In addition, we observed mutations in \textit{TP53}, \textit{UBR5} and \textit{IDH2}. Among these gene mutations, only those in \textit{STAT5B} and \textit{STAT3} have been described in HSTLs previously \cite{19,20}. We verified the accuracy of genetic variant identification from our deep sequencing data by performing Sanger sequencing in 78 individual variant events from the same cases (Supplementary Table S1). Representative plots from Sanger validation are shown in Supplementary Figure S1a–b. We found that Sanger sequencing agreed with our methods in 88\% of the variants, confirming that our laboratory methods for deep sequencing and computational methods for identifying genetic variants generated accurate results. We identified 13 somatically mutated HSTL driver genes (Supplementary Table S2) using a model that relies on the frequencies of non-synonymous events in HSTL, the size of the gene, the rate of non-synonymous variation occurring the in gene in healthy controls and the predicted impact of the altered amino acid(s) as we have described previously \cite{21–23}.

Estimates of cancer cell fraction based on observed allele frequencies for each driver gene are shown in Supplementary Figure S2a. \textit{STAT3}, \textit{PIK3CD}, and \textit{SETD2} show the highest proportion of clonal events among the driver genes. We compared cancer cell fractions (allele frequency adjusted for tumor purity) for pairs of mutations occurring in the same sample (Supplementary Figure S2b), and found considerable heterogeneity in the genes that arose as predominant drivers in different cases. In addition, we examined mutation data for the presence of well-characterized driver mutations, even in the absence of a clear somatic event in the discovery set. We identified rare mutations in a number of known driver genes including \textit{KRAS} (N=3) and \textit{CREBBP} (N=2) that have been described in other cancers. Additional somatic mutations found and putative driver mutations are shown in Supplementary Table S3.

We further examined chromosomal alterations in HSTLs (Figure 1d). Alterations in chromosome 7 were the most common genetic abnormality in HSTLs affecting 47\% (Supplementary Table S4, Supplementary Figures S3) of our cases. While the majority of cases manifested clearly as iso7q with losses of chromosome 7p and gains of 7q, there were five cases with only 7q amplifications. Supplementary Figure S3 shows the exact regions of chromosome 7 altered in all the affected cases, which are complete arm level alterations in the vast majority of cases. Trisomy 8 or amplifications of 8q were the next most frequent chromosomal aberration (31\%), frequently co-occurring with chromosome 7 alterations (18 overlapping cases). We also found losses in chromosome 10q (19\%) and gains in chromosome 1q (13\%) occurred in a significant proportion of HSTL cases. The only driver gene that is located within these copy number regions is \textit{UBR5}, on chromosome 8q. We did not find any mutations that clearly overlapped or occurred independently of these chromosomal alterations. The median number of genetic alterations per sample in these driver mutation and copy number events was three (Figure 1e).

We also sequenced the only described HSTL cell lines DERL2 and DERL7, which are derived from the same primary HSTL tumor \cite{24}. Their genetic profiles were nearly identical, and representative of the HSTL tumors, with mutations in \textit{STAT5B}, \textit{ARID1B}, \textit{Cancer Discov. Author manuscript; available in PMC 2018 April 01.
*SMARCA2* and *TP53*, as well as copy number alterations of isochromosome 7q, trisomy 8, loss of 10p, and gain of 1q (Supplementary Figure S4a–b).

**Clinical Characteristics of HSTL patients**

The patients’ gender, T cell receptor type and response to initial therapy are shown for comparison to the molecular features in Figure 1f. The clinical characteristics of our patients are summarized in Table 1.

Briefly, the median age was 42 (range 4–72). As expected, male patients comprised the majority (71%) of the cohort with a worse prognosis than females (p=0.05, log rank test). Elevated LDH levels and ECOG performance status (p=0.002 for both, log rank test) are the clinical variables most clearly associated with outcome. 80% of cases were of the γδ T cell receptor (TCR) type, but no survival association was seen with the TCR type. Patients with a complete response (CR) to initial treatment and those patients who received an autologous or allogenic transplant following initial treatment had significantly improved outcomes (p=0.001 and 0.014, respectively). Supplementary Table S5 contains complete clinicopathologic data, and Supplementary Figure S5 shows the Kaplan-Meier plots for these and additional clinical variables.

We also investigated the association of genetic mutations and copy number alterations with clinical outcome in an exploratory fashion. The most robust survival associations within the genetic features were found with the copy number alterations on chromosome 7 and 8 (p=0.051 and p=0.024, respectively), which were both associated with worse outcomes. Kaplan-Meier plots for these molecular covariates are shown in Supplementary Figure S6. Additional survival analyses are presented in the Supplementary Methods.

These data add clarity to other studies that have postulated links between clinicopathologic or molecular features and patient level outcomes in HSTL (Supplementary Table S6). Supplementary Table S7 shows the similarity of our patient cohort to those previously reported.

