Published Ahead of Print on September 5, 2019, as doi:10.3324/haematol.2019.226647.

haematologica
Journal of The Ferrata Storti Foundation

Defining the signatures of peripheral T-cell lymphoma, with a targeted 20-markers gene expression profiling assay (RT-MLPA)

by Fanny Drieux, Philippe Ruminy, Ahmad Abdel-Sater, François Lemonnier, Pierre-julien Viailly, Virginie Fataccioli, Vinciane Machand, Bettina Bisig, Audrey Letourneau, Marie Parrens, Céline Bossard, Julie Bruneau, Pamela Dobay, Liana Veresezan, Aurélie Dupuy, Anaïs Pujals, Cyrielle Robe, Nouhoum Sako, Christiane Copie-Bergman, Marie-Helene Delfau-Larue, Jean-Michel Picquenot, Hervé Tilly, Richard Delarue, Fabrice Jardin, Laurence de Leval, and Philippe Gaulard. Collaborative Groups: TENOMIC (Philippe Gaulard)

Haematologica 2019 [Epub ahead of print]

Citation: Fanny Drieux, Philippe Ruminy, Ahmad Abdel-Sater, François Lemonnier, Pierre-julien Viailly, Virginie Fataccioli, Vinciane Machand, Bettina Bisig, Audrey Letourneau, Marie Parrens, Céline Bossard, Julie Bruneau, Pamela Dobay, Liana Veresezan, Aurélie Dupuy, Anaïs Pujals, Cyrielle Robe, Nouhoum Sako, Christiane Copie-Bergman, Marie-Helene Delfau-Larue, Jean-Michel Picquenot, Hervé Tilly, Richard Delarue, Fabrice Jardin, Laurence de Leval, and Philippe Gaulard. Collaborative Groups: TENOMIC (Philippe Gaulard). Defining the signatures of peripheral T-cell lymphoma, with a targeted 20-markers gene expression profiling assay (RT-MLPA). Haematologica. 2019; 104:xxx doi:10.3324/haematol.2019.226647

Defining the signatures of peripheral T-cell lymphoma, with a targeted 20-markers gene expression profiling assay (RT-MLPA)

Running title: PTCL classification using a RT-MLPA assay

Fanny Drieux[1-3*], Philippe Ruminy[1*], Ahmad Abdel-Sater[1], François Lemonnier[3,4], Pierre-Julien Viailly[1], Virginie Fataccioli[3], Vinciane Marchand[1], Bettina Bisig[5], Audrey Letourneau[5], Marie Parrens[6], Céline Bossard[7], Julie Bruneau[8], Pamela Dobay[5], Liana Veresezan[1,2], Aurélie Dupuy[3], Anaïs Pujals[3,9], Cyrielle Robe[3], Nouhoum Sako[3], Christiane Copie-Bergman[3,9], Marie-Hélène Delfau-Larue[3,10], Jean-Michel Picquenot[1,2], Hervé Tilly[1], Richard Delarue[11], Fabrice Jardin[1*], Laurence de Leval[5*] and Philippe Gaulard[3,9*]

* equal contribution

[1]INSERM U1245, Centre Henri Becquerel, Rouen, France

[2]Service d'Anatomie et Cytologie Pathologiques, Centre Henri Becquerel, Rouen, France

[3]INSERM U955 and Université Paris-Est, Créteil, France

[4]Unité Hémopathies lymphoïdes, Groupe Hospitalier Henri Mondor, AP-HP, Créteil, France

[5]Institut de Pathologie, Centre Hospitalier Universitaire Vaudois (CHUV), Lausanne, Switzerland

[6]Service d'Anatomie et Cytologie Pathologiques, Hôpital Haut-Lévêque, CHU de Bordeaux, France

[7]Service d'Anatomie et Cytologie Pathologiques, CHU de Nantes, France

[8]Service d'Anatomie et Cytologie Pathologiques, Hôpital Universitaire Necker – Enfants Malades, Assistance Publique – Hôpitaux de Paris (APHP), Paris, France

[9]Département de Pathologie, Groupe Hospitalier Henri Mondor, AP-HP, Créteil, France

[10]Département d'Hématologie et Immunologie Biologique, Groupe Hospitalier Henri Mondor, AP-HP, Créteil, France

[11]Service Hématologie Adultes, Hôpital Universitaire Necker – Enfants Malades, Assistance Publique – Hôpitaux de Paris (APHP), Paris, France

Corresponding authors :
Philippe Gaulard
Département de Pathologie
Hôpital Universitaire Henri Mondor
94010 Créteil, France

E-mail : philippe.gaulard@aphp.fr
Phone : +33149812743/2728
Fax : +33149812733

and

Philippe Ruminy
INSERM U1245, Centre Henri Becquerel
Rue d'Amiens
76038 ROUEN Cedex1
E-mail : philippe.ruminy@chb.unicancer.fr
Phone : +33232882738

Abstract = 250 words
Text word count: 3830 words
5 Figures, 3 Tables
Supplemental: 1 file (Material, 4 Tables, 7 Figures)
Reference count: 30

Abstract

Peripheral T-cell lymphoma comprises a heterogeneous group of mature non-Hodgkin lymphomas. Their diagnosis is challenging, with up to 30% of cases remaining unclassifiable and referred to as “not otherwise specified”. We developed a reverse transcriptase-multiplex ligation-dependent probe amplification gene expression profiling assay to differentiate the main T-cell lymphoma entities and to study the heterogeneity of the “not specified” category.

