



Lymphoma

Targeted deep sequencing reveals clonal and subclonal mutational signatures in Adult T-cell leukemia/lymphoma and defines an unfavorable indolent subtype

Ambroise Marçais^{1,2} · Ludovic Lhermitte² · Maria Artesi³ · Cécile Laurent⁴ · Keith Durkin³ · Vincent Hahaut^{3,5} · Nicolas Rosewick^{3,5} · Felipe Suarez^{1,6} · David Sibon^{1,2} · Morgane Cheminant^{1,6} · Veronique Avettand-Fenoel⁷ · Julie Bruneau⁸ · Michel Georges³ · Claudine Pique⁹ · Anne Van den Broeke^{3,5} · Vahid Asnafi² · Olivier Hermine^{1,6}

Received: 13 January 2020 / Revised: 27 May 2020 / Accepted: 29 May 2020
© The Author(s), under exclusive licence to Springer Nature Limited 2020

Abstract

Adult T-cell leukemia/lymphoma (ATL) carries a poor prognosis even in indolent subtypes. We performed targeted deep sequencing combined with mapping of HTLV-1 proviral integration sites of 61 ATL patients of African and Caribbean origin. This revealed mutations mainly affecting TCR/NF-κB (74%), T-cell trafficking (46%), immune escape (29%), and cell cycle (26%) related pathways, consistent with the genomic landscape previously reported in a large Japanese cohort. To examine the evolution of mutational signatures upon disease progression while tracking the viral integration architecture of the malignant clone, we carried out a longitudinal study of patients who either relapsed or progressed from an indolent to an aggressive subtype. Serial analysis of relapsing patients identified several patterns of clonal evolution. In progressing patients, the longitudinal study revealed NF-κB/NFAT mutations at progression that were present at a subclonal level at diagnosis (allelic frequency < 5%). Moreover, the presence in indolent subtypes of mutations affecting the TCR/NF-κB pathway, whether clonal or subclonal, was associated with significantly shorter time to progression and overall survival. Our observations reveal the clonal dynamics of ATL mutational signatures at relapse and during progression. Our study defines a new subgroup of indolent ATLs characterized by a mutational signature at high risk of transformation.

These authors contributed equally: Anne Van den Broeke, Vahid Asnafi, Olivier Hermine

Supplementary information The online version of this article (<https://doi.org/10.1038/s41375-020-0900-3>) contains supplementary material, which is available to authorized users.

✉ Ambroise Marçais
ambroise.marcais@aphp.fr

- ¹ Service d'Hématologie Adultes, Hôpital Universitaire Necker-Enfants Malades, Assistance Publique Hôpitaux de Paris, Université Paris Descartes, 149-161 rue de Sèvres, 75015 Paris, France
- ² Laboratoire d'onco-hématologie, Hôpital Universitaire Necker-Enfants Malades, Assistance Publique Hôpitaux de Paris, Institut Necker-Enfants Malades, INSERM U1151, Université Paris Descartes, 149-161 rue de Sèvres, 75015 Paris, France
- ³ Unit of Animal Genomics, GIGA, Université de Liège (ULiège), Avenue de l'Hôpital 11, B34, 4000 Liège, Belgium
- ⁴ Institut Carnot CALYM, Pierre-Bénite, France

Introduction

Adult T-cell leukemia/lymphoma (ATL) is a rare mature T-cell malignancy associated with human T-cell leukemia-1 (HTLV-1) infection. The Southwestern region of Japan, Sub-Saharan Africa, South America, and the Caribbean area are the main endemic regions [1]. HTLV-1 infection mainly

- ⁵ Laboratory of Experimental Hematology, Institut Jules Bordet, Université Libre de Bruxelles (ULB), Boulevard de Waterloo 121, 1000 Brussels, Belgium
- ⁶ Institut Imagine, INSERM U1163, CNRS ERL8654, 24 Boulevard du Montparnasse, 75015 Paris, France
- ⁷ Laboratoire de Microbiologie clinique, Hôpital universitaire Necker-Enfants Malades, AP-HP, 149-161 rue de Sèvres, 75015 Paris, France
- ⁸ Département de Pathologie, Hôpital universitaire Necker-Enfants Malades, APHP, Université Paris Descartes, Sorbonne Paris Cité, 149-161 rue de Sèvres, Paris 75015, France
- ⁹ INSERM U1016, CNRS UMR 8104, Université Paris Descartes, Sorbonne Paris Cité, Institut Cochin, Paris, France

occurs during childhood and only ~5% of infected individuals will develop ATL following a long latency period [2]. Four subtypes are described: two indolent (smoldering and chronic) and two aggressive (acute and lymphoma) forms [3]. The prognosis is poor, especially in aggressive subtypes with a median overall survival (OS) of <1 year despite intensive treatments including allogeneic stem cell transplantation [4, 5]. Indolent forms also carry a relatively adverse outcome with a median OS of 36 and 56 months for chronic and smoldering subtypes, respectively, mainly related to their evolution toward an aggressive subtype [4, 6]. As a consequence, the identification of indolent patients at high risk of progression toward an advanced disease remains a major challenge in the management of ATL patients.

The characterization of the landscape of genetic abnormalities in a Japanese ATL cohort revealed the presence of recurrent somatic mutations and copy number variants, especially in three main pathways: TCR/NF- κ B signaling, T-cell trafficking, and immune escape mechanisms [7]. However, little is known as regards the dynamics of genomic alterations in patients who relapse or progress from an indolent subtype to an aggressive form of the disease.

The very poor outcome of aggressive but also indolent subtypes stresses the urgent need for new therapeutic approaches and tools to better understanding mechanisms governing relapse as well as identifying patients with indolent subtypes who are at risk of transformation toward an aggressive form. To this end, much effort has been devoted to identifying such patients. Previous studies have identified a subgroup of the indolent subtype called “unfavorable chronic” using routine biological variables associated with decreased progression-free survival (PFS) and decreased OS [8]. From a molecular point of view, two studies have highlighted the negative impact of *CDKN2A* deletions and *TP53* mutations [9–11]. More recently, the correlation of genomic alterations described by the Japanese consortium with corresponding clinical data reinforced the hypothesis that the profile of oncogenic changes, such as *IRF4* mutations, *CDKN2A* deletion, and *PDL-1* amplification, is associated with a shorter PFS and OS in indolent forms [12].

Here, we describe the genomic landscape of 61 ATL cases of Afro-Caribbean origin based on targeted deep sequencing (TDS) combined with genome-wide mapping of proviral integration sites. We explore the evolution of mutational signatures in two distinct sets of longitudinal samples. This revealed viral and mutational profile patterns specific of relapse. Furthermore, our study identifies clonal and subclonal mutational profiles that define an indolent ATL subgroup at high risk of transformation. Our study provides novel insights into the genetic basis of ATL, the molecular dynamics of relapse and progression, and

proposes a new prognosis classification scheme for indolent subtypes according to a specific unfavorable molecular signature.

Patients and methods

Patients

The study had a retrospective observational design. Sixty-one patients with an ATL at diagnosis were analyzed. Diagnosis of ATL was based on clinical and biological features including immunophenotyping, the presence of HTLV-1-specific antibodies, and the detection of HTLV-1 proviral DNA. Serial samples were available for 15 patients. These longitudinal cohort consisted of one sample collected at diagnosis and either a second or several samples corresponding to progression to an indolent to an aggressive (disease progression, i.e., from chronic to acute), or to a relapse after a period of complete clinical remission according to previously published response criteria [8]. This study was approved by the ethical committee (CPP Ile de France II, CNIL: number 1692254, registration number 000001072) and all surviving patients gave written informed consent.

Nucleic acid extraction

DNA and RNA extractions were performed on frozen peripheral blood mononuclear cells (PBMC) or frozen tumor tissue using the “all PREP DNA/RNA mini kit” (Qiagen) or “DNA Blood and Tissue kit” (Qiagen) according to the manufacturer’s instructions.