**Genetic differences between HSTL and other T cell and B cell lymphomas**

HSTL has a distinct clinical presentation and response to therapy compared to other B and T cell lymphomas. We next investigated whether HSTLs are also genetically distinct from these tumors. We compared the frequencies of the most recurrently mutated genes in HSTL to other lymphomas studied by whole exome sequencing including the T cell lymphomas PTCL not otherwise specified (PTCL NOS) (8), angioimmunoblastic T cell lymphoma (AITL) (8,12), cutaneous T cell lymphoma (CTCL) (13,16–18), adult T cell lymphoma (ATL) (15), anaplastic large cell lymphoma (ALCL) (14), and NK T cell lymphoma (NKTCL) (10), as well as the B cell lymphomas diffuse large B cell lymphoma (DLBCL) (23,25,26), Burkitt lymphoma (BL) (21,27,28) and mantle cell lymphoma (MCL) (22,29). The mutational frequencies for these lymphoma types are shown in Figure 1g and Supplementary Table S8, and additional HSTL mutation details shown in Supplementary Table S9.
We found that mutations in *SETD2, INO80, TET3* and *STAT5B* occurred almost exclusively in HSTLs (p<0.05 in each case, Fisher’s exact test), compared to other T and B cell lymphoma types. On the other hand, mutations in *RHOA*, a defining feature of PTCL NOS and AITL, never occurred in HSTL cases. Additionally, HSTLs shared some mutations that have mostly been observed in B cell lymphomas including *PIK3CD* (23), *UBR5* (30) and *SMARCA2* (22). These data indicate that HSTLs have a unique profile of genetic mutations that distinguishes them from other non-Hodgkin lymphoma subtypes and may underlie its distinct clinical behavior.

**Functional consequences of *SETD2* silencing in HSTL**

*SETD2* was the most frequently silenced gene in HSTLs, affecting almost a third of the cases. The mutations in *SETD2* occurred predominantly in exons 3, 10–12 and 19–21 (Figure 2a). The Set2-Rpb1 interacting domain (SRI) domain at the COOH-terminus of the *SETD2* protein product was the most frequently mutated protein domain, but a number of mutations were also interspersed among different domains. We noted that 67% (16/24) of the mutations in *SETD2* occurring in the 17 HSTL patients were either frameshift or nonsense mutations. 44% (7/17) of the *SETD2* mutated patients had more than one mutation in *SETD2*. We analyzed the potential biallelic nature of the *SETD2* mutations in one of these patients, by analysis of paired-end exome sequencing reads and Sanger sequencing. We found that the two *SETD2* mutations in that case occurred on mutually exclusive DNA strands suggesting a biallelic event (Supplementary Figure S7). Our genetic data thus indicate that *SETD2* has a tumor suppressor function in HSTLs with frequent loss-of-function mutations that may occur in biallelic fashion. No expression differences in *SETD2* were seen between wildtype and *SETD2* mutant samples (Supplementary Figure S8).

*SETD2* functions as a methyltransferase at lysine 36 of histone 3 and 4, catalyzing H3/4K36 trimethylation at dimethyl sites. The gene has diverse functions related to cancer including tumor proliferation, sensing and repair of DNA damage, and other oncogenic changes in gene expression (32–37). To determine the functional consequences of *SETD2* loss in HSTLs, we performed RNA interference experiments using lentiviral vectors expressing two different shRNAs targeting *SETD2* in DERL2 HSTL cells. We observed robust knockdown of *SETD2* mRNA expression in HSTL cells expressing the shRNAs targeting *SETD2* compared to non-silencing (scramble) control (Figure 2b). We further observed that loss of *SETD2* mRNA expression was accompanied by reduced *SETD2* protein expression along with loss of H3 lysine 36 trimethylation in the same cells (H3K36me3, Figure 2c).

We further measured the downstream effects of *SETD2* knockdown by performing RNAseq to define gene expression profiles of HSTL DERL2 cells with knockdown of *SETD2*. We compared the expression profiles of cells with *SETD2* knockdown through two different shRNAs and those expressing the non-silencing control shRNA in three independent replicates (Figure 2d). Through gene set enrichment analysis (38), we identified the pathways that were upregulated (FDR <0.25) in the setting of reduced *SETD2* expression (Supplementary Table S10). We found that eight of the ten most significantly altered gene sets were related to proliferation or cell cycle progression (including G2M checkpoint genes
and core cell cycle genes as shown in Figure 2d), identifying proliferation as the major oncogenic process affected by SETD2 loss in HSTL cells.

