Association of clonal hematopoiesis with angioimmunoblastic T-cell lymphoma: genetic evidence for stem cell origin, etiology and prediction of secondary hematologic malignancy

Shuhua Cheng[1], Wei Zhang[2], Giorgio Inghirami[1], Wayne Tam[1]*

[1]Department of Pathology and Laboratory Medicine, [2] Genomics Resources Core Facility, Weill Cornell Medicine, New York, NY 10021

***Correspondence:** Wayne Tam, MD, PhD
Department of Pathology and Laboratory Medicine
Weill Cornell Medicine
Room K507, 1300 York Avenue
New York, NY 10065
Phone: 212-746-6486
Fax: 212-746-8345
Email: wtam@med.cornell.edu

Running Title: CH-related evidence for cell origin, etiology and prediction of CHN in AITL

Abstract

Background Initiation and progression of Angioimmunoblastic T cell lymphoma (AITL) and peripheral T cell lymphoma, not otherwise specified (PTCL-NOS), remains poorly understood. A subset of AITL/PTCL-NOS patients develop concomitant hematologic neoplasms (CHN), including myeloid and B-cell neoplasms, and a biomarker to predict this risk is lacking. We explore these unknown areas by exploring potential relationship among clonal hematopoiesis (CH), AITL/PTCL-NOS and CHN using a genomics approach.

Methods We performed a retrospective study of patients with AITL (n=25) or PTCL-NOS (n=2) whose lymphoma and matched bone marrow (BM) or peripheral blood (PB) specimens were collected and subjected to targeted next-generation sequencing. Clinicopathologic data of these cases were retrieved from the medical records. CH-associated and lymphoma-specific mutations were identified, analyzed and compared among T-cell lymphoma, CH in BM/PB, and if present, CHN.

Findings 70.4% of the AITL/PTCL-NOS patients were found to harbor CH-associated genomic alterations, characterized by loss-of-function mutations in *TET2* and *DNMT3A*. The CH-associated alterations were shared among CH and T cell lymphoma, as well as concurrent myeloid neoplasms or diffuse large B-cell lymphoma (DLBC) that occurred before or after AITL development. Mutations acquired at a later stage of AITL/PTCL-NOS development were enriched for tobacco smoking-associated missense mutations. Kaplan-Meier survival analysis showed that the presence of ≥2 CH-associated *TET2* mutations with ≥ 15% allele burden was associated with significantly higher risk for CHN (hazard ratio = 10.81, CI 2.2 to 53.16, p value = 0.0022, PPV = 87.5%, NPV = 92.7%).

Interpretation We provided genetic evidence that AITL/PTCL-NOS, CH, myeloid neoplasms and even DLBCL can frequently arise from common mutated hematopoietic stem-cell clones. Our findings also support smoking exposure as a potential risk factor for AITL/PTCL-NOS development, and identify potential biomarker for CHN risk in AITL/PTCL-NOS patients.

Funding the National Cancer Institute.

Key words

Angioimmunoblastic T cell lymphoma, concurrent hematologic neoplasm, etiology, clonal hematopoiesis, mutated hematopoietic stem cells, TET2 mutation, Prediction.

Research in context

Evidence before this study

With no language and time restriction, we searched PubMed up to January 2018 using the search terms “angioimmunoblastic T cell lymphoma” (AITL), “peripheral t cell lymphoma, not otherwise specified”(PTCL-NOS), “clonal hematopoiesis” (CH), “mutated hematopoietic stem cell”, “concomitant hematologic neoplasm” and “biomarker”. This search aimed to identify studies that documented relationship among clonal hematopoiesis, AITL/PTCL-NOS and concomitant hematologic neoplasms, as well as those documented etiology of AITL and biomarkers for development of concomitant hematologic neoplasm. We found several published studies with sporadic cases in which the *TET2* or/and *DNMT3A* mutations detected in AITL were also found in their BM/PB compartments or concomitant myeloid neoplasms. At the end of May this year, when our studies was in progress and approached completion, a published study showed that clonal hematopoiesis associated mutations identified in the marrow were also found in the matched primary T cell lymphomas and concomitant myeloid neoplasms that occurred after AITL development.

Added value of this study

This retrospective study, with the largest cohort of patients reported to date, showed that AITL/PTCL-NOS, clonal hematopoiesis, concurrent myeloid or B cell neoplasms, were often initiated and progressed from common mutated hematopoietic stem cells. We also demonstrated that these concurrent hematologic neoplasms could occur after or before AITL development. This study discovered smoking exposure as a potential risk factor for AITL development in healthy adults who harbor clonal hematopoiesis, and identified two or more clonal

hematopoiesis-associated TET2 mutations with ≥ 15% of allele burden as a predictor for development of concurrent hematologic neoplasms in AITL patients.

Implications of all the available evidence

Our data suggests that clonal hematopoiesis may be considered as a pre-malignant state for both lymphoid and myeloid malignancies. Understanding of the mechanism of clonal hematopoiesis and its progression to frank hematopoietic malignancy will help development of effective intervention strategy. Public health measures like cessation of smoking might potentially reduce risk of developing AITL in healthy adults. Regularly surveilling TET2 mutation status in peripheral blood in AITL/PTCL-NOS patients might potentially assist to predict risk of second hematologic neoplasm, consequently optimizing clinical management to improve clinical outcome.

Introduction

Peripheral T cell lymphoma (PTCL) is a heterogenous group of lymphoid tumors and encompass PTCL-NOS, angioimmunoblastic T cell lymphoma (AITL) and several other entities of T-cell lymphoma [1], likely driven by an array of recurrent genomic defects[2]. Except for PTCL-NOS, AITL is the most common subtype of PTCL (18.5% of mature T cell lymphoma) and is believed to arise from a subset of peripheral mature CD4+ T cells corresponding to follicular helper T (T_{FH}) cells, characterized immunophenotypically by expression of a set of cellular markers like PD1, CXCR5, BCL-6, CD10, CXCL13 and ICOS-1 [3-7]. Morphologically, AITL is typically characterized by a polymorphous lymphoid infiltrate with a proliferation of medium-sized tumor cells with clear cytoplasm (clear cell immunoblasts), associated with prominent proliferation of high endothelial venules and follicular dendritic cells. A subset of PTCL, termed as PTCL with T_{FH} phenotype in the updated WHO classification, may be biologically related to AITL, sharing some clinic-pathologic features with AITL and is thought to be also derived from T_{FH} [8]. Although progress has been made in understanding AITL pathogenesis and developing new treatment[9], AITL remains as an aggressive lymphoid tumor, with low estimated rates of overall and failure-free survival at five years (33% and 18%, respectively) [10]. To develop more effective therapeutic agents against AITL and PTCL in general, with T_{FH} phenotype further understanding of the molecular pathogenic mechanisms of AITL is needed.

Genetically, AITL is characterized by a number of genomic mutations in *TET2*, *RHOA*, *DNMT3A* and *IDH2* [2,11-14]. Cell-intrinsic and/or extrinsic factors that facilitate the accumulation of these AITL-related mutations remain unclear. Mutations in *TET2* and *DNMT3A* are also frequently associated with myeloid malignancies, including acute myeloid leukemia (AML), myeloproliferative neoplasms (MPN) and myelodysplastic/myeloproliferative neoplasms

(MDS/MPN). *TET2* and *DNMT3A* are also the most commonly mutated genes associated with clonal hematopoiesis (CH) in healthy adults, especially those over 60 years of age. CH has been shown to be an aging-related process characterized by the clonal expansion of hematopoietic cells harboring one or more somatic mutations, as a result of selective advantage in the hematopoietic stem and progenitor cells due to enhanced self-renewal and inhibition of differentiation[15-18]. It has been noted that myeloid and lymphoid malignancies may co-occur in the same patients [19]. Considering the similarities in genomic mutation profiles of AITL, myeloid malignancies and CH, it has been postulated that there may be a biological link between these entities. To test this, the current study implemented next generation sequencing (NGS) approach to analyze neoplastic T cells and paired bone marrow/peripheral blood specimens from a cohort of 27 patients with AITL or PTCL-NOS, and explored the potential of using the genomic findings from this study to shed light into the etiology of AITL development and to predict clinical progression related to development of second hematologic neoplasms, which is one of the clinical challenges in the clinical management of these cancer patients.

