Distinctive genomic features of human T-lymphotropic virus type 1-related adult T-cell leukemia-lymphoma in Western populations

Caroline S. Myers,[1] Eli Williams,[2] Carlos Barrionuevo Cornejo,[3] Georgios Pongas,[4] Ngoc L. Toomey,[4] Jose A. Sanches,[5] Maxime Battistella,[6] Samuel Mo,[1] Melissa Pulitzer,[7,8] Cristopher A. Moskaluk,[2] Govind Bhagat,[8,9] Kenneth Ofori,[9] Jonathan J. Davick,[10] Octavio Servitje,[11] Denis Miyashiro,[5] Fina Climent,[11] Kimberley Ringbloom,[1] Daniela Duenas,[3] Calvin Law,[1] Sandro Casavilca Zambrano,[3] Luis Malpica,[12] Brady E. Beltran,[13] Denisse Castro,[13] Luciana Barreto,[4,14] Carlos Brites,[15] Jennifer R. Chapman,[4] Jaehyuk Choi,[1#] Alejandro A. Gru[16#] and Juan C. Ramos[4#]

[1]Department of Dermatology, Northwestern University Feinberg School of Medicine, Chicago, IL, USA; [2]University of Virginia School of Medicine, Charlottesville, VA, USA; [3]Instituto Nacional de Enfermedades Neoplasicas, Lima, Peru; [4]University of Miami, Miami, FL, USA; [5]Universidade de São Paolo, São Paolo, Brazil; [6]Université de Paris, Paris, France; [7]Memorial Sloan-Kettering Cancer Center, New York, NY, USA; [8]New York-Presbyterian Hospital, New York, NY, USA; [9]Columbia University School of Medicine, New York, NY, USA; [10]University of Iowa, Iowa City, IA, USA; [11]Hospital Universitari de Bellvitge, Barcelona, Spain; [12]University of Texas MD Anderson Cancer Center, Houston, TX, USA; [13]Hospital Nacional Edgardo Rebagliati Martins, Lima, Peru; [14]Instituto Nacional de Câncer José Alencar Gomes da Silva, Rio de Janeiro, Brazil; [15]Federal University of Bahia, Salvador, Brazil and [16]Columbia University Irving Medical Center/New York-Presbyterian, New York, NY, USA

[#]JC, AAG and JCR contributed equally as senior authors.

Correspondence: A.A. Gru
aag2222@cumc.columbia.edu

Received: February 8, 2024.
Accepted: July 29, 2024.
Early view: August 8, 2024.

https://doi.org/10.3324/haematol.2024.285233

Abstract

Adult T-cell leukemia-lymphoma (ATLL) is an aggressive malignancy driven by human T-cell leukemia virus type 1 (HTLV-1). Although patients from the Western hemisphere (Afro-Caribbean and South American) face worse prognoses, our understanding of ATLL molecular drivers derives mostly from Japanese studies. We performed multi-omic analyses to elucidate the genomic landscape of ATLL in Western cohorts. Recurrent deletions and/or damaging mutations involving *FOXO3, ANKRD11, DGKZ*, and *PTPN6* implicate these genes as potential tumor suppressors. RNA-sequencing, published functional data and *in vitro* assays support the roles of *ANKRD11* and *FOXO3* as regulators of T-cell proliferation and apoptosis in ATLL, respectively. Survival data suggest that *ANKRD11* mutation may confer a worse prognosis. Japanese and Western cohorts, in addition to acute and lymphomatous subtypes, demonstrated distinct molecular patterns. *GATA3* deletion was associated with chronic cases with unfavorable outcomes. *IRF4* and *CARD11* mutations frequently emerged in relapses after interferon therapy. Our findings reveal novel putative ATLL driver genes and clinically relevant differences between Japanese and Western ATLL patients.

Introduction

Adult T-cell leukemia-lymphoma (ATLL) is an aggressive hematologic malignancy caused by the human T-cell leukemia virus type I (HTLV-1), which is endemic in South America, the Caribbean, western Africa, and southern Japan.[1,2] Clinically, ATLL is often characterized by lymphadenopathy with or without lymphocytosis, organomegaly, multi-organ involvement (more commonly skin), and immunosuppression.[3,4] It can be classified into at least four clinical subtypes. Acute and lymphomatous are by far the most common and lethal variants, while chronic and smoldering forms tend to behave indolently until they ultimately progress to more aggressive subtypes. A chronic variant with unfavorable features ("unfavorable chronic") presents with lymphocytosis and elevated levels of lactate dehydrogenase, has a worse prognosis, and progresses to acute subtypes in a shorter period of time.[5]
The prognosis of ATLL is dismal. The 4-year survival rates for lymphomatous and acute forms are less than 20%, with median survival less than 11 months.[6,7] First-line treatment

options include multi-agent chemotherapy, biologics such as the anti-CCR4 antibody mogamulizumab,[8] zidovudine, and interferon-α.[4] However, disease relapse occurs in nearly all patients, and even those who undergo allogeneic stem cell transplantation have a median survival of less than 6 months.[6] Afro-Caribbean ATLL patients present with distinct, more severe clinical features than their Japanese counterparts in published cohorts.[9,10] These include a younger age at diagnosis by >10 years and a worse overall survival. Despite the clinical impact of these discrepancies, the molecular features of ATLL have not been well characterized among patients in the Western hemisphere. Comprehensive studies of ATLL using genome-wide approaches have come primarily from Japan.[11-13] Studies in the Western hemisphere have been limited in number, size, and breadth.[14,15] These limitations are due in part to the broad geographic spread of ATLL in the Western hemisphere, encompassing HTLV-1 endemic areas that have not traditionally participated in research programs.[16] More generally, these features reflect a widespread underrepresentation of Afro-Caribbean and Hispanic patients in genomic research.[16,17] Less than 1% of donors in the International Cancer Genome Consortium hail from the Caribbean or South American countries.[18] The National Cancer Institute's Genomic Data Commons includes <10% of donors who report African ancestry and <5% who report Hispanic ethnicity.[19] Underrepresentation of these populations in genomic research affects both patients and researchers: it impedes patients' inclusion in the growing benefits of personalized medicine and withholds valuable information from our collective knowledge of cancer biology.

Our group sought to overcome these challenges by assembling the largest cohort of Western hemisphere ("Western") ATLL patients to date. We acquired samples from underserved and indigenous populations in South America, the Caribbean, and immigrant communities in the United States. We undertook genome-wide characterization of these patients' molecular features to gain an unparalleled view into the mechanistic basis of ATLL in Western patients. Our multimodal genomic study employed whole-exome sequencing, copy number variation (CNV) data and RNA-sequencing corroboration to seek novel driver genes. We then looked for molecular drivers of clinical phenotypes including variation in geographic regions, clinical subtypes, and response to therapies.

Our approach is the first to use exome-wide analysis to identify population-based differences between the molecular landscape of Japanese and Western ATLL. Furthermore, the inclusion of underrepresented populations in our genome-wide analysis uncovered novel driver gene candidates that affected apoptosis and T-cell proliferation *in vitro*. Finally, we examined the relationships between molecular features and clinical outcomes in ATLL patients and newly elucidated distinct molecular features characterizing acute and lymphomatous subtypes. Together, these foundational analyses illustrate global patterns of ATLL molecular features. They also yield a novel genetic perspective on Western ATLL, thus providing a basis for future preclinical and clinical investigations.

Methods

Patients' characteristics and sample collection

One hundred sixty-five patients with a confirmed diagnosis of ATLL were included in this study after quality control (*Online Supplementary Table S9*). Twelve patients contributed both pre- and post-relapse samples to this study. Specimens for molecular study were obtained from the blood of patients with leukemic presentation or from formalin-fixed, paraffin-embedded or frozen tissue of patients who presented with solid tumors. Immunohistochemistry was performed on representative sections of 32 patients' samples, with scoring performed by two independent blinded dermatopathologists. All patients' samples were collected under protocols approved by the local internal review boards of the participating institutions in accordance with the Declaration of Helsinki. The diagnostic criteria for ATLL, subtype classification, and methodologies for immunohistochemistry, nucleic acid extraction and sequencing can be found in the *Online Supplementary Methods*.

Statistical analysis

Where possible, statistical analyses used are indicated in the text. Survival data were analyzed using the R package Survival.[20] Mutual exclusivity analyses were performed using the R package Discover.[21] Numerical values (e.g., log fold change, RNA-sequencing counts) were compared between two groups using a Student two-tailed *t* test. Frequencies of mutations or clinical outcomes between groups were compared using the Fisher exact test. Specific methodologies applied for mutational, CNV and RNA-sequencing analyses can be found in the *Online Supplementary Methods*.

T-cell proliferation assay pipeline

Human T cells were isolated from enriched leukapheresis products and transfected with caspase 9-sgRNA ribonucleoprotein (crRNP) complex as previously described (*Online Supplementary Methods*). After CRISPR knockdown, cells were stimulated with plate-bound anti-human CD3/CD28 and interleukin-2 as described in the *Online Supplementary Methods*. On day 13, cells were stained with carboxyfluorescein succinimidyl ester (CFSE) (*Online Supplementary Methods*) and cultured for 4 days in complete RPMI medium without interleukin-2 and with or without stimulation. After 4 days, cell proliferation was assessed by fluorescent-activated cell sorting. The gating strategy, including positive controls, is illustrated in *Online Supplementary Figure S7*.

FOXO3 gene overexpression and knockdown constructs in patient-derived adult T-cell leukemia-lymphoma cell lines

ATLL-84c and ATLL-97c are clonally-proven ATLL cell lines

derived from tumor cells carrying the typical $CD4^+CD25^+$ ATLL phenotype (established in Ramos' laboratory). ATLL lines were transduced as described in the *Online Supplementary Methods*. Protein levels were verified by western blot, and the nucleotide composition of *FOXO3* mutant vectors was verified by DNA polymerase chain reaction and sequencing. Transduced cells were exposed to etoposide, belinostat and/or dimethylsulfoxide vehicle. Annexin V staining, as an indicator of apoptosis, was then evaluated with flow cytometry (*Online Supplementary Methods*).

Results

Multimodal analysis of the adult T-cell leukemia-lymphoma genomic landscape

We compiled a multimodal dataset from 165 Western ATLL patients (*Online Supplementary Figure S1A*). ATLL diagnoses were confirmed by histopathological findings, HTLV-1 serological assays, clonal T-cell populations as determined by immunophenotyping and gene rearrangement studies, and HTLV-1 polymerase chain reaction validation in skin biopsy cases of limited quantities. The geographic coverage of this cohort was broad: countries of origin were largely South American (Brazil, Peru, Panama, and Ecuador; N=76) and the Caribbean (Haiti, Jamaica, Trinidad, Dominican Republic, Bahamas, Antigua, Martinique, St. Vincent, Tortola, West Indies and US Virgin Islands; N=76). We also included patients of reported African descent from France (N=6) and Miami, Florida, USA (N=7). For subsequent data analysis, the patients' ethnicity was categorized based upon single nucleotide polymorphisms using EthSeq[22] (Table 1).

Data modalities used in this study included Oncoscan CNV, whole-exome sequencing and RNA sequencing. We combined these data with whole-exome sequencing and CNV data from 83 and 426 Japanese patients, respectively.[12] Consistent with the literature,[9] Western patients presented with disease at a younger age (51 *vs.* 65 years, $P=6.62x10^{-14}$,

Table 1. Demographic characteristics of the cohorts of patient.

Characteristic	Total Western	WES	Oncoscan	RNA	Japanese WES	Japanese CNV
Patients, N	165	122	128	92	83	426
Disease subtype, N						
Acute	84	63	70	39	39	194
Chronic with:						95
Favorable features	9	5	7	2	26	
Unfavorable features	17	15	11	13	0	
Lymphomatous	45	31	31	31	13	97
Smoldering	9	7	8	7	5	24
Unclassified	1	1	1	-	-	-
Sex, N						
Female	82	64	68	48	43	-
Male	83	58	60	44	40	-
Age in years, mean	51.48	51.73	51.42	52.18	65.23	-
Region, N						
Caribbean	76	69	61	38		
Europe	6	4	5	5	NA	NA
North America	7	7	7	3		
South America	76	42	55	46		
Ethnicity, N						
African	78	78	63	42		
Native South American	29	29	17	21		
South American/Asian	11	9	9	8	NA	NA
European	2	2	1	-		
European/South American	3	3	2	1		
South Asian Islander	42	1	36	3		
Unclassified	-	-	-	17		
Survival in weeks, mean	83.74	93.03	83.42	97.15	NA	NA

The patients' demographics are summarized ("total Western") as well as broken down by data type. Numbers represent distinct patients, some of whom may have contributed both pre- and post-relapse samples. The demographics of the Japanese cohort are also summarized (right two columns). Demographic Japanese copy number variation data were obtained from published descriptions[12] rather than in-house review. WES: whole-exome sequencing; CNV: copy number variation; NA: not available.