We next tested the proliferative ability of HSTL cells with intact and reduced SETD2 expression by performing colony-forming assays in cells expressing shRNAs targeting SETD2 as well as the non-silencing control. We found that the cells post SETD2 knockdown had more than two-fold increased colony formation (p<0.01, paired Wilcoxon-rank test) compared to control non-silencing cells with intact SETD2 expression (Figure 2e). We also examined the effects of SETD2 on cell proliferation using alamar blue cell quantitation assays and found that, compared to non-silencing control (Figure 2f), the HSTL cells expressing lower SETD2 have significantly higher proliferation and increased cell counts (p<0.01, paired Wilcoxon-rank test), consistent with its role as a tumor suppressor gene.

Our genetic and functional data thus implicate SETD2 as a novel tumor suppressor gene in HSTL and demonstrates that loss of function SETD2 mutations serve to increase proliferation in HSTL cells.

**STAT5B, PIK3CD mutations and downstream signaling**

STAT5B and STAT3 mutations have been previously reported in HSTL (19,20). We found somatic STAT5B and STAT3 mutations occurring mostly in a mutually exclusive manner (Supplementary Figure S9). Mutations in STAT5B and STAT3 occurred predominantly in their SH2 domains, consistent with previous reports showing mutations in this gene in HSTL and other malignancies. Many of the PIK3CD mutations (Figure 3a) occurred in analogous fashion to constitutively activating mutations in homologous regions of PIK3CA, which activates PI-3 kinase/AKT signaling in non-hematologic cancers (39).

Given this pattern of mutations, we hypothesized that STAT5B, STAT3, PIK3CD mutations in HSTL would constitutively activate downstream signaling pathways, and that mutant STAT5B and PIK3CD may potentially cooperate to maintain proliferation pathways within HSTL cells. To test this hypothesis, we first examined the effects of the STAT5B mutations through their overexpression in serum starved 293T cells (the HSTL cell line DERL2 is STAT5B-mutant). We found that the dominant hotspot mutations N642H and V712E were particularly efficacious in maintaining STAT5 phosphorylation, as were many of the other STAT5B mutants compared to the wild type or empty vector control cells (Figure 3b). Phosphorylated AKT or total AKT were not altered in these cells, indicating the STAT5B mutations directly enable phosphorylation of STAT5 that has been shown to have a number of downstream oncogenic effects (19).

We further investigated the impact of PIK3CD mutations on signaling in HSTL cells. We overexpressed three PIK3CD mutants (R38C, K111Q and N334K) found in HSTL tumors in DERL2 HSTL cells. In DERL2 cells overexpressing PIK3CD mutants, we found greatly increased phosphorylated AKT. In addition, we noted increased phosphorylated STAT5 (pSTAT5) compared to cells with control constructs (Figure 3c), indicating potential cooperativity between PI-3 kinase and JAK-STAT signaling in maintaining downstream cell survival/proliferation signals.
Next, we tested the ability of PI-3 kinase mutants to protect DERL2 HSTL cells from cell death under IL-2 deprivation/serum starvation conditions. After 24 hours of serum starvation and cytokine withdrawal, DERL2 cells bearing mutant PIK3CD constructs had significantly higher viability compared to those expressing wild type or empty vector controls (Figure 3d). These data confirm that mutant PIK3CD could maintain cell viability and proliferation.

Finally, we tested the effects of PI-3 kinase and STAT5B signaling inhibition as a potential targeted strategy for treating HSTL. In DERL2 HSTL cells, we tested the effects of a MAP kinase inhibitor (selumetinib, an off-target control) as well as STAT5B inhibitor (with CAS 285986-31-4), PI-3 kinase inhibitor (idelalisib), and combinations of the STAT5B and PI-3 kinase inhibitors. We found that the MAP kinase inhibitor had no effect on cell growth, but the STAT5B inhibitor significantly attenuated cellular growth in vitro at 96 hours versus vehicle treatment control (Figure 3e). Notably, the addition of PI-3 kinase inhibition to STAT5B inhibition resulted in further reduction of cellular viability than either STAT5B or PI-3 kinase inhibition alone, indicating potential synergism interaction when targeting STAT5B/PI-3 kinase cooperativity in HSTL cells.

DISCUSSION

HSTL remains among the direst diagnoses any cancer patient can face. The young age of the patients and the general lack of progress in the treatment options or outcomes further compounds the grimness of the prognosis (40). A recognized risk factor for developing HSTL is the use of immune suppressive and immune modulatory therapies such as infliximab and azathioprine (41). The growing use of these therapies in rheumatologic and inflammatory conditions suggests that HSTL might become an increasingly larger concern in the future.

Our work indicates that the distinct clinical behavior of HSTL may originate from its distinct genetic profile. Compared to other T cell lymphomas, mutations in genes such as RHOA, CD28, CCR4 and other genes that are defining features of other T cell lymphomas are notably absent in HSTLs. Conversely, the most frequent HSTL mutations including SETD2, INO80, and PIK3CD, occur infrequently in other T cell lymphomas. STAT5B mutations have been described in primary cutaneous γδ PTCL (19), but are not mutated in most other T cell lymphomas derived from αβ T cells. Interestingly, STAT3 mutations are a frequent feature of a number of lymphomas including NKTCL (10) and diffuse large B cell lymphoma (25).