The test evaluates the expression of 20 genes, including 17 markers relevant to T-cell immunology and lymphoma biopathology, one EBV-related transcript, and variants of *RHOA (*G17V) and *IDH2 (*R172K/T).

By unsupervised hierarchical clustering, our assay accurately identified 21/21 ALK-positive anaplastic large cell lymphomas, 16/16 extranodal NK/T-cell lymphomas, 6/6 hepatosplenic T-cell lymphomas, and 13/13 adult T-cell leukemia/lymphomas. ALK-negative anaplastic lymphomas (n=34) segregated into one cytotoxic cluster (n=10) and one non-cytotoxic cluster expressing Th2 markers (n=24) and enriched in *DUSP22*-rearranged cases. The 63 T_{FH}-derived lymphomas divided in two subgroups according to a predominant T_{FH} (n=50) or an enrichment in Th2 (n=13) signatures. We next developed a support vector machine predictor which attributed a molecular class to 27/77 not specified T-cell lymphomas: 17 T_{FH}, 5 cytotoxic ALK-negative anaplastic, and 5 NK/T-cell lymphomas. Among the remaining cases, we identified two cell-of-origin subgroups corresponding to cytotoxic/Th1 (n=19) and Th2 (n=24) signatures. A reproducibility test on 40 cases yielded a 90% concordance between 3 independent laboratories.

This study demonstrates the applicability of a simple gene expression assay for the classification of peripheral T-cell lymphomas. Its applicability to routinely-fixed samples makes it an attractive adjunct in diagnostic practice.

Introduction

Peripheral T-cell lymphomas (PTCLs) are a diverse group of neoplasms, representing 10-15% of all lymphomas worldwide, with large geographic variation. According to the 2017 revision of the World Health Organization (WHO) classification of lymphoid neoplasms, PTCLs comprise up to 30 entities derived from various subsets of mature T or NK cells (1). The heterogeneity and rarity of these tumors, combined with their complex immunophenotypic profile and partially overlapping features across different entities, make their diagnosis particularly challenging. In addition, there is a high variability in the diagnostic workup among pathologists, which may account for relatively poor reproducibility of the diagnoses (2–4). Although most cases can be ascribed to specific disease entities, approximately one third of PTCLs not fulfilling the criteria for other entities remain unclassifiable and are categorized "by default" as PTCL-NOS.

The classification of PTCLs has undergone major changes over the past years, due to the incorporation of many new information on their genetic background and taking into account the notion that PTCL arise from discrete subsets of normal T cells. In recent years, the description of the signature and mutational landscape of PTCLs has generated novel molecular biomarkers to refine the diagnostic criteria for some entities. Notably, the expression of T_{FH} markers and the presence of genetic lesions associated with angio-immunoblastic T-cell lymphoma (AITL) (such as *RHOA*, *TET2*, *DNMT3A*, and *IDH2* mutations), found in a significant proportion of PTCL-NOS (5–10), led to the reclassification of these as "nodal PTCL with a T_{FH} phenotype" (T_{FH}-PTCL) in the revised WHO classification (1). Among anaplastic large cell lymphoma (ALCL), the identification of recurrent rearrangements of the *ALK* gene led to individualize ALK-positive ALCL as a definitive entity (ALCL ALK+), and to reconsider ALCL without ALK rearrangement as a distinct but genetically heterogeneous group comprising subtypes characterized by alterations of the *DUSP22/IRF4* or *TP63* genes with distinct clinical, pathological and biological features (11). Among the remaining PTCL-NOS category, two molecular subgroups defined by the expression of the TBX21 and GATA3 transcription factors have been proposed (12,13), with a worse prognosis suggested for GATA3-positive cases (13–16). In daily diagnostic practice, however, high-throughput technologies are difficult to integrate. Moreover, the immunohistochemical surrogates are not fully validated and require an increasingly large panel of antibodies, and their evaluation

may present some difficulties or limitations (3,17).

Here, we designed a simple targeted mRNA expression profiling assay based on reverse transcriptase-multiplex ligation-dependent probe amplification (RT-MLPA), using a panel of molecular markers relevant to the characterization of PTCL. We first assessed the accuracy of this assay in the classification of PTCL entities other than PTCL-NOS, and then used the assay to study the heterogeneity of PTCL-NOS. Our findings support this RT-MLPA assay as a robust and useful tool, suitable for the routine classification of PTCLs and therefore an optimal clinical management of PTCL patients.

Methods

Patients and tumor samples

A series of 270 lymphoma samples were selected within the framework of the multicentric T-cell lymphoma consortium (TENOMIC), of the Lymphoma Study Association (LYSA). All cases had been reviewed by at least two expert hematopathologists, according to the criteria of the recently updated WHO classification (1). The series was enriched in nodal T_{FH}-PTCL (T_{FH}-PTCL) defined by the expression of at least 2 T_{FH} markers among CD10, BCL6, CXCL13, PD1, ICOS and in PTCL-NOS defined as a diagnosis of exclusion of any well-defined entity. The design of the study is summarized in Figure 1. Briefly a classification cohort (n=230) was used to train a SVM classifier and a diagnostic cohort (n=40) was used to evaluate its inter-laboratory reproducibility on FFPE samples. The study was approved by the local ethics committee (CPP Ile de France IX 08-009).