Results

Mutation profiling of AITL/PTCL-NOS and matched BM/PB supports an origin of AITL/PTCL-NOS from mutated HSCs associated with clonal hematopoiesis

For mutation profiling, we sequenced 27 pairs of AITL or PTCL-NOS samples using a 538-gene targeted NGS panel that covered recurrently mutated genes associated with T-cell lymphomas[20]. Of the genomic regions targeted by the panel, 90% had a coverage depth of >1,000. Those sequenced samples included 27 diagnostic lymph node (LN) specimens from patients with AITL (n = 25) or PTCL-NOS (n = 2) and their corresponding bone marrow (BM, n=23) or peripheral blood samples (PB, n =4) from our archived specimens (hereafter denoted as AITL/PTCL-NOS). The overall genomic and pathologic findings showed that of the 27 BM/PB samples, 10 had no detectable involvement by AITL or PTCL-NOS (37%), while 17 were involved by the neoplastic T cells (63%) of variable abundance. One BM sample showed concomitant diagnostic involvement by a myeloproliferative neoplasm (MPN) (Patient #20) (**Suppl. Table 1** and **Suppl. Fig. 1C**).

The genomic alterations found in the matched BM/PB can be due to (i) BM/PB involvement by AITL/PTCL-NOS; (ii) CH; or both (i) and (ii). To accurately characterize the mutation spectrum in the BM/PB, we distinguished the CH-associated mutations from those attributed to the BM/PB involvement by the T neoplastic cells according to the following algorithm. The tumor burden (TB) was estimated for each of the BM/PB specimens involved by the lymphomas (16 AITL and 1 PTCL-NOS) based on their histologic, immunophenotyping and T-cell receptor gamma (*TCRG*) gene rearrangement findings (**Suppl. Fig. 1C**, **Suppl. Table 1**), and compared to the variant allele frequencies (VAFs) of the somatic alterations of T-cell lymphoma-associated genes, for example *RHOA* p. G17V, a molecular characteristic of AITL

[11,12,14,21]. The AITL/PTCL-NOS-related variants present in the BM/PBs are highlighted with rectangles on the heat map (**Suppl. Fig. 1C)** or marked in yellow in the **Suppl. Table 3**, and their VAFs (VAF^{inv}) were found to be 1.06% on average (median, 0.565%; range: 0.1 - 5.64%) (**Suppl. Fig. 1D)**. The variants whose VAFs could not be attributed to AITL/PTCL-NOS involvement alone in the BM/PB, or the variants detected in BM/PB uninvolved by lymphoma were presumed to correspond to variants related to CH, which was confirmed by the presence of these CH-associated mutations in purified neutrophils in the peripheral blood of one patient (#24, **Suppl. Fig.4**). Compared to the AITL/PTCL-NOS-related variants, the VAFs of the CH-associated variants (VAF^{CH}) were significantly higher ($p<0.0001$), ranging from 0.24% to 51.5% with a mean of 22.5% (median: 16.5%), approximately 21.2 times higher than the VAF^{inv} on average (**Suppl. Fig. 1D**).

We identified in the matched BM/PB specimens 44 variants from 14 genes, excluding the variants attributed to the BM/PB involvement by AITL/PTCL-NOS as described above (**Suppl. Fig. 1**). These alterations included 17 missense (38.64%) and 12 nonsense (27.27%) SNVs, 6 frameshift (13.64%) deletions, 7 frameshift (15.91%) and 1 in-frame insertions (**Fig. 1B**, **Suppl. Fig. 3**, **Suppl. Table 2** and **3**). Among these 44 somatic mutations, 37 mutations, identified in 19 of the 27 (70.4%) cases, were shared with those found in the primary lymphoma, and seven were BM/PB specific (**Fig. 1B**, **1C**, **1E)**. The recurrent shared mutations were primarily restricted to *TET2* (74% of the cases) and *DNMT3A* (37% of the cases), consistent with the top CH-associated genes previously reported (**Fig. 1B, 1C**) [22]. All but two (Pt #5, #20) of the cases with CH-associated mutations in the BM/PB did not have a dx of an overt myeloid neoplasm. These CH-associated mutations presumably were acquired very early in the common ancestral hematopoietic stem cells (HSC) from which both the myeloid and T-cell lineages are

derived. Consequently, we defined these shared CH-associated variants as early mutations as seen below. In 8 patients (Patient #7, #9, #11, #12, #17, #21, #23, #26) (29.6%), no CH-associated mutations were detected in the BM or PB.

In the 27 diagnostic LN samples, we identified a total of 102 non-synonymous somatic mutations in 37 genes with a median of ~3 variants per sample, including 62 missense (60.78%) and 20 nonsense (19.61%) single-nucleotide variants (SNVs), 1 in-frame (0.98%) and 9 frameshift (8.8%) deletions, 1 in-frame and 7 frameshift (6.9%) insertions **(Fig. 1B**, **Suppl. Table 2, Suppl. Fig. 2**). Of these 102 mutations, 37 were associated and shared with CH (Fig. 1B, 1C, 1E). In more than half of the T lymphoma cases, not only could we detect early CH-associated mutations, we also identified 65 mutations that are likely acquired during the later stage of AITL/PTCL-NOS development (referred as late mutations hereinafter) (**Fig.1B, 1C**). The recurrent late mutations were limited to several oncogenes and tumor suppressor genes, including the well-known driver genes like *RHOA* (67% of the cases), *TET2* (48%), *IDH2* (33%), *PLCG1*(10%), *TP53*(10%), *VAV1* (10%), and are characterized both by the absence of *DNMT3A* mutations (**Fig. 1C, 1D**) and by the enrichment of missense mutations, which were increased from 36.1% in the CH-associated mutations to 75.2% in the late mutations (proportion test, P value < 0.0001) (**Fig. 1C** and **1E**). The mutations in *IDH2*, *PLCG1*, *TP53*, were found exclusively as late mutations and not CH-associated mutations (**Fig. 1C, 1D**).

Fig.1A shows 4 representative AITL cases and 2 PTCL-NOS cases with their matched BM/PB, where red dots indicate the CH-associated variants present in both the primary lymphoma and BM/PB, and black dots represent the variants associated with AITL/PTCL-NOS (also highlighted with rectangles). Detailed description of these illustrative cases is provided in the supplemental text. There are a couple of notable findings: first, more than one CH clone can

be present in the BM, and their clonal representations in the BM may not reflect those in the lymphoma, as seen in the *DNMT3A*-mutated clones in patient #4. These results suggest that the same *DNMT3A* mutation can have differential effect depending on the cell lineage affected. Second, findings in patient #29 raise the possibility that besides the neoplastic T cells, reactive lymphocytes in these two cases might also harbor the CH-associated mutations.

Our results support a tumor model in which AITL/PTCL, NOS emerges from mutated and expanded HSC clones that are associated with CH in the BM as well as serving as the lymphoma precursors. The latter often acquires additional missense mutations during the course of development to frank lymphomas.

Late mutations in AITL/PTCL, NOS are enriched for C>A transversion substitutions possibly associated with smoking

We investigated whether there might be an etiologic difference between the CH-related mutations and the late mutations by analyzing the type of substitutions for these mutations.

