CDKN2A deletion is a frequent event associated with poor outcome in patients with peripheral T-cell lymphoma not otherwise specified (PTCL-NOS)


Haematologica 2020 [Epub ahead of print]

Citation: Francesco Maura, Anna Dodero, Cristiana Carniti, Niccolò Bolli, Martina Magni, Valentina Monti, Antonello Cabras, Daniel Leongamornlert, Federico Abascal, Benjamin Diamond, Bernardo Rodriguez-Martin, Jorge Zamora, Adam Butler, Inigo Martincorena, Jose M.C. Tubio, Peter J. Campbell, Annalisa Chiappella, Giancarlo Pruneri, and Paolo Corradini. CDKN2A deletion is a frequent event associated with poor outcome in patients with peripheral T-cell lymphoma not otherwise specified (PTCL-NOS).
Haematologica. 2020;105:xxx

Publisher's Disclaimer.
E-publishing ahead of print is increasingly important for the rapid dissemination of science. Haematologica is, therefore, E-publishing PDF files of an early version of manuscripts that have completed a regular peer review and have been accepted for publication. E-publishing of this PDF file has been approved by the authors. After having E-published Ahead of Print, manuscripts will then undergo technical and English editing, typesetting, proof correction and be presented for the authors' final approval; the final version of the manuscript will then appear in print on a regular issue of the journal. All legal disclaimers that apply to the journal also pertain to this production process.
**CDKN2A deletion is a frequent event associated with poor outcome in patients with peripheral T-cell lymphoma not otherwise specified (PTCL-NOS)**


1 Myeloma Service, Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY, USA.
2 Department of Oncology and Hemato-Oncology, University of Milan, Milan, Italy;
3 The Cancer, Ageing and Somatic Mutation Programme, Wellcome Sanger Institute, Hinxton, Cambridgeshire, United Kingdom;
4 Weill Cornell Medical College, New York, NY, USA.
5 Department of Medical Oncology and Hematology, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy
6 Department of Pathology and Laboratory Medicine, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy;
7 CIMUS - Molecular Medicine and Chronic Diseases Research Centre, University of Santiago de Compostela, Spain
8 Present address: Department of Medical Oncology and Hematology, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy.
Department of Hematology Azienda Ospedaliera Città della Salute e della Scienza, Turin, Italy.

**Key words**: PTCL-NOS, Whole Genome Sequencing, Survival, PTEN, CDKN2A

**Running title**: CDKN2A and PTEN co-deletion in PTCL-NOS

**Corresponding Author:**

Prof Paolo Corradini
Division of Hematology and Bone Marrow Transplantation,
Fondazione IRCCS Istituto Nazionale dei Tumori, Via Venezian 1, 20133 Milan, Italy
Phone.: +39 0223903227; Fax: +39 0223902092
E-mail: paolo.corradini@unimi.it

Francesco Maura, MD,
Myeloma Service, Department of Medicine
Memorial Sloan Kettering Cancer Center, 1275 York Ave, New York, NY 10065
Phone: 212.639.5126 ; Fax: 646.277.7116
E-mail: maurof@mskcc.org
Abstract

Nodal peripheral T-cell lymphoma not otherwise specified (PTCL-NOS) remains a diagnosis encompassing a heterogenous group of PTCL cases not fitting criteria for more homogeneous subtypes. They are characterized by a poor clinical outcome when treated with anthracycline-containing regimens. A better understanding of their biology could improve prognostic stratification and foster the development of novel therapeutic approaches. Recent targeted and whole exome sequencing studies have shown recurrent copy number abnormalities (CNAs) with prognostic significance.

Here, investigating 5 formalin-fixed, paraffin embedded cases of PTCL-NOS by whole genome sequencing (WGS), we found a high prevalence of structural variants and complex events, such as chromothripsis likely responsible for the observed CNAs. Among them, \textit{CDKN2A} and \textit{PTEN} deletions emerged as the most frequent aberration, as confirmed in a final cohort of 143 patients with nodal PTCL. The incidence of \textit{CDKN2A} and \textit{PTEN} deletions among PTCL-NOS was 46% and 26%, respectively. Furthermore, we found that co-occurrence of \textit{CDKN2A} and \textit{PTEN} deletions is an event associated with PTCL-NOS with absolute specificity. In contrast, these deletions were rare and never co-occurred in angioimmunoblastic and anaplastic lymphomas. \textit{CDKN2A} deletion was associated with shorter overall survival in multivariate analysis corrected by age, IPI, transplant eligibility and GATA3 expression (adjusted HR =2.53; 95% CI 1.006-6.3; p=0.048). These data suggest that \textit{CDKN2A} deletions may be relevant for refining the prognosis of PTCL-NOS and their significance should be evaluated in prospective trials.
Introduction