RT-MLPA assay gene expression profiling

RNA extracted from frozen and/or FFPE tumor samples was applied to RT-MLPA (for details, see reference 18 and Supplemental Methods). Briefly, this targeted multiplex assay consists of the hybridization and ligation of specific probes on cDNA, followed by PCR amplification. We designed 41 probes (Eurofins MWG Operon, Ebersberg, Germany) targeting 20 genes, selected for their relevance to PTCL classification (Table 1). RTMLPA results were compared to Affymetrix HG-U133-plus-2.0 gene expression data in 72 previously reported cases(18,19).

Bioinformatic analysis

A web interface was developed for the complete analysis of the RT-MLPA results (https://bioinfo.calym.org/RTMLPA). A support vector machine (SVM) was developed to classify PTCL samples: two-thirds of the 184 PTCLs of the classification cohort which clustered in defined molecular branches according to the clustering (Figure 4), were randomly selected to train the classifier, which was validated in the remaining third cases. A bootstrap resampling process was used to build 100 independent training and validation series. A definitive SVM predictor was thus trained using the 184 cases. This supervised learning model assigns a class to every PTCL sample. Therefore, we integrated the distance to the centroid of the predicted class for each sample to avoid classifying distant samples into the same group. The analytical process is detailed in Supplemental Methods.

Histopathology and molecular validation

RT-MLPA signatures were correlated to immunochemical data, including expression of GATA3 and TBX21. The cut off for positive immunohistochemical staining was 10% of presumed neoplastic cells (detailed in supplemental methods). FISH for *DUSP22/IRF4* rearrangement was performed in 20 ALCLs. Mutations were validated using PCR allele-specific and/or targeted deep sequencing (20,21). Technical details are presented in Supplemental Methods.

Data analysis

Affymetrix and RT-MLPA gene expression values were correlated using Spearman's correlation test. Correlations between immunohistochemical results and RT-MLPA gene expression values were evaluated using Wilcoxon's rank-sum test. Unsupervised hierarchical analysis was performed using the Ward method.

Overall and progression free survival analysis were performed using Kaplan Meier method and the log-rank test. The Mann-Whitney test was used to analyze continuous data and the Fisher exact test to analyze categorical data.

Results

Design and validation of the RT-MLPA assay

The design of the study is presented in Figure S1. The gene set and sequences of the RT-MLPA probes are shown in Tables 1 and Table S1, respectively. The panel was designed to include several genes encoding immunohistochemical or genetic markers routinely used for the diagnosis of PTCL and genes of interest selected from previous transcriptomic and genomic studies(9,10,12,13,18). It includes genes related to the major CD4 and CD8 T-cell subsets, genes defining the main subsets of Th cells [T_{FH} (CXCL13, CXCR5, ICOS, BCL6), Th1 (TBX21, IFN□), Th2 (GATA3, CCR4), and Treg (FOXP3)], as well as genes encoding cytotoxic molecules (PRF, GZMB). CD30 and ALK were chosen to identify ALCL and CD56 and EBER1 (Epstein-Barr virus encoding small RNAs) were selected to identify HSTL and NKTCL. We also included the *RHOA*G17V and *IDH2*R172K/T variants, as the most prevalent hotspot mutations of T_{FH}-derived PTCL.

We obtained RT-MLPA profiles for all 230 PTCLs of the classification cohort. Representative RT-MLPA profiles for each entity are shown in Figure S2. T_{FH}.PTCL profiles were characterized by the expression of a combination of T_{FH} genes (*CXCL13, CXCR5, ICOS,* and *BCL6*), together with frequent *RHOA* and/or *IDH2* variants. ATLLs expressed Th2 markers (*GATA3* and *CCR4*) and *ICOS*, with variable levels of *FOXP3*. NKTCL showed high expression of EBER1 and *GZMB*, as well as Th1 markers (*TBX21* and *IFNγ*). HSTL expressed *CD56, GATA3, TBX21*, and *BCL6*. ALK-positive ALCL expressed *ALK*, *CD30*, *PRF*, and *GZMB*. ALK-negative ALCLs comprised two distinct profiles, with or without expression of *PRF* and *GZMB*. The non-cytotoxic ALCLs showed high expression of *CD30* and Th2 markers (*GATA3* and *CCR4*) but not *PRF* or *GZMB*. Unexpectedly, RT-MLPA identified *ALK* expression in a case of ALCL initially considered ALK-negative (based on negative immunostaining with the ALK1 clone), leading to reclassification to ALK-positive ALCL. This was further confirmed by IHC using an alternative antibody (D5F3 clone) (Figure S3).