Targeted deep sequencing (TDS)

TDS was applied to genomic DNA extracted from 61 ATL samples at diagnosis and 28 additional serial samples. A first approach consisted of a custom capture-based next generation sequencing (NGS) panel representing 80 genes relevant to hematological malignancies. A TruSeq Amplicon-based NGS panel of 36 genes was then designed and further used to confirm the genes most frequently affected in our cohort (Supplementary Tables 1, 2a, b). To this confirmatory panel, we added 19 genes that were recently shown altered with a frequency of $\geq 3\%$ of the cases in the Japanese ATL cohort [7]. In total, 99 genes were sequenced. Libraries were prepared according to Illumina instructions and sequenced on a Miseq instrument (ILLUMINA, 500 \times with mean coverage 95% for the capture panel and 2000 \times with mean coverage 90% for the TruSeq panel). The TruSeq panel allowed the detection of mutations at a subclonal level defined by variant allele frequency

(VAF) <5%. Clonal and subclonal mutations were thus defined as mutations with VAF >5% and <5%, respectively.

Sequences were aligned to the human reference genome hg19 using BWA. Downstream processing was carried out with the Genome Analysis Toolkit (GATK), SAMtools, and Picard, following documented best practices. Variant calls were made with the GATK Unified Genotyper. The annotation process was based on the latest release of the Ensemble number version GRCch37 database. Variants were annotated and analyzed using the Polyweb software interface designed by the Bioinformatics platform of the University Paris Descartes. Prioritization of the variants was performed using the same interface. Relevant somatic mutations were called and filtered by excluding synonymous SNVs and known variants listed in SNP databases (dbSNP131, 1000 Genomes Project, 5000 Exomes project). All variants showing <5 alternative reads were excluded. All other variants were manually checked and kept if they were found either in the COSMIC database or estimated to be relevant according to the Polyphen score, which was used to estimate the impact of the mutation on protein structure or function [13]. Sequencing data are deposited in the European Genome-Phenome Archive (www.ebi.ac.uk/ega), under accession number EGAS00001004483.

Evaluation of duplication and large deletion events with TDS

In order to evaluate duplication and large deletion events, we assessed the copy numbers of each targeted region in a given individual. For each individual, the relative read counts for each targeted region was determined as the ratio of the read count for that region divided by the total absolute read counts of all targeted regions of the design. The ratio of the relative read count of a region in a given individual over the average relative read counts in other individuals of the run resulted in the estimated copy number for that region in that individual.

SNP array karyotyping

Genome-wide copy number analysis was performed on 32 samples (27 samples at diagnosis and 5 samples at progression/relapse) using the Affymetrix CYTOSCAN HD and the Illumina HUMAN-OMNIEXPRESS according to the manufacturers' instructions. Details are available in Supplementary methods.

Next generation sequencing (NGS) genome-wide mapping of HTLV-1 integration sites

To map HTLV-1 integration sites and measure the abundance of the corresponding clones, DNA extracted from

PBMCs of ATL patients was sheared and processed as described [14, 15]. Briefly, sheared DNA fragments were selectively end-repaired and A-tailed by Taq polymerase mediated linear extension from LTR-specific primers located in both the 5' and 3' LTR. Biotinylated dTTPs were incorporated during this extension step. Adapters containing UMIs were ligated to the A overhangs and streptavidin bead-mediated selection was carried out, followed by PCR amplification using off-the-shelf Nextera indexes to enable multiplexing. Libraries were sequenced on a MiSeq instrument (kit V2, 2 × 150 bp). Removal of PCR duplicates combining UMI, shear site, and genomic integration site information ensured accurate estimation of clone abundance. HTLV-1 genomic integration information are deposited in the Retroviral Integration Database [16].

Quantification of HTLV-1 proviral load and RNA sequencing

HTLV-1 proviral loads and ATL transcriptomes were examined as previously reported [14, 15, 17]. Details are available in Supplementary methods.

Statistical analysis

The end-points studied were OS and PFS. Progression was defined as disease progression to an aggressive subtype for patients with an indolent subtype who achieved partial response (PR) or remained in stable disease (SD). OS and PFS curves were estimated using Kaplan–Meier product-limit estimator. Data were censored at time of allogeneic stem cell transplantation. All analyses were carried out using the R software.

Fish plot

For clonal evolution, variants were filtered based on sequencing depth (>20 reads of total depth and >10 reads for alternative reads) and further filtered by excluding synonymous SNV, UTR, and intronic, and if listed as SNPs in public databases (Supplementary Table 3). Fish plots and clonality charts were generated in R using SciClone, ClonEvol, and Fish plot packages [18, 19].

Results

Genomic landscape of ATL

Sixty-one cases of ATL at diagnosis were analyzed by TDS. The median age at diagnosis was 51 years (range 17–73 years). Among them, 15 presented an indolent (smoldering

or chronic) and 46 had an aggressive (39 acute and 8 lymphoma) subtypes (Supplementary Table 4).

A total of 248 clonal somatic mutations (defined as mutations showing a VAF > 5%, Methods) in 48 genes were identified including missense ($n = 137$), nonsense ($n = 61$), splice ($n = 17$), frameshift ($n = 26$), non-frameshift insertion/deletion ($n = 5$), or phase ($n = 2$) mutations. Mutations were detected in 57 patients (92%) with a median of three mutations per patient (range 0–10). The most commonly mutated genes were *CCR4* ($n = 23$, 38%), *PLCG1* ($n = 18$, 29%), *PRKCB* ($n = 13$, 21%), *CARD11* ($n = 13$, 21%), *TBLIXR1* ($n = 11$, 18%), *TP53* ($n = 11$, 18%), *FAS* ($n = 10$, 16%), *GATA3* ($n = 10$, 16%), *NOTCH1* ($n = 9$, 15%), *CSNK2B* ($n = 9$, 15%), *STAT3* ($n = 8$, 13%), *RHOA* ($n = 7$, 12%), and *VAV1* ($n = 7$, 12%) (Figs. 1a, 2, Supplementary Table 5). Forty-five (74%) patient samples harbored gene deletions, which were confirmed by SNP array analysis in 27 cases with available material (Supplementary Fig. 1). The most commonly deleted genes were *CDKN2A* ($n = 22$, 36%), *ATXN1* ($n = 13$, 21%), *HLA-B* ($n = 11$, 18%), and *TP53* ($n = 10$, 17%). Nine patients (15%) showed two mutations in the same gene: *CCR4* ($n = 3$), *PLCG1* (3), *CARD11* ($n = 1$), *HNRNPA2B1* ($n = 1$), *TBLIXR1* ($n = 1$), and *HLA-B* ($n = 1$). Ten patients had mutations in one allele and a deletion in the other allele. Genes affected by such bi-allelic inactivation were *TP53* ($n = 5$), *HLA-B* (1), *CD58* ($n = 1$), *FAS* ($n = 1$), *RBI* ($n = 1$), and *GPR183* ($n = 1$).

Genomic alterations were clustered in four main pathways. Forty-five (74%) patients carried mutations affecting the TCR/NF- κ B pathway (*CD28*, *PLCG1*, *CARD11*, *PRKCB*, *CBLB*, *IRF4*, *CSNK1A1*, *FYN*, *RHOA*, and *VAV1*), among which 31 (51%) affecting the restricted NF- κ B/NFAT pathway (*PLCG1*, *CARD11*, *PRKCB*, and *IRF4*) [7, 20]. Twenty-eight (46%) ATLS harbored mutations affecting the T-cell trafficking pathway (*CCR4*, *CCR7*, and *GPI83*). Eighteen (29%) ATLS showed mutations affecting genes involved in immune escape (*FAS*, *HLA-B*, *B2M*, and *CD58*). Sixteen (26%) patients showed mutations in genes involved in cell cycle regulation (*TP53*, *POT1*, and *RBI*). In addition, 31 (51%) patients showed mutations in genes involved in regulation of transcription and/or epigenetic regulation and 13 (21%) in the JAK/STAT signaling pathway. The mutational landscape identified in our study is consistent with the ATL mutational profiles described in the previously published Japanese study (Fig. 1b, c). In addition, we did not observe differences between African ($n = 26$) and Caribbean ($n = 22$) descend in terms of genomic landscape. However, we observed a high rate of mutations in the cell cycle pathway in the “Guyana” subgroup (6/7) mainly due to a high number of mutations in the *TP53* locus. These results should be interpreted with caution due to the small number of patients in the Guyana subgroup (Supplementary Table S4 and Supplementary Fig. 2).