For the CH-associated mutations, overall transition (Ti) and transverse (Tv) substitution rates are comparable (Ti vs Tv, median, 50% vs 50%, mean, 54.76% vs 45.24%, P value >0.05) (**Fig. 2A**). At the base substitution level, C>T is found most frequently (44%), followed by C>G (20%) (**Fig. 2A**), which is consistent with the reported aging and/or overactivity of APOBEC family members-associated mutational process through replicative mutagenesis [23]. In the late mutations, Ti and Tv are also not significantly different (Ti vs Tv, median, 38.53 % vs 61.43 %, mean, 49.31% vs 50.68%, P value >0.05), At the base substitution level, however, besides C>T (35%), C>A emerges as one of the predominant mutant forms (36.7%), which is believed to be associated with the smoking-induced mutational process (**Fig. 2B**) [23]. On a case basis, C>A substitutions are enriched in late mutations compared to CH mutations (mean, 38.8% vs 5.95%;

median, 33.3% vs 0%; t test, P value = 0.0024) (**Fig. 2C**), and the fraction of the cases with C>A substitution in the late mutations is 4-5 times that with C>A in the CH associated mutations (67% vs 16%), suggesting a potential role of the smoking-associated C>A mutational process for acquisition of additional driver mutations in the pathogenesis of AITL.

Enrichment of smoking-associated C>A substitution in the late mutations raises the possibility that patients with AITL/PTCL-NOS might have a higher risk for lung cancer or other smoking associated cancers. To test this hypothesis, we compared the prevalence of lung cancer in three populations, including patients with AITL/PTCL-NOS (the current study cohort, n=28, including one additional PTCL-NOS case without matched PB/BM matched control, Patient #27, see Suppl. Table 1), patients with chronic lymphocytic leukemia (CLL, n= 1329) and aging-matched general US population (n= 153798075, estimated, age, 40 to 80+ years old) as control group. The data for CLL group was published [24], and the data for the control group was downloaded from the CDC website (https://gis.cdc.gov/Cancer/USCS/DataViz.html). Analysis shows that the prevalence of lung cancer in AITL/PTCL-NOS is 27.8 times higher than that in the aged-matched general population (10.71% vs 0.37%, P<0.00001), and 5.47 times higher than that in the CLL group (10.71% vs 1.96%, P<0.00001) (**Fig. 2D**), supporting the association between AITL/PTCL-NOS and smoking-induced tumorigenesis.

AITL with hematologic neoplasms of other lineages arise from common mutated hematopoietic stem cell precursors

Three patients with AITL presented with additional hematologic neoplasms of other lineages. We present here the clonal evolution patterns of these tumors based on the results of the mutation profiling (**Fig. 3**). One of these cases provides genetic evidence for the progression of CH to overt myeloid malignancy through acquisition of additional mutations (patient #5, Fig.

3A). Patient #20 illustrates that the AITL does not necessarily have to be the initially diagnosed malignancy in patients with both AITL and a second malignancy. The third case was an unusual case in which the patient (#14) had CH, AITL as well as DLBCL, the latter was associated with acquisition of an *EZH2* hotspot mutation. Detailed descriptions of these illustrative cases are provided in the supplementary text.

Together, our data further provide evidence that AITL can be associated with the development of a hematopoietic neoplasm of different lineages, i.e. myeloid or B-lymphoid, either preceding or subsequent to the diagnosis of AITL. In all cases, truncal mutations common to all lineages are seen, with late mutations seen in specific tumors (e.g. *SRSF2* in myeloid, *EZH2* in DLBCL).

Impact of destructive *TET2* mutations on development of multiple hematologic malignancies

One of the features shared among the 3 cases with concomitant hematologic neoplasms is that they all had multiple (>1) pathogenic mutations in *TET2*. This observation prompted us to investigate the relationship between *TET2* mutation status and occurrence of multiple hematologic malignancies, specifically through assessing effects of *TET2* mutation status on concomitant hematologic malignancy-free survival in AITL patients. To increase the power of the statistical analysis, the patients included in our study were combined with an outside cohort of AITL patients whose relevant genomic and survival data were recently published [25], leading to the total number of 49 cases for survival analysis.

The patients were initially divided into two groups: wild type *TET2* group (no *TET2* mutation found in the BM/PB samples) and pathogenic *TET2* mutant groups (one or more *TET2* mutation detected in the BM/PB samples). Although there was a trend that AITL patients with

the pathogenic *TET2* mutations detected in the BM/PB had a worse clinical outcome, no statistically significant differences in the second hematologic malignancy-free survival were observed (P value = 0.3273, stratified hazard ratio = 0.29), consistent with the literature [26].

We further stratified the patients into the high *TET2* mutation burden and no or low *TET2* mutation burden subgroups. The criteria for inclusion in the first subgroup (n =8) are as follows: (1) the CH identified in the BM/PB harbored two or more pathogenic mutations in *TET2*, including pathogenic SNVs, nonsense and frameshift mutations interpreted as "Tier 1" or "Tier 2" mutations according to a published professional guideline in molecular pathology [27]; (2) The VAF of each of the pathogenic *TET2* variants described in (1) was ≥ 15%. The cases that did not meet these two criteria were assigned to the second subgroup (n=41). Kaplan-Meier survival analysis showed that AITL patients carrying two or more pathogenic *TET2* mutations with high allelic burden (the first subgroup) had significantly shorter second hematologic malignancy-free survival (P value = 0.0034) (**Fig. 4**). Cox proportional hazards model also estimated a significantly higher hazard ratio (HR) in the first subgroup (stratified hazard ratio, 10.81, 95% CI of ratio, 2.2 to 53.2). Further analysis shows that specificity and sensitivity of this CHN biomarker reach 97.44%, 70%, respectively, with 87.5% of positive predictive value (PPV) and 92.7% of negative predictive value (NPV). This survival analysis indicates that harboring two or more pathologic mutations in *TET2* with relatively high allele burden (≥15%) is an independent risk factor to predict the second hematologic malignancies in AITL patients.

Discussion

In the current study, we examined the landscape of the genomic alterations in AITL/PTLC-NOS and their paired BM or PB using a large-panel targeted sequencing approach in the largest cohort of the AITL patients reported to date. We demonstrated that in about 60% of AITL/PTLC-NOS patients, identical pathogenic *TET2* and/or *DNMT3A* mutations were shared between AITL/PTCL-NOS and CH found in the BM or PB. Studies of large cohorts have demonstrated an increased risk of hematologic malignancy for CH [22,28], but no definitive link has been established between CH and AITL/PTCL-NOS from those studies. Our findings suggest that these *TET2* and/or *DNMT3A* mutations occur very early in the HSC before they give rise to the common lymphoid progenitors and common myeloid progenitors, and propose a possible link between CH and development of AITL[2]. Interestingly, the VAF of the CH-associated mutations is 22.5% on average in our cohort, and is higher compared to the average VAF of CH-related mutations in the general population [22]. This observation is in line with the higher risk of hematopoietic malignancy associated with increased VAF (>10%)[29]. It is conceivable that certain *TET2* or *DNMT3A* mutations are stronger drivers which can result in more expanded CH and/or higher efficient T-cell lymphoma development. For example, as seen in patient #28, there were 3 *TET2* mutations identified in the LN, each of which appears to be present in separate clones and have different capacity to generate CH based on VAF in the BM (0%, 5.47% and 10.89%, respectively). In addition, our study supports a mutated HSC origin for AITL. As the *TET2* and/or *DNMT3A* mutations are propagated to the lymphoid and myeloid progeny of the mutated HSC, it can be speculated that in the lymphoid compartment, the impacts of these mutations vary depending on the developmental and differentiation stage of the T-cells, and may be most felt in the T-cells of follicular helper cell origin (TFH). Lastly, our interesting

case of an AITL patient with CH and subsequent development of DLBCL and the sharing of the same TET2 mutation among all three lesions suggest that a subset of DLBCL, possibly the molecular subtype characterized by mutated *TET2* [30] may originate from mutated HSC. To our knowledge, this is the first reported case in which the mutated HSC developed into three distinct tumors of diverse lineages.