Student two-tailed *t* test) and with a greater proportion of aggressive (acute and lymphomatous) subtypes (78% *vs.* 63%, *P*=0.015, Fisher exact test).

We first sought to identify novel genomic drivers in our Western cohort using previously validated methods.[23-26] Our methods included multiple orthogonal analyses. First, we analyzed whole-exome sequencing data from 122 patients. We approached mutation calling conservatively, employing previously validated approaches to remove possible sources of sequencing or alignment errors or ambiguity[23-26] (*Online Supplementary Methods; Online Supplementary Figures S1* and *S2*). To further validate these data against independent datasets as well as to identify low-prevalence driver mutations shared with other cancers, we integrated our data with data from Japanese ATLL patients (N=83),[12] a published cohort of patients with T-cell lymphoma (N=699 patients),[24,25] and publicly available libraries of mutation data across cancer types (>1.4x10^6 tumor samples).[27]

The distribution of mutations within cancer driver genes is not random; we have leveraged these patterns to identify novel cancer-promoting mutations in other T-cell lymphomas.[23-26,28] Thus, we prioritized mutations that demonstrated patterns characteristic of oncogenes and tumor suppressors (*Online Supplementary Figure S2*). To identify putative oncogenes, we looked for characteristic recurrent "gain-of-function" mutations at amino acid hotspots. Twenty-seven genes harbored recurrent non-synonymous amino acid alterations (*Online Supplementary Tables S1* and *S2*). Tumor suppressors generally contain recurrent damaging mutations, i.e. mutations with a high likelihood of inducing loss of function (e.g., stop-gain, start-loss, or splice-site mutations). Thirteen genes had a statistically significant burden of damaging mutations (*Online Supplementary Table S3*).

Next, we analyzed exome-wide CNV data from 128 Western ATLL patients in combination with previously published data from 426 Japanese patients.[12,13] We chose Oncoscan, a clinically utilized and industry-validated commercial CNV assay, for analysis of our Western datasets.[29] In the combined Japanese and Western dataset, we identified 17 and 35 chromosomal regions subject to statistically significant rates of duplication or deletion, respectively, by methods previously described[24,25,30] (Figure 1A). Identifying potential gene targets of CNV can be difficult, in part because these CNV are, by nature, polygenic. To identify putative tumor suppressors and oncogenes in these recurrent CNV regions, we employed a previously validated stepwise hierarchical algorithm[23-26] (*Online Supplementary Figure S3*). This analysis identified a putative target gene in 39 chromosomal regions (*Online Supplementary Tables S4* and *S5*).

In total, we identified 64 putative ATLL driver genes distributed across 15 biologically relevant pathways (Figure 2A) (*Online Supplementary Table S6*). Twelve genes were significant by both orthogonal whole-exome sequencing and CNV analyses (*GATA3, CCR4, TP53, ARID2, CSNK2B, NOTCH1, CBLB, CD58, ANKRD11, IRF4, CARD11, CD28*) (Figure 1B-E). Twenty-six genes were implicated by point mutation analysis alone (*Online Supplementary Tables S1-S3*) and 26 by copy number mutation analysis alone (*Online Supplementary Tables S4* and *S5*). Our results confirmed 42 driver genes reported in Japanese ATLL populations.[12,13,31,32] Twenty-two genes were newly implicated in ATLL by this analysis (*Online Supplementary Table S6*). Four novel putative driver genes were mutated at frequencies of 10% or greater (*FOXO3, APC, WNK1, ANKRD11*). Collectively, mutations in the 22 novel putative driver genes were found in 80% of samples.

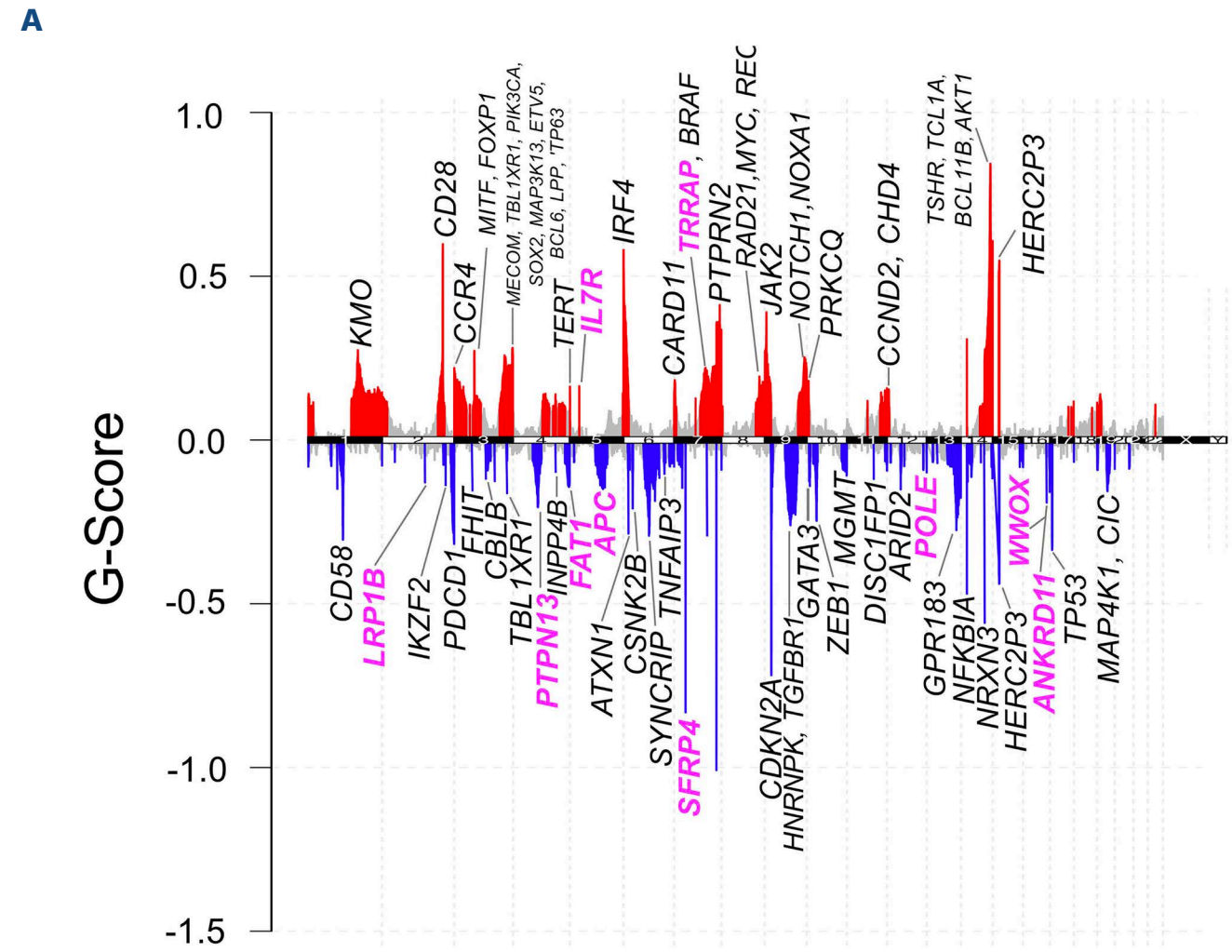
Of the 22 genes newly implicated in ATLL, 14 had mutations that have been functionally validated as driver genes in cell or animal models of other types of cancer. For an additional three genes, we identified new cancer-associated hotspots: in *WNK1* (p.Ala372), *RBBP4* (p.Arg131) and *FOXO3* (p.Asp199) (Figure 1G; *Online Supplementary Figure S4A, C*). Five genes had not been previously recognized as recurrently mutated in cancer (*CD3E, ANKRD11, DGKZ, PTPN6, RHOB*) (Figure 1B-F; *Online Supplementary Figure S4B*). Collectively, 44% of samples had at least one mutation in a gene not previously reported in cancer.

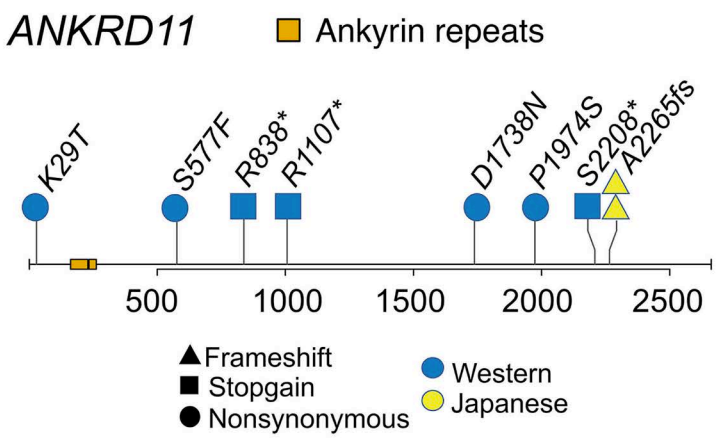
The pathways most commonly affected by driver gene mutations included CD28/PI3K-AKT signaling (*CD28, VAV1, PLCG1, PRKCQ, FOXO3*, and negative regulators *CBLB, INPP4B, PTPN6*), T-cell receptor (TCR)/NF-κB signaling (*CD3E, PLCG1, VAV1, PRKCQ, PRKCB, PTPRN2, CARD11, IKBKB, RLTPR, CSNK2B, IRF4*, and negative regulators *CBLB, PTPN6, DGKZ, NFKBIA, TNFAIP3, TRAF3*), and cell migration (*CCR7, CCR4, RHOA, RHOB, GPR183, NRXN3, VAV1, WNK1*). The TCR- and PI3K-mediated CD28 co-receptor pathways intersect at the activation of PKCθ (*PRKCQ*) via PCLγ1 (*PLCG1*) and PDK1, respectively, which connect proximal TCR and CD28 co-receptor signaling events, ultimately leading to NF-κB activation (Figure 2B).

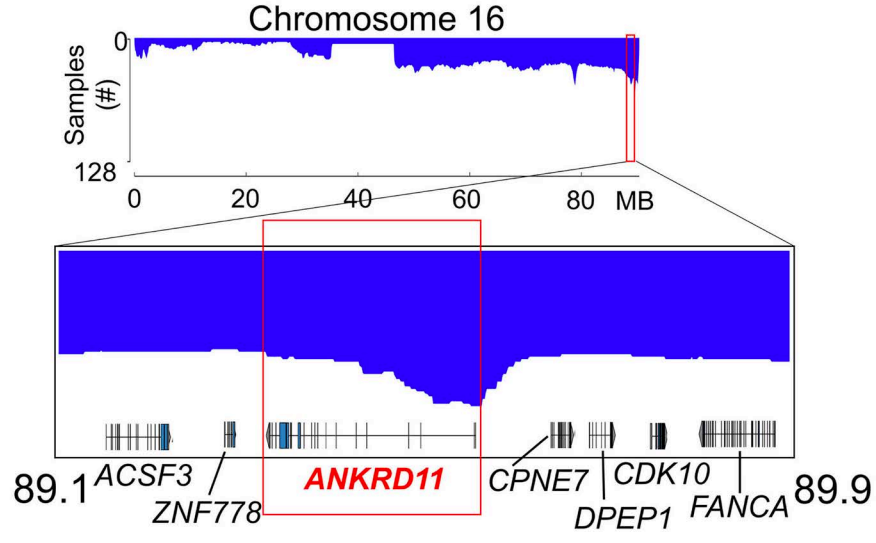
Genomic and functional validation of novel putative driver genes

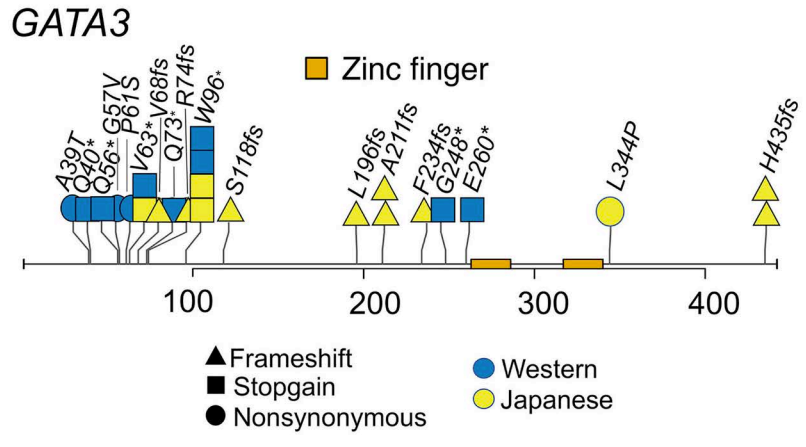
Our mutational analyses allowed us to identify both known and novel putative driver genes; however, their biological roles in ATLL are not implicit. For example, despite its recurrent damaging mutations (often characteristic of tumor suppressors[28]), *CCR4* has been shown to be an oncogene with functionally validated gain-of-function truncated variants.[33] For this reason, we classified putative driver genes as suspected oncogenes or tumor suppressors based upon their pattern of point mutations leveraged against CNV mutational patterns and published functional studies (*Online Supplementary Table S6*). We then sought to validate the functional consequences of these putative oncogenes or tumor suppressors *in vitro*.