Indolent and aggressive subtypes differ in their mutational profiles

The average number of mutations per sample at diagnosis was significantly higher in aggressive subtypes than in indolent forms (5 versus 2, $p = 0.000084$, Fig. 3a, b). This mainly resulted from a higher number of genomic alterations in genes involved in the TCR/NF- κ B pathway. More specifically, 40/46 (87%) patients with an aggressive subtype showed at least one activating mutation in this pathway compared with 5/15 (33%) in indolent ATLS ($p = 0.00015$). This higher rate of mutations may be biased by the over-representation of genes belonging to the TCR/NF- κ B pathway in our TDS panel ($n = 10$). However, the difference was also significant when comparing the proportion of activating mutations in the restricted NF- κ B/NFAT pathway: 28/46 (61%) versus 3/15 (27%) ($p = 0.00776$), while in other highlighted pathways with similar gene numbers in the panel—immune escape ($n = 4$), cell cycle regulation ($n = 3$), T-cell trafficking ($n = 3$)—it was not (Fig. 3c, $p = ns$).

Genomic landscape and proviral integration architecture in ATLS at relapse

In order to explore the potential contribution of the temporal acquisition of driver mutations at relapse, we analyzed seven patients (ATLS 1, 11, 44, 60, 100, 105, and 110) who achieved complete remission but relapsed after a median time of 32.9 months (range 6.2–75.7 months) (Fig. 4 and Supplementary Table 6). With the exception of two patients (ATLS 11 and 110) who showed a viral clone switch between diagnosis and relapse (a viral clone being defined as a sub-population of cells carrying the HTLV-1 provirus at the same position in the genome), all other cases ($n = 5$) showed the same dominant viral malignant clone at diagnosis and relapse, with a similar relative abundance of the corresponding clones. Viral RNA-seq data were also examined for ATLS 1, 44, 60, and 100 at both time points. This confirmed that HTLV-1 transcriptional profiles did not significantly differ between diagnosis and relapse, with consistent production of the viral antisense transcripts HBZ and absence of robust HTLV-1-dependent transcription from the positive strand (Supplementary Fig. 3) [14, 15].

Acquisition of de novo mutations in ATLS at relapse

In relapsing patients who showed an identical viral malignant clone at diagnosis and relapse ($n = 5$), the median number of mutations at diagnosis was 4 (range 3–7). As expected, we identified additional mutations at relapse. Among these de novo mutations, we mainly found genes affecting the TCR/NF- κ B pathway and

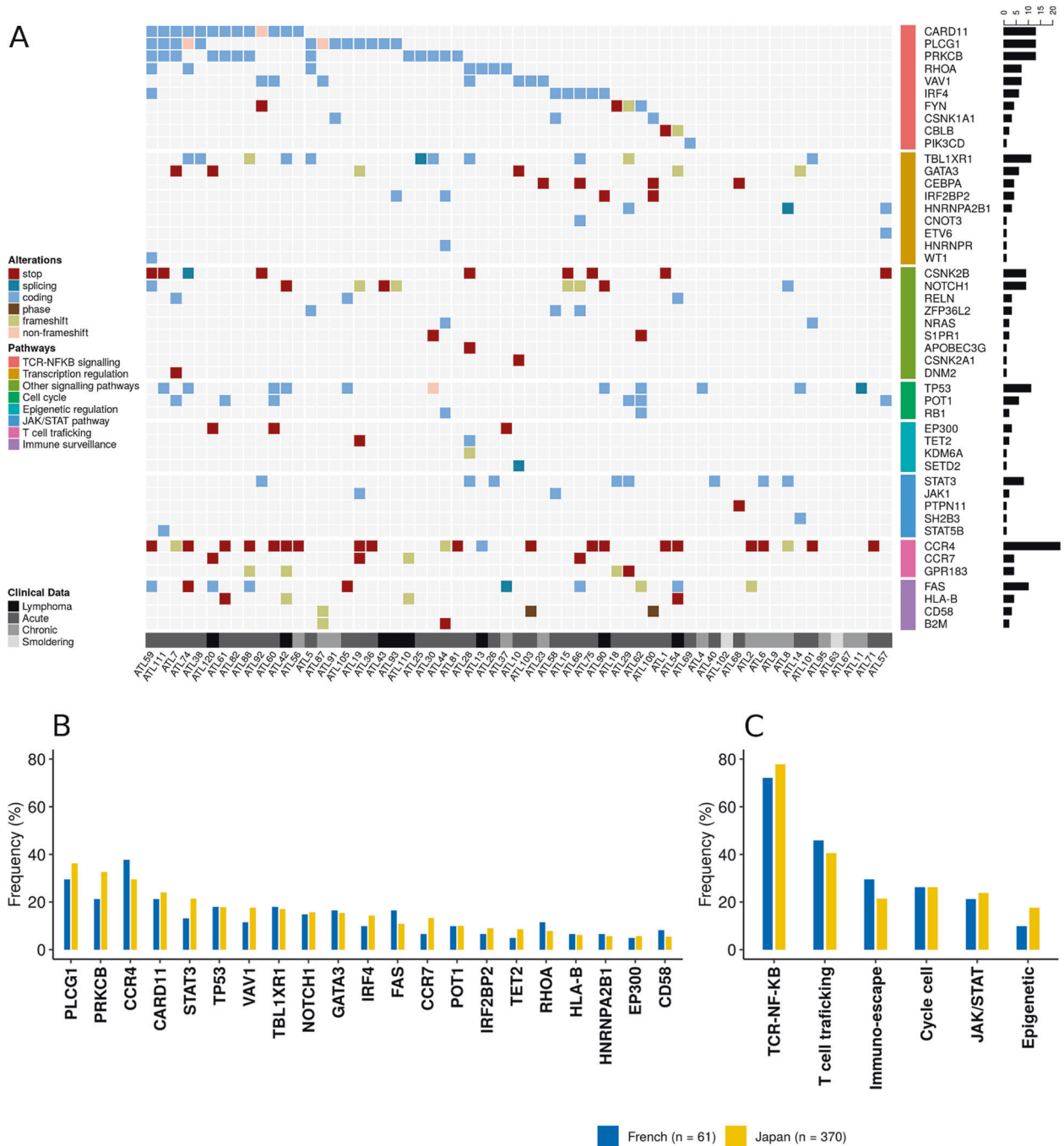


Fig. 1 Mutational landscape of ATL. **a** Distribution of clonal somatic mutations in 61 ATL cases classified according to the affected signaling pathway. Mutational profiles obtained by targeted sequencing of 99 genes in 61 ATLs including 8 lymphoma (black), 38 acute (dark gray), 13 chronic (gray), and 2 smoldering (light gray) subtypes. Four cases did not show clonal mutations. Clonal mutations are defined as mutations with a VAF >5% and are ordered according to the signaling pathway they affect. Mutation hierarchization was performed

according to previously published studies [7, 20]. The mutation *PLCG1* R48W was not considered as a TCR/NF-κB pathway damaging mutation [20]. **b** Frequency of main mutated genes in our ATL cohort and patients from the Japanese ATL analyzed by targeted exome sequencing (370 cases) [7]. **c** Frequency of main pathways affected by mutations in our ATL cohort and patients from the Japanese ATL cohort analyzed by targeted exome sequencing.

immune escape mechanisms (Fig. 5a, b and Supplementary Fig. 3a). More specifically, three cases acquired activating mutations in the NF-κB/NFAT pathway at relapse: ATLs 44, 60, and 105 showed mutations in *IRF4*,

PRKCB, and *PLCG1* loci, respectively. Three cases acquired loss of function mutations in *HLA-B*, a gene affecting immune-surveillance (ATLs 1, 44, and 100). In two patients (ATLs 44 and 60), some of the mutations