Paired RT-MLPA profiles and Affymetrix gene expression data available in 72 cases (23 AITL and 49 PTCL-NOS) were compared(18,19). There were significant correlations for each evaluable gene (*TNFRSF8/CD30, PRF, GZMB, GATA3, CXCL13, ICOS, CD8, BCL6, CD4, FOXP3, CCR4, CXCR5*, and *TBX21*) (Figure S4). RT-MLPA and immunohistochemical data scores also showed significant correlations for the 12 evaluable markers (Figure S5). There

was also a good correlation with the EBER *in situ* hybridization results, showing the capacity of the assay to correctly detect EBV infection. RT-MLPA profiles performed in duplicates in 20 PTCLs on RNAs extracted from both frozen and FFPE samples, showed a strong correlation (rho>0.7, Spearman correlation test, Figure S6).

RT-MLPA identified 33/33 *RHOA*G17V and 9/10 *IDH2* R172K/T mutations, detected by either AS-qPCR and/or NGS studies. The only RT-MLPA failure corresponded to an AITL with an *IDH2R172K* mutant with a 2.8% allele frequency, which was only detected by NGS (see details in Table S2).

Unsupervised analysis highlights heterogeneity among ALK-negative ALCL and TFH-PTCL

Given the expected heterogeneity of PTCL-NOS, we first restricted our analyses to specified PTCL entities (not taking into account PTCL-NOS). Unsupervised hierarchical clustering performed on 153 such cases (30 AITLs, 33 T_{FH}-PTCL, 16 NKTCLs, 13 ATLLs, 6 HSTLs, 21 ALK-positive ALCLs and 34 ALK-negative ALCLs) identified 2 main branches separating cytotoxic and non-cytotoxic entities, as shown in Figure 1.

The cytotoxic branch is divided into 2 clusters, one very homogeneous cluster (C5) comprising the 16 NKTCLs, and a second cluster (C6) composed of 31 cytotoxic ALCLs (21 ALK-positive and 10 ALK-negative). The other branch is divided into four clusters (C1-4). The C1 cluster contained the 6 HSTL. The 63 PTCLs with a T_{FH} phenotype (AITL and T_{FH}-PTCL) distributed along the two clusters C2 and C3. The C2 cluster comprised a major group of AITL/T_{FH}-PTCL characterized by a T_{FH} signature (C2, n=50). In addition to 12 ATLLs, 13 T_{FH}-PTCL and one AITL in the C3 cluster showed an enrichment in T_{FH} and Th2 markers. Interestingly, *RHOA* mutations were identified in 26/50 (52%) and 2/13 (15%) of the C2 and C3 clusters respectively (p=0.027). By immunohistochemistry, these T_{FH}-PTCL in the C3 cluster showed expression of GATA3 (in more than 50% of tumor cells) in 9/12 (75%) contributive cases (Figure 2). The C4 cluster contained 24 ALK-negative non-cytotoxic ALCLs with Th2 signature, with 8/16 contributive cases showing *DUSP22* rearrangement by FISH.

PTCL-NOS distribute among distinct clusters using unsupervised clustering

When applied to all 230 PTCL samples (including 77 PTCL-NOS), unsupervised clustering showed that the majority of PTCL-NOS (n = 48/77, 62.3%) clustered within four of the six previous clusters as they showed gene signatures in common with molecular T_{FH}-PTCL (C2, n

= 6), T_{FH}/Th2 PTCLs (C3, n = 19), NKTCL (C5, n = 5), and cytotoxic ALCL (C6, n = 18) (Figure 3). Despite a variable expression of CD30 by immunohistochemistry, 18 PTCL-NOS distributed within the cluster of cytotoxic ALCL based on the expression of cytotoxic markers plus *TBX21* and *IFN*, consistent with a possible Th1 origin. Accordingly, 8/12 of these cases tested for TBX21 by immunohistochemistry were positive. This molecular subgroup is further referred as "cytotoxic/Th1 PTCL" according to its signature. In addition, 19 other PTCL-NOS cases, all with negative HTLV-1 serologies, clustered with ATLL, based on an enrichment in Th2 molecules GATA3 and CCR4 and are referred to as molecular "Th2 PTCL". Finally, 29 PTCL-NOS did not cluster within any of the defined branches, and segregated with 19 other cases (4 AITLs, 12 T_{FH}-PTCL, 1 ATLL).

A SVM classifier accurately classifies specified PTCLs and identifies subgroups within PTCL-NOS