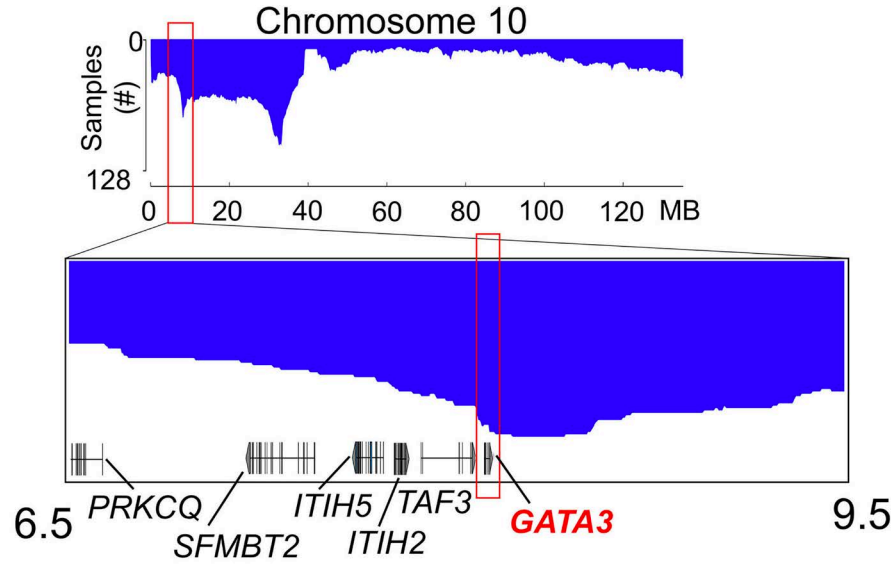
To do this, we first utilized previously published, publicly available genome-wide CRISPR screens for T-cell activation. These included two CRISPR interference screens for genes mediating TCR-independent cytokine production,











Continued on following page.

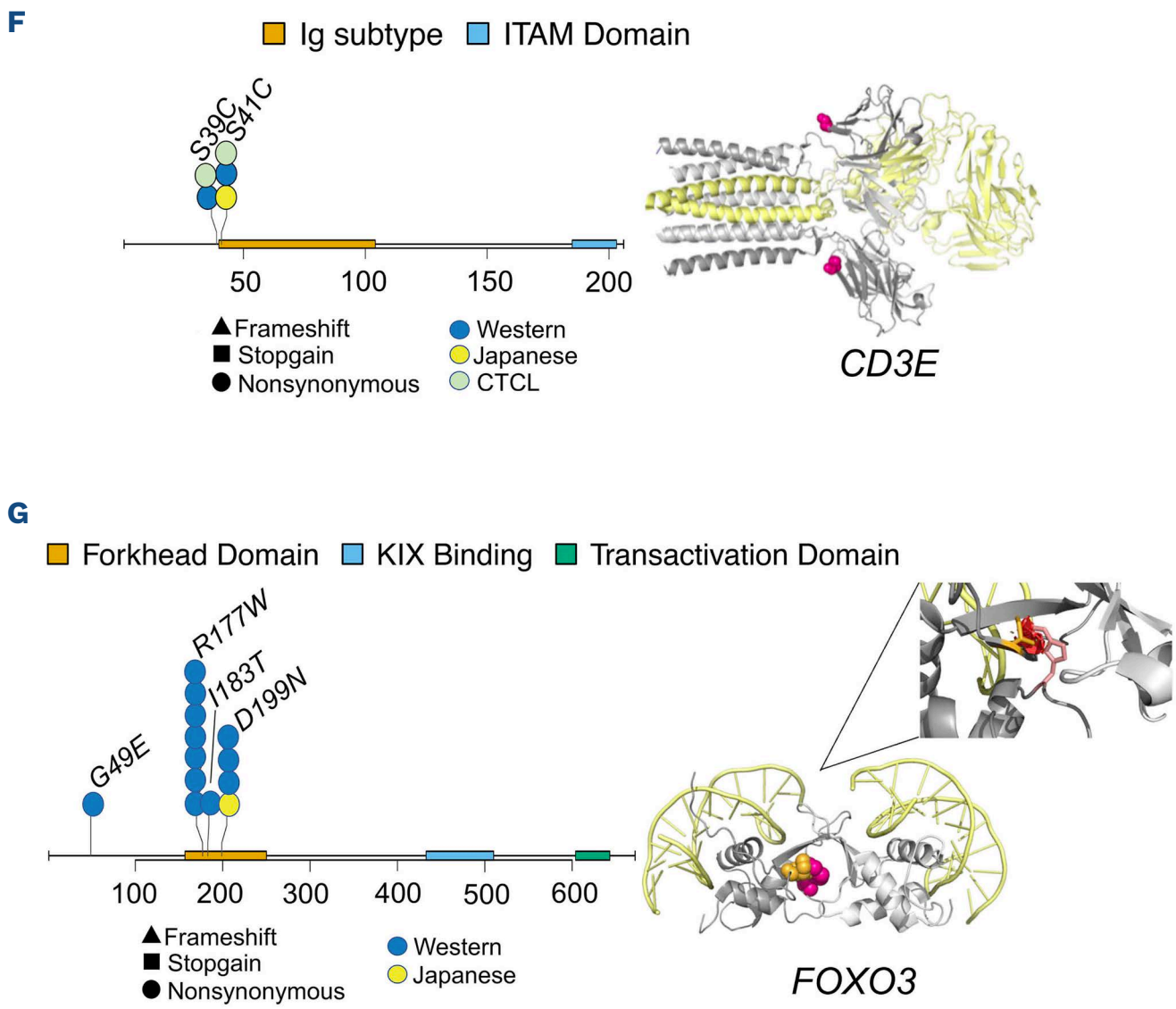


Figure 1. Putative adult T-cell leukemia-lymphoma driver genes discovered through mutation analysis. (A) Plot of GISTIC amplification/deletion peaks labeled with their corresponding driver genes. Amplifications are depicted in red and deletions in blue. Peaks are displayed along the genome represented lengthwise. Peak heights indicate GISTIC-calculated significance (G-score). Genes newly implicated in adult T-cell leukemia-lymphoma (ATLL) and/or cancer are highlighted in magenta. Both Western and Japanese patients are included in this analysis (N=554). (B, D, F, G) Point mutation burden among Western ATLL patients (N=122), Japanese ATLL patients (N=83), and published T-cell lymphoma patients (N=800) across *ANKRD11* (B), *GATA3* (D), *CD3E* (F) and *FOXO3* (G). Shapes and colors indicate mutation type and cohort, respectively, as indicated in the legend. (F, G) Right panels show recurrently mutated residues in magenta color within the context of their respective protein structures; ligands are shown in yellow. (F) CD3E is shown in dark gray. CD3ζ and CD3γ are shown in light gray. Bound T-cell receptor is shown in yellow. (G) A novel FOXO3 variant is shown with mutant tryptophan in hot pink and nearby isoleucine shown in orange. The zoomed-in panel shows predicted steric clashing between these residues as red discs, suggesting that a tryptophan mutation may be energetically unfavorable. The other subunit in the FOXO3 homodimer is shown in light gray. DNA is shown in yellow. (C, E) Histograms indicating segments of overlapping deletion in *ANKRD11* (C) and *GATA3* (E). The depicted cohort includes both Japanese and Western patients (N=554). Surrounding genes are indicated at the bottom of the figure. KIX: kinase-inducible domain interacting; ITAM: immunoreceptor tyrosine-based activation motif; CTCL: cutaneous T-cell lymphoma.

two amplification screens for TCR-independent cytokine production, and an interference screen for TCR-dependent cell proliferation.[34,35] We predicted that single-guide RNA (sgRNA) for oncogenes should show patterns of alteration in CRISPR screens consistent with the promotion of T-cell proliferation or cytokine production. sgRNA for tumor suppressors should show the opposite patterns.

Sixty-two (97%) of the putative driver genes showed a pattern of sgRNA alteration in one or more CRISPR screens consistent with their predicted tumor suppressor/oncogenic roles. The only exceptions were *TP53* and *PDCD1*. Presumably the effects of these genes cannot be easily modeled in short-term cultures (*TP53*) or without ligands (*PDCD1*). Fifty-five (86%) of the genes displayed the predicted patterns across multiple orthogonal CRISPR screens. Twenty-five (39%) showed the predicted sgRNA up- or down-regulation at magnitudes significantly greater than chance (false discovery rate [FDR] <0.05) (Figure 3A, B; *Online Supplementary Figure S5A*).

We next examined novel putative driver genes individually, including *ANKRD11, DGKZ, PTPN6* and *CD3E*, and novel hotspots in the cancer-associated gene *FOXO3*. We examined the function of each of these genes in the previously published CRISPR screens described above and performed orthogonal validation assays where possible.

Implicated by both point mutation and copy number mutation analysis, *ANKRD11* encodes a chromatin scaffolding protein binding histone deacetylases involved in the dif-

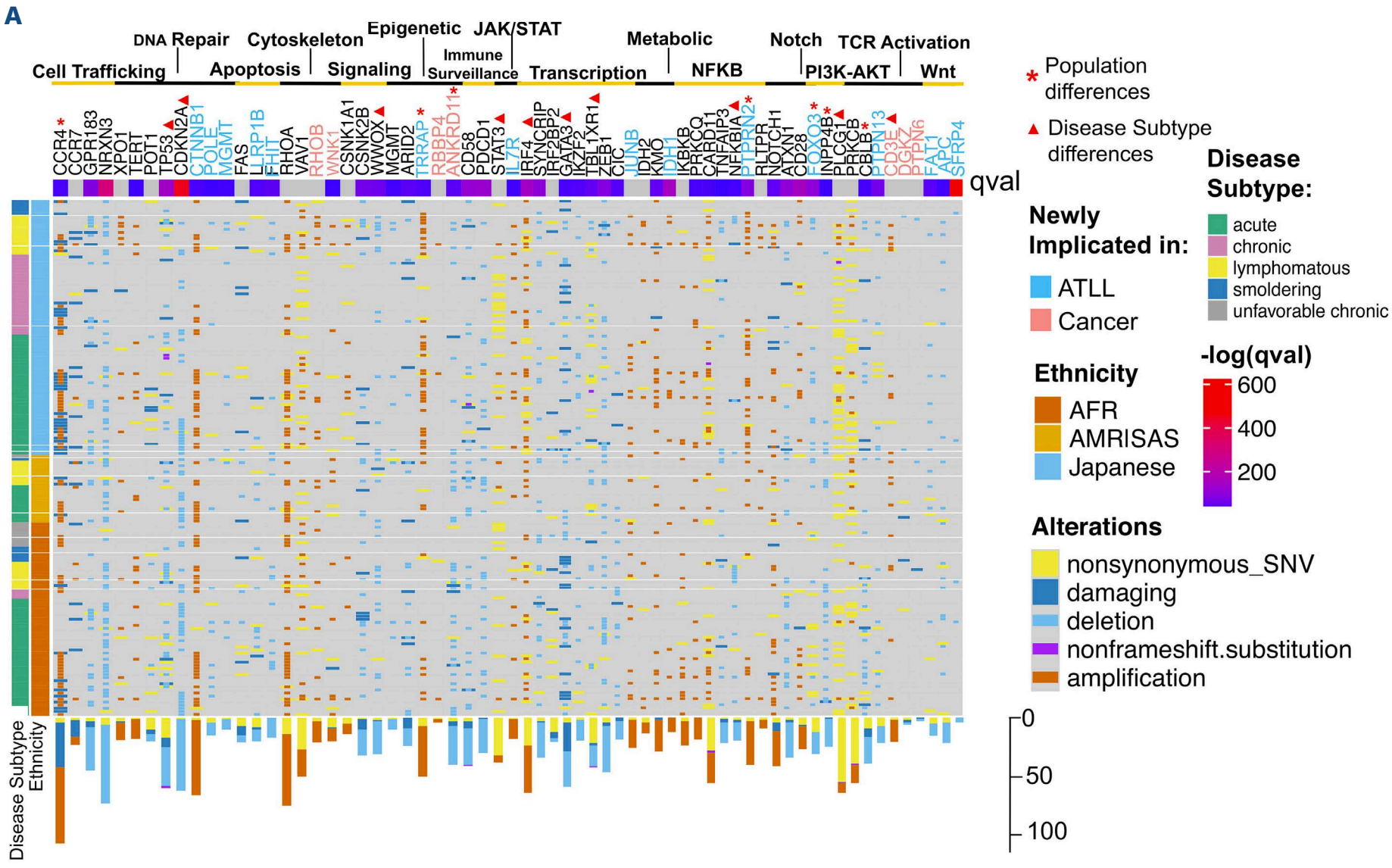


Figure 2. Key pathways involving recurrent mutations identified in Western and Japanese samples. (A) Frequencies and types of copy number variation (CNV) and point mutations in Western and Japanese samples. Only samples with both CNV and whole-exome sequencing data available are portrayed in this plot (N=168). Mutation type is indicated by the color of the squares. Bars at the bottom indicate the gross number of mutations. Colored bars to the left indicate single nucleotide polymorphism-determined ethnic classification of patients and categorization of their disease subtype. The significance of the GISTIC peak in which the