The findings from this investigation confirm and extend the results previously published regarding the cellular origin of AITL [12,21,25,31-34]. Most of these previous studies presented sporadic AITL cases in which the *TET2* or/and *DNMT3A* mutations present in AITL were also found in their BM/PB compartments. Two reports showed that AITL shared the same *TET2* mutations with the isolated $CD20^+/CD19^+$ (B cells) or $CD34^+$ cells [31,33]. These studies also pointed to a mutated HSC that gives rise to lymphoid and myeloid cells harboring the same mutations. A high-risk CH was also documented as the cellular origin of AITL and *NPM1*-mutated AML in a patient [21]. While our manuscript was under preparation, the results of a study conceptually similar to ours regarding the cellular origin of AITL was reported [25]. Consistent with our observations, the report showed that the mutations related to CH (i.e. *TET2* or *DNMT3A*) were detected in both the neoplastic T-cell and myeloid compartments in 15 out of 22 AITL patients (68%), and associated with second myeloid neoplasm development after the diagnosis of AITL in 4 cases. However, in their cohort, no cases were reported where AITL developed subsequent to myeloid neoplasms. Our study presented one such case (Patient #20) whose AITL developed after 10 years of PV and the two hematologic neoplasms shared three identical *JAK2* and *TET2* mutations (**Fig. 3B**). The identification of cases in which myeloid neoplasms precede the diagnosis of AITL provides additional supportive evidence to the postulation that the mutated HSCs are the common origin for these hematologic neoplasms,

which develop independently and divergently in tumor evolution. Whether ATIL precedes or develops subsequent to the myeloid neoplasms may depend on the stochastic dynamics of the clonal evolution.

Additional late non-CH mutations are found in 68.4% of the AITL/PTCL-NOS in our cohort, consistent with the belief that CH-associated *TET2* and *DNMT3A* mutations are insufficient for tumorigenesis and additional genetic alterations are required. Consistent with this notion, in AITL animal models, *TET2* disruption or $RHOA^{G17V}$ expression alone failed to induce AITL development; however, AITL-like lymphoma developed once *TET2* disruption and $RHOA^{G17V}$ expression were combined[35-37]. For the development of myeloid neoplasms, additional mutations beyond CH-associated *TET2* and *DNMT3A* mutations drive further clonal expansion from CH. These mutations may be acquired early (patient #20, *JAK2***, Fig, 3B**) or late during tumor development (patient # 24, *JAK2*, Suppl. Fig. 4).

We discovered that the late non-CH mutations are enriched for the missense mutations and the C>A substitutions (**Fig. 1E**, **Fig. 2**). In the current cohort, the C>A base substitution is associated with critical mutations in a number of oncogenic genes. This finding may have implications on treatment and prevention of AITL. It is believed that the C>A mutation is likely caused by mis-replication or mis-repairing of DNA damage induced by tobacco carcinogens [23,38], which largely result in missense mutations[39]. Consistent with this causative link is our findings that the majority of the late non-CH mutations identified in our cohort were missense mutation (75.2%, Fig.1E), and that the patients with AITL/PTCL-NOS have a 27.8-fold increased risk for development of lung cancer compared to the age-matched general population (Fig. 2D). Consequently, our findings suggests that cessation of smoking may be a potential effective

intervention to prevent AILT development in higher risk population, particularly those already found to harbor CH.

On the contrary, the CH-associated genetic alterations in the current cohort are characterized primarily by the C>T and C>G mutations (64% of all the mutations) in *TET2* and *DNMT3A* (**Fig.2A**, **2C**). This pattern was reported to be associated with the AID/APOBEC family of cytidine deaminases or aging-dependent function decline of base-excision repair machinery [23]. This mutational mechanism might also play a role in the non-CH late mutations, as 35% of the non-CH mutations were C>T substitutions (**Fig. 2B**, **2C**). Interestingly, reduced accumulation of the C to T mutations by inactivation of *AID* blocked development of B cell malignancies in aging *TET2* deficient mice[40], implying that AID might be a therapeutic targeting candidate for lymphoma, including AITL.

Furthermore, we found that CH associated with multiple-hit *TET2* (defined as ≥2 pathogenic *TET2* mutations with VAFs of ≥ 15%) is an independent risk factor for development of concurrent hematologic malignancies (**Fig. 4**). Recently, certain features of CH predictive of hematopoietic malignancy development were identified[29]. These features include >1 mutated gene, VAF≥ 10% and mutations in specific genes and variants, for example *TP53* and *IDH1/2*Our *TET2* biomarker includes 2 or more mutations and VAF of at least 15%. However, *TET2* has not been previously implicated as a marker for increased risk of hematologic malignancy in CH in general. It is possible that this multiple-hit TET2 biomarker is specific and only relevant in the setting of patients with AITL and CH. Mechanistically, it is possible that two or more *TET2* mutations each with relatively high mutation burden (>15%) correlates with increased clonal expansion and/or more severe disruption of TET2 activity, thereby increasing the global chance of acquiring additional driver mutations and hence increased risk for

development of second hematologic neoplasms. Consistent with this hypothesis, aging *TET2* deficient mice develop diverse hematologic malignancies [41]. A reliable predictor for concurrent hematologic malignancies may be helpful for clinical stratification and management for this subset of the AITL patients. For myeloid malignancies with the double-hit *TET2* mutations, more intensive therapeutic regimen like bone marrow transplantation might be warranted for reducing the risk of AITL development.

In summary, our study provides genomic evidence of an origin of AITL/PTCL-NOS from a mutated HSC clone, which can be associated with CH as well as development of myeloid and even B-cell malignancies. The development of these hematopoietic malignancies of different lineages occur via divergent evolution from the mutated HSC clone, often with acquisition of additional mutations frequently induced by mutagenic agents like tobacco. We also identified a potential biomarker: two or more pathogenic TET2 mutations with high mutation burden for the development of second hematologic neoplasm in AITL patients. Single cell methodology will help definitively determine the clonal architecture in lymphoma and BM with multiple mutations and enhance our understanding of the initiation of tumor induced by CH-associated mutations.

Methods

The detailed methods were described in the supplemental text.

Contributors

SC contributed to the study design and conceptualization of the project, optimized bioinformatics pipeline, analyzed and interpreted the data, and wrote the manuscript. WZ performed the experiment and contributed to the data analysis. GI provided critical study materials and contributed to the revision of the manuscript. WT conceptualized and directed the project, designed the experiments, analyzed and interpreted the data, and wrote the manuscript. All authors read the manuscript and approved it before submission.

Declaration of interests

This study is supported by a grant (R01 CA194547) from the National Cancer Institute to WT.

All the authors declare no conflicts of interest.

References

1. Swerdlow S, Campo E, Harris N, et al. WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues. Vol. 2 (ed Revised 4th). Lyon, France: IARC; 2017.
2. Fiore D, Cappelli LV, Broccoli A, Zinzani PL, Chan WC, Inghirami G. Peripheral T cell lymphomas: from the bench to the clinic. *Nat Rev Cancer*. 2020;20(6):323-342.
3. Chiba S, Sakata-Yanagimoto M. Advances in understanding of angioimmunoblastic T-cell lymphoma. *Leukemia*. 2020.
4. Marafioti T, Paterson JC, Ballabio E, et al. The inducible T-cell co-stimulator molecule is expressed on subsets of T cells and is a new marker of lymphomas of T follicular helper cell-derivation. *Haematologica*. 2010;95(3):432-439.
5. Mourad N, Mounier N, Briere J, et al. Clinical, biologic, and pathologic features in 157 patients with angioimmunoblastic T-cell lymphoma treated within the Groupe d'Etude des Lymphomes de l'Adulte (GELA) trials. *Blood*. 2008;111(9):4463-4470.
6. Dupuis J, Boye K, Martin N, et al. Expression of CXCL13 by neoplastic cells in angioimmunoblastic T-cell lymphoma (AITL): a new diagnostic marker providing evidence that AITL derives from follicular helper T cells. *Am J Surg Pathol*. 2006;30(4):490-494.
7. Attygalle A, Al-Jehani R, Diss TC, et al. Neoplastic T cells in angioimmunoblastic T-cell lymphoma express CD10. *Blood*. 2002;99(2):627-633.
8. Swerdlow SH, Campo E, Pileri SA, et al. The 2016 revision of the World Health Organization classification of lymphoid neoplasms. *Blood*. 2016;127(20):2375-2390.
9. Lemonnier F, Dupuis J, Sujobert P, et al. Treatment with 5-azacytidine induces a sustained response in patients with angioimmunoblastic T-cell lymphoma. *Blood*. 2018;132(21):2305-2309.