We next built a support vector machine (SVM) model to assign each case to a class based on the RT-MLPA data (Figure 4A). One hundred and eighty-four PTCLs corresponding to the molecular groups defined according to the latter clustering (Figure 3) were used for the construction of the model and to define the molecular classes: 45 T_{FH}-PTCL/AITLs, 21 NKTCLs, 42 Th2 PTCL-NOS/ATLL, 50 cytotoxic-ALCLs, 20 non-cytotoxic ALCLs, and the 6 HSTLs. The 46 PTCLs which did not cluster within these 6 defined branches were not considered to train the classifier. The SVM algorithm accurately assigned 140/153 specified PTCLs to the correct categories: 16/16 NKTCL, 13/13 ATLL, 6/6 HSTL, 31/31 cytotoxic ALCL, 24/24 non-cytotoxic ALK-negative ALCL cases and 50/63 AITL/T_{FH}-PTCL. To note, 11 T_{FH}-PTCL without *RHOA* mutation were classified as molecular Th2 PTCL, 1 as molecular cytotoxic/Th1 and one AITL was distant from the barycenter of the T_{FH}/AITL class and not classified. Of the 77 remaining PTCL-NOS, 70 (91%) were classified by the SVM as T_{FH}/AITL (C2, n=17), cytotoxic/Th1 PTCL (C6, n=19), ALK-negative ALCL (C6, n=5), NKTCL (C5, n=5) while 24 cases were molecularly designated Th2 PTCL (C3). Finally, 7 cases, which were distant to the barycenter of their predicted SVM class (2 Th2, 3 T_{FH}/AITL, 2 cytotoxic/Th1) could not be attributed a molecular class by the SVM. Figure 4B illustrates the subgroups of PTCL-NOS as evidenced in the principal component analysis (PCA). A correlation of the SVM class with the histopathological data of the 77 PTCL-NOS is presented in Figure S7.

Clinico-pathological correlations of the SVM classes

Survival data were available for 88.7% (204/230) of the patients. The median duration of follow-up was 122 months (interquartile range IQR = 80.5-173). Among the 132/175 patients with follow-up data available who received anthracyclin-based chemotherapy, the median overall survival (OS) and progression free survival (PFS) were 15 months (IQR = 6-51.5) and nine months (IQR = 3-36), respectively. The outcome of PTCL patients was poor (5y OS = 27%), except for those with ALK-positive ALCL (5y OS = 70%) as shown in Figure 6A.

Considering the two clusters observed among PTCL with a T_{FH} phenotype (Figure 1), we failed to demonstrate any significant clinical difference (see Table 3). Interestingly, *RHOAG17V* mutations were found in 52% and 15.4% respectively among the C2 (T_{FH}) and C3 (Th2) clusters (p=0.02). *TET2* mutations were observed in 60.6% and 14.3% of the C2 and C3 clusters (p=0.039). *DNMT3A* and *IDH2* mutations were found respectively in 25% and 20% of the C2 cluster but none within the C3 (p=ns).

Among ALCL, non-cytotoxic ALK-negative ALCLs were characterized by the expression of Th2 mRNAs (*GATA3* and *CCR4*) and GATA3 expression by immunohistochemistry in 11/12 informative cases. As shown in figure 5A, non-cytotoxic and cytotoxic ALK-negative ALCL disclosed similar PFS and OS. However, despite similar main clinical characteristics (Table S4), patients with a *DUSP22*-rearrangement (n=8) tended to have a favorable outcome (5y OS = 62.5%) close to that of ALK-positive ALCL patients (70% five-year OS), compared to the very poor prognosis of patients without *DUSP22* rearrangement (n=8) (5y OS = 12.5%, p=0.07) (Figure 5B). No *TP63* rearrangement was detected in this series.

Finally, within the sized limit of the current series, among PTCL-NOS, there was no significant difference in patient's outcome of with respect to their "Th1/cytotoxic" or "Th2" molecular signatures (Figure 5C), or immunohistochemical profiles (Figure 5D).

Reproducibility of the RT-MLPA assay among three centers in routine practice

We evaluated the reproducibility of the RT-MLPA assay in the diagnostic setting by testing 40 FFPE PTCL samples in three independent centers. A concordance in the diagnostic class

proposed by the classifier was observed between the three centers for 36 (90%) samples (Table 2), with a strong correlation between the RT-MLPA values for each gene (Table S3). Among the 36 concordant samples, the SVM class was in accordance with the pathological diagnosis for 32 cases (89%). The four discrepancies consisted of two tumor-cell rich AITL assigned to the Th2/ATLL-like group, as previously observed in the classification cohort, one AITL with a prominent cytotoxic T cell environment assigned to the cytotoxic/Th1-like group, and one ATLL with a double CD4/CD8 phenotype that was not classified.

Discussion

The classification of PTCLs is often challenging and poorly reproducible, with a recent study showing a 31.5% rate of discrepancy between the referral and expert diagnoses (4), likely due to the complexity of these rare neoplasms and the wide range of practices among pathologists and laboratories (3). Hsi et al. pointed the limited number of immunohistochemical markers assessed in the routine practice, especially the T_{FH} markers, resulting in a poor characterization of PTCLs and a high frequency of PTCL-NOS diagnosed in the US (22). The ligation-dependent RT-PCR assay has been reported to be a simple and robust assay applicable to FFPE samples, that can be used to classify DLBCL into GCB or ABC subtypes (23,24). Here, we expanded that RT-MLPA can contribute in the routine setting to classify the main specified categories of non-cutaneous PTCL. This assay, which can be performed in a limited time (48 hours), requires only commonly used equipment, including a thermocycler and genetic fragment size analyzer. The profiles are publicly accessible and easy to interpret through a dedicated website. Finally, the assay is cost-effective (approximately five dollars per sample) (23). RT-MLPA is a useful tool in combination with pathological evaluation to characterize PTCLs, especially when immunohistochemistry is flawed or incomplete. In addition to evaluate the expression of Th-differentiation antigens and markers suitable for immunohistochemistry, the current RT-MLPA assay also provides genetic information, such as *RHOA* and *IDH2* mutations, which are highly relevant for the diagnosis of PTCL of T_{FH} origin (19), even though the *RHOA* G17V mutation has also been reported in a small minority of ATLL (25), as observed in one case of our series. The accurate classification of specified PTCLs other than NOS entities in most cases corroborates the relevance of the designed gene panel. Altogether, though some markers in our RT-MLPA

assay might not be useful in every PTCL case, this "one fits all" assay evaluates diagnostic markers covering the different PTCL entities in a systematic and cost-effectiveness way.