Continued on following page.

mutation is found, if applicable, is indicated by the heatmap on top (values indicate the negative logarithm of the q-value). Novel mutations are indicated by gene names in blue (newly implicated in adult T-cell leukemia-lymphoma [ATLL]) or red (newly implicated in cancer). Genes with mutational frequencies differing between subtypes or population cohorts are indicated with a triangle or a star, respectively. (B) An illustration of CD28- and T-cell receptor-initiated signaling pathways and their associated molecules based on a review of the current literature. Cytoplasmic and nuclear compartments are separated by the dashed red line representing the nuclear membrane. Downstream DNA-bound transcription factors (FOXO3a, AP-1, NFAT) and pathway- inducible genes (*BIM, FOXP3, IFNG, IL2, IRF4*) are shown. Stars mark putative driver mutations. HBZ (HTLV-1 bZIP factor) viral protein is shown as a dotted oval shape. TCR: T-cell receptor; AFR: African; AMRISAS: Native American/Southeast Asian; SNV: single nucleotide variation.

ferentiation of neural cells. It has also been postulated to interact with p53 in breast cancer models.[36,37] Its role in T cells is unknown. In the Western dataset, 2.5% of samples were affected by recurrent *ANKRD11* damaging mutations (p.Arg838*, p.Arg1007*, p.Ser2208*; P=4.31x10^{-4}) (*Online Supplementary Table S3*). By orthogonal CNV analysis, we found that *ANKRD11* falls within a 2.5 Mb region of significant deletion on chromosome 16 (q=2.4x10^{-13}, GISTIC2.0) deleted in 15% of patients (N=84) (Figure 1A-C; *Online Supplementary Table S5*). Among all the genes in this region, *ANKRD11* has an outsized proportion (likelihood ratio >5) of gene-localizing mutations, suggesting that it is the target gene of this recurrently deleted chromosomal segment. In 3.5% (N=6) of samples with both CNV and whole-exome sequencing data, both *ANKRD11* alleles were mutated either via biallelic deletion or mutation plus loss of heterozygosity. By RNA sequencing, samples with *ANKRD11* deletions had significantly decreased expression of *ANKRD11* transcripts (P=0.004, Student two-tailed *t* test) (Figure 3C). Finally, we decided to examine the role of *ANKRD11 in vitro*. We performed CRISPR knockout of *ANKRD11* in human T cells. CFSE proliferation assays showed significant increases in cell division in *ANKRD11* knockout cells upon TCR stimulation (P=0.0045; Student two-tailed *t* test) (Figure 3D, E).

DGKZ is a protein kinase responsible for dampening the effects of TCR stimulation by catalyzing breakdown of the downstream signaling molecule phosphatidic acid.[38] Consistent with its putative tumor suppressor function, *DGKZ* significantly inhibited TCR-mediated proliferation in previously published CRISPR screens (FDR=5.5x10^{-4}). In tumor samples from Western ATLL patients, we saw reduced or absent DGKZ protein in 10/10 (100%) patients regardless of *DGKZ* mutation status, suggesting that it could be transcriptionally downregulated in ATLL (*Online Supplementary Figure S5B*).

PTPN6 and *CD3E* are two novel putative driver genes implicated by recurrent point mutations. *CD3E* encodes a subunit of the CD3/TCR complex.[39] Samples in Western, Japanese, and publicly available T-cell lymphoma cohorts were recurrently mutated at p.Ser41Cys (N=1 Western, 1 Japanese, 1 T-cell lymphoma). Additional samples contained mutations at the nearby p.Ser39Cys (N=1 Western, 1 Japanese, 1 T-cell lymphoma) (Figure 1F). In genome-wide CRISPR screens, *CD3E* was a significant positive regulator of TCR-mediated cell proliferation (FDR=3.5x10^{-5}) (Figure 3B). *PTPN6* encodes for the protein SHP-1, a negative regulator of T-cell activation and PI3K signaling.[40] Three samples (2.5%) contained damaging mutations in *PTPN6*. By genome-wide CRISPR screen, *PTPN6* was a significant inhibitor of TCR-independent interleukin-2 production (FDR=7.2x10^{-5}). By western blot analysis, 50% of Western ATLL samples did not show significant PTPN6 expression. (*Online Supplementary Figure S5C*).

We noted recurrent mutations at a novel hotspot in the cancer-associated gene *FOXO3* (Figure 1G) (*Online Supplementary Table S2*). *FOXO3* encodes a transcription factor regulating T-cell differentiation.[41] It has been functionally validated as a tumor suppressor in several solid-organ malignancies.[42] Mechanistic studies have shown that HTLV-1 viral proteins HBZ and Tax suppress *FOXO3* protein function;[43, 44] however, *FOXO3* genetic mutations have not been directly implicated in ATLL by published genomic studies. Ten percent of the Western cohort of patients had *FOXO3* hotspot mutations (*Online Supplementary Tables S1* and *S2*). While 31% of these samples contained mutations at the cancer hotspot p.Asp199Asn, 54% (N=7) of the Western *FOXO3* mutations occurred in a previously unidentified hotspot, p.Arg177Trp. Both hotspots fall within the *FOXO3* DNA-binding forkhead domain in proximity to the validated dominant negative variant p.Ser256Ala, suggesting that they could act as dominant negative mutations (Figure 1G). By orthogonal CNV analysis, FOXO3 also falls within a broad region (51 Mb) of significant deletion (q=3.35x10^{-40}, GISTIC2.0). It was deleted in 9.4% (N=52) of samples. By western blot analysis of a random selection of ATLL patients, FOXO3 protein expression was reduced or absent in 11/13 patients (85%) (*Online Supplementary Figure S5C, D*), including 5/6 (83%) of patients with deletions or mutations. Weighted gene correlation analysis[45] of RNA-sequencing data suggests that *HBZ* expression relates to expression of *FOXO3* as well as *ANKRD11* (*Online Supplementary Figure S5E, F*). Consistent with prior large-scale studies on ATLL tumors, we did not find significant expression of *Tax* in the samples analyzed.[12]

We performed CRISPR and shRNA-mediated knockdown of *FOXO3* in our patient-derived ATLL cell lines. We subjected these cells to anti-neoplastic chemotherapy (etoposide) and a biological agent (belinostat) used to treat T-cell lymphomas. Knockdown of Foxo3 protein expression in ATLL cell lines by *FOXO3*-specific sgRNA

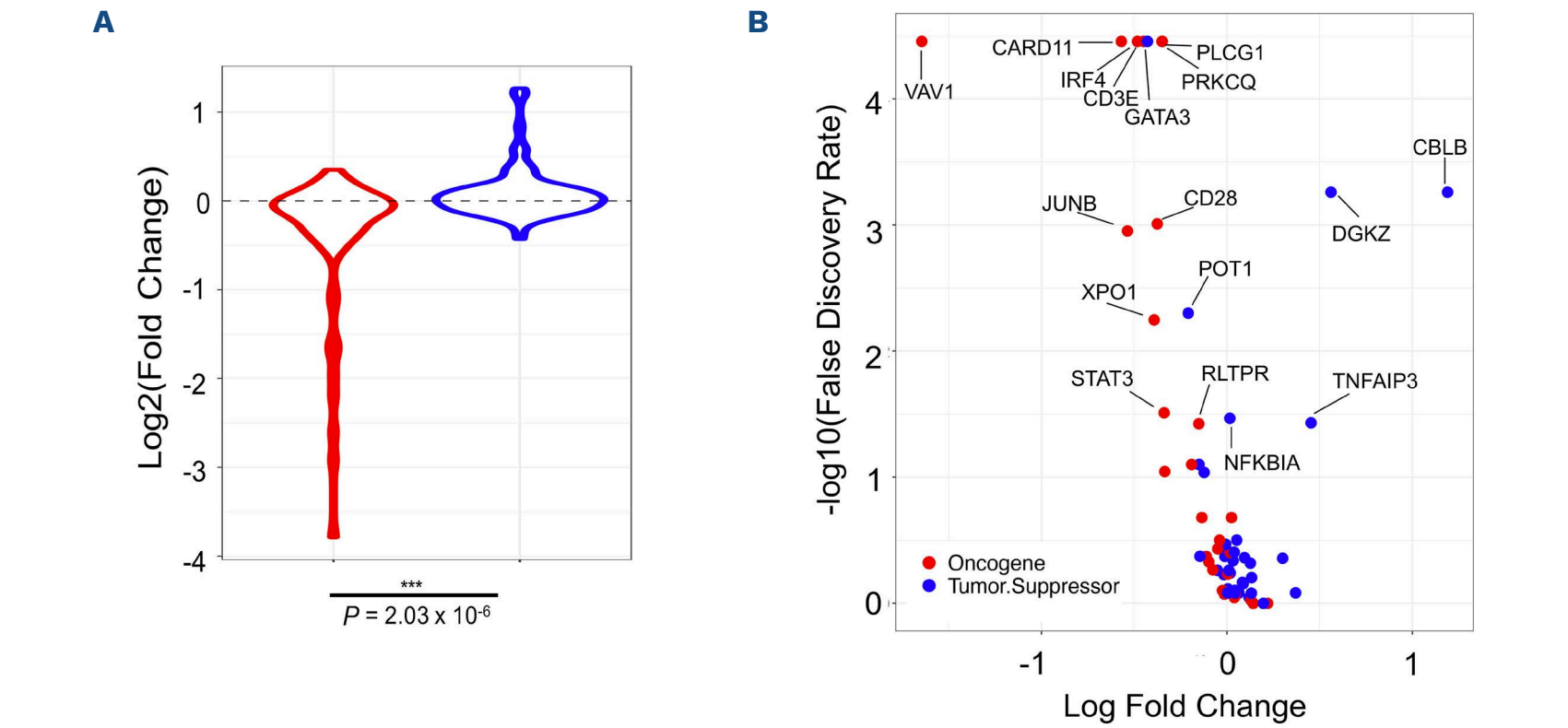
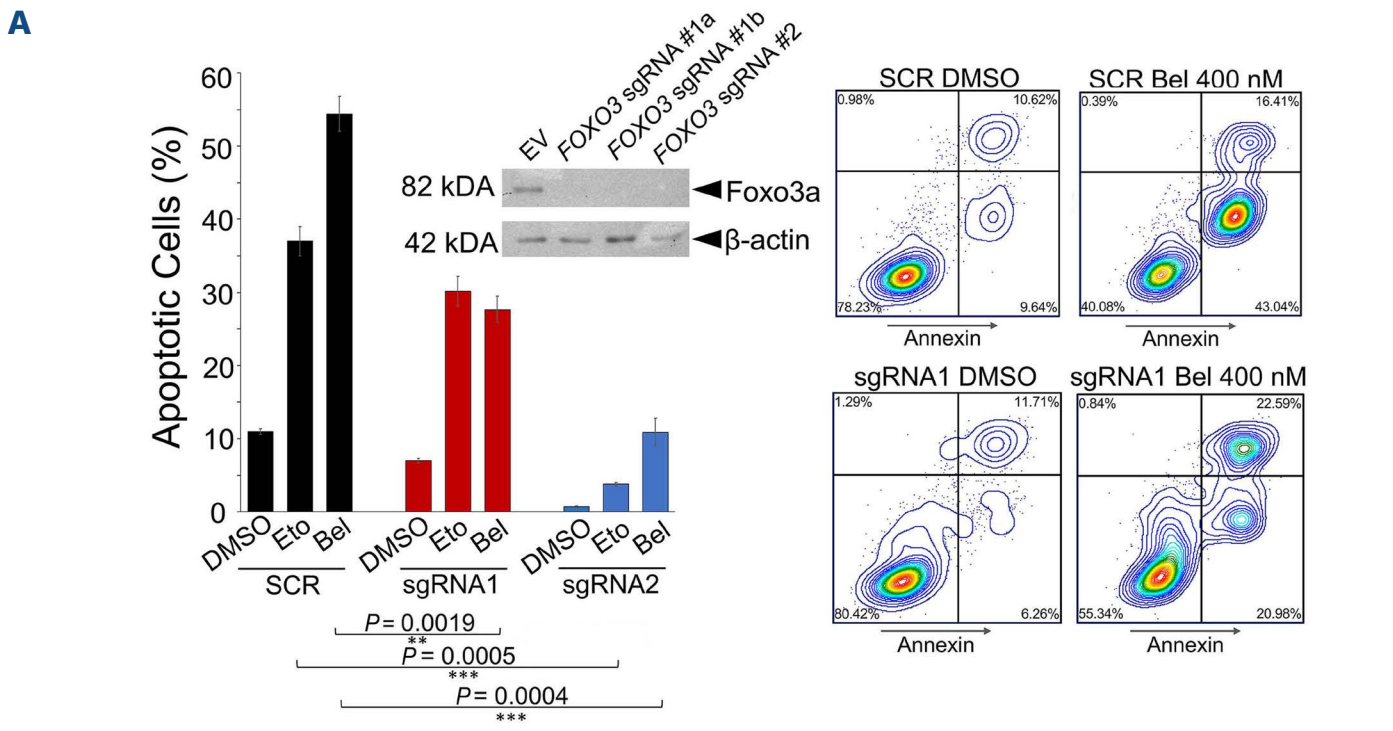