Within the heterogeneous categories of nodal peripheral T-cell lymphoma (PTCL) subtypes currently recognized by the WHO classification system\(^1\text{--}^4\), many subtypes have defining molecular and phenotypic features. For example, within the anaplastic large cell lymphoma (ALCL) subtype, translocations of the ALK oncogene define a homogeneous subgroup, while ALK-negative cases carry mutations or translocations resulting in the activation of the JAK/STAT pathway\(^5\). In angioimmunoblastic T-cell lymphoma (AITL), mutations in *DNMT3A*, *TET2*, *IDH2* and *RHOA* have been described in a significant fraction of patients\(^6\text{--}^8\). These features are increasingly used for diagnostic and prognostic purposes, and may translate into targeted or rationale treatment approaches\(^9,10\). A subgroup of PTCL not otherwise specified (PTCL-NOS) carries a T-cell follicular helper (TFH) phenotype and some phenotypic and genetic features of AITL including mutations affecting *TET2*, *DNMT3A*, and *RHOA* genes\(^6,11\). Therefore, nodal peripheral T-cell lymphomas with TFH phenotype have been included in the revised WHO classification as a provisional entity\(^4\). On the contrary, the vast majority of PTCL-NOS are a large yet ill-defined subgroup that lack defining genetic of phenotypic features\(^6,11\text{--}^13\). They are characterized by a poor clinical outcome when treated with anthracycline-containing regimens and some non-randomized studies suggest that high-dose chemotherapy can offer some survival benefit in young patients\(^14,15\). However, a better understanding of their biology could improve prognostic stratification and foster the development of novel therapeutic approaches.

The gene expression profile of PTCL-NOS suggests that two major subtypes can be identified, one characterized by *GATA3* and the other by *TBX21* overexpression, the former carrying worse prognosis\(^16\text{--}18\). More recently, two groups
identified recurrent copy-number alterations (CNAs) in the tumor suppressors TP53 and CDKN2A with adverse prognostic significance, and generally associated with genomic instability\textsuperscript{19,20}. These efforts have been based on next-generation sequencing (NGS) of targeted gene panels and single-nucleotide polymorphism (SNP) arrays. Contrary to the above studies, whole genome sequencing (WGS) can interrogate the full repertoire of somatic mutations, CNAs, SV and even mutational processes involved in cancer pathogenesis\textsuperscript{21}. However, WGS in PTCL have been hampered by the availability of high-quality, tumor-rich DNA. On the other end, novel findings must be validated and applicable to everyday diagnostic routine before the benefits of additional information are translated into clinical improvements.

With the hypothesis that WGS may reveal the genomic bases of recurrent genomic alterations in PTCL-NOS, here we decided to test this approach in formalin-fixed, paraffin embedded (FFPE) samples. Furthermore, using an extended cohort of cases, we validate our and published results by FISH and immunohistochemistry, and analyze prognostic correlates.

**Methods**

**Sample selection**

Eleven PTCL-NOS patients treated at the Department of Hematology, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy were selected for the WGS analysis (Supplementary Table 1). The study protocol was approved by the institutional review board (n. INT 19/13) and was conducted in accordance with the Declaration of Helsinki. The diagnosis was made according to the 2016 World Health Organization (WHO) classification\textsuperscript{4} and subtype assignment was independently
reviewed by two expert hemato-pathologists. Only PTCL-NOS case were selected. Nodal lymphomas with T-follicular helper (TFH) cell origin, defined according to the 2017 WHO update⁴ by the association of a CD4 + phenotype with the expression of at least 2 TFH markers among ICOS, PD1, CXCL13, BCL6 and CD10 by, were excluded on morphology and immunohistochemical bases. Additionally a comprehensive morphologic assessment was performed evaluating features commonly associated with nodal lymphomas with T-follicular helper cell origin such as the presence of large Reed-Sternberg-like, CD30 +/CD15 + and PAX5+ cells, EBV positivity without a pabulum and with a growth pattern that mimics either a follicular lymphoma or progressively transformed germinal centers. Samples were chosen according to tumor cellularity and amount of extracted DNA (>500 ng). DNA was extracted from formalin-fixed paraffin-embedded (FFPE) blocks in 10 patients, as previously described²². For one patient (PD30774a), DNA was extracted from a fresh frozen sample (peripheral blood mononuclear cells harvested during leukemic progression). For one patient, we sequenced 2 samples: one collected at diagnosis (PD30770a) and one at relapse (PD30770c). Therefore, all but 2 samples (PD30774a and PD30770c) were collected at diagnosis. DNA from buccal swabs was used as normal match for all samples.