SETD2, a histone lysine methyltransferase, was the most frequently silenced gene in HSTLs. SETD2 has been described as a tumor suppressor gene in other malignancies including clear cell renal carcinoma (42), DLBCL (23), and in a subset of acute leukemias (43). Our genetic and experimental data identify SETD2 as a novel tumor suppressor gene in HSTL. Our experiments indicate that SETD2 mutations serve to increase proliferation in HSTL tumors. More work is needed to define the precise role of SETD2 in HSTL oncogenesis as well as its potential role in regulating T cell development.
The recurrent genetic mutations in HSTL also suggest new therapeutic targets in the disease with new hope for these patients. The activating mutations in PI3KCD and STAT signaling genes provide a framework for identifying new therapies in this disease. The rareness of the disease makes it unlikely that a comprehensive clinical trial can be performed; thus the clinical experience with potential off-label use of these targeted therapies should be carefully reported to inform future therapies in HSTL patients.

Thus, our data define the genetic landscape of HSTLs and implicates novel gene mutations linked to HSTL pathogenesis and therapeutic targets.

METHODS

Study design and sample selection

HSTL tumors and normal tissues were obtained from institutions that constitute the Hematologic Malignancies Research Consortium (44) and Tenomic Consortium (45). All cases were reviewed to verify the accuracy of the pathology diagnosis. Tumor samples were derived from FFPE tissues, and normal samples were from unaffected bone marrow, when available. Archival patient tumor and normal samples, as well as clinical data, were collected according to a protocol approved by the Duke University Institutional Review Board, which was exempt from informed consent, and in accordance with the Declaration of Helsinki.

Exome sequencing and derivation of genomic data

Briefly, genomic DNA was sheared to 250 bp and size/concentration were verified using Bioanalyzer (Agilent Technologies). Sheared DNA was end-repaired, A-tailed, and ligated to Illumina paired-end adapters. The resulting libraries were amplified using Illumina PE specific primers. Samples were column purified and the size and quantity of the final libraries was determined using Bioanalyzer (Agilent Technologies). The resulting libraries were then hybridized for 24 hours to DNA baits provided in the SureSelect Human All Exon 50MB kit (Agilent Technologies). The captured libraries were amplified and sequenced using the Illumina platform. FASTQs were aligned to the hg19 reference genome using BWA-MEM. Indel realignment and base quality recalibration are done with GATK tools. Variants were called using Mutect. Variants were annotated with Annovar. Average sequencing depth throughout samples was 80x. Additional bioinformatic pipeline details as well as detailed, complete supplemental methods are in the Supplementary Methods.

Cell lines and cell culture

Human embryonal kidney 293T cells were obtained from the American Type Culture Collection (ATCC, obtained Sept. 2013) and maintained as a monolayer in DMEM with 10% fetal bovine serum. Modified 293T cells (GP2 cells from Clontech, obtained June 2012) were cultured in identical conditions. DERL2 and DERL7 cells (24) were obtained from Dr. Philippe Gaulard in Oct. 2013 and subjected to exome sequencing to confirm isochromosome 7q and their other known genetic markers. Cell lines were not otherwise authenticated and assumed to be authentic as they were obtained from the original supplier. Unless otherwise noted, DERL2 and DERL7 HSTL cells were cultured in 20% human AB
positive serum in 50 ng/mL IL2 (Peprotech). For gene overexpression experiments serum starvation was performed by washing cells with PBS twice then re-introducing cells to media without serum or cytokines followed by signaling or viability experiments.

**Western blotting**

For each sample, $10^6$ cells were washed once in ice cold PBS and lysed in 150 microliters of RIPA buffer (Sigma) containing protease, phosphatase inhibitors (Roche) and 2 mM EDTA. Crude lysates were sonicated at 10% duty cycle, intensity 4, 200 cycles/burst on a Covaris S-series device then centrifuged at 13,000 g for 10 minutes and supernatants were used for downstream experiments. 10–30 micrograms of protein were separated on 4–18% Bis-Tris gradient gels and transferred overnight to Immobilon PVDF membranes (Millipore) in Tris-glycine buffer with 20% methanol at 30 V. Primary antibodies (anti-SETD2, anti-H3K36me3, anti-STAT3, anti-phospho-STAT3, anti-STAT5, anti-phospho-STAT5, anti-phospho-AKT, anti-GAPDH, all from Cell Signaling, and anti-p110 from abCam) were diluted 1:1000 in 5% BSA/TBS-T and bound overnight followed by washing and binding of HRP conjugated anti-rabbit antibody (Santa Cruz Biotechnology). Blots were developed using SuperSignal Pico chemiluminescent substrate (Thermo Scientific).