Figure 3. Functional validation of putative driver genes. (A) Violin plots showing log fold change of sgRNA (N=19,114) for putative oncogenes and putative tumor suppressors in genome-wide CRISPR interference screens for T-cell receptor (TCR)-dependent proliferation and cytokine production. sgRNA changes for putative oncogenes are shown on the left; those for putative tumor suppressors are shown on the right. The *P* value represents the significance of the difference between these distributions (Student two-tailed *t* test). A similar analysis of CRISPR amplification screens is included in *Online Supplementary Figure S5*. (B) Patterns of sgRNA alteration for putative oncogenes/tumor suppressors in a genome-wide CRISPR interference screen of TCR-dependent proliferation. The x axis represents the logarithm of the fold change, the y axis represents the negative logarithm of the false discovery rate. Putative oncogenes are shown in red and putative tumor suppressors in blue. CRISPR interference and amplification screens for TCR-independent cytokine production were analyzed similarly. (C) *ANKRD11* RNA transcript levels in *ANKRD11* mutant (blue) and wild-type (red) samples. Dots represent individual values, central horizontal bar represents the mean, and error bars represent the standard error. ***P*<0.01, comparison performed using the Student two-tailed *t* test. (D) Targeted CRISPR knockout screens in normal T cells for *ANKRD11*. Carboxyfluorescein succinimidyl ester (CFSE) dilution progresses (CFSE dye diminishes) from right to left. Stimulated cells are shown in red and unstimulated in blue. (E) Division index in targeted CRISPR knockout screens in normal T cells for *ANKRD11 versus* control. Error bars represent the standard error. ***P*<0.01, comparison performed using a paired-ratio *t* test. WT: wild-type; NT.CTRL: non-transduced control.



Continued on following page.

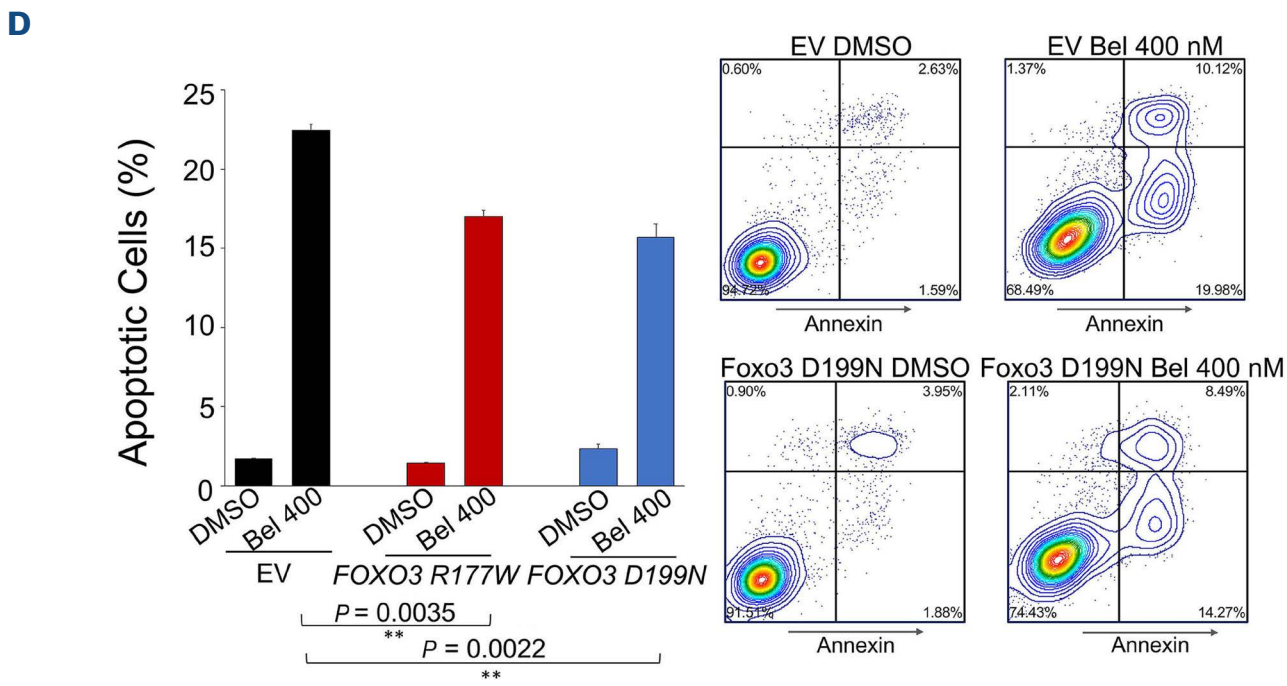


Figure 4. Functional validation of the pro-apoptotic role of FOXO3 in adult T-cell leukemia-lymphoma cells. (A-D) Left panels show the percentage of apoptotic cells, as determined by the percentage of cells staining positively for annexin V by fluorescence-activated cell sorting (FACS) analysis. Brackets represent comparisons made after subtracting dimethyl sulfoxide vehicle (DMSO) control values. The statistical significance of differences was determined by a Student two-tailed *t* test (**P*<0.05, ***P*<0.01, ****P*<0.001). Error bars represent the standard error from separately treated triplicate samples. Right panels show gated data from FACS analysis, with anti-annexin V on the x axis and propidium iodine on the y axis. (A-C) Western blots validating protein levels in knockdown/overexpression constructs are shown at the top of the lefthand panels. β-actin was used as the protein loading control. (A) Percentage of apoptotic cells in ATLL-97c *FOXO3* knockdown constructs exposed to DMSO vehicle, etoposide 200 nM or belinostat 200 nM. Two *FOXO3* knockdown constructs were established using distinct *FOXO3*-specific single-guide RNA (sgRNA1 and sgRNA2). A construct established from a non-specific scrambled single-guide RNA was used as a control. (B) Percentage of apoptotic cells in ATLL-97c *FOXO3* overexpression constructs exposed to belinostat at concentrations of 200, 400 or 800 nM. m-CAT-1-expressing ATLL-97c cells were transduced with pseudoviral particles containing one of two overexpression constructs (#1 or #5). Cells transduced with empty vectors were used as a control. (C) Percentage of apoptotic cells in ATLL-84c doxycycline-inducible constructs exposed to DMSO vehicle, belinostat 200 nM (Bel 200) or belinostat 400 nM (Bel 400). Dox (+) cells were treated with doxycycline 1 μg/mL at least 72 h before and at the start of drug treatment experiments. Dox (-) cells were not exposed to doxycycline. (D) Percentage of apoptotic cells after exposure to belinostat 400 nM (Bel 400) in ATLL-84c cells transduced with lentiviruses containing either p.Arg177Trp (R177W) or p.Asp199Asn (D199N) *FOXO3* mutant constructs. Cells transfected with an empty lentivirus were used as controls. Vector nucleotide sequences were verified by sequencing. EV: empty vector; Eto: etoposide; Bel: belinostat, SCR: non-specific scrambled single-guide RNA; Wt: wild-type; Dox: doxycycline.

(sgRNA1 and sgRNA2) conferred resistance to drug-induced apoptosis, as compared to cells transfected with scrambled (SCR) sgRNA (Figure 4A). Similarly, knockdown of *FOXO3* via tetracycline-inducible shRNA protected ATLL cells from dose-dependent drug-induced apoptosis (Figure 4C). In contrast, mCAT-1^{+} ATLL cells transduced with *FOXO3*-overexpressing pseudoviral particles showed higher rates of drug-induced apoptosis in comparison to cells transduced with empty controls (Figure 4B). Finally, we examined cells transfected with mutant *FOXO3* p.Asp199Asn and p.Arg177Trp-expressing lentiviral vectors. These cells resembled *FOXO3* knockdown in their phenotype: compared to cells transfected with empty vectors, they showed resistance to drug-induced apoptosis (Figure 4D). Because the p.Arg177Trp and p.Asp199Asn mutations both map to the Foxo3a DNA-binding domain, we hypothesized that *FOXO3* mutations could affect T-cell apoptosis by altering the transcription of pro-apoptotic target genes. However, knockdown of Foxo3a did not appear to affect protein levels of BIM and p21, both encoded by known pro-apoptotic Foxo3a target genes (*BCL2L11* and *CDKN1A*, respectively) (*Online Supplementary Figure S6B*). Collectively, these data support the pro-apoptotic role of *FOXO3* in ATLL, although the mechanisms by which it exerts its effects remain to be identified.

The adult T-cell leukemia-lymphoma genomic landscape differs between Western and Japanese patients

Previous exome-wide studies of ATLL have been conducted in Japanese populations.[11,12] We sought to determine whether there are molecular differences between Western and Japanese cohorts that could be associated with the known differences in clinical presentation.[9,10] We observed broad similarities in commonly mutated genes, but several were mutated at significantly different frequencies. For example, Japanese patients had a significantly greater burden of point mutations in *CCR4*, the most commonly mutated putative driver gene overall (34% *vs*. 16% of patients in the Western populations, *P*=0.006, Fisher exact test) (Figure 5). *PTPRN2* (34% *vs*. 13%, *P*=0.002) and *TRRAP* (43% *vs*. 14%, *P*=3.2x10^{-5}) were also both more commonly mutated in Japanese patients than Western ones (Fisher