Whole genome sequencing data analysis

WGS libraries were prepared with the TruSeq DNA PCR-Free Library Preparation Kit (Illumina San Diego, CA, USA) from 500ng of genomic DNA, aiming for an average target insert of 300bp. Sequencing was performed on the Illumina X10 platform at the Wellcome Sanger Institute (WSI) on a 150bp-paired end protocol to a target depth of 40x for tumor samples and 30x for normal controls. The
sequence data described here will be made available from the EGA repository (EGAS00001002057). FASTQ files were aligned to the reference genome (GRCh37/hg19) using BWAmem and deduplicated aligned BAM files were analyzed using the following published tools available at the WSI: ASCAT and Battenberg to measure clonal and subclonal copy number changes and to estimate the tumor cell fraction of each sample\textsuperscript{23}, Caveman and Pindel for Single Nucleotide Variants (SNVs) and small insertion-deletions (indels)\textsuperscript{24,25}; BRASS for SVs (large inversions and deletions, translocations, internal tandem duplications)\textsuperscript{23}. Complex events were defined, as previously described\textsuperscript{26}. TraFIC was used to described the somatic L1 retrotransposition landscape\textsuperscript{27}.

The repertoire of mutational processes operative in PTCL-NOS was analyzed by extracting the corresponding mutational signatures using the 96-class matrix of all possible substitutions in their 5’ and 3’ context with non-negative matrix factorization (NNMF), as previously described\textsuperscript{28,29}. Mutations were classified as drivers based on the COSMIC census catalogue of cancer genes (http://cancer.sanger.ac.uk/cosmic/).

Validation cohorts

To expand our analysis on mutations in driver genes, we included published exomes data from sixty-three PTCL patients (30 PTCL-NOS ,15 AITL, 23 ALCL ALK neg and 16 EATL-II \textsuperscript{6,7,5,30,31,13}. COSMIC census was used to create the catalogue of genes potentially involved by nonsynonymous mutations (https://cancer.sanger.ac.uk/census).

\textit{CDKN2A} and \textit{PTEN} status was validated by dual-colour FISH on paraffin-embedded selected tumor areas using commercially available DNA probes: for p16 (\textit{CDKN2A}) we used a spectrum orange-labelled locus-specific \textit{CDKN2A} (9p21)
probe and spectrum green-labelled chromosome 9 centromeric probe (LSI CDKN2A/CEP 9); for PTEN, a spectrum orange-labelled locus-specific PTEN (10q23) probe and a spectrum green-labelled chromosome 10 centromeric probe (LSI PTEN/CEP 10) – Vysis Inc., Downers Grove, IL, USA. A CDKN2A deletion was defined in presence of either a homozygous deletion in >10% of cells or a hemizygous deletion in >40% of cells.27 A cut off or 40% was applied to define the hemizygous PTEN loss. CDKN2A/p16 and GATA3 protein expression was tested by immunohistochemistry (IHC), as previously described16–18. To define GATA3 IHC positivity we used the recently proposed 50% cut off18.

For further validation, we included published data from 20 AITL samples with copy number status defined using different NGS approaches7,31 and 81 PTCLs investigated by SNP array32,33. For this last series, we downloaded all the available CEL files from GEO (accession numbers GSE15842 and GSE50252) and then LogR and BAF data were extracted using pennCNV34. We then applied the ASCAT algorithm to perform segmentation and retrieve allele-specific absolute copy number alterations. The description of entire cohort used to validate CDKN2A and PTEN status can be found in Supplementary Table 2.

All contingency analyses were performed by Fisher’s exact test. Estimated progression-free survival (PFS) and overall survival (OS) were calculated with the Kaplan-Meier method and groups were compared with the log-rank test. Multivariate analysis was performed with Cox regression. All the analyses were performed using appropriate functions in the R 3.4.2 software (www.r-project.org).

Results

PTCL-NOS show a complex genomic architecture
We performed WGS on 12 tumors and 11 matched normal samples from 11 patients affected by PTCL-NOS, achieving an average depth of 27X (Supplementary Table 1). For one patient (PD30770) a diagnosis and a relapse tumor sample were available. As expected, analysis was severely hampered by the nature of samples analyzed, so that samples from 6 patients did not pass quality control. Causes ranged from insufficient cluster generation due to low quality DNA (1 sample), low cancer cell fraction (CCF) (1 sample), FFPE-induced artifacts (4 samples). More data on quality control, FFPE artifacts and mutational signatures in PTCL-NOS WGS samples can be found in the Supplemental Data section.