**RNA Knockdown**

Lentiviral pTRIPZ constructs (Open Biosystems) containing either a scrambled non-silencing sequence as control, or a hairpin sequence targeting the SETD2 gene, were mixed with pVSV-G (46) and pPAX2 (Addgene #12259, from Didier Trono) and transfected into TLA-HEK-293T cells. Virus was harvested 72 hours post-transfection and concentrated by centrifugation at 90,000 g for 90 minutes at 4°C. DERL2 HSTL cells were infected with virus in the presence of 4 mg/mL of polybrene, using $5\times10^6$ cells per well in a 6-well plate by centrifugation at 1000 g for 60 minutes. Cells were selected by puromycin for 96 hours; then shRNA production was induced by doxycycline (1 g/mL replaced every 48 hours) for 96 hours prior to downstream experimentation. Successful knockdown of SETD2 was confirmed by quantitative PCR (qPCR) and Western blot. RNA sequencing methods and derivation of gene expression profiles of HSTL cells with SETD2 or control shRNAs is described in the Supplementary Methods.

**Lentiviral gene overexpression**

mRNA from DERL2 HSTL cells was extracted and a corresponding cDNA library was used to amplify full length PIK3CD and STAT5B sequences which were ligated downstream of the CMV promoter at the PacI restriction site of the pQCXIP retroviral vector using PacI restriction sites on the flanking cloning primer used for gene amplification. cDNA constructs in pQCXIP plasmids were verified using Sanger sequencing and mutagenesis to yield desired mutant constructs was performed by PCR-technique using the NEB mutagenesis kit. Mutant cDNA expression clones were again verified by Sanger sequencing. Lentiviral particles were produced via pQCXIP co-transfection with VSV-G plasmid in GP2 293 cells and virus was harvested and cells transfected as described above of lentiviral shRNA constructs.
**Colony forming assays**

DERL2 HSTL cells, exposed to doxycycline for >96 hours, were plated in 1000/mL in IMDM-based methylcellulose semi-solid media containing 50 ng/mL IL2, 25 ng/mL SCF and 1 mg/mL doxycycline. Cells were cultured for 4 weeks at 37°C and 5% CO2 10% of media volume containing 10x concentrations of IL2, SCF and doxycycline was added every 5 days. Once visible colonies formed, plates were photographed, and colonies were identified and counted.

**Cellular proliferation assays**

To measure cellular proliferation over time, cells were plated in 96-well plates, 5,000 cells/well, in 220 μl of media in replicates with conditions varied according to the experimental design. Quantitation of the number of cells at each desired timepoint was accomplished by measured by adding 25 μl alamar blue to replicate wells and measuring fluorescence (excitation 430 nM, emission 490 nM) on a Tecan plate reader after 6 hours incubation of alamar blue with cells. Fluorescence values obtained in this manner were normalized to values obtained by adding alamar blue at time = 0 or at the commencement of the experiment in order to obtain values relative to proliferation over time.

**Pharmacologic treatment of cell lines**

The viability of DERL2 HSTL cells exposed to small molecule inhibitors was performed as previously described (47) with the exception that alamar blue fluorescence was measured instead of absorbance values post MTT and detergent incubation. Selumetinib was obtained from Selleck chemicals, STAT5B inhibitor (CAS 285985-31-4) was obtained from Millipore and idelalisib was obtained from Gilead biosciences. Controls included carrier control and 10% DMSO control for 100% and 0% proliferation controls, respectively. After 3 days of incubation, 25 μl of alamar blue reagent was added to the cells, which were then incubated at 37 °C for 12h. Fluorescence values at 590 nm were assessed with a plate reader device (Tecan). Normalized proliferation (% viable cells) for each well was calculated as Normalized proliferation = (well absorbance − 10% DMSO control absorbance)/(carrier control absorbance − 10% DMSO control) or values were normalized to fluorescence values at baseline for cell line or drug assay experiments.

**Statistics**

Fisher’s exact test was used to compare mutation frequencies between lymphoma types and between subgroups of HSTL. The Cox proportional hazards model was used to test significance in survival associations with clinical and molecular covariates. Survival curves were drawn using the Kaplan-Meier method. Individual methods used for analysis of clinical, pathologic and mutational features and survival are listed in the Supplementary Methods.