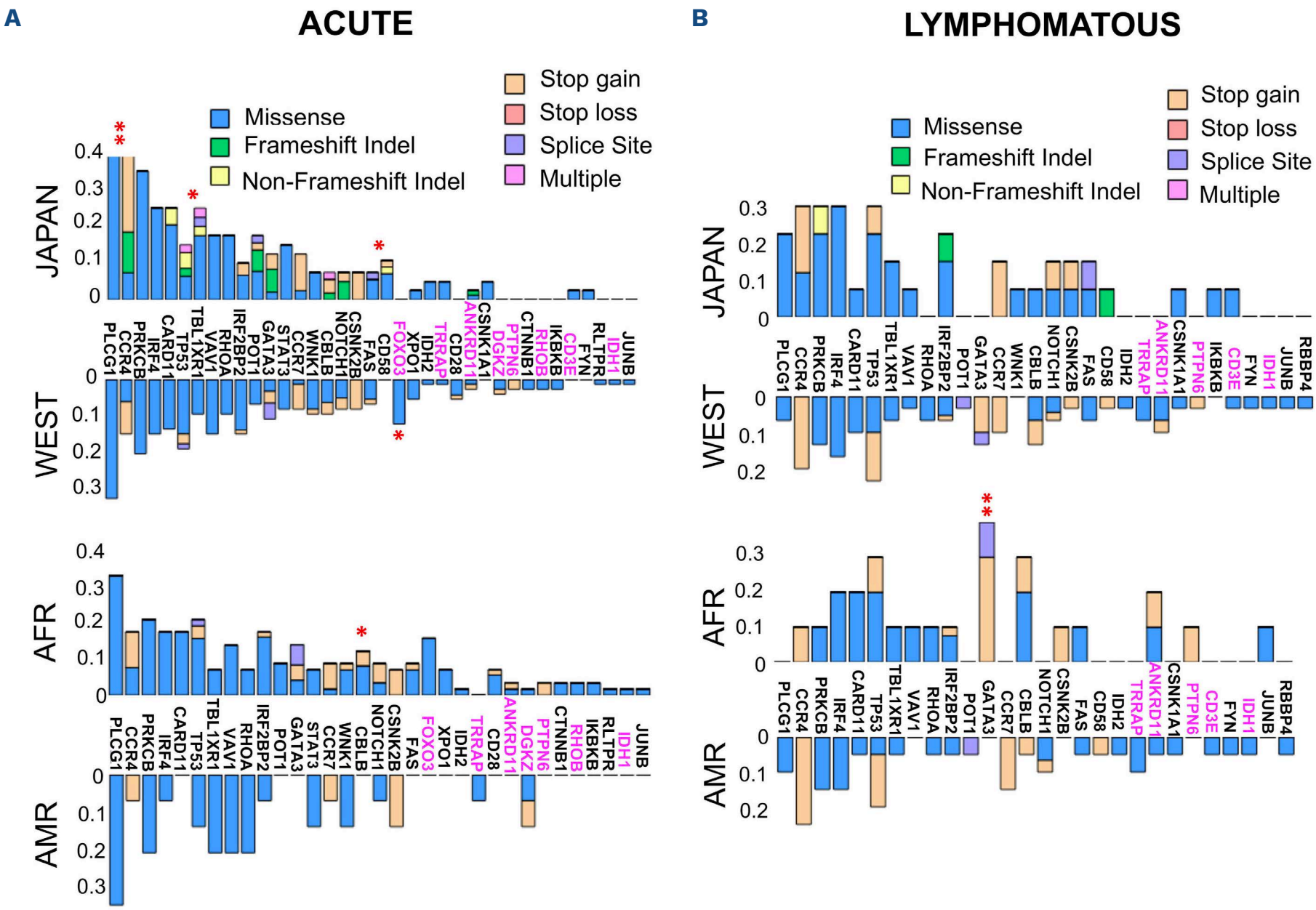


Figure 5. Differential patterns in mutational frequency between Japanese and Western adult T-cell leukemia-lymphoma cohorts. (A, B) Comparison of point mutation frequencies between Japanese/Western and Afro-Caribbean/South American populations. The color of the bars represents mutation type, while the height of the bars represents the frequency with which a gene is mutated in the specified population. Statistical comparisons were performed with the Fisher exact test. **P*<0.05, ***P*<0.01. (A) Comparisons in acute cases (N=62 Western *vs.* 39 Japanese and 48 Afro-Caribbean *vs.* 14 South American). (B) Comparisons in lymphomatous cases (N=29 Western *vs.* 13 Japanese and 9 Afro-Caribbean *vs.* 20 South American). (C, D) Histograms of amplifications (C) and deletions (D) across the genome in Japanese (top) and Western (bottom) cohorts. The y axis represents the proportion of cohorts with amplification/deletion in a given region. Genes significantly different between cohorts, as discussed in the text, are highlighted with arrows. Genes newly implicated in adult T-cell leukemia-lymphoma. and/or cancer are highlighted in magenta. AFR: Afro-Caribbean; AMR: South America

exact test) (Figure 5C).
Several putative driver genes were deleted or mutated with significantly greater frequency in the Western cohort. These included *INPP4B* (21% *vs*. 6%, *P*=0.006), *ANKRD11* (34% *vs*. 11%, *P*=4.10 x 10^{-4}) and *CBLB* (27% *vs*. 13%, *P*=0.035) (Figure 5D) (Fisher exact test). Of Western patients, 8/13 (62%) had clearly reduced CBLB expression; 50% of these had mutation or deletions (*Online Supplementary Figure S5C, D*). *FOXO3* point mutations were eight times more common in Western patients (10% *vs*. 1.2%, *P*=0.017, Fisher exact test) (Figure 5A). Mutations at the novel hotspot *FOXO3* p.Arg177Trp occurred only in Western patients of African descent with aggressive leukemic ATLL subtypes (acute N=7, unfavorable chronic N=1).

Different clinical subtypes of adult T-cell leukemia-lymphoma are characterized by distinct mutational patterns

We next sought to identify molecular drivers of clinical behavior in different ATLL subtypes. Consistent with observations from the Japanese cohort, in the Western cohort we identified enrichment of *STAT3* mutation in "indolent" ATLL (chronic and smoldering) (37% *vs*. 11%, *P*=1.2x10^{-4}), enrichment of total number of mutations in "aggressive" ATLL (acute and lymphomatous) (119 *vs*. 47 mutations, *P*=4.3x10^{-7}), and enrichment of *TP53*, *CDKN2A* and *IRF4* mutations in aggressive ATLL (31% *vs*. 12%, *P*=0.01; 40% *vs*. 16%, *P*=0.002 and 38% *vs*. 18%, *P*=0.02, respectively) (Fisher exact test).[11]
The enrichment of aggressive clinical subtypes in our Western dataset allowed us to distinguish molecular features of acute *versus* lymphomatous cases in the Western cohort. *NRXN3* and *CCR4* were commonly mutated (>30% of samples) in both subtypes. *CDKN2A* mutations (46% *vs*. 25%, *P*=0.009) and *PLCG1* amplification/mutation (37% *vs*. 14%, *P*=0.002) were significantly more frequent in acute cases than in lymphomatous cases. In contrast, *TP53*, *WWOX, CD3E, TBL1XR1* and *NFKBIA* were genetically altered significantly more often in lymphomatous cases (*Online Supplementary Table S7*). Consistent with their enrichment in distinct disease subtypes, *TP53* and *PLCG1* mutations showed significant mutual exclusivity (q=0.004).[21]
There is little molecular information available about the chronic ATLL subtype with unfavorable features.[9] In our dataset, "unfavorable chronic" cases (N=17) resembled aggressive cases in their mutational and CNV burden (Figure 6A, B). They were characterized by an increased frequency of heterozygous *GATA3* deletions. Of the 11 unfavorable chronic cases with CNV data, 45% (N=5) had *GATA3* deletions (3.8-fold enrichment compared to other cases, *P*<0.001, Fisher exact test). Immunohistochemistry analysis of GATA3 expression in ATLL patient samples confirmed that GATA3 protein levels were lower in the unfavorable chronic cases than in other aggressive subtypes (*P*=0.05, Student two-tailed *t* test) (Figure 6D).

The molecular features of adult T-cell leukemia-lymphoma are associated with survival and response to therapy

In addition to clinical subtype and mutational status, we annotated our dataset with patients' survival and response to chemotherapy and/or zidovudine-interferon therapy. We used a Cox multivariate regression analysis to test the effect of gene mutation on survival. We limited our analysis to patients with well-annotated clinical data who were seen for at least one follow-up visit (*Online Supplementary Table S9*). As expected, indolent subtypes were associated with significantly longer survival times than aggressive subtypes (hazard ratio = 0.15, 0.26 and 0.26 for smoldering, unfavorable chronic and chronic subtypes, respectively). Treatment type (chemotherapy, zidovudine-interferon, both or neither) did not influence survival. Consistent with previous studies, *STAT3* mutation trended towards being associated with significantly decreased mortality in dataset-wide analyses. However, in analyses controlled for clinical subtype and treatment modality and corrected for multiple comparisons, only *ANKRD11* and *TP53* mutation were significantly associated with increased mortality (hazard ratio = 2.70 and 2.67, respectively) (Figure 6C, E). Pairwise analysis did not show significantly different outcomes for patients with both mutations than for patients with either driver mutation alone.
One mutation was associated with response to therapy in the Western cohort: Patients with *CDKN2A* loss were significantly more likely to experience a complete response to chemotherapy (24% *vs*. 0%, *P*=0.025) (Figure 6F). Unfortunately, this mutation was not associated with decreased mortality.

Newly acquired *IRF4* and *CARD11* mutations in relapsed adult T-cell leukemia-lymphoma

Consistent with the literature, disease relapse was common in our cohort.[6] Our dataset included ten patients who experienced disease relapse after receiving zidovudine-interferon treatment as their initial therapy. Two of these patients had multiple relapses. We investigated whether the variability in mortality within our dataset could be related to genomic patterns conferring a high likelihood of relapse. We first examined patterns of malignant clonality in relapsed samples. As expected, relapse samples shared the same HLA genotype as initial samples, except for one patient who underwent loss of heterozygosity in HLA-A upon relapse. Comparison of transcription data-derived TCR clonotypes in initial *versus* relapsed samples demonstrated that all relapses harbored the same TCR as the initial clone (Figure 7A). Interestingly, relapsed samples generally shared few point mutations with initial tumor samples but did share most CNV mutations. Patients generally accumulated additional CNV mutations with relapse (Figure 7B). We then sought to determine whether there were any specific genomic changes associated with disease relapse. We found

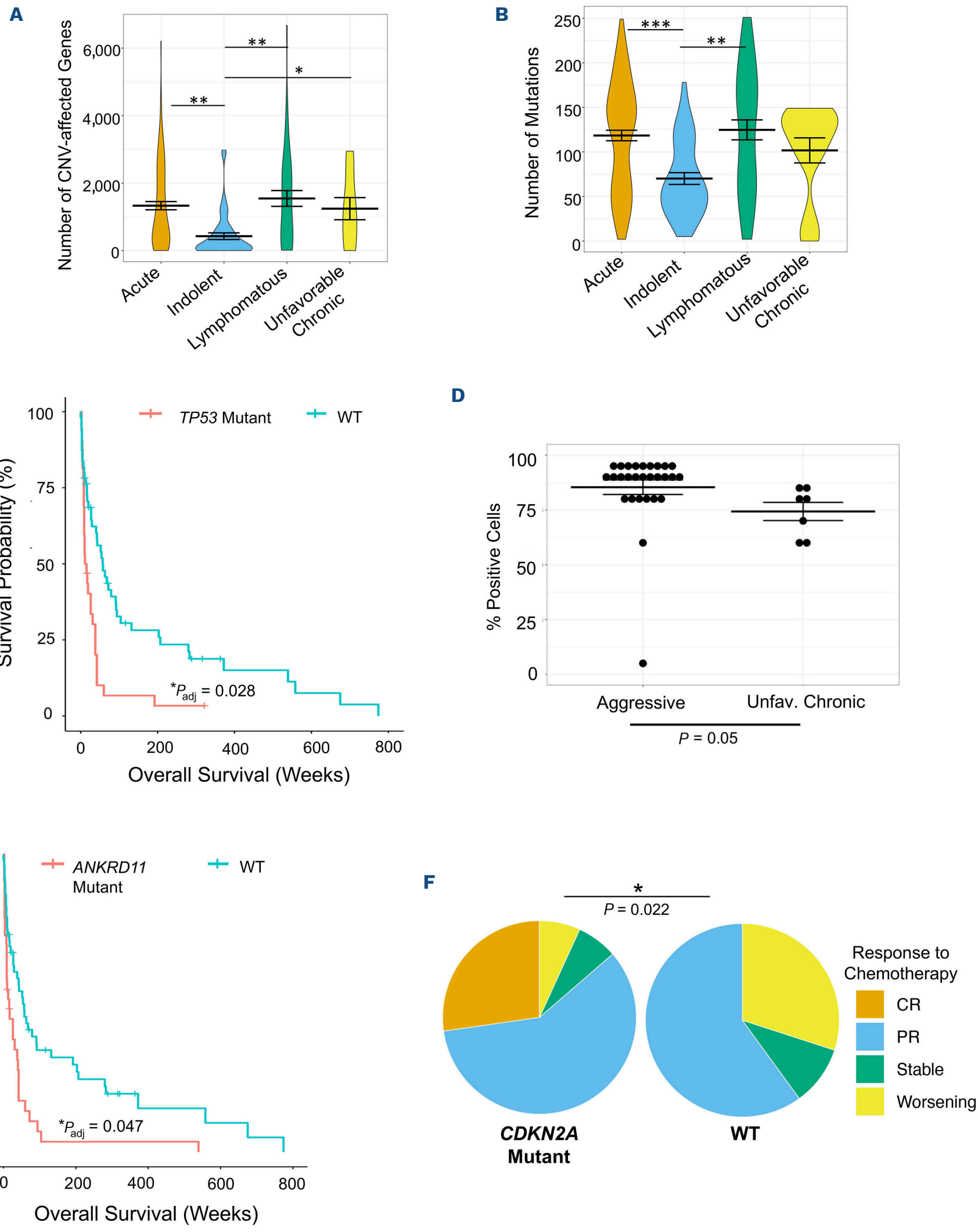


Figure 6. Genomic alterations associated with clinical subtype and mortality. (A, B) Violin plots demonstrating the similarity of the number of genes deleted or amplified per sample (A) and the number of point mutations per sample (B) between the unfavorable chronic subtype of adult T-cell leukemia-lymphoma (N=15), and acute (N=63) and lymphomatous (N=31) subtypes. Bars represent the mean (central bar) and standard error. A Student two-tailed *t* test was used to determine the statistical significance of differences, *P<0.05, **P<we identified ***P<0.001. (C) Comparison of GATA3 protein levels by immunohistochemistry in aggressive (acute and lymphomatous) and unfavorable chronic cases, Student two-tailed *t* test. Dots represent individual values,

Continued on following page.

the central horizontal bar represents the mean, and the error bars represent the standard error. (D, E) Kaplan-Meier curves indicating the effect of *TP53* (D) and *ANKRD11* (E) mutation on overall survival. Red lines indicate overall mortality for patients with a mutation in the specified gene. Blue lines indicate overall mortality for patients without a mutation in that gene. Significance was determined using a Cox multivariate analysis of the overall cohort controlling for disease subtype and treatment modality.[45] P_{adj} = *P* value after Bonferroni correction for multiple comparisons. (F) Frequencies of chemotherapy responses in patients with a *CDKN2A* deletion/damaging mutation (left) compared to patients with wild-type *CDKN2A* (right). Fisher exact test, **P*<0.05. Unfav.: unfavorable; WT: wild-type; CR: complete response; PR: partial response.

recurrent mutations in *CARD11* (3 samples) and *IRF4* (4 samples) that were acquired in disease relapse samples after zidovudine-interferon therapy. Although uncommon in the general cohort, *IRF4* mutations were significantly more common in disease relapse (33% *vs*. 10% of samples, *P*=0.038, Fisher exact test) (Figure 7C). Of 12 samples taken after disease relapse, four had new *IRF4* mutations (3/9 patients). One patient who experienced relapse twice developed two different *IRF4* mutations in each relapse instance. Manual review of sequencing data confirmed that none of the *IRF4* mutations in relapsed samples was present even at a subclonal level in parent samples. We further analyzed seven additional relapsed samples for p.Leu70Val or p.Lys59Arg, the most common *IRF4* mutations in our cohort, by standard Sanger sequencing (*Online Supplementary Table S8*) (Figure 7D). We found p.Lys59Arg mutations in two samples from relapsed patients for whom no baseline samples were available. Within the entire cohort, no patient with an *IRF4* mutation experienced complete response to zidovudine-interferon therapy, suggesting that this mutation may be associated with primary or acquired resistance (*Online Supplementary Table S9*).