In the remaining 5 patients, no recurrent point mutations or indels in onco-driver genes were identified, in line with recent observations\textsuperscript{6,11}. The rare involvement of onco-driver gene in PTCL-NOS compared to other PTCL subtypes was confirmed in additional 84 published whole exome sequencing cases\textsuperscript{5–7,30,31} (Supplementary Figure 1-2). We therefore focused the investigation on structural variants (SVs: defined as inversions, translocations, internal tandem duplications (ITD) and deletions) and CNAs, which could be investigated in depth in our 5 WGS samples. Three hundred and seventy-two SVs were extracted with a median of 74 per sample (range 56-86) (Figure 1A). Intriguingly, at least one complex event was observed in all but one patient, including five chromothripsis events in three patients (Figure 1B)\textsuperscript{35}. Importantly, several cancer genes were affected either by SVs directly or by CNAs caused by SVs and complex events (Figure 1C). For example, sample PD30771a was characterized by 3 different chromothripsis events on multiple chromosomes, including one causing a homozygous deletion of CDKN2A (Figure 1C). The genomic landscape was even more complex in relapse samples: PD30774a showed a whole genome duplication, and a large chromothripsis event
was responsible for disruption of several known oncogenes resulting, for instance, in ARID1B and ARID2 losses (Figure 1C); analysis of serial diagnosis and relapse samples in patient PD30770 showed acquisition of numerous SVs and CNAs, and a complex event in chromosome 16 (Supplementary Figure 3). Overall, this confirms how driver events in PTCL-NOSs may often go beyond coding gene mutations and involve complex structural events that can only be investigated by whole genome sequencing.

**CDKN2A and PTEN are frequently co-deleted in PTCL-NOS, but not in other PTCLs**

Most of CNA and SV breakpoints were not recurrent, but we found the tumor suppressor CDKN2A deleted in 4 out of 5 patients. Collectively, we found a simple homozygous deletion (HD), and a second HD caused by chromothripsis. In two further cases, breakpoints around the copy-number loss could not be resolved by a distinct SV event, so the cause remains unknown (Figure 1C). Interestingly, PTEN loss was observed in 2 out of 4 CDKN2A-deleted patients. FISH analysis on sections from the 5 cases analyzed by WGS confirmed the CDKN2A and the PTEN deletion in all cases (Figure 2A), and the diploid status of the genes in the ones where no deletions were found by NGS. The high prevalence of these deletions in our small dataset confirmed recently published data and prompted us to extend our observations. To this end, we evaluated by FISH 37 PTCL archive cases from our institution (30 PTCL-NOS, 5 ALKneg ALCL and 2 AITL cases, Supplementary Table 2). The prevalence of CDKN2A deletions among PTCL-NOS was 50%, and HDs accounted for two third of cases. Overall, 31% of cases showed PTEN deletions.

To compare the prevalence of CDKN2A and PTEN deletion across the main PTCL histologies, we included 101 cases from published datasets7,31–33 for a total of
143 PTCLs of which 63 PTCL-NOS (Supplementary Table 2; Supplementary Figure 4). *CDKN2A* was deleted in 29/63 (46%) PTCL-NOS cases, including 19/29 (65%) HDs. This was significantly more frequent than what found in other histologies (Figure 2B-C). *PTEN* was deleted in 17/63 (26%) PTCL-NOS cases, 13 (76%) of which also carried a *CDKN2A* loss. Importantly, the co-occurrence of *CDKN2A* and *PTEN* was specific for PTCL-NOS (Figure 2B-C), representing the first genetic abnormality of such kind.

**CDKN2A genetic status does not correlate with p16 nor with GATA3 expression**

We next evaluated *CDKN2A* expression (p16) by IHC in 31 PTCL-NOS samples for whose archival material was available. Overall, the average percentage of tumor cells expressing p16 in WT cases was 23.44% (range, 0-90%), while it was 6.21% (range 0-30%) in deleted cases, either mono- or bi-allelic. This difference was significant (*p* = 0.01, Student t-test) (Figure 3A-C). Using the previously described cutoff of 20% of tumor cells to define a categorical present/absent p16 expression status, only one case was p16 positive when *CDKN2A* was deleted, and no cases of bi-allelic deletions showed p16 positivity. However, 9/14 cases were p16 negative in spite of a WT *CDKN2A* locus (Figure 3A). This suggests that reduced expression of p16 can stem from non-genetic events, including epigenetic or post-transcriptional mechanisms as reported in other cancer types36,37.

Next, we correlated the genetic status of *CDKN2A* and GATA3 expression. Again, the picture was far from clear: the average percentage of GATA3-expressing cells was 23.12% in CDKN2A WT cases, and 33.72% in CDKN2A-deleted cases. This difference was not significant (*p* = 0.31, Student t-test) (Figure 3D). Even after
setting the threshold for GATA3 expression at 50%\textsuperscript{18}, the number of GATA3 positive cases did not differ between CDKN2A WT and deleted cases.

Finally, taking advantage of the combined WGS-SNP array cohort of 33 PTCL-NOS cases, we explored additional structural events in this subgroup. We observed 17p13 (TP53) deletions in 11 patients (33%), with partial overlap with CDKN2A deletions (Supplementary Figure 5).

**CDKN2A deletions carry prognostic significance**

Clinical data were available for all PTCL-NOS patients from our WGS and FISH cohort except one that was not treated as a PTCL-NOS due to an initial diagnostic misclassification. The median age of the cohort was 59 years (range, 22-85), with 79% (27/34) transplant eligible and 43% (14/32) with IPI above 2. Most of the patients (n=31, 94.4%) received an induction therapy with an anthracycline. During the course of the disease, 14 patients (41%) received autologous stem cell transplantation (SCT) and/or allogeneic SCT (Table 1)\textsuperscript{38,14}.