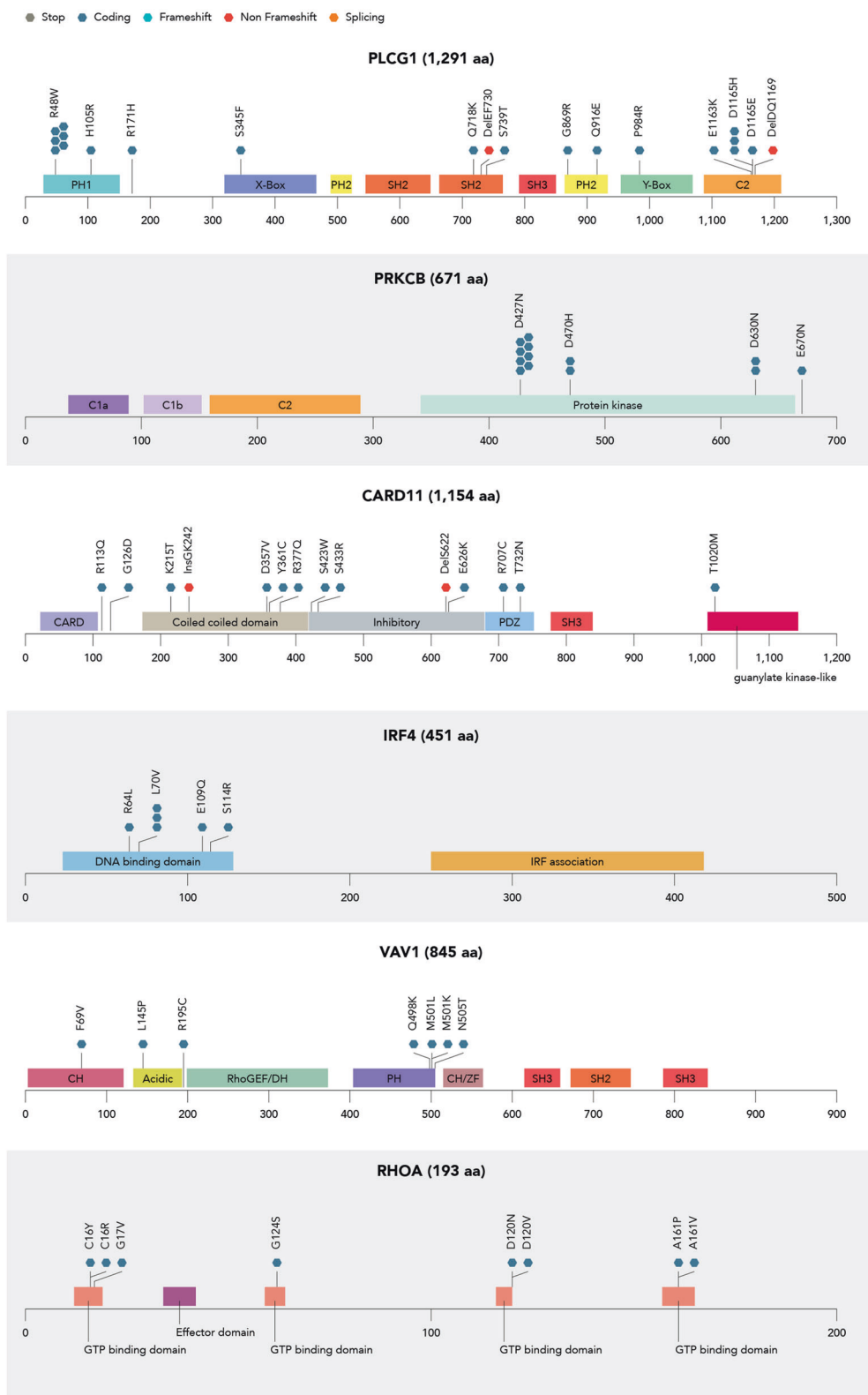


Fig. 2 Mapping of variants identified in genes belonging to the NF-kB/NFAT pathway and coding for GTPase proteins. Distribution of clonal (VAF > 5%) mutations found in 41/61 diagnostic samples in main mutated genes belonging to the TCR/NF-kB pathway: *PLCG1*,

PRKCB, *CARD11* and *IRF4* (NF-kB/NFAT pathway) and *VAV1* and *RHOA* (GTPase proteins). Nonsense mutations are marked in grey, missense in blue grey, frameshift in turquoise, non-frameshift in red and splicing in orange.

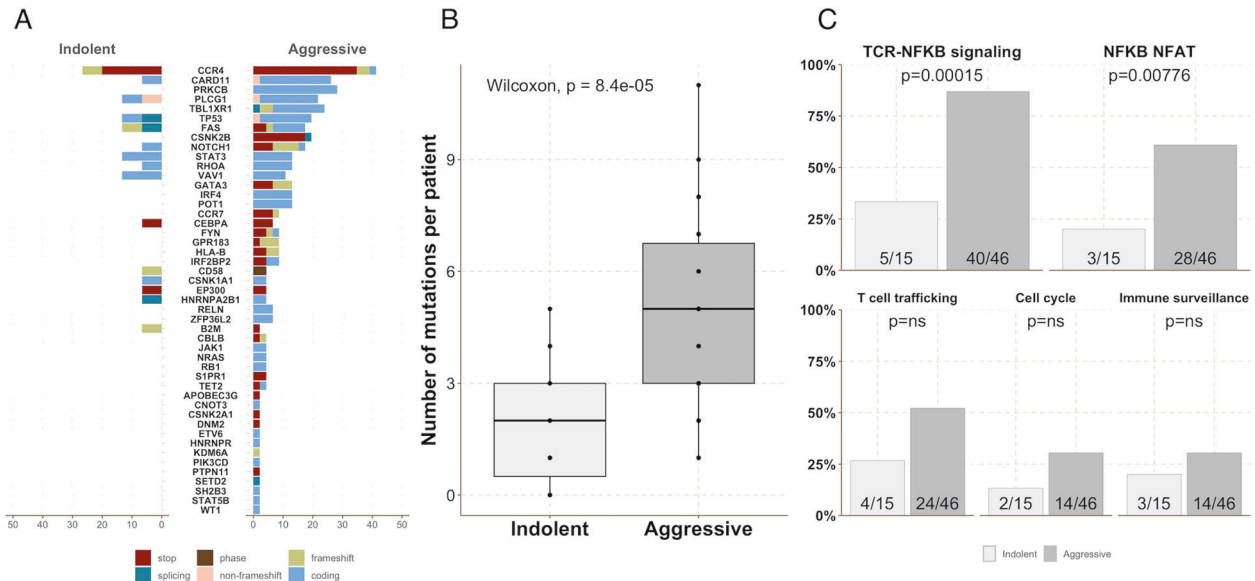


Fig. 3 Aggressive ATL subtypes carry a higher number of mutations. **a** Frequency of each mutation among indolent and aggressive subtypes. **b** Number of mutations observed per ATL sample in indolent and aggressive subtypes. **c** Number of ATL cases with somatic

mutations affecting genes belonging to the main mutated pathways in indolent (light gray) and aggressive (dark gray) subtypes, respectively.