Discussion

Through multimodal exome-wide analysis, we have identified novel genomic features of ATLL that highlight differences between the genomic landscapes of Japanese and Western cohorts of patients.

Newly identified putative driver genes include *FOXO3* and *ANKRD11*. *FOXO3* encodes a transcription factor that has been postulated to be a tumor suppressor and mediator of apoptosis via the PI3K pathway.[46] Our *in vitro* experiments using patient-derived ATLL cell lines demonstrated the acquisition of antineoplastic drug resistance after *FOXO3* knockdown and the augmentation of drug-induced apoptosis with *FOXO3* overexpression, suggesting that FOXO3 may have a pro-apoptotic role in ATLL. Similar to the p.Asp199Asn variant previously observed in solid malignancies,[40] expression of the novel p.Arg177Trp variant observed in our cohort also conferred resistance to antineoplastic drugs *in vitro*. *ANKRD11* is a novel driver gene identified as significant through both mutational and CNV analysis and validated through *in vitro* proliferation assays. *ANKRD11* is especially relevant due to its association with increased mortality. While Japanese studies have noted recurrent deletion of *ANKRD11* in ATLL, it was previously dismissed as breakage at a putative fragile site.[13] Our analysis is the first to distinguish *ANKRD11* and *FOXO3* as putative driver genes through mutational analysis and functional assays. Because both *ANKRD11* and *FOXO3* mutations were significantly enriched in Western cohorts of patients, their elucidation as putative driver genes was only possible after the inclusion of underrepresented Western patients in our analysis.

Previous studies have demonstrated that Western ATLL patients have more severe outcomes than Japanese patients.[9,10] While differences in medical infrastructure likely contribute to some discrepancies, healthcare inequalities alone appear inadequate to explain all observed clinical disparities (e.g., the diagnosis of ATLL by >10 years earlier in patients from medical resource-poor regions of the Caribbean).[9] It is plausible that differences in the distribution of ATLL driver mutations may also contribute to the disparate clinical outcomes in Western *versus* Japanese ATLL cases.[9,15] Our data support the presence of molecular differences in ATLL that correspond to geography and clinical outcome. For example, mutations in *INPP4B* and *ANKRD11* were more common in Western patients and predicted worse survival, revealing potential tumor-associated causes of disparate clinical outcomes in the West. In contrast, *CCR4* mutations were more common in Japanese patients. Such a finding might account for the better therapeutic responses to mogamulizumab seen in Japanese patients compared to the rates seen in the USA.[47]

These data also suggest new clinical targets for ATLL. The collective enrichment of *FOXO3, INPP4B, CBLB, DGKZ*, and *PTPN6* mutations in Western ATLL patients suggest a potentially mounting role of the PI3K-AKT pathway in ATLL oncogenesis. PI3K inhibitors are therapies approved by both the Food and Drug Administration and the National Comprehensive Cancer Network for chronic lymphocytic leukemia[48] and non-Hodgkin lymphomas.[49,50] In recent trials, they have shown to be promising agents in patients with peripheral T-cell lymphoma[51] and in preclinical studies of ATLL.[52] These results support further investigation of the potential of the PI3K-AKT pathway as a therapeutic target in ATLL, particularly in Western patients.

Our combined Western and Japanese dataset constitutes the largest cohort of exome-wide sequencing data to date. Because of this, we were able to make novel observations regarding the patterns of genomic alterations in different clinical subtypes. *TP53* and *CDKN2A* alterations have previously been shown to be increased in aggressive cases

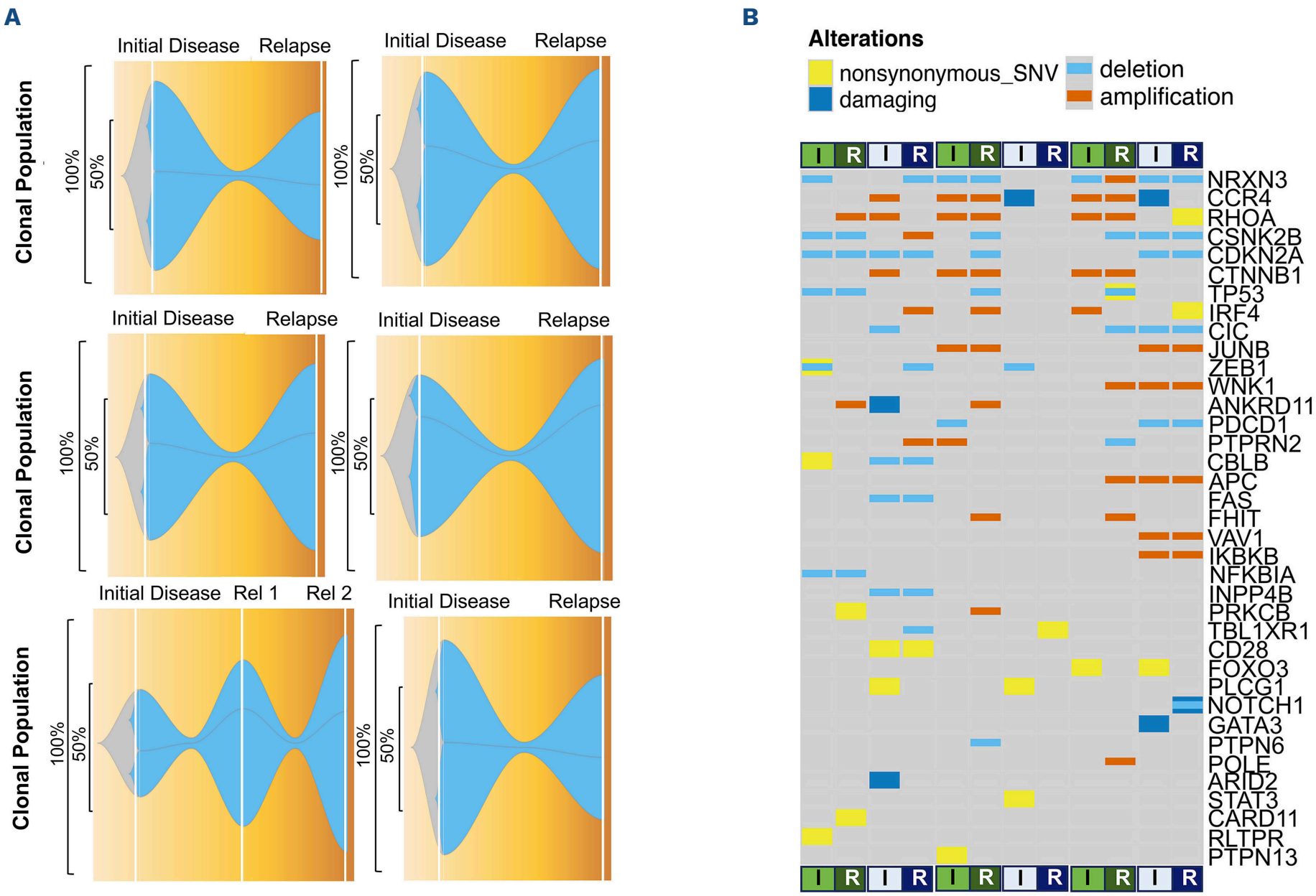


Figure 7. Association of *IRF4* mutations with relapse in adult T-cell leukemia-lymphoma. (A) Clonality of tumor populations as determined by T-cell receptor (TCR) sequence identification from RNA-sequencing data. Patients shown here contributed samples from initial disease as well as subsequent relapse. Clonal populations were defined as cells with identical TCR. Each blue shape represents the dominant T-cell clone, as determined by α and β TCR subunit identity. The vertical axis represents the size of clonal populations as determined by the frequencies of the dominant α and β subunit. Relapses are indicated by vertical white lines. Multiple relapses are shown in sequential order, from left to right. (B) Comparison of point and copy number variation (CNV) mutations in initial (I) *versus* relapsed (R) samples. Pairs of initial/relapsed samples from the same patient are indicated in light and dark shades, respectively, of alternating blue and green. Mutation types include deletions, amplifications, damaging mutations and nonsynonymous (missense) mutations, as shown in the legend. Putative driver genes are listed on the left. (C) Frequency of new *IRF4* mutations in first-time disease (N=121) compared to relapsed samples (N=12). Fisher exact test, *P<0.05. (D) Distribution

Continued on following page.

of mutations across *IRF4*. Mutations shown in blue occurred in Western samples (N=121), those shown in yellow occurred in the Japanese cohort (N=83). Especially frequent mutations are represented as pie markers indicating the proportion of variants contained in either the Western (blue) or Japanese (yellow) cohort. Marker stem heights represent the number of cases as depicted on the y axis. SNV: single nucleotide variation; WT: wild-type.

(acute and lymphomatous combined). Our analysis examined acute and lymphomatous cases separately and found that *TP53* loss is more characteristic of lymphomatous cases while *CDKN2A* loss is more characteristic of acute cases. These molecular differences could account for the distinct clinical presentations of acute and lymphomatous ATLL. They might also guide treatment in the future as more targeted therapies become available for ATLL. We also identified *GATA3* deletion as a putative tumor suppressor associated with chronic ATLL with unfavorable features, as it is characterized by recurrent deletion and damaging mutation in our cohort. Consistent with this, it has been implicated in constraining regulatory T-cell proliferation.[53] In our dataset, disease relapse after zidovudine-interferon therapy was associated with the acquisition of *IRF4* mutations affecting the DNA binding domain. This mutational pattern could be the result of zidovudine-interferon treatment-driven selection. IRF4, also known as MUM1, is a transcription factor that regulates the expression of interferon-inducible genes in leukocytes.[54] Lack of IRF4 protein expression has been previously associated with response to zidovudine-interferon therapy in patients with ATLL.[55]

It has become increasingly evident that different ethnic groups have disparate prognoses for the same diseases. However, our understanding of diversity in disease pathogenesis is limited by the underrepresentation of certain patient populations in genomic research. To address these challenges in the context of ATLL, we generated the largest dataset of untargeted genomic information in Western ATLL patients so far, including diverse African descendants from North and South America, the Caribbean, and South American indigenous populations. Through the analysis of this dataset, we identified novel molecular ATLL features, some of which are associated with more aggressive disease. Several of these genes, including *ANKRD11, INPP4B* and *FOXO3*, were mutated significantly more often in Western patients. Our functional assays support the roles of these genes as tumor suppressors *in vitro*. Other mutations associated with treatment outcomes (e.g., *CCR4*) were more common in Japanese cohorts. Our present findings augment current knowledge and reflect the diversity of the molecular landscape of ATLL. Furthermore, they emphasize the need for the inclusion of underrepresented populations in genomic research to better understand and address disparate factors in patients' outcomes.

Disclosures

JC is a co-founder of Moonlight Bio, Inc. The other authors have no conflicts of interest to disclose.