After a median follow-up of 70 months, 11 patients (32%) were still alive. We thus evaluated whether CDKN2A and PTEN genomic status affected the response to therapy and outcome. PTEN deletion did not impact the clinical outcome in terms of both PFS and OS. In contrast, CDKN2A-deleted cases were more frequent in primary refractory than wild-type cases [15/20 (75%) vs 5/13 (38%); p=0.06]. Furthermore, 16/20 (80%) CDKN2A-deleted patients died in disease progression already within the first year of diagnosis. Consequently, PTCL-NOS patients carrying CDKN2A deletions had a significantly shorter PFS compared to wild type patients [5-year PFS 7.5% (95%CI: 1.3%-42%) vs 24%, (95%CI: 9%-63.5%) p=0.045] and OS [5-year OS 22.5% (95%CI: 9.4%-54%) vs 52%, (95%CI: 30%-90%) p=0.048].
The association between short survival and CDKN2A deletion retained its significance after multivariate correction by age, IPI, transplant eligibility and GATA3 IHC expression (Figure 4C-D).

**Discussion**

In this paper we performed the first WGS experiment on PTCL-NOS samples. Despite the limited series, WGS showed ubiquitous instances of complex rearrangements and prevalent inactivation of classical tumor suppressor genes by SVs and CNAs. Chromothripsis in particular, a common mechanism underlying complex rearrangements, was never described in PTCL-NOS and frequently seen in our samples thanks to the comprehensive analysis of the genome allowed by WGS. This observation likely provides the mechanistic bases of the frequent CNAs observed in PTCL-NOS cases in larger cohorts, including CDKN2A deletions\(^{19,20}\). In our study, we could also observe how CDKN2A deletions were often focal and biallelic. This inactivation pattern is highly specific for driver tumor suppressors, and has been described for CDKN2A in other hematological malignancies\(^{39-41}\). We went on to confirm this observation in a large validation cohort, showing how other PTCL subtypes did not share these abnormalities, pointing at a specific driver role of CDKN2A in PTCL-NOS pathogenesis. What was most striking though, was that PTEN and CDKN2A co-deletion had a 100% positive predictive value for the diagnosis of PTCL-NOS, providing the first genetic abnormality that provides absolute specificity for this PTCL subgroup. The integration of these genomic aberrations with point mutations recently reported in distinct subtypes might allow the generation of an accurate diagnostic flow based on NGS as described for other
malignancies\textsuperscript{42–44}, potentially able to overcome historical classification difficulties in PTCL.

Importantly, deletions of \textit{PTEN}, a negative regulator of the PI3K pathway were also enriched in PTCL-NOS. From a mechanistic point of view, this suggests a possible epistatic relationship between the two pathways, and in fact the role of \textit{PTEN} in T-lymphomagenesis has been postulated in vivo\textsuperscript{45}.

A distinction between PTCL-NOS cases has recently been proposed, such that some cases would show a more complex genomic architecture with frequent CNAs involving tumor suppressor genes and worse prognosis, while others would show fewer abnormalities and carry a better prognosis\textsuperscript{19,20}. A correlation between the former group and higher GATA3 expression has also been postulated by some groups\textsuperscript{19}, but not confirmed by others\textsuperscript{20}. High GATA3 expression is itself an adverse prognostic marker in PTCL-NOS\textsuperscript{17}. However, in our series, allelic status of CDKN2A did not correlate with GATA3 expression by IHC. Furthermore, we found that p16 expression by IHC was a poor predictor of \textit{CDKN2A} genetic status, likely suggesting that, similarly to other cancers, non-genetic mechanisms converge towards p16 down-regulation as a universal driver of PTCL-NOS. Given the high heterogeneity of the PTCL-NOS subgroup, and the variability of IHC staining of different proteins, further studies will be needed to validate these findings. However, our multivariate analysis of survival confirmed that CDKN2A allelic status was a significant predictor of inferior PFS and OS in PTCL-NOS. Importantly, this was independent of age, IPI, transplant eligibility and GATA3 expression by multivariate analysis. Future studies on larger number of patients will be required to confirm our findings.
While the molecular categorization of PTCL-NOS still remains controversial, our study confirms the utility of WGS in the study of malignancies where recurrent gene mutations are lacking and the classification itself is often ambiguous. Moving from the complexities of WGS, we also show that an old-fashioned FISH test can be of relevance in PTCL-NOS. In the future, the diagnostic value of $CDKN2A$ and $PTEN$ co-deletion, and the prognostic value of $CDKN2A$ deletions will need to be assessed in prospective studies.