In addition, the RT-MLPA assay highlighted the heterogeneity in the gene signature of ALK-negative ALCL and PTCL of T_{FH} origin as defined in the updated WHO classification. We observed that a significant proportion of ALK-negative ALCLs display a distinct signature, with expression of CD30 and Th2 genes, but no cytotoxic markers. These cases showed a dense and cohesive pattern of CD30-positive anaplastic large cells but, in contrast to common ALCL, only few hallmark cells, absence of EMA and a frequently preserved T-cell programme with most T-cell antigens being retained. Genetically, this group appeared heterogeneous but was enriched in cases *DUSP22/IRF4* rearrangement (in 8/16 non-cytotoxic cases, versus in only 1/10 cytotoxic ALK-negative ALCLs p=0.09). Despite absence of significant morphological or immunophenotypic difference between cases with or without *DUSP22* rearrangement, we further expand here that the *DUSP22* status is of clinical relevance with a better survival of *DUSP22*-rearranged ALCL (11,26). In the absence of *TP63* rearrangement (data not shown), further investigations are needed to explore the genetic abnormalities in the *DUSP22*-non rearranged cases.

Within the umbrella category of nodal PTCL of T_{FH} derivation, comprising AITL and a number of nodal PTCL previously classified under PTCL-NOS but expressing 2 or more T_{FH} markers, our study interrogates the biological and clinical significance of two subgroups (1). Indeed, whereas the majority (79%) was attributed to the T_{FH}/AITL class by the molecular classifier, a minority disclosed, in addition to T_{FH} markers, enrichment in Th2 genes. We confirmed a "mixed" T_{FH}/Th2 immunophenotype for nine T_{FH}-PTCLs showing a Th2 signature by RT-MLPA. It has been reported that Th1 and Th2 cells can express T_{FH} markers, and conversely that T_{FH} cells have the capacity to express Th1 or Th2 cytokines (23,24). It has also been suggested that a subset of TFH cells may originate from Th2 lymphocytes in the presence of B cells and that T_{FH} cells can acquire GATA3 expression (25,26). Overall, these data raise the question of Th cell plasticity and the specificity of the criteria required to assess T_{FH}-derived PTCLs. Indeed, the current definition of a T_{FH} phenotype based on two or more T_{FH} markers may have some limitations in certain cases (1) and the integration of genetic markers, made possible by the RT-MLPA assay, may be helpful. In this respect, when comparing the "TFH/AITL" class to the other SVM classes, it appeared that, among the four TFH genes in

our RT-MLPA panel (BCL6, CXCL13, CXCR5, ICOS), ICOS and then BCL6 appeared less specific than CXCL13 and CXCR5 (data not shown). Further studies however are needed to determine whether cases with a "mixed" T_{FH}-Th2 RNA profile or immunophenotype are better considered as T_{FH}-PTCL or Th2 PTCL-NOS with associated T_{FH} markers. The quasi-absence of *RHOA* and *TET2* mutation in the cases with a Th2-like profile may support the latter hypothesis.

Our targeted RT-MLPA panel confirmed two subgroups among PTCL-NOS, based on the expression of *TBX21* and cytotoxic markers or *GATA3* and *CCR4* in 39% and 31% of cases, respectively, expanding previous studies (13,14). The prognostic relevance of these 2 groups is controversial (27). Within the limits of our retrospective study, we failed to demonstrate any significant correlation with outcome between the Th1 and Th2 molecular or phenotypic subgroups. Additionally, all TBX21 cases in our series had a cytotoxic profile by RT-MLPA and immunohistochemistry. Among PTCL-NOS with a cytotoxic phenotype, the RT-MLPA assay highlighted a small group of cases with an EBV signature, now referred to as "nodal T/NK lymphoma EBV-positive" according to the revised WHO classification(1). Whether these cases, confirmed to be EBV-positive in almost all neoplastic cells by EBER ISH, are related to extranodal NK/T lymphoma nasal-type warrants further investigation. Of note, the Th2 signature based on the expression of *GATA3* and *CCR4* in our panel was characteristic, although not specific, of ATLL. The expression of *GATA3* and *CCR4* together with variable expression of *FOXP3* in ATLL is in accordance with a recent study showing that the HBZ transcript induces *CCR4* expression in CD4 T cells by enhancing *GATA3* expression, whereas *FOXP3* expression was inconsistent in ATLL. However, the distinction between PTCL-NOS with a Th2-like signature and ATLL requires an investigation for HTLV1 serology and/or viral integration (28,29).