Contributions

AAG, JC and JCR designed and supervised the study, developed methodology, performed experiments, analyzed and interpreted data, and prepared the manuscript. CSM developed methodology, analyzed and interpreted genomic and clinical data, and prepared the manuscript. EW performed the Oncoscan copy number variation analysis and interpreted data. KR and SM analyzed and interpreted whole-exome sequencing data. GP and NLT designed CRISPR constructs for FOXO3. NLT and LB provided technical support, processed tumor samples, performed standard gene sequencing and immunoblotting, and interpreted data. CL performed and interpreted targeted CRISPR screens. AAG, CCB, GB, JD, SC, CSM, MPP, KO and JRC performed immunohistochemistry and immunophenotyping studies and analyzed and interpreted data. CCB, MB, DM, SC, JAS, BB, DC, LM, CB, OS, FC and JCR collected clinical data and developed the databases of the adult patients with T-cell leukemia-lymphoma at local institutions. All authors reviewed the manuscript.

Acknowledgments

We acknowledge the patients who consented to inclusion in this study, Admera Health, the University of Virginia Comprehensive Cancer Center, the University of Miami Sylvester Comprehensive Cancer Center and its Flow Cytometry Shared Resource and data analysis by Patricia Guevara, and the Northwestern University Flow Cytometry Core and Research Computing Services for their invaluable contributions.

Funding

JC was supported in part by the National Institutes of Health (R01CA260064-01A1), the Bakewell Foundation and the Leukemia and Lymphoma Society grant 1377-21. JCR was supported in part by the National Institutes of Health/National Cancer Institute (R01CA223232) and the University of Miami-Sylvester Comprehensive Cancer Center (P30CA240139). CSM was supported by the Monticello College Foundation Olin Fellowship. The content of this paper is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Data-sharing statement

Demographic data, mutational counts and copy number variation peaks may be found in the Online Supplementary Data file available with the online version of this article. For original data, please contact Dr. Jaehyuk Choi at jaehyuk.choi@northwestern.edu

References

1. Poiesz BJ, Ruscetti FW, Gazdar AF, Bunn PA, Minna JD, Gallo RC. Detection and isolation of type C retrovirus particles from fresh and cultured lymphocytes of a patient with cutaneous T-cell lymphoma. Proc Natl Acad Sci U S A. 1980;77(12):7415-7419.
2. Gessain A, Cassar O. Epidemiological aspects and world distribution of HTLV-1 infection. Front Microbiol. 2012;3:388.
3. Shimoyama M. Diagnostic criteria and classification of clinical subtypes of adult T-cell leukaemia-lymphoma. A report from the Lymphoma Study Group (1984-87). Br J Haematol. 1991;79(3):428-437.
4. Cook LB, Fuji S, Hermine O, et al. Revised adult T-cell leukemia-lymphoma International Consensus Meeting Report. J Clin Oncol. 2019;37(8):677-687.
5. Ishitsuka K, Tamura K. Human T-cell leukaemia virus type I and adult T-cell leukaemia-lymphoma. Lancet Oncol. 2014;15(11):e517-526.
6. Katsuya H, Ishitsuka K, Utsunomiya A, et al. Treatment and survival among 1594 patients with ATL. Blood. 2015;126(24):2570-2577.
7. Imaizumi Y, Iwanaga M, Nosaka K, et al. Prognosis of patients with adult T-cell leukemia/lymphoma in Japan: a nationwide hospital-based study. Cancer Sci. 2020;111(12):4567-4580.
8. Subramaniam JM, Whiteside G, McKeage K, Croxtall JC. Mogamulizumab: first global approval. Drugs. 2012;72(9):1293-1298.
9. Malpica L, Pimentel A, Reis IM, et al. Epidemiology, clinical features, and outcome of HTLV-1-related ATLL in an area of prevalence in the United States. Blood Adv. 2018;2(6):607-620.
10. Shah UA, Shah N, Qiao B, et al. Epidemiology and survival trend of adult T-cell leukemia/lymphoma in the United States. Cancer. 2020;126(3):567-574.
11. Kataoka K, Iwanaga M, Yasunaga JI, et al. Prognostic relevance of integrated genetic profiling in adult T-cell leukemia/lymphoma. Blood. 2018;131(2):215-225.
12. Kataoka K, Nagata Y, Kitanaka A, et al. Integrated molecular analysis of adult T cell leukemia/lymphoma. Nat Genet. 2015;47(11):1304-1315.
13. Kogure Y, Kameda T, Koya J, et al. Whole-genome landscape of adult T-cell leukemia/lymphoma. Blood. 2022;139(7):967-982.
14. Marcais A, Lhermitte L, Artesi M, et al. Targeted deep sequencing reveals clonal and subclonal mutational signatures in adult T-cell leukemia/lymphoma and defines an unfavorable indolent subtype. Leukemia. 2021;35(3):764-776.
15. Shah UA, Chung EY, Giricz O, et al. North American ATLL has a distinct mutational and transcriptional profile and responds to epigenetic therapies. Blood. 2018;132(14):1507-1518.
16. Alvarez-Gomez RM, De la Fuente-Hernandez MA, Herrera-Montalvo L, Hidalgo-Miranda A. Challenges of diagnostic genomics in Latin America. Curr Opin Genet Dev. 2021;66:101-109.
17. Spratt DE, Chan T, Waldron L, et al. Racial/ethnic disparities in genomic sequencing. JAMA Oncol. 2016;2(8):1070-1074.
18. Zhang J, Bajari R, Andric D, et al. The International Cancer Genome Consortium data portal. Nat Biotechnol. 2019;37(4):367-369.
19. Grossman RL, Heath AP, Ferretti V, et al. Toward a shared vision for cancer genomic data. N Engl J Med. 2016;375(12):1109-1112.
20. Therneau T. A package for survival analysis in R 2022 [R package version 3.3-1]. https://CRAN.R-project.org/package=survival. Accessed April 11, 2023.
21. Canisius S, Martens JW, Wessels LF. A novel independence test for somatic alterations in cancer shows that biology drives mutual exclusivity but chance explains most co-occurrence. Genome Biol. 2016;17(1):261.
22. Romanel A, Zhang T, Elemento O, Demichelis F. EthSEQ: ethnicity annotation from whole exome sequencing data. Bioinformatics. 2017;33(15):2402-2404.
23. Daniels J, Doukas PG, Escala MEM, et al. Cellular origins and genetic landscape of cutaneous gamma delta T cell lymphomas. Nat Commun. 2020;11(1):1806.
24. Park J, Daniels J, Wartewig T, et al. Integrated genomic analyses of cutaneous T cell lymphomas reveal the molecular bases for disease heterogeneity. Blood. 2021;138(14):1225-1236.
25. Park J, Yang J, Wenzel AT, et al. Genomic analysis of 220 CTCLs identifies a novel recurrent gain-of-function alteration in RLTPR (p.Q575E). Blood. 2017;130(12):1430-1440.
26. Choi J, Goh G, Walradt T, et al. Genomic landscape of cutaneous T cell lymphoma. Nat Genet. 2015;47(9):1011-1019.
27. Tate JG, Bamford S, Jubb HC, et al. COSMIC: the Catalogue Of Somatic Mutations In Cancer. Nucleic Acids Res. 2019;47(D1):D941-D947.
28. Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA Jr, Kinzler KW. Cancer genome landscapes. Science. 2013;339(6127):1546-1558.
29. Foster JM, Oumie A, Togneri FS, et al. Cross-laboratory validation of the OncoScan(R) FFPE Assay, a multiplex tool for whole genome tumour profiling. BMC Med Genomics. 2015;8:5.
30. Mermel CH, Schumacher SE, Hill B, Meyerson ML, Beroukhim R, Getz G. GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. Genome Biol. 2011;12(4):R41.
31. Boons E, Nogueira TC, Dierckx T, et al. XPO1 inhibitors represent a novel therapeutic option in adult T-cell leukemia, triggering p53-mediated caspase-dependent apoptosis. Blood Cancer J. 2021;11(2):27.
32. Uchida Y, Yoshimitsu M, Hachiman M, et al. RLTPR Q575E: a novel recurrent gain-of-function mutation in patients with adult T-cell leukemia/lymphoma. Eur J Haematol. 2021;106(2):221-229.
33. Nakagawa M, Schmitz R, Xiao W, et al. Gain-of-function CCR4 mutations in adult T cell leukemia/lymphoma. J Exp Med. 2014;211(13):2497-2505.
34. Schmidt R, Steinhart Z, Layeghi M, et al. CRISPR activation and interference screens decode stimulation responses in primary human T cells. Science. 2022;375(6580):eabj4008.
35. Shifrut E, Carnevale J, Tobin V, et al. Genome-wide CRISPR screens in primary human T cells reveal key regulators of immune function. Cell. 2018;175(7):1958-1971.
36. Noll JE, Jeffery J, Al-Ejeh F, et al. Mutant p53 drives multinucleation and invasion through a process that is suppressed by ANKRD11. Oncogene. 2012;31(23):2836-2848.
37. Sirmaci A, Spiliopoulos M, Brancati F, et al. Mutations in ANKRD11 cause KBG syndrome, characterized by intellectual disability, skeletal malformations, and macrodontia. Am J Hum Genet. 2011;89(2):289-294.
38. Chen SS, Hu Z, Zhong XP. Diacylglycerol kinases in T cell tolerance and effector function. Front Cell Dev Biol. 2016;4:130.
39. Birnbaum ME, Berry R, Hsiao YS, et al. Molecular architecture of the alphabeta T cell receptor-CD3 complex. Proc Natl Acad Sci U S A. 2014;111(49):17576-17581.
40. Plas DR, Johnson R, Pingel JT, et al. Direct regulation of ZAP-70 by SHP-1 in T cell antigen receptor signaling. Science.

1996;272(5265):1173-1176.
41. Ouyang W, Beckett O, Ma Q, Paik JH, DePinho RA, Li MO. Foxo proteins cooperatively control the differentiation of Foxp3+ regulatory T cells. Nat Immunol. 2010;11(7):618-627.
42. Liu Y, Ao X, Ding W, et al. Critical role of FOXO3a in carcinogenesis. Mol Cancer. 2018;17(1):104.
43. Olagnier D, Sze A, Bel Hadj S, et al. HTLV-1 Tax-mediated inhibition of FOXO3a activity is critical for the persistence of terminally differentiated CD4+ T cells. PLoS Pathog. 2014;10(12):e1004575.
44. Tanaka-Nakanishi A, Yasunaga J, Takai K, Matsuoka M. HTLV-1 bZIP factor suppresses apoptosis by attenuating the function of FoxO3a and altering its localization. Cancer Res. 2014;74(1):188-200.
45. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008;9:559.
46. Li Z, Bridges B, Olson J, Weinman SA. The interaction between acetylation and serine-574 phosphorylation regulates the apoptotic function of FOXO3. Oncogene. 2017;36(13):1887-1898.
47. Phillips AA, Fields PA, Hermine O, et al. Mogamulizumab versus investigator's choice of chemotherapy regimen in relapsed/refractory adult T-cell leukemia/lymphoma. Haematologica. 2019;104(5):993-1003.
48. National Comprehensive Cancer Network. NCCN Clinical Practice Guidelines in Oncology. Chronic lymphocytic leukemia/small lymphocytic lymphoma version 2.2022. https://www.nccn.org/professionals/physician_gls/pdf/cll.pdf. Accessed January 18, 2022.
49. National Comprehensive Cancer Network. NCCN Clinical Practice Guidelines in Oncology. B-cell lymphomas 2.2022. https://www.nccn.org/professionals/physician_gls/pdf/b-cell.pdf. Accessed March 21, 2022.
50. Richardson NC, Kasamon Y, Pazdur R, Gormley N. The saga of PI3K inhibitors in haematological malignancies: survival is the ultimate safety endpoint. Lancet Oncol. 2022;23(5):563-566.
51. Horwitz SM, Koch R, Porcu P, et al. Activity of the PI3K-δ,γ inhibitor duvelisib in a phase 1 trial and preclinical models of T-cell lymphoma. Blood. 2018;131(8):888-898.
52. Katsuya H, Cook LBM, Rowan AG, Satou Y, Taylor GP, Bangham CRM. Phosphatidylinositol 3-kinase-δ (PI3K-δ) is a potential therapeutic target in adult T-cell leukemia-lymphoma. Biomark Res. 2018;6:24.
53. Yu F, Sharma S, Edwards J, Feigenbaum L, Zhu J. Dynamic expression of transcription factors T-bet and GATA-3 by regulatory T cells maintains immunotolerance. Nat Immunol. 2015;16(2):197-206.
54. Yamagata T, Nishida J, Tanaka S, et al. A novel interferon regulatory factor family transcription factor, ICSAT/Pip/LSIRF, that negatively regulates the activity of interferon-regulated genes. Mol Cell Biol. 1996;16(4):1283-1294.
55. Ramos JC, Ruiz P Jr, Ratner L, et al. IRF-4 and c-Rel expression in antiviral-resistant adult T-cell leukemia/lymphoma. Blood. 2007;109(7):3060-3068.