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

Francesco Maura,[1-4] Anna Dodero,[5] Cristiana Carniti,[5] Niccolò Bolli,[2,5] Martina Magni,[5] Valentina Monti,[6] Antonello Cabras,[6] Daniel Leongamornlert,[3] Federico Abascal,[3] Benjamin Diamond,[1] Bernardo Rodriguez-Martin,[7] Jorge Zamora,[7] Adam Butler,[3] Inigo Martincorena,[3] Jose M. C. Tubio,[7] Peter J. Campbell,[3] Annalisa Chiappella,[8°] Giancarlo Pruneri[2,6] and Paolo Corradini[2,5]

[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, Aging and Somatic Mutation Program, Wellcome Sanger Institute, Hinxton, Cambridgeshire, UK; [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 Center, University of Santiago de Compostela, Santiago de Compostela, Spain and [8]Department of Hematology Azienda Ospedaliera Città della Salute e della Scienza, Turin, Italy.

°Current address: Department of Medical Oncology and Hematology, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy.

 doi:10.3324/haematol.2020.262659

Received: June 12, 2020.

Accepted: September 2, 2020.

Pre-published: September 10, 2020.

Correspondence: *PAOLO CORRADINI* - paolo.corradini@unimi.it

FRANCESCO MAURA - francesco.maura85@gmail.com

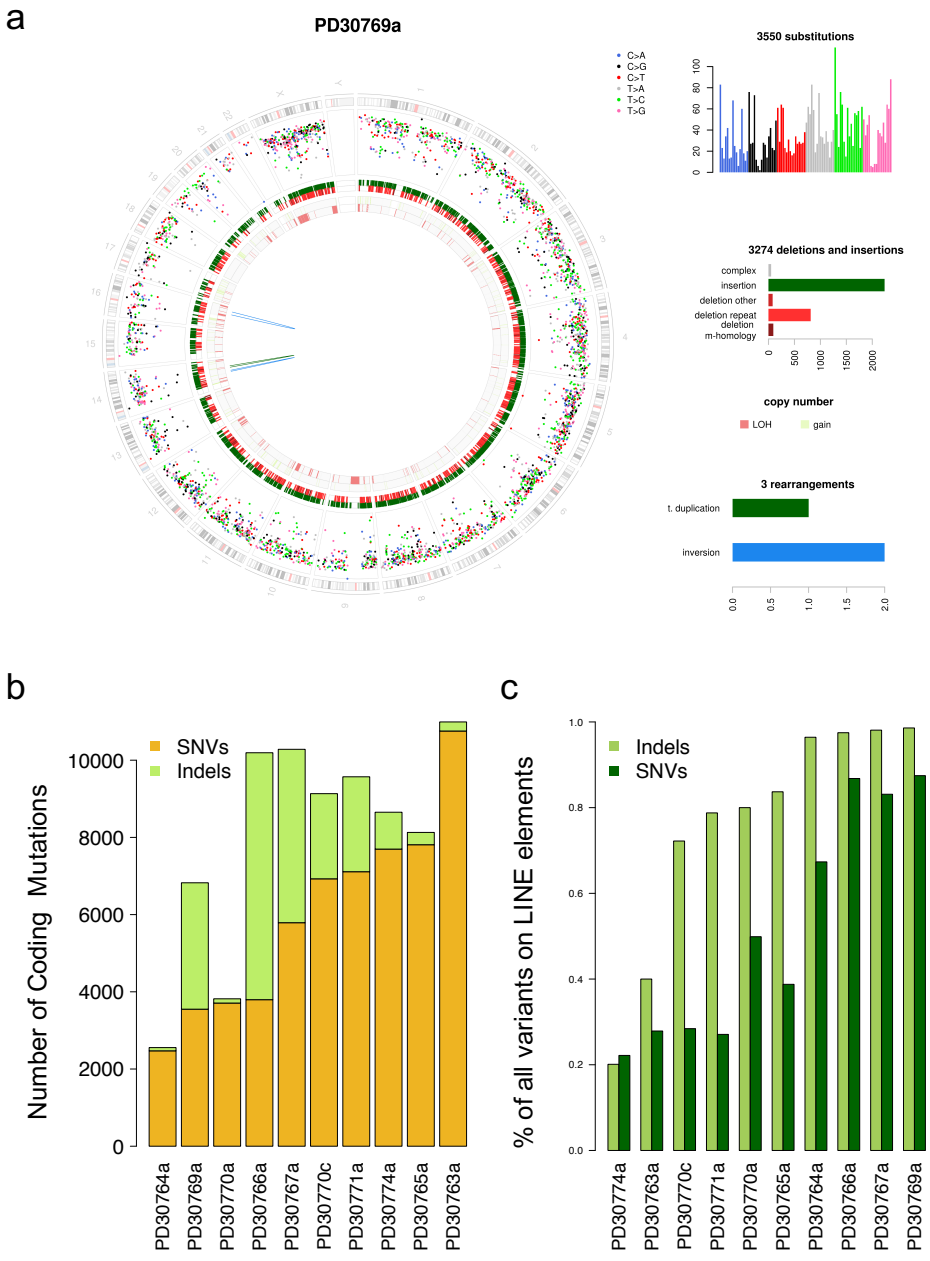
Supplementary Data

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.

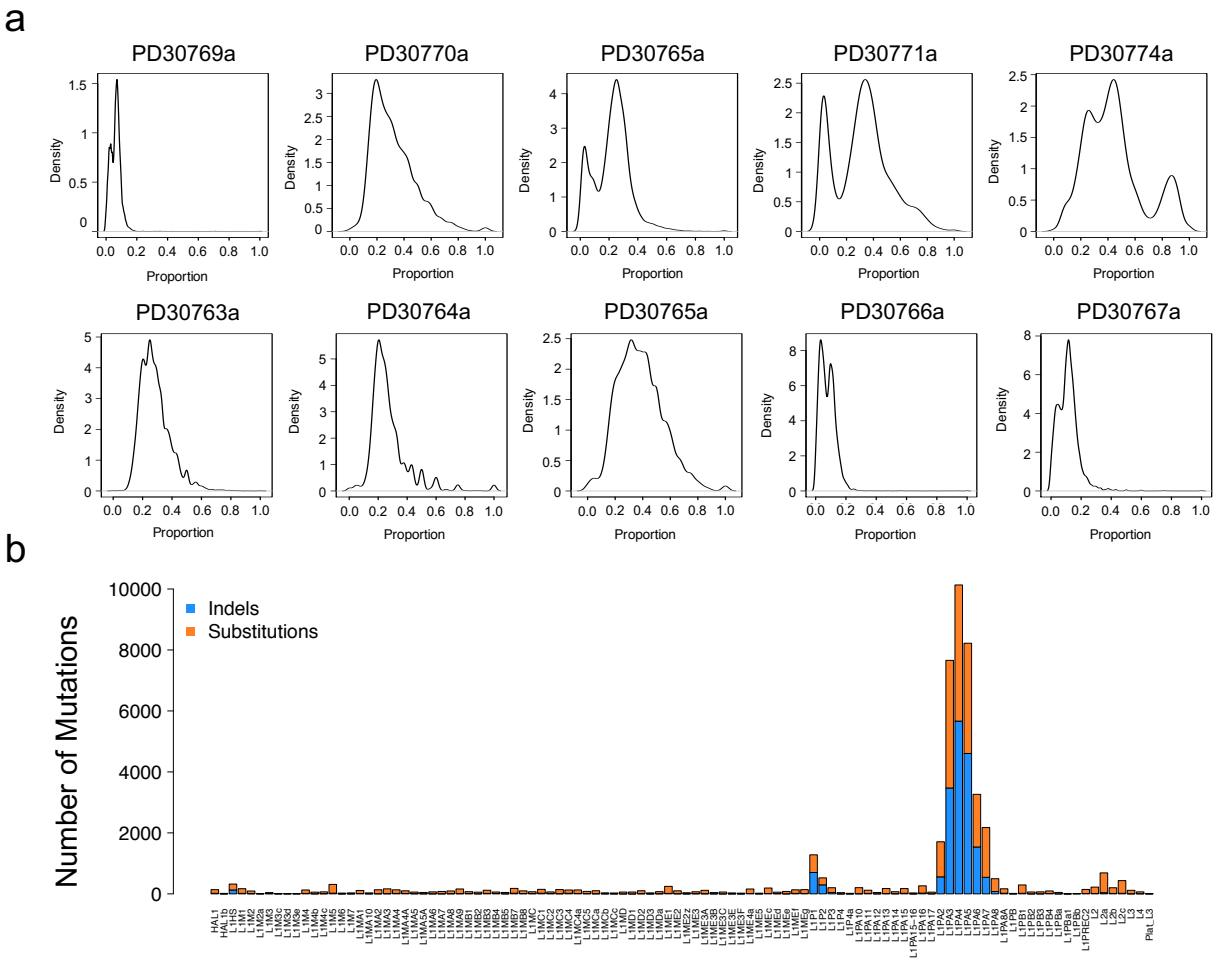
Initial quality control of FFPE-WGS data