**Disclosure of conflicts of interest:**

No conflict of interests to declare

**Authorship:**

P.C., F.M. designed the study, collected and analyzed the data and wrote the paper; D.L., A.D., C.C, N.B. analyzed the data and wrote the paper; B.R.M., J.T., J.Z., F.A., A.B., P.J.C., B.D., and I.M. analyzed the data; A.T., A.C., M.F. and G.P. performed the IHC and FISH validations; G.B., M.M. and A.C. collected the data;

**Acknowledgements:**

This work is supported by AIRC (Associazione Italiana Ricerca sul Cancro) and A.I.L. (Associazione Italiana Contro le Leucemie-Linfomi e Mieloma ONLUS).

This work is supported by the Memorial Sloan Kettering Cancer Center NCI Core Grant (P30 CA 008748)

FM is supported by the American Society of Hematology, the International Myeloma Foundation and The Society of Memorial Sloan Kettering Cancer Center
References


### Tables

**Table 1.** Demographic and clinical characteristics of the PTCL-NOS cohort used for survival analysis

<table>
<thead>
<tr>
<th>Variable</th>
<th>All patients</th>
<th>CDKN2 deleted</th>
<th>CDKN2 wt</th>
<th>n/e</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>59.9 (22-85)</td>
<td>59.5 (22-84)</td>
<td>59.9 (34-85)</td>
<td></td>
<td>p=0.7</td>
</tr>
<tr>
<td>Gender (female)</td>
<td>17/34 (50%)</td>
<td>7/20 (35%)</td>
<td>10/14 (71%)</td>
<td>0</td>
<td>p=0.07</td>
</tr>
<tr>
<td>Bone Marrow Infiltration</td>
<td>9/30 (30%)</td>
<td>4/18 (22%)</td>
<td>5/12 (41%)</td>
<td>4</td>
<td>p=0.4</td>
</tr>
<tr>
<td>Extra-nodal Disease</td>
<td>12/34 (35%)</td>
<td>9/20 (45%)</td>
<td>3/14 (21%)</td>
<td></td>
<td>p=0.27</td>
</tr>
<tr>
<td>IPI &gt;2</td>
<td>14/32 (43%)</td>
<td>7/18 (39%)</td>
<td>8/14 (57%)</td>
<td>2</td>
<td>p=0.47</td>
</tr>
<tr>
<td>SCT eligible</td>
<td>27/34 (79%)</td>
<td>16/20 (80%)</td>
<td>11/14 (78%)</td>
<td></td>
<td>p=1</td>
</tr>
<tr>
<td>Anthracycline Induction CT</td>
<td>31/34 (91%)</td>
<td>18/20 (90%)</td>
<td>13/14 (9%)</td>
<td></td>
<td>p=0.7</td>
</tr>
<tr>
<td>Response to first line1 (CR)</td>
<td>13/33 (39%)</td>
<td>5/20 (25%)</td>
<td>5/13 (38%)</td>
<td>1</td>
<td>p=0.06</td>
</tr>
<tr>
<td>ASCT or AlloSCT</td>
<td>14/34 (41%)</td>
<td>5/20 (25%)</td>
<td>9/14 (64%)</td>
<td></td>
<td>p=0.035</td>
</tr>
<tr>
<td>GATA3 &gt;50%</td>
<td>9/34 (26%)</td>
<td>6/20 (30%)</td>
<td>3/14 (21%)</td>
<td></td>
<td>p=0.7</td>
</tr>
</tbody>
</table>

SCT = stem cell transplantation  
ASCT = autologous stem cell transplantation  
AlloSCT = allogeneic stem cell transplantation  
CR = complete remission  
PT = partial remission  
n/e = not evaluable
Figure Legends

Figure 1. **PTCL-NOS structural variants and complex events.** Prevalence of structural variants (A) and complex events (B) among PTCL-NOS. C) heatmap summarizing the oncogenes whose disruption was caused by at least one structural variants. The heatmap color highlighted the type of SV involved. Yellow = not determinable, i.e. the oncogene loss could not be mapped to a specific structural variant. For three events, the chromothripsis event responsible for the oncogene loss is drawn on the right.

Figure 2. **Prevalence of CDKN2A and PTEN deletions among different PTCL subtypes.** A) Heatmap showing the prevalence of CDKN2A and PTEN deletions across the main PTCL subtypes. B) representative FISH pictures of PTEN (top) and CDKN2A (bottom) deletions. C) frequency table of CDKN2A deletions, mono- and bi-allelic, and co-occurrence with PTEN deletions across PTCL subtypes. All statistical comparisons were performed by Fisher’s exact test. PTCL-NOS showed a significant enrichment for these three genomic aberrations compared to all the other PTCL histologies (p<0.01).