**Data Availability**

Primary sequence data are available at the European Genome-phenome Archive (EGA) under accession number EGAS00001002182.
Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Authors

Matthew McKinney\textsuperscript{1,*}, Andrea B. Moffitt\textsuperscript{2,*}, Philippe Gaulard\textsuperscript{3}, Marion Travert\textsuperscript{3}, Laurence De Leval\textsuperscript{4}, Alina Nicola\textsuperscript{e}\textsuperscript{5}, Mark Raffeld\textsuperscript{5}, Elaine S. Jaffe\textsuperscript{5}, Stefania Pittaluga\textsuperscript{5}, Liqiang Xi\textsuperscript{5}, Tayla Heavican\textsuperscript{6}, Javeed Iqbal\textsuperscript{6}, Karim Belhadj\textsuperscript{3}, MH Delfau-Larue\textsuperscript{3}, Virginie Fataccioli\textsuperscript{3}, Magdalena B. Czader\textsuperscript{7}, Izidore S. Lossos\textsuperscript{8}, Jennifer R. Chapman-Fredricks\textsuperscript{9}, Kristy L. Richards\textsuperscript{9}, Yuri Fedoriw\textsuperscript{9}, Sarah L. Ondrejka\textsuperscript{10}, Eric D. Hsi\textsuperscript{10}, Lawrence Low\textsuperscript{11}, Dennis Weisenburger\textsuperscript{11}, Wing C. Chan\textsuperscript{11}, Neha Mehta-Shah\textsuperscript{12}, Steven Horwitz\textsuperscript{12}, Leon Bernal-Mizrachi\textsuperscript{13}, Christopher R. Flowers\textsuperscript{13}, Anne W. Beaven\textsuperscript{1}, Mayur Parihar\textsuperscript{14}, Lucile Baseggio\textsuperscript{15}, Marie Parrens\textsuperscript{16}, Anne Moreau\textsuperscript{17}, Pierre Sujobert\textsuperscript{18}, Monika Pilichowska\textsuperscript{19}, Andrew M. Evens\textsuperscript{19}, Amy Chadburn\textsuperscript{20}, Rex K.H. Au-Yeung\textsuperscript{21}, Gopesh Srivastava\textsuperscript{21}, William W. L. Choi\textsuperscript{21}, John R. Goodlad\textsuperscript{22}, Igor Aurer\textsuperscript{23}, Sandra Basic-Kinda\textsuperscript{23}, Randy D. Gascoyne\textsuperscript{24}, Nicholas S. Davis\textsuperscript{1}, Guojie Li\textsuperscript{1}, Jenny Zhang\textsuperscript{1}, Deepthi Rajagopalan\textsuperscript{1}, Anupama Reddy\textsuperscript{1}, Cassandra Love\textsuperscript{1}, Shawn Levy\textsuperscript{25}, Yuan Zhuang\textsuperscript{1}, Jyotishka Datta\textsuperscript{26}, David B. Dunson\textsuperscript{26}, and Sandeep S. Dave\textsuperscript{1,2}

Affiliations

\textsuperscript{1}Duke Cancer Institute, Duke University Medical Center, Durham NC, USA \textsuperscript{2}Duke Center for Genomics and Computational Biology, Duke University, Durham NC, USA \textsuperscript{3}Hôpital Henri Mondor, Department of Pathology, AP-HP, Créteil, France, INSERM U955, Créteil, France, and University Paris-Est, Créteil, France \textsuperscript{4}Pathology Institute, CHUV Lausanne, Switzerland \textsuperscript{5}Laboratory of Pathology, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA \textsuperscript{6}University of Nebraska, Omaha, NE, USA \textsuperscript{7}Indiana University, Indianapolis IN, USA \textsuperscript{8}University of Nebraska, Omaha, NE, USA \textsuperscript{9}University of North Carolina, Chapel Hill, NC USA \textsuperscript{10}Cleveland Clinic, Cleveland OH, USA \textsuperscript{11}City of Hope Medical Center, Duarte CA, USA \textsuperscript{12}Memorial Sloan Kettering Cancer Center, New York, NY USA \textsuperscript{13}Emory University, Atlanta GA, USA \textsuperscript{14}Tata Medical Center, Kolkata, India \textsuperscript{15}Centre Lyon-Sud, Pierre-Bénite, France \textsuperscript{16}Hôpital Pessac, Bordeaux, France \textsuperscript{17}Pathology, Hôpital Hôtel-Dieu, Nantes, France \textsuperscript{18}Faculté de Médecine Lyon-Sud Charles Mérieux, Université Claude Bernard, Lyon, France \textsuperscript{19}Tufts University Medical Center, Boston MA, USA \textsuperscript{20}Presbyterian Hospital, Pathology & Cell Biology, Cornell University, New York, NY \textsuperscript{21}University of Hong Kong, Queen Mary Hospital, Hong Kong, China \textsuperscript{22}Department of Pathology, Western General Hospital, Edinburgh, UK \textsuperscript{23}University Hospital Centre Zagreb, Zagreb, Croatia \textsuperscript{24}British Columbia Cancer Agency, University of British Columbia, Vancouver, Canada \textsuperscript{25}Hudson Alpha Institute for Biotechnology, Huntsville AL, USA \textsuperscript{26}Department of Statistical Science, Duke University, Durham NC, USA
Acknowledgments

This work was supported by Janssen and grants from the National Institutes of Health (CA195173 and CA136895). Matthew McKinney was supported by the Lymphoma Research Foundation and Andrea Moffitt was supported by the Hertz Foundation Fellowship and the National Science Foundation Graduate Research Fellowship.