10. Federico M, Rudiger T, Bellei M, et al. Clinicopathologic characteristics of angioimmunoblastic T-cell lymphoma: analysis of the international peripheral T-cell lymphoma project. *J Clin Oncol*. 2013;31(2):240-246.
11. Yoo HY, Sung MK, Lee SH, et al. A recurrent inactivating mutation in RHOA GTPase in angioimmunoblastic T cell lymphoma. *Nat Genet*. 2014;46(4):371-375.
12. Sakata-Yanagimoto M, Enami T, Yoshida K, et al. Somatic RHOA mutation in angioimmunoblastic T cell lymphoma. *Nat Genet*. 2014;46(2):171-175.
13. Odejide O, Weigert O, Lane AA, et al. A targeted mutational landscape of angioimmunoblastic T-cell lymphoma. *Blood*. 2014;123(9):1293-1296.
14. Nakamoto-Matsubara R, Sakata-Yanagimoto M, Enami T, et al. Detection of the G17V RHOA mutation in angioimmunoblastic T-cell lymphoma and related lymphomas using quantitative allele-specific PCR. *PLoS One*. 2014;9(10):e109714.
15. Challen GA, Goodell MA. Clonal hematopoiesis: mechanisms driving dominance of stem cell clones. *Blood*. 2020;136(14):1590-1598.
16. Steensma DP, Ebert BL. Clonal hematopoiesis as a model for premalignant changes during aging. *Exp Hematol*. 2020;83:48-56.
17. Gondek LP, DeZern AE. Assessing clonal haematopoiesis: clinical burdens and benefits of diagnosing myelodysplastic syndrome precursor states. *Lancet Haematol*. 2020;7(1):e73-e81.
18. Jaiswal S, Ebert BL. Clonal hematopoiesis in human aging and disease. *Science*. 2019;366(6465).

19. Holst JM, Plesner TL, Pedersen MB, et al. Myeloproliferative and lymphoproliferative malignancies occurring in the same patient: a nationwide discovery cohort. *Haematologica*. 2019.

20. Fiore D, Cappelli LV, Zumbo P, et al. A Novel JAK1 Mutant Breast Implant-Associated Anaplastic Large Cell Lymphoma Patient-Derived Xenograft Fostering Pre-Clinical Discoveries. *Cancers (Basel)*. 2020;12(6).

21. Tiacci E, Venanzi A, Ascani S, et al. High-Risk Clonal Hematopoiesis as the Origin of AITL and NPM1-Mutated AML. *N Engl J Med*. 2018;379(10):981-984.

22. Genovese G, Kahler AK, Handsaker RE, et al. Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. *N Engl J Med*. 2014;371(26):2477-2487.

23. Alexandrov LB, Nik-Zainal S, Wedge DC, et al. Signatures of mutational processes in human cancer. *Nature*. 2013;500(7463):415-421.

24. Parekh K, Rusch V, Kris M. The clinical course of lung carcinoma in patients with chronic lymphocytic leukemia. *Cancer*. 1999;86(9):1720-1723.

25. Lewis NE, Petrova-Drus K, Huet S, et al. Clonal hematopoiesis in angioimmunoblastic T-cell lymphoma with divergent evolution to myeloid neoplasms. *Blood Adv*. 2020;4(10):2261-2271.

26. Lemonnier F, Couronne L, Parrens M, et al. Recurrent TET2 mutations in peripheral T-cell lymphomas correlate with TFH-like features and adverse clinical parameters. *Blood*. 2012;120(7):1466-1469.

27. Li MM, Datto M, Duncavage EJ, et al. Standards and Guidelines for the Interpretation and Reporting of Sequence Variants in Cancer: A Joint Consensus Recommendation of the

Association for Molecular Pathology, American Society of Clinical Oncology, and College of American Pathologists. *J Mol Diagn*. 2017;19(1):4-23.

28. Jaiswal S, Fontanillas P, Flannick J, et al. Age-related clonal hematopoiesis associated with adverse outcomes. *N Engl J Med*. 2014;371(26):2488-2498.
29. Warren JT, Link DC. Clonal hematopoiesis and risk for hematologic malignancy. *Blood*. 2020;136(14):1599-1605.
30. Lacy SE, Barrans SL, Beer PA, et al. Targeted sequencing in DLBCL, molecular subtypes, and outcomes: a Haematological Malignancy Research Network report. *Blood*. 2020;135(20):1759-1771.
31. Schwartz FH, Cai Q, Fellmann E, et al. TET2 mutations in B cells of patients affected by angioimmunoblastic T-cell lymphoma. *J Pathol*. 2017;242(2):129-133.
32. Nguyen TB, Sakata-Yanagimoto M, Asabe Y, et al. Identification of cell-type-specific mutations in nodal T-cell lymphomas. *Blood Cancer J*. 2017;7(1):e516.
33. Couronne L, Bastard C, Bernard OA. TET2 and DNMT3A mutations in human T-cell lymphoma. *N Engl J Med*. 2012;366(1):95-96.
34. Quivoron C, Couronne L, Della Valle V, et al. TET2 inactivation results in pleiotropic hematopoietic abnormalities in mouse and is a recurrent event during human lymphomagenesis. *Cancer Cell*. 2011;20(1):25-38.
35. Nguyen TB, Sakata-Yanagimoto M, Fujisawa M, et al. Dasatinib Is an Effective Treatment for Angioimmunoblastic T-cell Lymphoma. *Cancer Res*. 2020;80(9):1875-1884.
36. Ng SY, Brown L, Stevenson K, et al. RhoA G17V is sufficient to induce autoimmunity and promotes T-cell lymphomagenesis in mice. *Blood*. 2018;132(9):935-947.

37. Cortes JR, Ambesi-Impiombato A, Couronne L, et al. RHOA G17V Induces T Follicular Helper Cell Specification and Promotes Lymphomagenesis. *Cancer Cell*. 2018;33(2):259-273 e257.

38. Alexandrov LB, Ju YS, Haase K, et al. Mutational signatures associated with tobacco smoking in human cancer. *Science*. 2016;354(6312):618-622.

39. Blackford A, Parmigiani G, Kensler TW, et al. Genetic mutations associated with cigarette smoking in pancreatic cancer. *Cancer Res*. 2009;69(8):3681-3688.

40. Mouly E, Ghamlouch H, Della-Valle V, et al. B-cell tumor development in Tet2-deficient mice. *Blood Adv*. 2018;2(6):703-714.

41. Pan F, Wingo TS, Zhao Z, et al. Tet2 loss leads to hypermutagenicity in haematopoietic stem/progenitor cells. *Nat Commun*. 2017;8:15102.