Finally, our SVM model proposed a class for most cases, with few discrepancies. It may be a useful tool in combination with the pathological evaluation, especially when immunohistochemistry is not conclusive or not available. In this series, 23% of cases diagnosed by default as PTCL-NOS due to incomplete or failure in immunohistochemistry were classified as T_{FH}/AITL by our assay. Unclassified or misclassified cases by RT-MLPA were limited to T_{FH}-PTCL or AITL rich in reactive CD8-positive cytotoxic cells known to be abundant in a proportion of AITL (30), or cytotoxic PTCL with various reactive T_{FH} cells. These

cases illustrate the contribution of the microenvironment to the molecular signature, especially when tumor cell content is low or heterogeneous, a common problem encountered in all gene-expression methods. Therefore, RT-MLPA results should be interpreted in the light of clinical context, as well as biological and histopathological findings. It is to note that our assay does not represent the final answer for PTCL classification, but rather a step requiring an extensive reworking.

Overall, this study demonstrates the applicability of a robust and dedicated RT-MLPA assay which is easily transposable to the diagnostic workflow. Its simplicity of use, applicability to FFPE and frozen samples, integration of genetic features, and cost-effectiveness make it an attractive alternative to high-throughput technologies in the routine practice. Implementation of RT-MLPA in further large studies, especially in the setting of clinical trials, may confirm how this adjunct tool can contribute to a better classification of PTCLs and therefore improve the management of these patients in the era of personalized medicine.

Acknowledgments

The authors would like to thank the LYSA-Pathology and the Plateforme de Ressources Biologiques from Henri Mondor Hospital for its technical assistance and the participants of the Tenomic consortium (see Appendix). This work was supported in part by institutional grants from the Institut National du Cancer (INCA), the Fondation pour la Recherche Médicale (FRM, Equipe Labellisée DEQ20160334875), the Leukemia Lymphoma Society (LLS), the Fondation Force Hémato and the Institut Carnot (CALYM), MEDIC Foundation.

Authorship Contribution:

F. Drieux, P. Ruminy, F. Jardin, L. de Leval and P. Gaulard designed the study. F. Drieux, P. Ruminy, F Jardin developed the RT-MLPA methodologies. A. Abdel-Sater, F. Lemonnier, P-J Viailly, V. Fataccioli, V. Marchand, B. Bisig, A Letourneau, M. Parrens, C. Bossard, J. Bruneau, P. Dobay, L. Veresezan, A. Dupuy, A. Pujals, C. Robe, N. Sako, C. Copie-Bergman, M-H Delfau-Larue, J-M Picquenot, H. Tilly, R. Delarue, F. Jardin acquired the data. A. Abdel-Sater, F. Lemonnier, P-J Viailly, P. Ruminy, A Dupuy, L. de Leval, P. Gaulard analysed and interpreted the data. F. Drieux conducted the study and wrote the manuscript. P. Ruminy, L. de Leval and P. Gaulard supervised the study and wrote the manuscript. All authors have read and agreed with the manuscript's content.

Conflict-of-interest disclosure: the authors have no conflict of interest to disclose.

Appendix

Participants of the Tenomic consortium

A. Martin, Hôpital Avicenne, Bobigny, France; I. Soubeyran, P. Soubeyran, Institut Bergonié, Bordeaux, France; A. Pilon, O.Tournilhac, CHU Estaing, Clermont-Ferrand, France; P. Gaulard, C. Copie-Bergman, M.H. Delfau, J. Moroch, E. Poullot, F. Lemonnier, F. Le Bras, J. Dupuis, C. Haioun, Hôpital Henri Mondor, Créteil, France; T. Petrella, L. Martin, JN. Bastié, O. Casasnovas CHU,Dijon, France; B. Fabre, R. Gressin, CHU, Grenoble, France; L. de Leval,B. Bisig, E. Missiaglia, MP Dobay, A. Cairoli, CHUV, Lausanne, Suisse; C. Bonnet, J. Somja, CHU Sart-Tilman,Liège, Belgique; M.C. Copin, B. Bouchindhomme, F. Morschhauser, CHU, Lille, France; B.Petit, M. Delage, A. Jaccard, Hôpital Dupuytren, Limoges, France; F. Berger, B. Coiffier, A. Traverse-Glehen, L. Genestier, E.Bachy, CHU Sud, Lyon, France; T. Rousset, G. Cartron, V. Szablewski, Hôpital Gui de Chauliac-St Eloi, Montpellier, France; S. Thiebault, B. Drenou, Hôpital E. Muller, Mulhouse, France; K. Montagne, C. Bastien, S.Bologna, CHU de Brabois, Nancy, France; C. Bossard, S. Le Gouill, Hôtel-Dieu, Nantes, France; J. Brière, V. Meignin, C. Gisselbrecht, J. Soulier, Hôpital St Louis, Paris, France; B. Fabiani, A.Aline-Fardin, P. Coppo, Hôpital Saint-Antoine, Paris, France; F. Charlotte, J. Gabarre, Hôpital Pitié-Salpétrière, Paris, France; T. Molina, J. Bruneau, D. Canioni, E. Macintyre, V. Asnafi, D. Sibon, R. Delarue, JP Jaïs, Hôpital Necker, Paris, France; M. Parrens, J.P.Merlio, K. Bouabdallah, Hôpital Haut Lévêque, Bordeaux, France; S. Maugendre-Caulet, P. Tas,F. Llamas-Gutierrez T. Lamy, CHU Pontchaillou, Rennes, France; J.M. Picquenot, E.L Veresezan, F. Drieux, P. Ruminy, F. Jardin, C.Bastard, Centre H Becquerel, Rouen, France; M. Peoch', J. Cornillon, CHU, Saint Etienne, France; L. Lamant, C. Laurent, L. Ysebaert, Hôpital Purpan, Toulouse, France; J.Bosq, P. Dartigues, V. Ribrag, P. Dessen, G Meurice, Institut G Roussy, Villejuif, France; M. Patey, A. Delmer, Hôpital R. Debré, Reims, France; J.F. Emile, K. Jondeau, Hôpital Ambroise Paré, Boulogne, France; M.C. Rousselet, M. Hunault, A. Clavert, CHU, Angers, France ; C. Legendre, S. Castaigne, A.L.Taksin, CH Versailles, Le Chesnay, France; J. Vadrot, B. Joly, A. Devidas, CH Sud Francilien, Corbeil, France; G. Damaj, CHU Caen, France; F Radvanyi, E. Chapeaublanc, Institut Curie, Paris, France; S. Spicuglia, CIML, Marseille, France; C. Thibault, IGBMC, Illkirsch, France; V. Fataccioli, project coordinator, Hôpital Henri Mondor, Créteil, France.