References


30. McCoey et al. Cancer Discov. Author manuscript; available in PMC 2018 April 01.
STATEMENT OF SIGNIFICANCE

We report the first systematic application of whole exome sequencing to define the genetic basis of hepatosplenic T cell lymphoma (HSTL), a rare but lethal disease. Our work defines SETD2 as a novel tumor suppressor gene in HSTL and implicates novel genes in the disease including INO80 and PIK3CD.
Figure 1. Characterization of mutations, copy number, clinical data, and survival in HSTL cases
A. Kaplan-Meier curve for all HSTL cases with available survival data (n=47). Median survival is 11.9 months. Median follow-up is 4.2 years.
B. Heat map of mutated genes in HSTL (n=68). Each row represents a mutated gene in HSTL. Each column represents a patient sample. Blocks are color coded by functional type of mutation (orange: stopgain SNV; green: frameshift indel; purple: missense SNV; pink: nonframeshift indel; teal: synonymous). Samples are separated into discovery set (n=20)
with paired normal and validation set (n=48). Samples with more than one variant per gene are indicated with black dots on the block.

C. Number of cases affected per gene. Bars colored by most damaging mutation in each gene-sample pair.

D. Copy number alterations of chromosome 7p and 7q, 8p and 8q, 10p, and 1q. (light blue/dark blue: chromosome 7, light red/dark red: chromosome 8, dark green: chromosome 10p, yellow: chromosome 1q)

E. Number of events per sample, including significant mutations and arm level copy number changes.


Figure 2. Discovery and characterization of HSTL SETD2 mutations and function
A. Diagram of SETD2 gene exon model, protein domains, and HSTL mutations. Exon model shows 21 exons from the canonical transcript, which spread across 147 KB of genomic space. 24 SETD2 mutations are indicated at their amino acid position, color-coded by type of mutation (yellow: frameshift; red: nonsense; black: missense; blue: synonymous). Protein domains are indicated along the gene, color-coded by domain name (purple: AWS; pink: SET; green: PostSET; blue: Low Charge; orange: WW; dark red: SRI).
B. *SETD2* mRNA expression in *SETD2* shRNA knockdown and non-silencing control in DERL2. p <0.01 for both *SETD2* knockdowns vs. non-silencing (NS) control, average knockdown of >70%.

C. Immunoblot blot showing knockdown of *SETD2* and H3K36me3 loss after *SETD2* shRNA induction in DERL2 cells. Results representative of three repeat experiments.

D. Heatmap of genes differentially expressed in DERL2 HSTL after *SETD2* knockdown compared to non-silencing controls (left panel). Genes shown to have p<0.05 by Student’s t test of *SETD2* knockdown versus control. Right panel depicts two pathways significantly upregulated in *SETD2* knockdown based on gene set enrichment analysis (GSEA). Gene rank is based on t-statistic between control and knockdown samples. Enrichment score is shown on the y-axis as the gene list is traversed on the x-axis.

E. Results of methylcellulose colony formation assays at 15 days of incubation compared between two cell lines expressing *SETD2* shRNA constructs (left panel) or non-silencing (NS) control (p < 0.01 by two-tailed student’s t-test for comparison of colony number between control and each of the two *SETD2* constructs). The right panel depicts representative MTT counter-stained plates at day 15.

F. Quantitation of *in vitro* proliferation versus time for DERL2 HSTL cells bearing *SETD2* shRNA constructs versus non-silencing control. P=0.011 by two-tailed student’s t-test between shRNA #2 and NS control and P=0.0018 between SETD2 shRNA #1 and NS control.

All error bars throughout the figure represent the standard error of the mean.
Figure 3. STAT5B, STAT3 and PIK3CD mutations in HSTL
A. Protein domains and HSTL mutations in STAT5B, STAT3 and PIK3CD in HSTL tumors are shown, with each individual mutation event denoted by a shaded circle.
B. Immunoblots of phosphorylated and total STAT5B/AKT in 293T cells after overexpression of STAT5B HSTL mutant constructs (N642H, V712E, P702A, I704L, Q706L, Y665F, T628S, E438K, G492C), wild type, or GFP-overexpression/empty vector controls. β-tubulin was used as loading control. Blots are representative of three independent experiments.
C. Immunoblots of phosphorylated and total STAT5B and AKT in DERL2 HSTL cells bearing \textit{PIK3CD} mutant constructs (R38C, K111Q, N334K), wild type \textit{PIK3CD} retroviral expression construct, as well as empty vector retroviral vector transfection. GAPDH was used as loading control. Blots are representative of three independent experiments.