Figure 3. **CDKN2A protein expression in PTCL-NOS.** A) Expression of CDKN2A/p16 by immunohistochemistry. B-C) Representative examples of CDKN2/p16 IHC results in cases with normal (B) and deleted (C) alleles. D) GATA3 protein expression evaluated by IHC in CDKN2A deleted and wild type cases. P-value was estimated using a Wilcoxon test.
**Figure 4. CDKN2A deletion clinical impact.** Progression-free survival (PFS. A, B) and overall survival (OS. C, D) according to CDKN2A status. A, C) Kaplan-Meier plots of survival. P values were estimated using a log-rank test. B, D) multivariate analysis of survival, performed using Cox regression.
A

Histology

CDKN2A

PTEN

Mono-allelic deletion  Bi-allelic deletion

B

PD30765a

C

<table>
<thead>
<tr>
<th>Histology</th>
<th>CDKN2A deletion</th>
<th>CDKN2A bi-deletion</th>
<th>PTEN and CDKN2A co-deletions</th>
</tr>
</thead>
<tbody>
<tr>
<td>PTCL-NOS</td>
<td>29/63 (46%)</td>
<td>19/63 (30%)</td>
<td>13/63 (20%)</td>
</tr>
<tr>
<td>ALCL ALK neg</td>
<td>4/28 (14%)</td>
<td>1/28 (3.5%)</td>
<td>0/28 (0%)</td>
</tr>
<tr>
<td>ALCL ALK pos</td>
<td>0/25 (0%)</td>
<td>0/25 (0%)</td>
<td>0/25 (0%)</td>
</tr>
<tr>
<td>AITL</td>
<td>2/22 (9%)</td>
<td>1/22 (4.5%)</td>
<td>0/22 (0%)</td>
</tr>
<tr>
<td>p.value</td>
<td>p&lt;0.01</td>
<td>p&lt;0.01</td>
<td>p&lt;0.01</td>
</tr>
</tbody>
</table>
CDKN2A deletion is a frequent event associated with poor outcome in patients with peripheral T-cell lymphoma not otherwise specified (PTCL-NOS)

Maura et al.
**Supplementary Figure 1.** Summary of whole exome sequencing data used to investigate the prevalence of mutations in driver genes in PTCLs.
Supplementary Figure 2. Number of driver mutations per patient across different PTCLs.
**Supplementary Figure 3.** Genome plots at diagnosis (PD30770a) and at relapsed (PD30770c) after treatment in the only patient where 2 samples were collected at different time points. In this patient, at relapse, several novel SV were observed, including complex events (i.e. on chromosome 16). From the external ring to the internal: mutations, (vertically plotted according to their inter-mutational distance and where the color of each dot represents the mutation class), indels (dark green = insertion; and brown = deletion); copy number variants (red = deletions, green = gain), rearrangements (blue = inversion, red = deletions, green = ITD, black=translocations).
Supplementary Figure 4. Summary of copy number data used to investigate the prevalence of CDKN2A and PTEN deletion across different PTCLs.
Supplementary Figure 5. The prevalence of *TP53*, *PTEN* and *CDKN2A* deletions among PTCL-NOS samples with available ASCAT copy number data (SNP array or WGS data).
Supplementary Table 1. Sample characteristics and summary of the WGS

Supplementary Table 2. All PTCL patients included in the study and evaluated for CDKN2A and PTEN status.
CDKN2A deletion is a frequent event associated with poor outcome in patients with peripheral T-cell lymphoma not otherwise specified (PTCL-NOS)

Maura et al.

Initial quality control of FFPE-WGS data

We performed WGS on 12 tumors and 11 matched normal samples from 11 patients affected by PTCL-NOS, achieving an average depth of 27X. Two FFPE samples were removed from analysis: one due to low cancer cell fraction (CCF) as retrieved by ASCAT (PD30772a) and the other based on cluster generation issues during sequencing likely caused by a hyper-fragmented DNA (PD30768a). Among the remaining 10 tumor samples, we extracted 59,617 somatic base substitutions (range 2,471-10,756, median 6,358 per patient) and 20,531 small insertion-deletions (indels) (range 84-6,397, median 1,580) (Supplementary Data Fig. 1). Interestingly, four samples (PD30764a, PD30766a, PD30767a and PD30769a) were characterized by a similarly low CCF but also by a unique genomic profile, with a very low prevalence of coding mutations and a high indel/SNVs ratio in three of them (PD30766a, PD30767a and PD30769a).
**Supplementary Data Figure 1.** FFPE-artefact characterization. A) Circos plot of one sample heavily involved by FFPE artefact. From the external ring to the internal: mutations, (vertically plotted according to their inter-mutational distance and where the colour of each dot represents the mutation class), indels (dark green = insertion; and brown = deletion); copy number variants (red = deletions, green = gain), rearrangements (blue = inversion, red = deletions, green = ITD, black=translocations). B) Barplot showing the number of SNVs and indels for each sample. C) Barplot with the percentage of SNVs and indels extracted within LINE elopements. The PD30774a is the only samples where the DNA was extracted by fresh frozen material.
Investigating the genome-wide distribution of mutations in these four patients, we observed that more than 80% of indels and SNVs occurred within reference LINE-1 (L1) elements, predominantly of the L1PA family, and most of these were detected at low variant allele frequency (Supplementary Data Figure 2).