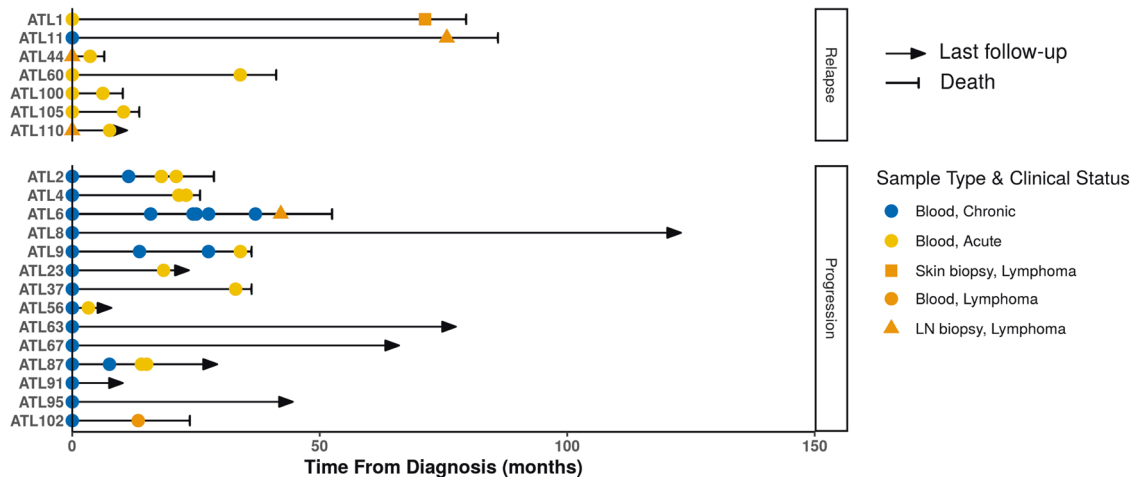


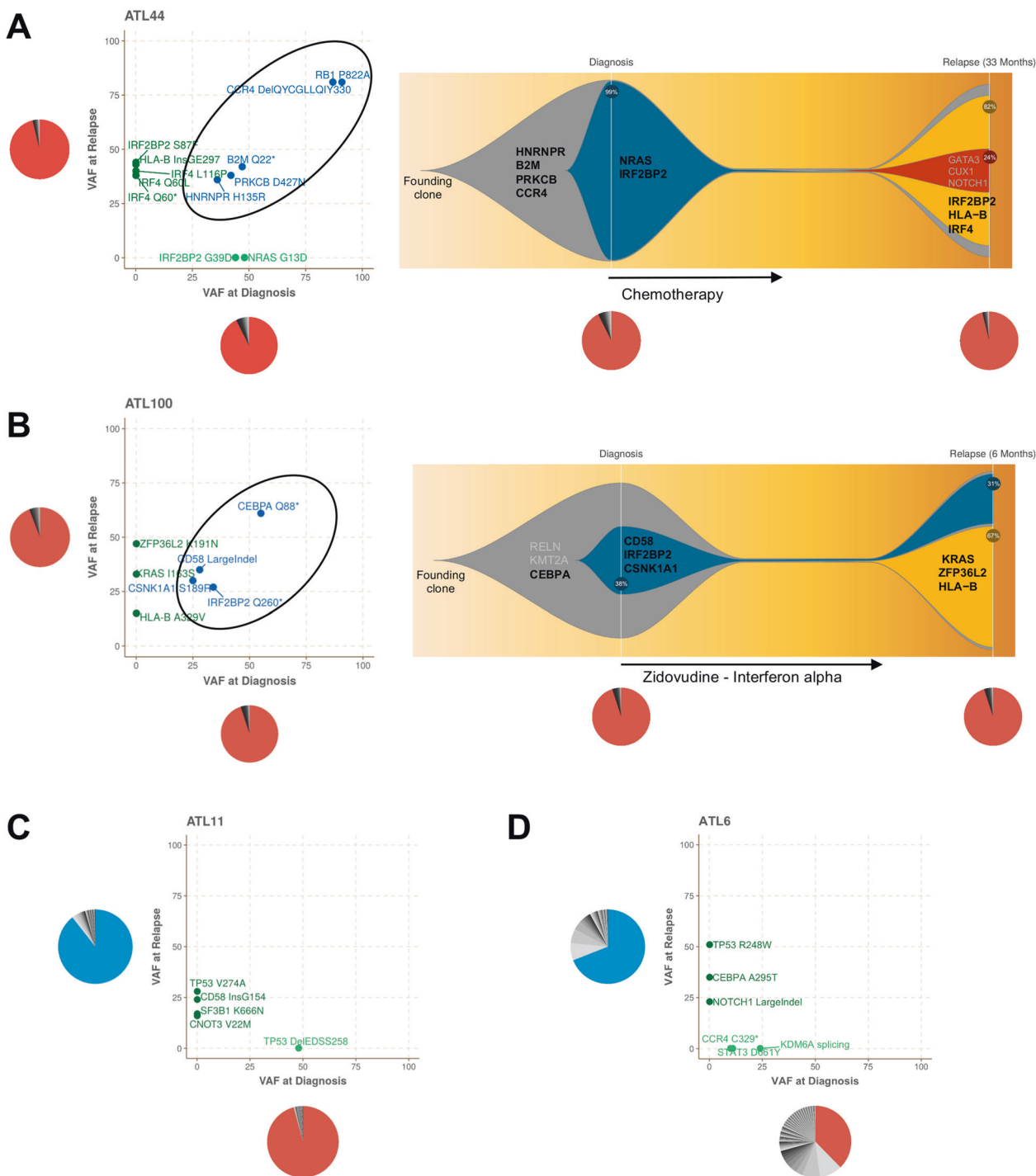
Fig. 4 Scaled schematic timeline of ATL cases examined in longitudinal studies. The line diagrams summarize sample collection time points for each longitudinally studied ATL patient. The top panel shows patients ($n = 7$) who achieved CR after first-line therapy and relapsed. The bottom panel represents patients with an indolent

subtype ($n = 14$) who were managed with various therapeutic approaches and who obtained PR or remained in SD. Among them, nine patients progressed to an aggressive subtype (acute or lymphoma) and five remained in PR or SD until the last follow-up.

identified at diagnosis were no longer identified at relapse. Interestingly, the acquisition of these de novo mutations emerged in a new distinct subclone within the founding clone. This new subclone, not detectable at diagnosis outcompeted the accompanying subclones after induction therapy. More specifically, in ATL 44 and ATL 60, the original major subclone had been eradicated (Fig. 5a and Supplementary Fig. 4a), while in ATL 100 and ATL 105 it persisted at similar frequency (Fig. 5b and Supplementary Fig. 4a).

Viral clone switch is not a rare event in the course of ATL

In addition to two above-mentioned cases that were characterized by a viral clone switch at relapse (ATLs 11 and 110), a third ATL patient with an indolent subtype showed a viral clone switch at progression (ATL 6). This patient developed a lymphoma subtype while being treated for an indolent form of the disease. For all three cases, as expected, the mutational pattern was different



between diagnosis and relapse, in line with the emergence of a novel clone. The new malignant clone identified at relapse consistently carried an inactivating mutation in the *TP53* locus and additional mutations such as alterations affecting immune escape mechanisms (ATLs 11 and 110), the Notch 1 pathway (ATL 6), and the TCR/NF- κ B pathway (ATL 110) (Fig. 5c, d and Supplementary Fig. 4b).

Genomic architecture of longitudinal samples reveals subclonal mutations in indolent subtypes

We then analyzed indolent subtypes, and more specifically those ($n = 9$) who progressed to an aggressive disease (Fig. 4, Supplementary Fig. 5, Supplementary Table 6). Three of these ATLs were also examined by RNA-seq and CGH array (ATLs 2, 4, and 37).

◀ **Fig. 5 Evolution of somatic mutations and viral integration architecture between diagnosis and relapse.** Examples of mutational evolution within the dominant viral malignant clone for four representative ATL cases studied at diagnosis and relapse. Pie charts represent the HTLV-1 proviral integration architecture of the corresponding ATLs. Scatter graphs show mutations identified at diagnosis (x-axis) and relapse (y-axis) with their corresponding VAF. Mutations depicted in light green are specific of diagnosis. Mutations in dark green are associated with relapse. Mutations in blue are shared between diagnosis and relapse. Mutational evolution of four additional ATLs examined are shown in Supplementary Fig. 3. **a, b** Diagnosis/relapse sample pairs from two patients who relapsed with an identical viral dominant clone. ATL 44 (**a**) and ATL 100 (**b**) show the same single dominant proviral integration site (IS) with a relative abundance of >90% at diagnosis and relapse, corresponding to the presumed malignant clone. The fish plots illustrate models of mutational evolution during the course of the disease. Damaging and non-damaging mutations are represented in black bold and gray, respectively. The remaining genes and the clones to which they are assigned are available in the supplement. In both cases, the founding clone (in gray) observed in the primary tumor at diagnosis carries specific damaging mutations. The major subclone represented in blue is either eradicated (mutations in *NRAS* and *IRF2BP2*, ATL 44) or persists at a similar frequency (mutations in *CD58*, *IRF2BP2*, and *CSNK1A1*, ATL 100) at relapse. In both cases, a novel subclone (yellow) derived from the founding clone emerges and expands, becoming the dominant subclone at relapse. That clone acquired additional mutations (*HLA-B*, *IRF4*, and *IRF2BP2* for ATL 44, *KRAS*, *ZFP36L2*, and *HLA-B* for ATL 100). Data from three additional patients available in Supplementary Fig. 3a. **c, d** Diagnosis/relapse sample pairs from patients who relapsed (ATL 11) or progressed (ATL 6) with a novel viral dominant tumor clone (clone switch). Data from an additional patient available in Supplementary Fig. 3b.