Figure legend

Fig. 1 Analysis of genomic alterations by target sequencing panel for primary lymphomas and paired BM/PB in patients with AITL and PTCL-NOS. (A) Presence of CH in patients with AITL and PTCL-NOS. Dot plots showing the detected variants and their VAFs in the AITL and PTCL-NOS (LN) and their matched BM/PB in representative AITL and PTCL-NOS cases with CH. The black circles indicate variants specific to the lymphomas, and the variants shared between the primary lymphomas and CH are highlighted in red. The variants attributed to lymphoma only are boxed. VAF, variant allele frequency; LN, lymph node; BM, bone marrow; PB, peripheral blood; CH, clonal hematopoiesis. Additional detailed descriptions of these illustrative cases are provided in the supplementary text. (B) Venn diagram showing the distribution of the shared, lymphoma or BM/PB-specific variants identified in the diagnostic LN and paired BM/PB samples. The shared variants are defined as variants identified in both the primary lymphoma and the BM/PB, the latter as CH-related variants. The variants predicted to be due only to lymphoma involvement in BM/PB have been excluded (see also **Suppl. Fig.1** for the distribution of all variants). (C) Summary of the CH-associated mutations in the BM/PB and LN, and the mutations postulated to accumulate at a later stage of lymphoma development (late mutations). The CH-associated mutations are shared between the primary lymphomas and the BM/PB and can be considered as early lesions in AITL/PTCL. The heat maps show the top recurrent mutations in both categories. Stacked bar plots show the type of variants and the mutation frequency (relative to our cohort) for each of the major mutated genes in the LN and BM/PB samples. (D) Forest plot showing differentially affected genes between the CH-associated and late mutations. (E) Comparison of the distribution of disruptive and missense

mutations in the CH-associated and late mutations. ∗ $p < 0.05$; ∗∗ $p < 0.01$; ∗∗∗ $p < 0.001$; ∗∗∗∗ $p < 0.0001$; NS, not significant.

Fig. 2 Late mutations in AITL/PTCL, NOS are enriched for C>A transversion substitutions possibly associated with smoking. Titv plot showing overall distribution of the six types of substitutions in the CH (A) and late (B) missense mutations acquired during AITL/PTCL development, as well as fraction of these substitutions in each sample. The median is indicated by a horizontal line. Bar plot on the left showing SNV classes and fraction of each substitution class among all missense mutations. Ti, transitions, Tv, transversions, SNVs, single nucleotide variants. (C) Side by side comparison of transition and transversion base substitutions acquired between the early CH associated and late mutations. (D) Bar plot comparing the prevalence of lung cancer among three populations indicated. ∗ $p < 0.05$; ∗∗ $p < 0.01$; ∗∗∗ $p < 0.001$; ∗∗∗∗∗ $p < 0.00001$; NS, not significant.

Fig. 3 AITL and concomitant hematologic neoplasms develop from common mutated hematopoietic stem cells. (A), (B), (C) Dot plots comparing VAFs of the mutations identified in the AITL and the concomitant hematological malignancies. Red dots show the variants shared between different hematologic neoplasms or entity in the same patient. Dark blue dots in (A) indicate the variants specially related to CMML, and the black dots in (A) denote the AITL-specific variants. Schematic diagrams depicting hypothetical clonal evolution models of the tumors deriving from mutated HSC are also presented. In patient #5 (A), additional mutations besides the CH-associated mutations were identified and implicated in the disease progression to AITL and CMML, respectively. In patient #20, no additional mutations besides those mutated in HSC are identified. In patient #14, a mutated *EZH2*, indicated by green dot, is implicated in the progression to DLBCL. In all 3 cases, there are mutations that are shared between the AITL

and the concomitant myeloid or B lymphomas, supporting evolution of these neoplasms from a common precursor. CMML, chronic myelomonocytic leukemia ; PV, polycythemia vena; Post-PV PMF, post-PV primary myelofibrosis; DLBCL, diffuse large B cell lymphoma; ITP, immune thrombocytopenia.

Fig.4 Pathogenic *TET2* mutation status in the BM/PB samples is a predictive biomarker for concomitant hematologic neoplasms in AITL/PTCL-NOS patients. Kaplan-Meier survival curve analysis of concomitant hematologic neoplasm-free survival in AITL or AITL-related patients based on *TET2* mutation status in the BM/PB. Concomitant hematologic neoplasm-free survival of AITL patients can be stratified based on absent/low or high TET2 mutation burden subgroups. p-value was calculated by log-Rank test. In one case, the second hematologic malignancy (PV) preceded the development of AITL.