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**).

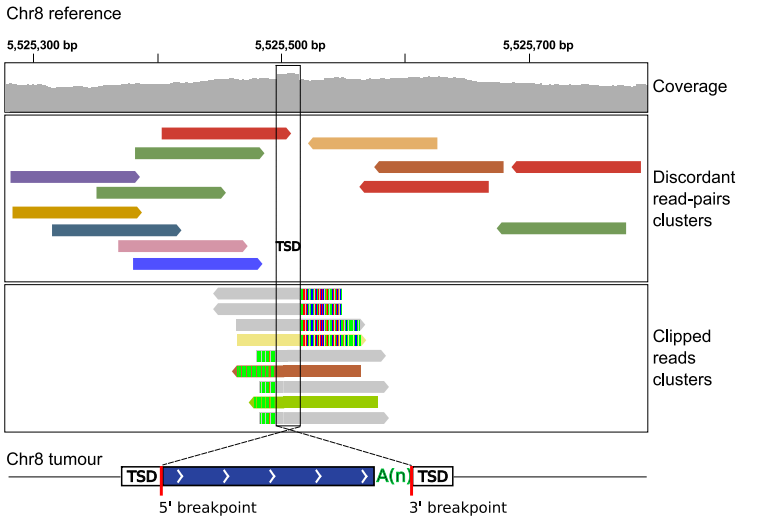
Supplementary Data Table 1. Summary of retrotransposons extracted by TraFic.

Tumor ID	nbTotal	nbSoloL1	nbL1TD	nbAlu	nbSVA	nbERVK	nbPSD	nbL1DEL	nbL1DUP
PD30763a	0	0	0	0	0	0	0	0	0
PD30764a	0	0	0	0	0	0	0	0	0
PD30765a	3	3	0	0	0	0	0	0	0
PD30766a	26	25	1	0	0	0	0	0	0
PD30767a	25	25	0	0	0	0	0	0	0
PD30769a	12	12	0	0	0	0	0	0	0
PD30770a	0	0	0	0	0	0	0	0	0
PD30770c	7	7	0	0	0	0	0	0	0
PD30771a	13	13	0	0	0	0	0	0	0
PD30774a	0	0	0	0	0	0	0	0	0

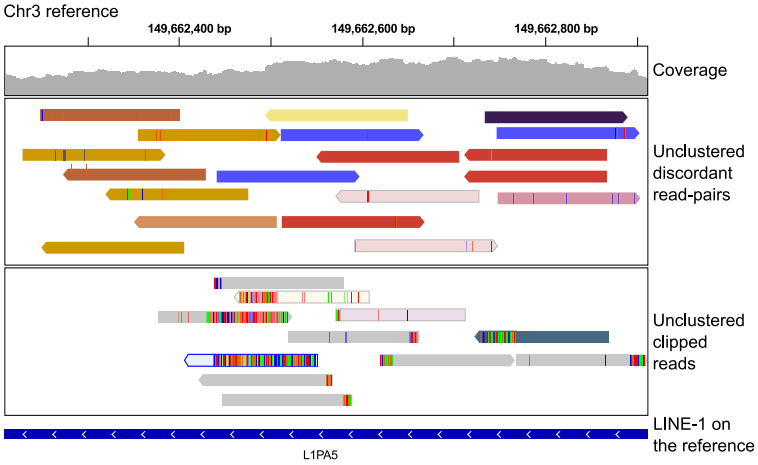
The vast majority of L1 events (~91%, 78/86) were consistently located over repeats of the L1PA subfamily. Manual curation revealed how all of these indels and retrotranspositions were compatible with sequencing and/or alignment artifacts rather than genuine events (**Supplementary Data Figure 3**).