Patients were managed with different therapeutic approaches including ZIDOVUDINE-INTERFERON therapy and watch and wait strategy (Supplementary Table 6). Of the 14/15 patients who achieved PR or showed SD, nine progressed to an aggressive subtype with a median progression time of 25 months (range 3.6–42.1 months), while five patients remained in PR or SD until the last follow-up (median follow-up 32 months).

With the exception of ATL 6 who showed a viral clone switch at progression, all cases retained the same viral malignant clone with a similar relative abundance between sequential paired samples. The viral transcriptional profiles did not show significant differences between time points. Like for ATL samples at relapse, we did not observe robust transcription from the positive strand while antisense transcripts were consistently produced, in agreement with previous observations [7, 14].

Genomic profiling of ATL samples at progression revealed the acquisition of mutations in seven of the nine patients examined. De novo mutations were identified in two main pathways: NF-kB/NFAT ($n = 4$, ATLs 2, 4, 37, and 102) and cell cycle ($n = 3$, ATLs 6, 9, and 23). Interestingly, the novel clonal mutations affecting the NF-kB/NFAT pathway at progression were identified at a subclonal level in the matched indolent ATL samples (VAF < 5%,

range 0.7–5%) (Fig. 6a, b and Supplementary Figs. 6, 7). We identified two mutations in the *CARD11* locus (ATLs 2 and 102, hot spot D345E), one in the *PLCG1* locus (ATL 4, hot spot D1165H) and one in the *PRKCB* locus (ATL 37, hot spot D427N). Analysis of dynamic evolution revealed a strong enrichment of NF-kB/NFAT pathway-associated mutations at progression, uncovering the clonal outgrowth of a preexisting rare subclone present as a sub-population embedded in the main tumor clone at the indolent stage (Fig. 6a, b).

Finally, of the two patients who progressed without TCR/NF-kB mutations, one progressed to a lymphoma subtype with a viral clone switch (ATL 6, Fig. 5c) as previously mentioned, while the other (ATL 9) retained the same dominant viral clone with additional mutations in *TP53* and *FAS* loci at progression (Supplementary Fig. 6). These mutations were not identified at earlier time points (diagnosis and 6 months before progression), even at a subclonal level, suggesting their de novo acquisition.

Of note, we also verified that the increase in VAFs observed upon ATL progression were effectively revealing clonal or subclonal outgrowth rather than an overall increase in the proportion of HTLV-1 infected cells. To this end, we examined proviral loads of the corresponding samples and found no significant difference between matched sample pairs.

Subclonal and clonal mutations have predictive value in indolent subtypes

ATL patients with an indolent subtype and at least one clonal or subclonal mutation in the TCR/NF-kB pathway ($n = 8$) had a significantly shorter PFS compared with patients who did not show mutations ($n = 6$) (median 20 months versus not reached, 2 year PFS 33% versus 100%, $p = 0.0034$) and a shorter OS (median 28.5 months versus not reached; 3 year OS 40% versus 100%, $p = 0.0005$), suggesting these mutations are the dominant driver of progression. At progression, the disease became refractory to any treatment and, with the exception of one patient who received allogeneic stem cell transplantation, all patients died shortly after progression (Fig. 6c).

Discussion

This study describes the mutational landscape of ATL in a cohort of 61 patients mainly originating from Africa and the Caribbean area, broadening the genomic characterization of this heterogeneous disease to a distinct genetic background. Our observations are in agreement with those reported in a larger Japanese cohort of 426 ATLs [7]. We confirm the significant proportion of genomic alterations in genes

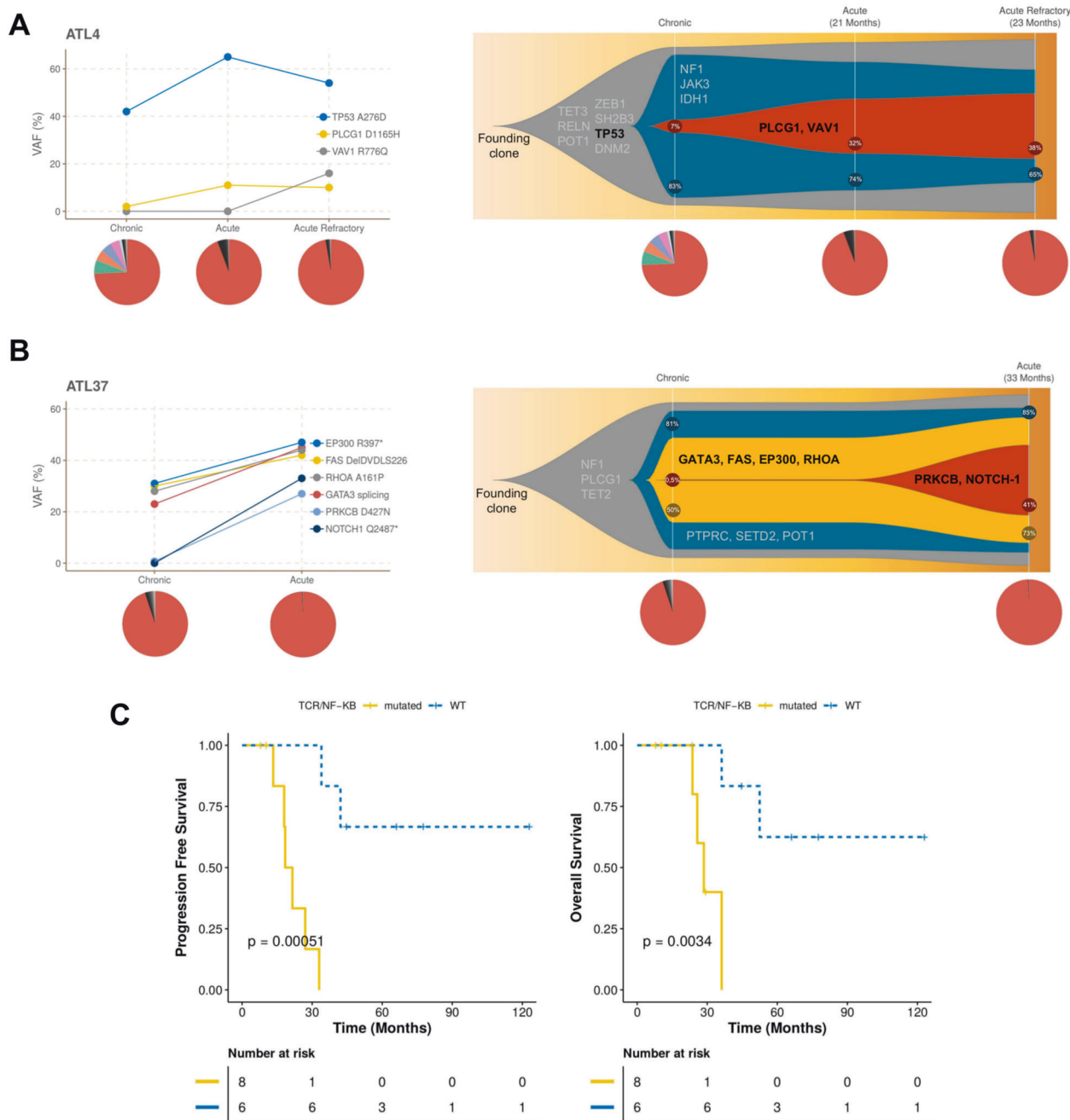


Fig. 6 Evolution of subclonal and clonal somatic mutations in indolent ATLs progressing toward an aggressive subtype. a, b Longitudinal samples from two ATL patients with an initial indolent subtype who progressed to an aggressive subtype (ATL 4 and ATL 37). Pie charts represent the proviral integration architecture. Both ATLs show a single dominant viral insertion site corresponding to the presumed malignant clone. Fish plots illustrate models of mutational

evolution. The founding clone is represented in gray. The subclone that carried the mutation in the TCR/NF-kB pathway (*PLCG1* locus for ATL 4, *PRKCB* locus for ATL 37) is represented in red. Data from three additional patients available in Supplementary Fig. 6. **c** Kaplan–Meier curve for progression-free survival (PFS) and overall survival (OS) according to the presence of at least one activating mutation of the TCR/NF-kB pathway at a clonal or subclonal level.