Fig. 1

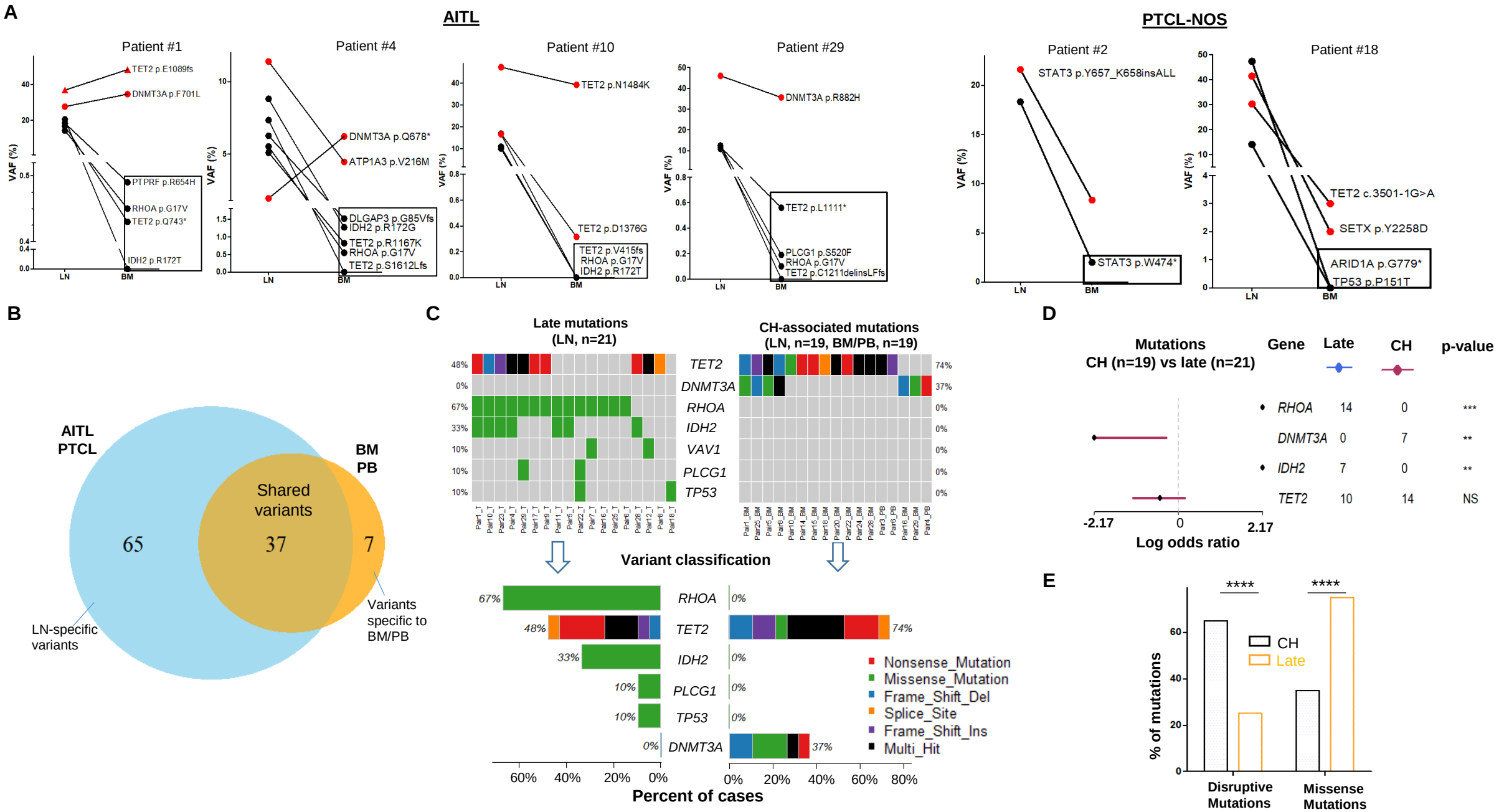

A
AITL
PTCL-NOS
Patient #1
TET2 p.E1089fs
DNMT3A p.F701L
PTPRF p.R654H
RHOA p.G17V
TET2 p.Q743*
IDH2 p.R172T
Patient #4
DNMT3A p.Q678*
ATP1A3 p.V216M
DLGAP3 p.G85Vfs
IDH2 p.R172G
TET2 p.R1167K
RHOA p.G17V
TET2 p.S1612Lfs
Patient #10
TET2 p.N1484K
TET2 p.D1376G
TET2 p.V415fs
RHOA p.G17V
IDH2 p.R172T
Patient #29
DNMT3A p.R882H
TET2 p.L1111*
PLCG1 p.S520F
RHOA p.G17V
TET2 p.C1211delinsLFfs
Patient #2
STAT3 p.Y657_K658insALL
STAT3 p.W474*
Patient #18
TET2 c.3501-1G>A
SETX p.Y2258D
ARID1A p.G779*
TP53 p.P151T
VAF (%)
LN
BM
B
AITL
PTCL
BM
PB
Shared variants
65
37
7
LN-specific variants
Variants specific to BM/PB
C
Late mutations (LN, n=21)
CH-associated mutations (LN, n=19, BM/PB, n=19)
TET2
DNMT3A
RHOA
IDH2
VAV1
PLCG1
TP53
Variant classification
Percent of cases
Nonsense_Mutation
Missense_Mutation
Frame_Shift_Del
Splice_Site
Frame_Shift_Ins
Multi_Hit
D
Mutations CH (n=19) vs late (n=21)
Gene
Late
CH
p-value
RHOA 14 0 ***
DNMT3A 0 7 **
IDH2 7 0 **
TET2 10 14 NS
-2.17
0
2.17
Log odds ratio
E
CH
Late
% of mutations
Disruptive Mutations
Missense Mutations
****
****

Fig. 2

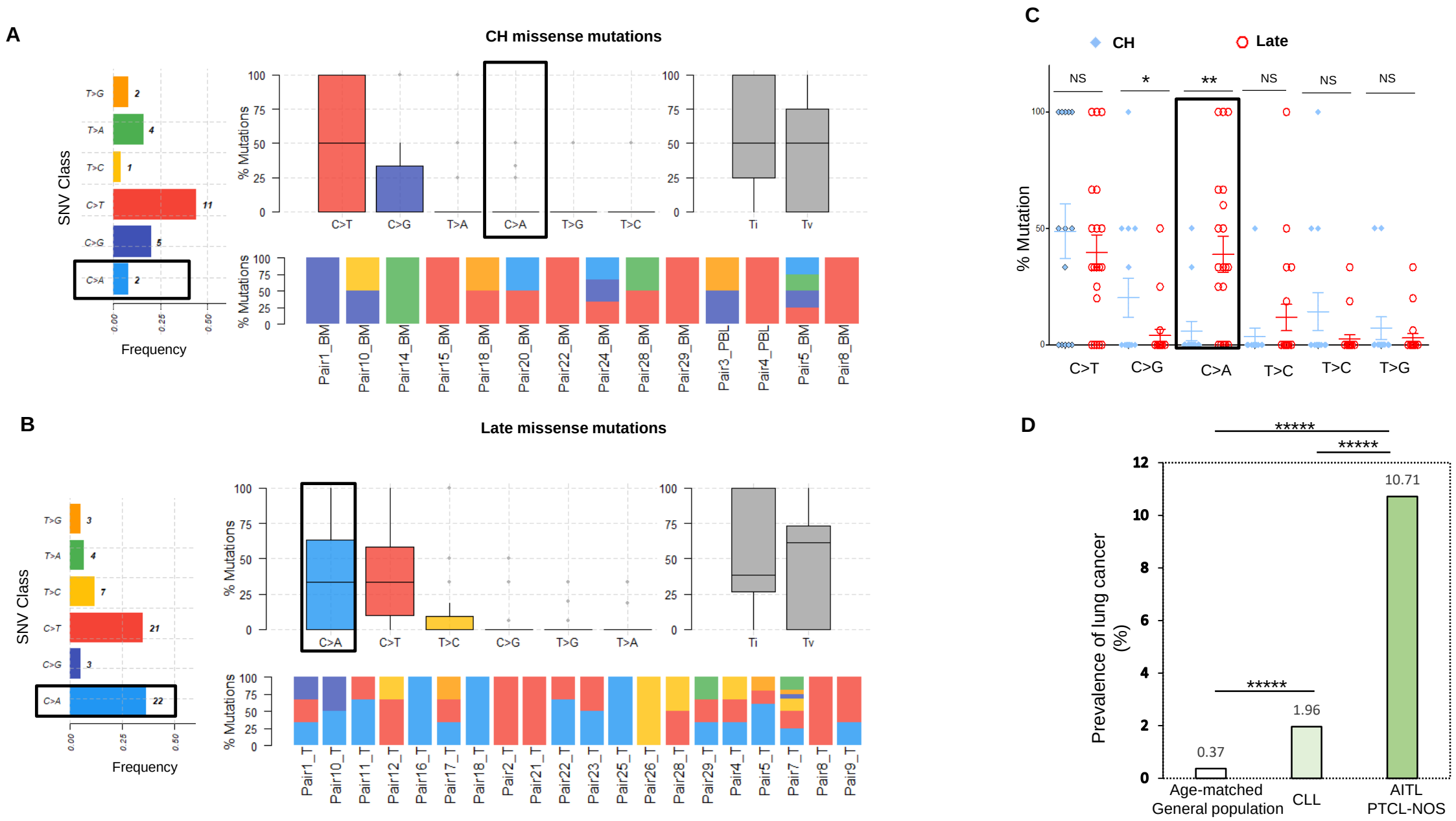

A
CH missense mutations
SNV Class
T>G 2
T>A 4
T>C 1
C>T 11
C>G 5
C>A 2
Frequency
% Mutations
C>T C>G T>A C>A T>G T>C
Ti Tv
Pair1_BM Pair10_BM Pair14_BM Pair15_BM Pair18_BM Pair20_BM Pair22_BM Pair24_BM Pair28_BM Pair29_BM Pair3_PBL Pair4_PBL Pair5_BM Pair8_BM
B
Late missense mutations
SNV Class
T>G 3
T>A 4
T>C 7
C>T 21
C>G 3
C>A 22
Frequency
% Mutations
C>A C>T T>C C>G T>G T>A
Ti Tv
Pair1_T Pair10_T Pair11_T Pair12_T Pair16_T Pair17_T Pair18_T Pair2_T Pair21_T Pair22_T Pair23_T Pair25_T Pair26_T Pair28_T Pair29_T Pair4_T Pair5_T Pair7_T Pair8_T Pair9_T
C
CH
Late
NS * ** NS NS NS
% Mutation
C>T C>G C>A T>C T>C T>G
D
*****
*****
*****
Prevalence of lung cancer (%)
0.37
1.96
10.71
Age-matched General population
CLL
AITL PTCL-NOS