D. Cell viability measurements by alamar blue fluorescence quantitation of DERL2 HSTL cells bearing mutant (R38C, K111Q, N334K) or wild type \textit{PIK3CD} retroviral expression constructs (as well as empty vector retroviral vector transfection). *(p<0.05 for comparison between each mutant construct and wild type control by Student’s t-test).

E. DERL2 HSTL cell viability after treatment with MEK inhibitor (selumetinib 2.5 μM), PI3 kinase inhibitor (idelalisib 10 μM), STAT5B inhibitor (CAS 285985-31-4, 50 μM) and vehicle control (defined as media containing the maximum concentration of DMSO as diluent from drug experiments) or STAT5B/PI3-kinase inhibitor combination. *(p = 0.001 between vehicle and STAT5B inhibitor) and **(p = 0.044 between STAT5B inhibitor and combination inhibitor) at 96 hours of incubation.

All error bars throughout the figure represent the standard error of the mean.
# Table 1

## Summary of clinical variables and survival association

Clinical characteristics, the most common copy number aberrations, treatment, and outcome statistics for the cohort are summarized. The number of cases with each feature are listed, as well as a percentage of the cohort, excluding samples with missing data for that variable. Association with survival is tested with a univariate log-rank test and p-values are provided for the significant tests (p<0.1).

![Table](https://example.com/table1.png)

<table>
<thead>
<tr>
<th>Clinical Feature</th>
<th>Clinical/Pathological Characteristic</th>
<th>Number of Cases (Percent)</th>
<th>Log-Rank Univariate Test for Survival Association P-value (NS=not significant)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age</strong></td>
<td>Median (years)</td>
<td>42</td>
<td>NS</td>
</tr>
<tr>
<td></td>
<td>0–35 years</td>
<td>23 (42)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>35–50 years</td>
<td>16 (30)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>&gt;50 years</td>
<td>15 (28)</td>
<td></td>
</tr>
<tr>
<td><strong>Gender</strong></td>
<td>Male</td>
<td>48 (71)</td>
<td>0.054</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>20 (29)</td>
<td></td>
</tr>
<tr>
<td><strong>Race</strong></td>
<td>Caucasian</td>
<td>50 (74)</td>
<td>NS</td>
</tr>
<tr>
<td></td>
<td>African</td>
<td>13 (19)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Asian</td>
<td>4 (6)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hispanic</td>
<td>1 (1)</td>
<td></td>
</tr>
<tr>
<td><strong>ECOG Performance Status</strong></td>
<td>0–1</td>
<td>21 (62)</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>2–4</td>
<td>13 (38)</td>
<td></td>
</tr>
<tr>
<td><strong>LDH Level</strong></td>
<td>Elevated</td>
<td>24 (67)</td>
<td>0.002</td>
</tr>
<tr>
<td><strong>TCR Type</strong></td>
<td>Gamma-Delta</td>
<td>37 (80)</td>
<td>NS</td>
</tr>
<tr>
<td></td>
<td>Alpha-Beta</td>
<td>9 (20)</td>
<td></td>
</tr>
<tr>
<td><strong>Chr 7 Aberration</strong></td>
<td>Present</td>
<td>33 (49)</td>
<td>0.051</td>
</tr>
<tr>
<td></td>
<td>Absent</td>
<td>35 (51)</td>
<td></td>
</tr>
<tr>
<td><strong>Chr 8 Aberration</strong></td>
<td>Present</td>
<td>21 (31)</td>
<td>0.024</td>
</tr>
<tr>
<td></td>
<td>Absent</td>
<td>47 (70)</td>
<td></td>
</tr>
<tr>
<td><strong>Treatment</strong></td>
<td>Chemotherapy only</td>
<td>37 (93)</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>Autologous or Allogenic Transplant</td>
<td>10 (26)</td>
<td>0.014</td>
</tr>
<tr>
<td><strong>Response to Initial Treatment</strong></td>
<td>CR</td>
<td>11 (38)</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>PR/SD</td>
<td>7 (24)</td>
<td>(CR vs. non-CR)</td>
</tr>
<tr>
<td></td>
<td>NR/PD</td>
<td>11 (38)</td>
<td></td>
</tr>
<tr>
<td><strong>Survival</strong></td>
<td>Median Survival (months)</td>
<td>11.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Median Follow-Up (months)</td>
<td>49.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Alive at 6 months</td>
<td>35 (75)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Alive at 12 months</td>
<td>24 (50)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Alive at 24 months</td>
<td>17 (37)</td>
<td></td>
</tr>
</tbody>
</table>