coding for proteins involved in the TCR/NF-kB signaling pathway, associated with T-cell trafficking and tumor immune-evasion, but also cell cycle integrity. These results support the conclusion that the patients' genetic background and environment do not have a significant impact on the mutational landscape of ATL. Although there is

undisputable evidence that the virus plays a critical role in the transformation process of HTLV-1-infected T-cells, we observed a significant overlap between the ATL mutational landscape and mutations reported in mature T- or B-cell malignancies of nonviral etiology [20–23]. We also found that compared with indolent forms, aggressive subtypes

carry a higher rate of clonal mutations mainly associated with a higher proportion of mutations in the TCR/NF- κ B pathway including the NF- κ B/NFAT pathway.

A recent study has reported the mutational landscape of a smaller cohort of 30 ATL patients from Caribbean descent [24]. The authors observed a higher rate of mutations in genes implicated in epigenetic regulation such as *EP300* (20%) and a lower rate in the TCR/NF- κ B pathway compared with the study published by the Japanese consortium. A possible explanation of this discrepancy is the composition of the targeted NGS panels. More specifically, the American and the Japanese panels only share 17 genes. When comparing our cohort with the cohort of Shah et al., even if the panels shared 45 genes, only 16 of the genes mutated in our panel were present in the American panel. In addition, the American panel did not include most of the main mutated genes identified in the Japanese cohort and ours, notably those belonging to the TCR/NF- κ B pathway (*PLCG1*, *PRKCB*, and *IRF4*) and to the T-cell trafficking (*CCR4*, which is mutated in 1/3 of the cases in our cohort).

Unlike previous genomic studies, we analyzed serial samples of ATL patients to explore the dynamics of mutational profiles observed at diagnosis and better understand disease evolution. The longitudinal study revealed the evolution of mutational signatures observed in the original HTLV-1-infected malignant T-cell clone. HTLV-1 clonality analysis indicated that the main insertion site identified at diagnosis was retained upon progression or relapse in the majority of the patients. This provided indirect evidence that the mutational changes originated from the founding malignant clone rather than a newly emerged tumor clone with a different proviral integration site.

The observations in patients with aggressive subtypes who relapsed with the same main viral insertion site systematically revealed changes in the genomic landscape of the founding malignant clone through the acquisition of de novo mutations. Additional mutations were clustered in two main pathways: immune escape and NF- κ B/NFAT pathway. In addition, at relapse we observed clonal heterogeneity within the founding tumor clone with the emergence of a novel dominant subclone that was not identified at diagnosis and carried specific somatic mutations, while the main subclone at diagnosis had been either eradicated or maintained at comparable levels. Our observations in ATL are consistent with findings in acute myeloid leukemia (AML) [18, 19]. In agreement with earlier reports [7], we consistently observed the absence of Tax expression in ATLs whether at diagnosis, relapse, or upon progression [14, 15]. This supports the conclusion that in malignant T-cell, the acquisition of somatic mutations is independent of Tax expression although the contribution of viral products to early steps of T-cell transformation is undeniable [25, 26]. We also observed three cases with a

viral clone switch at progression/relapse, which consistently carried a mutation in the *TP53* locus. Malignant transformation of a newly emerging tumor clone may be facilitated by genomic instability promoted by therapeutic intervention, in agreement with previous observations in myeloid malignancies and clonal hematopoiesis [27, 28]. However, the contribution to the malignant development of these clones by the transient expression of the Tax oncoprotein cannot be excluded [29].

Deep sequencing of patients with indolent ATL who progressed to aggressive forms revealed genomic alterations mainly in the TCR/NF- κ B pathway. In contrast to patients who relapsed, aggressive ATLs progressing from indolent subtypes showed an enrichment in TCR/NF- κ B-associated mutations from rare preexisting subclones found within the main original tumor clone at the indolent stage. These subclonal mutations are clustered in genes belonging to the NF- κ B/NFAT pathway (*PLCG1*, *CARD11*, and *PRKCB*) and are known to cause a gain of function.

Mutational analysis suggests that, irrelevant of their clonal or subclonal nature, mutations in at least one of four genes associated with the TCR-NF- κ B pathway (*PLCG1*, *CARD11*, *PRKCB*, and *VAV1*) carry prognostic significance for ATL patients with an indolent subtype. We found that these patients experienced significantly shorter time to progression and survival rates. The prognostic value of subclonal mutations observed in ATL is consistent with recent reports demonstrating the clinical relevance of *TP53* subclonal mutations in chronic lymphoid leukemia and AML [30–32]. Our findings reveal the prognostic impact of low-frequency TCR/NF- κ B-mutated subclones on the outcome of ATL patients, suggesting that this phenomenon occurs for the well-known TCR/NF- κ B-driver pathway. We envision that the characterization of the subclonal architecture and its dynamics in the evolution of the disease will be relevant for the management of ATL patients. The observation of subclonal outgrowth at ATL progression and the prognostic value of the corresponding subclonal mutations identified in our discovery cohort will require validation in a larger patient cohort. Previous findings from a Japanese genomic study identified mutations in *IRF4* and/or amplification of *PDL-1* and/or deletion of *CDKN2A*, defining an unfavorable subgroup of indolent ATLs [12]. We suggest implementing this “classifier” with the four genes identified in our study. Targeting this specific set of mutations in an integrated NGS workflow may become a common approach to assess ATL’ genomic alterations and improve risk stratification for indolent ATLs.

In conclusion, our study describes the genomic landscape of ATL in patients of Afro-Caribbean origin, revealing a genomic architecture similar to that previously identified in a Japanese cohort, ruling out the contribution of the genetic background to HTLV-1-dependent tumorigenesis.

Furthermore, our longitudinal study identifies clonal evolution and dynamics as important and previously unappreciated contributors to disease progression in relapsing and progressing ATL patients from a mutational and a viral point of view. A particularly significant observation is the identification of mutations in the TCR/NF- κ B pathway in low-frequency subclones defining a new unfavorable subgroup within the indolent subtype with a high risk of progression to an aggressive disease.

Acknowledgements This study was supported by grants from the Institut National du cancer (INCA), the Fonds de la recherche scientifique (FRS), Télévie (Belgium), les Amis de l'Institut Bordet, and the International Brachet Stiftung (IBS). AM received a doctoral fellowship from the Institut National du Cancer (INCA), NR and KD are Scientific Research Worker of Télévie, VH was supported by a Télévie grant and les Amis de l'Institut Bordet, and MA by a post-doctoral fellowship from the FRS. We thank Wouter Coppieters, Latifa Karim, Manon Deckers, and the GIGA Genomics Platform for sequencing services and excellent technical support.