Fig.3

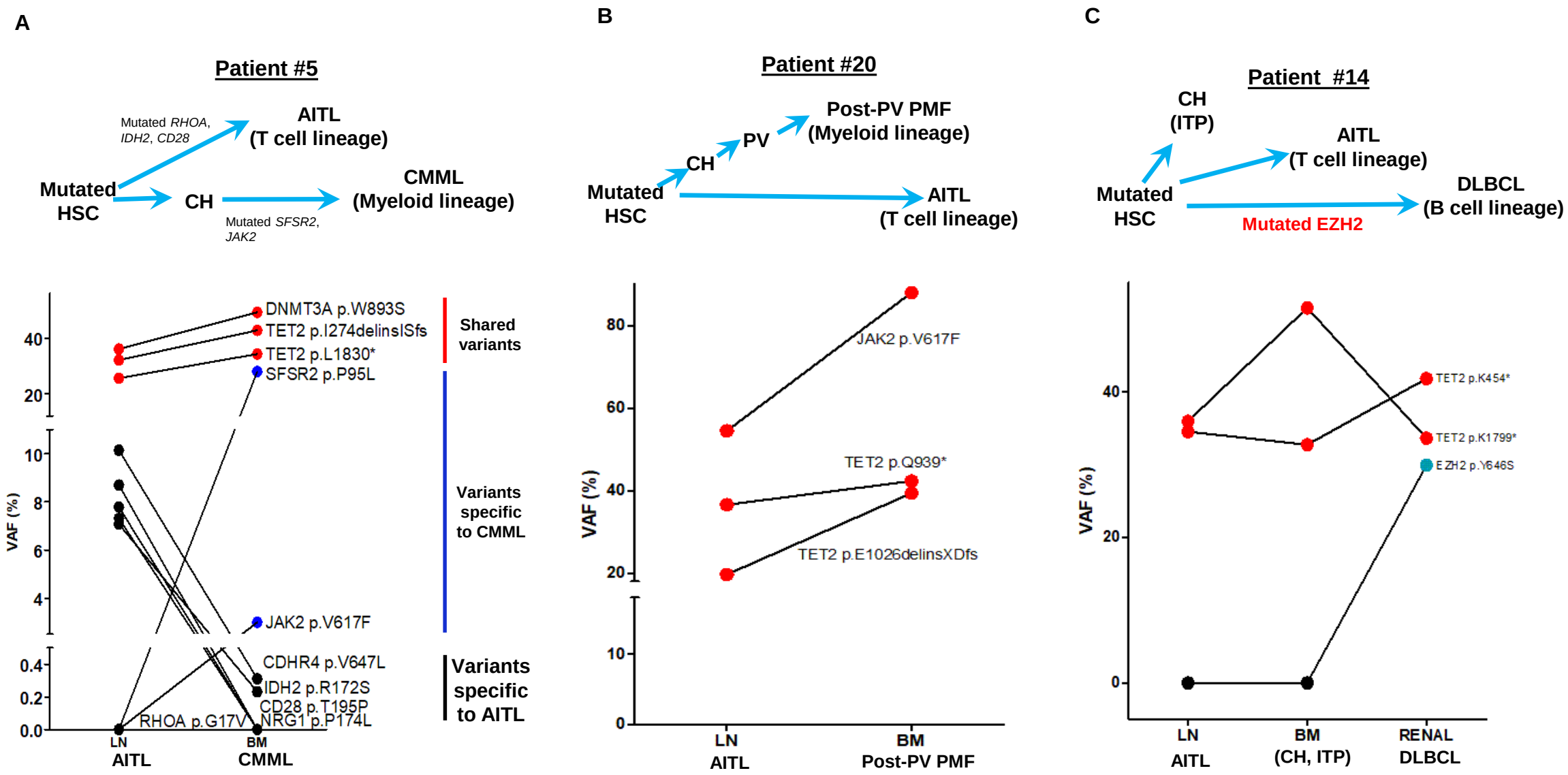
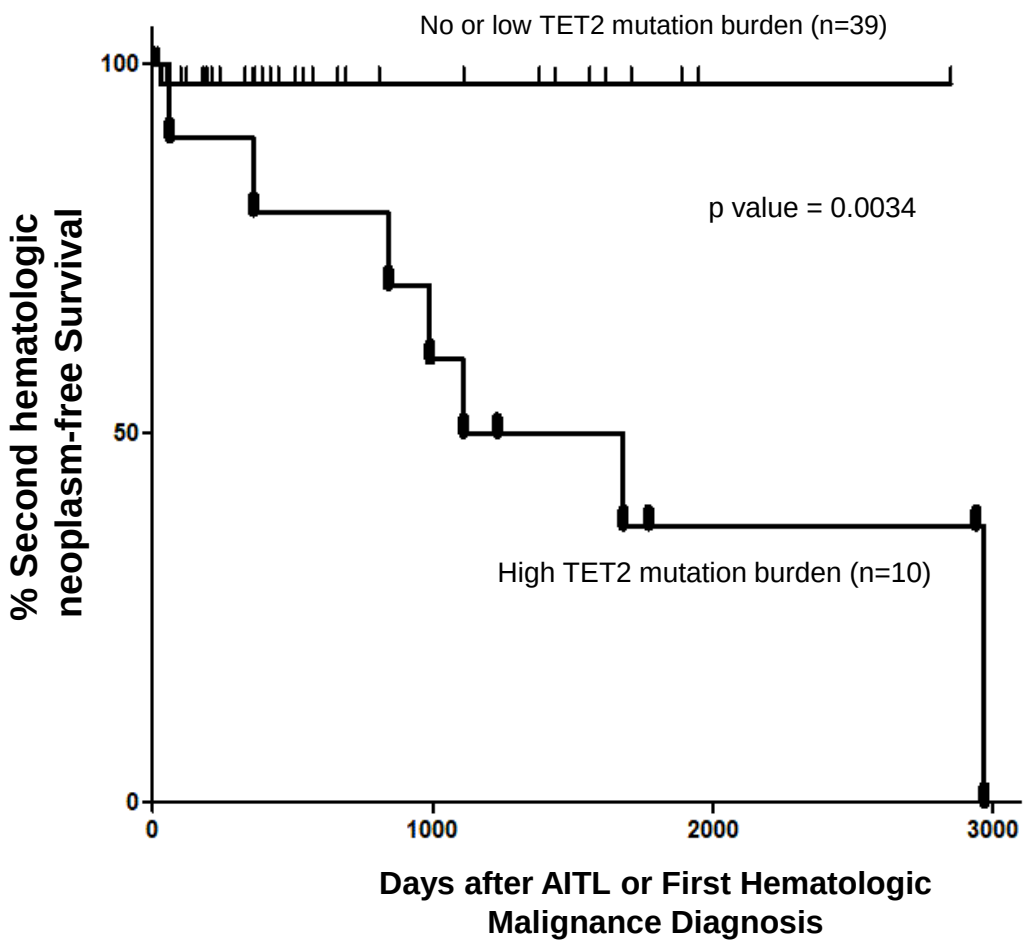

A
Patient #5
Mutated RHOA, IDH2, CD28
AITL (T cell lineage)
Mutated HSC
CH
CMML (Myeloid lineage)
Mutated SFSR2, JAK2
DNMT3A p.W893S
TET2 p.I274delinsISfs
TET2 p.L1830*
SFSR2 p.P95L
JAK2 p.V617F
CDHR4 p.V647L
IDH2 p.R172S
CD28 p.T195P
NRG1 p.P174L
RHOA p.G17V
Shared variants
Variants specific to CMML
Variants specific to AITL
VAF (%)
LN
AITL
BM
CMML
B
Patient #20
Post-PV PMF (Myeloid lineage)
PV
CH
Mutated HSC
AITL (T cell lineage)
JAK2 p.V617F
TET2 p.Q939*
TET2 p.E1026delinsXDfs
VAF (%)
LN
AITL
BM
Post-PV PMF
C
Patient #14
CH (ITP)
AITL (T cell lineage)
Mutated HSC
DLBCL (B cell lineage)
Mutated EZH2
TET2 p.K454*
TET2 p.K1799*
EZH2 p.Y646S
VAF (%)
LN
AITL
BM
(CH, ITP)
RENAL
DLBCL

Fig. 4


No or low TET2 mutation burden (n=39)
p value = 0.0034
High TET2 mutation burden (n=10)
% Second hematologic neoplasm-free Survival
100
50
0
0
1000
2000
3000
Days after AITL or First Hematologic Malignance Diagnosis