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.

a


Chr8 reference
5,525,300 bp
5,525,500 bp
5,525,700 bp
Coverage
Discordant read-pairs clusters
TSD
Clipped reads clusters
Chr8 tumour
TSD
A(n)
TSD
5' breakpoint
3' breakpoint

b


Chr3 reference
149,662,400 bp
149,662,600 bp
149,662,800 bp
Coverage
Unclustered discordant read-pairs
Unclustered clipped reads
LINE-1 on the reference
L1PA5

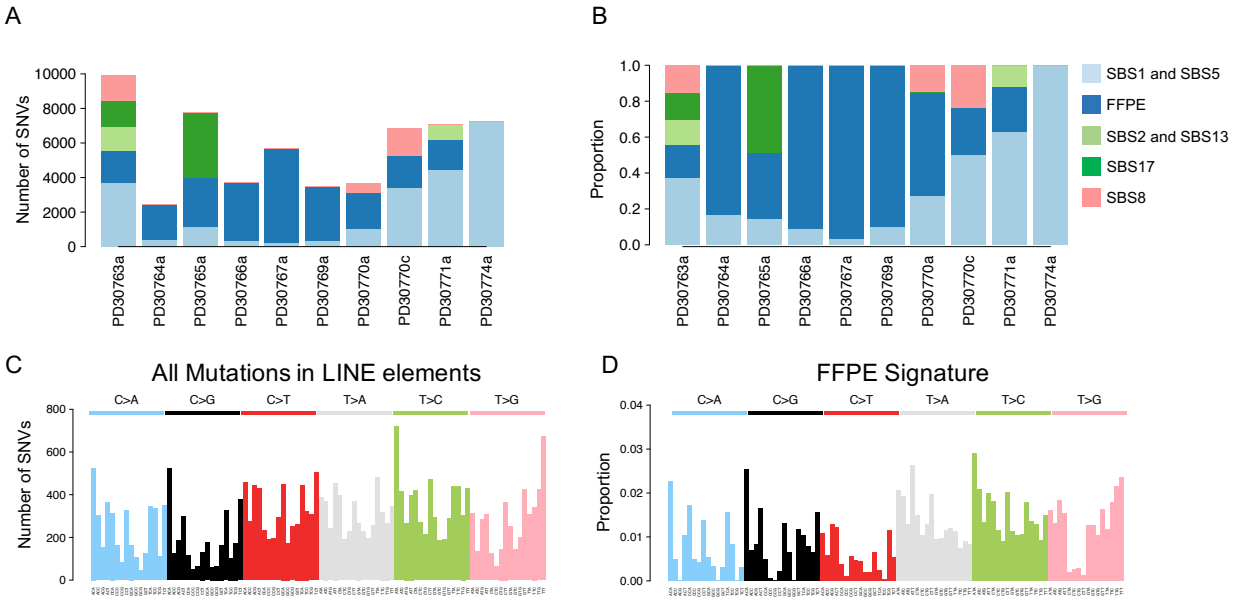
Supplementary Data Figure 3. The hallmarks of artefactual somatic LINE-1 (L1) insertion calls in FFPE treated samples**.** A) Illustrative example of a canonical L1 integration though target-primed reverse transcription (TPRT). Two well defined clusters composed by discordant reads in opposite orientations 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.

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**

4A-B), 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

1. PCAWG Mutational Signatures Working Group, PCAWG Consortium, Alexandrov LB, et al. The repertoire of mutational signatures in human cancer. Nature 2020;578(7793):94–101.

2. Bolli N, Maura F, Minvielle S, et al. Genomic patterns of progression in smoldering multiple myeloma. Nature Communications 2018;9(1):3363.

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