Author contributions AM designed the study and conducted the analyses. LL conducted TDS analysis. MA, KD, MG, and AVdB conducted the HTS and clonality analysis. NR and VH conducted bioinformatics analysis of HTS and viral expression. KD performed CGH/SNP array analysis. CL conducted bioinformatics analysis of NGS and mutational clonal evolution. AM, FS, DS, MC, and OH took care of the patients. VAF determined patients' proviral loads. JB stored histological samples, CP supervised acid nucleic extraction, AM, AVdB, VA, and OH wrote the manuscript and all authors critically reviewed and approved the manuscript prior to submission.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

- Gessain A, Cassar O. Epidemiological aspects and world distribution of HTLV-1 infection. *Front Microbiol.* 2012;3:388.
- Iwanaga M, Watanabe T, Yamaguchi K. Adult T-cell leukemia: a review of epidemiological evidence. *Front Microbiol.* 2012;3:322.
- 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;79:428-37.
- Katsuya H, Ishitsuka K, Utsunomiya A, Hanada S, Eto T, Moriuchi Y, et al. Treatment and survival among 1594 patients with ATL. *Blood.* 2015;126:2570-7.
- Cook LB, Fuji S, Hermine O, Bazarbachi A, Ramos JC, Ratner L, et al. Revised Adult T-Cell leukemia-lymphoma international consensus meeting report. *J Clin Oncol J Am Soc Clin Oncol.* 2019;37:677-87.
- Takasaki Y, Iwanaga M, Imaizumi Y, Tawara M, Joh T, Kohno T, et al. Long-term study of indolent adult T-cell leukemia-lymphoma. *Blood.* 2010;115:4337-43.
- Kataoka K, Nagata Y, Kitanaka A, Shiraishi Y, Shimamura T, Yasunaga J-I, et al. Integrated molecular analysis of adult T-cell leukemia/lymphoma. *Nat Genet.* 2015;47:1304-15.
- Tsukasaki K, Hermine O, Bazarbachi A, Ratner L, Ramos JC, Harrington W Jr, et al. Definition, prognostic factors, treatment, and response criteria of adult T-cell leukemia-lymphoma: a proposal from an international consensus meeting. *J Clin Oncol J Am Soc Clin Oncol.* 2009;27:453-9.
- Yamada Y, Hatta Y, Murata K, Sugawara K, Ikeda S, Mine M, et al. Deletions of p15 and/or p16 genes as a poor-prognosis factor in adult T-cell leukemia. *J Clin Oncol J Am Soc Clin Oncol.* 1997;15:1778-85.
- Nosaka K, Maeda M, Tamiya S, Sakai T, Mitsuya H, Matsuoka M. Increasing methylation of the CDKN2A gene is associated with the progression of adult T-cell leukemia. *Cancer Res.* 2000;60:1043-8.
- Tawara M, Hogerzeil SJ, Yamada Y, Takasaki Y, Soda H, Hasegawa H, et al. Impact of p53 aberration on the progression of Adult T-cell leukemia/lymphoma. *Cancer Lett.* 2006;234:249-55.
- Kataoka K, Iwanaga M, Yasunaga J-I, Nagata Y, Kitanaka A, Kameda T, et al. Prognostic relevance of integrated genetic profiling in adult T-cell leukemia/lymphoma. *Blood.* 2018;131:215-25.
- Bond J, Graux C, Lhermitte L, Lara D, Cluzeau T, Leguay T, et al. Early response-based therapy stratification improves survival in adult early thymic precursor acute lymphoblastic leukemia: a group for research on adult acute lymphoblastic leukemia study. *J Clin Oncol J Am Soc Clin Oncol.* 2017;35:2683-91.
- Rosewick N, Durkin K, Artesi M, Marçais A, Hahaut V, Griebel P, et al. Cis-perturbation of cancer drivers by the HTLV-1/BLV proviruses is an early determinant of leukemogenesis. *Nat Commun.* 2017;8:15264.
- Artési M, Marçais A, Durkin K, Rosewick N, Hahaut V, Suarez F, et al. Monitoring molecular response in adult T-cell leukemia by high-throughput sequencing analysis of HTLV-1 clonality. *Leukemia.* 2017;31:2532-5.
- Shao W, Shan J, Kearney MF, Wu X, Maldarelli F, Mellors JW, et al. Retrovirus integration database (RID): a public database for retroviral insertion sites into host genomes. *Retrovirology.* 2016;13:47.
- Marçais A, Waast L, Bruneau J, Hanssens K, Asnafi V, Gaulard P, et al. Adult T cell leukemia aggressiveness correlates with loss of both 5-hydroxymethylcytosine and TET2 expression. *Oncotarget.* 2017;8:52256-68.
- Ding L, Ley TJ, Larson DE, Miller CA, Koboldt DC, Welch JS, et al. Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. *Nature.* 2012;481:506-10.
- Christen F, Hoyer K, Yoshida K, Hou H-A, Waldhueter N, Heuser M, et al. Genomic landscape and clonal evolution of acute myeloid leukemia with t(8;21): an international study on 331 patients. *Blood.* 2019;133:1140-51.
- Vallois D, Dobay MPD, Morin RD, Lemonnier F, Missiaglia E, Juillard M, et al. Activating mutations in genes related to TCR signaling in angioimmunoblastic and other follicular helper T-cell-derived lymphomas. *Blood.* 2016;128:1490-502.
- Lenz G, Davis RE, Ngo VN, Lam L, George TC, Wright GW, et al. Oncogenic CARD11 mutations in human diffuse large B cell lymphoma. *Science.* 2008;319:1676-9.
- Challa-Malladi M, Lieu YK, Califano O, Holmes AB, Bhagat G, Murty VV, et al. Combined genetic inactivation of β 2-Microglobulin and CD58 reveals frequent escape from immune recognition in diffuse large B cell lymphoma. *Cancer Cell.* 2011;20:728-40.
- Mareschal S, Dubois S, Viailly P-J, Bertrand P, Bohers E, Maingonnat C, et al. Whole exome sequencing of relapsed/

- refractory patients expands the repertoire of somatic mutations in diffuse large B-cell lymphoma. *Genes Chromosomes Cancer*. 2016;55:251–67.
24. Shah UA, Chung EY, Giricz O, Pradhan K, Kataoka K, Gordon-Mitchell S, et al. North American ATLL has a distinct mutational and transcriptional profile and responds to epigenetic therapies. *Blood*. 2018;132:1507–18.
 25. Matsuoka M, Jeang K-T. Human T-cell leukaemia virus type 1 (HTLV-1) infectivity and cellular transformation. *Nat Rev Cancer*. 2007;7:270–80.
 26. Matsuoka M, Jeang K-T. Human T-cell leukemia virus type 1 (HTLV-1) and leukemic transformation: viral infectivity, Tax, HBZ and therapy. *Oncogene*. 2011;30:1379–89.
 27. Shih AH, Chung SS, Dolezal EK, Zhang S-J, Abdel-Wahab OI, Park CY, et al. Mutational analysis of therapy-related myelodysplastic syndromes and acute myelogenous leukemia. *Haematologica*. 2013;98:908–12.
 28. Coombs CC, Zehir A, Devlin SM, Kishtagari A, Syed A, Jonsson P, et al. Therapy-related clonal hematopoiesis in patients with non-hematologic cancers is common and associated with adverse clinical outcomes. *Cell Stem Cell*. 2017;21:374–82.e4.
 29. Mahgoub M, Yasunaga J-I, Iwami S, Nakaoka S, Koizumi Y, Shimura K, et al. Sporadic on/off switching of HTLV-1 Tax expression is crucial to maintain the whole population of virus-induced leukemic cells. *Proc Natl Acad Sci USA*. 2018;115: E1269–78.
 30. Nadeu F, Delgado J, Royo C, Baumann T, Stankovic T, Pinyol M, et al. Clinical impact of clonal and subclonal TP53, SF3B1, BIRC3, NOTCH1, and ATM mutations in chronic lymphocytic leukemia. *Blood*. 2016;127:2122–30.
 31. Nadeu F, Clot G, Delgado J, Martín-García D, Baumann T, Salaverria I, et al. Clinical impact of the subclonal architecture and mutational complexity in chronic lymphocytic leukemia. *Leukemia*. 2018;32:645–53.
 32. Prochazka KT, Pregartner G, Rücker FG, Heitzer E, Pabst G, Wölfler A, et al. Clinical implications of subclonal TP53 mutations in acute myeloid leukemia. *Haematologica*. 2019; 104:516–23.