Supplementary Data Figure 2. Genomic features associated with FFPE-artefact Indels and SNVs V-allelic frequency (VAF) density plots for each sample (A). PD30764a, PD30766a, PD30767a and PD30769a SNVs and indels were mostly characterized by very low VAF (<20%). B) Distribution of SNVs and indels across different LINE elements. A significant enrichment was observed within the L1PA for both SNVs and indels among FFPE samples.
Consistent with this observation, the TraFIC pipeline highlighted a particularly increased somatic L1 retrotransposition activity among the three cases with a high indel rate (Supplementary Data Table 1).

### Supplementary Data Table 1. Summary of retrotransposons extracted by TraFic.

<table>
<thead>
<tr>
<th>Tumor ID</th>
<th>nbTotal</th>
<th>nbSoloL1</th>
<th>nbL1TD</th>
<th>nbAlu</th>
<th>nbSVA</th>
<th>nbERVK</th>
<th>nbPSD</th>
<th>nbL1DEL</th>
<th>nbL1DUP</th>
</tr>
</thead>
<tbody>
<tr>
<td>PD30763a</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>PD30764a</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>PD30765a</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>PD30766a</td>
<td>26</td>
<td>25</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>PD30767a</td>
<td>25</td>
<td>25</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>PD30769a</td>
<td>12</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>PD30770a</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>PD30770c</td>
<td>7</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>PD30771a</td>
<td>13</td>
<td>13</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>PD30774a</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

The vast majority of L1 events (~91%, 78/86) were consistently located over repeats of the L1PA subfamily. Manual curation revealed how all of these indels and retrotranspositions were compatible with sequencing and/or alignment artifacts rather than genuine events (Supplementary Data Figure 3).

After removing these artifacts, no significant retrotransposition activity was highlighted by TraFiC across the entire series, in line with the general absence of this somatic process among different hematological malignancies.
Mutational Signatures in PTCL-NOS

Our WGS data allowed the investigation of the mutational signature landscape in PTCL-NOS for the first time. Using the NNMF framework algorithm seven mutational signatures were extracted - six known and one previously undescribed (Supplementary Data Figure 4A-B). Among known signatures, SBS1 and SBS5 are related to cell aging and were observed in all samples, confirming their ubiquitous activity across normal and tumor tissues\(^1\). We demonstrated a contribution of the APOBEC family of DNA deaminases (SBS2 and SBS13) to the mutational spectrum of PTCL, adding yet another disease entity to the list of lymphoid neoplasms where this process is operative. Within hematological cancers, SBS8 was described in myeloma\(^2\) (39), SBS17 was described in B-cell lymphomas\(^3\). While the etiology of these signatures remains unknown, our findings show that PTCL-NOS share similar mutational processes with other lymphoid malignancies. The novel signature was particularly enriched among the four samples heavily affected by FFPE and artefactual mutations in retrotransposable elements, and absent in the unique fresh frozen sample (Supplementary Data Figure 3).

Supplementary Data Figure 3. The hallmarks of artefactual somatic LINE-1 (L1) insertion calls in FFPE treated samples. A) Illustrative example of a canonical L1 integration though target-primed reverse transcription (TPRT). Two well defined clusters composed by discordant reads in opposite orientations and whose mates align on the body of a L1 element elsewhere support the integration of a L1 element. Two additional clusters of clipped reads reveal the insertion breakpoints, the length of the target site duplication (TSD) and the insertion DNA strand. B) Artefactual L1 insertion call in PD30765a, a PTCL-NOS sample treated with FFPE before paired-end whole-genome sequencing. The high number of scattered discordant read-pairs leads to the artefactual insertion call. On the bottom, abundant clipped reads but not organized into breakpoint clusters in the region. Most of these artifacts (~91%, 78/86) are located over repeats of the L1 family, especially from L1PA subfamilies.
potentially relating the signature to formalin-induced DNA degradation. Furthermore, the 96 classes profile around mutations in LINE elements was highly similar to this novel FFPE signature (cosine similarity = 0.91) (Supplementary Data Figure 4C-D), suggesting a link between formalin fixation and the mutational process represented by this signature.
**Supplementary Data Figure 4.** Barplot showing the absolute (A) and relative (B) contribution of different extracted mutational processes for each sample. C) The 96-mutational classes of all mutations occurred within LINE elements in FFPE-samples. D) The FFPE-related mutational signature profile.
Considering the low cancer purity, the low number of coding mutations, the high number of FFPE artifacts in both SNVs and indels, the samples PD30764a, PD30766a, PD30767a and PD30769a were removed from the subsequent analysis.
Supplementary references