REFERENCES

1. Swerdlow S., Campo E., Harris N.L et al. WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues (IARC WHO Classification of Tumours). Revised 4th Edition. Vol. 2. 2017.

2. Herrera AF, Crosby-Thompson A, Friedberg JW, et al. Comparison of referring and final pathology for patients with T-cell lymphoma in the National Comprehensive Cancer Network: Second-Opinion Pathology Review of TCL. Cancer. 2014;120(13):1993–1999.

3. Hsi ED, Said J, Macon WR, Rodig SJ, et al. Diagnostic Accuracy of a Defined Immunophenotypic and Molecular Genetic Approach for Peripheral T/NK-cell Lymphomas: A North American PTCL Study Group Project. Am J Surg Pathol. 2014;38(6):768–775.

4. Laurent C, Baron M, Amara N, et al. Impact of Expert Pathologic Review of Lymphoma Diagnosis: Study of Patients From the French Lymphopath Network. J Clin Oncol. 2017;35(18):2008–2017.

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

6. Couronné L, Bastard C, Bernard OA. TET2 and DNMT3A mutations in human T-cell lymphoma. N Engl J Med. 2012;366(1):95–96.

7. Cairns RA, Iqbal J, Lemonnier F, et al. IDH2 mutations are frequent in angioimmunoblastic T-cell lymphoma. Blood. 2012;119(8):1901–1903.

8. Manso R, Sánchez-Beato M, Monsalvo S, et al. The RHOA G17V gene mutation occurs frequently in peripheral T-cell lymphoma and is associated with a characteristic molecular signature. Blood. 2014;123(18):2893–2894.

9. Palomero T, Couronné L, Khiabanian H, et al. Recurrent mutations in epigenetic regulators, RHOA and FYN kinase in peripheral T cell lymphomas. Nat Genet. 2014;46(2):166–170.

10. Sakata-Yanagimoto M, Enami T, Yoshida K, et al. Somatic RHOA mutation in angioimmunoblastic T cell lymphoma. Nat Genet. 2014;46(2):171–175.

11. Parrilla Castellar ER, Jaffe ES, Said JW, et al. ALK-negative anaplastic large cell lymphoma is a genetically heterogeneous disease with widely disparate clinical outcomes. Blood. 2014;124(9):1473–1480.

12. Iqbal J, Weisenburger DD, Greiner TC, et al. Molecular signatures to improve diagnosis in peripheral T-cell lymphoma and prognostication in angioimmunoblastic T-cell lymphoma. Blood. 2010;115(5):1026–1036.

13. Iqbal J, Wright G, Wang C, et al. Gene expression signatures delineate biological and prognostic subgroups in peripheral T-cell lymphoma. Blood. 2014 May;123(19):2915-2923.

14. Wang T, Feldman AL, Wada DA, et al. GATA-3 expression identifies a high-risk subset of PTCL, NOS with distinct molecular and clinical features. Blood. 2014;123(19):3007–3015.

15. Zhang W, Wang Z, Luo Y, Zhong D, Luo Y, Zhou D. GATA3 expression correlates with poor prognosis and tumor-associated macrophage infiltration in peripheral T cell lymphoma. Oncotarget. 2016;7(40):65284-65294.

16. Heavican TB, Bouska A, Yu J, et al. Genetic drivers of oncogenic pathways in molecular subgroups of peripheral T-cell lymphoma. Blood. 2019;133(15):1664–1676.

17. de Jong D, Rosenwald A, Chhanabhai M, et al. Immunohistochemical Prognostic Markers in Diffuse Large B-Cell Lymphoma: Validation of Tissue Microarray As a Prerequisite for Broad Clinical Applications—A Study From the Lunenburg Lymphoma Biomarker Consortium. J Clin Oncol. 2007;25(7):805–812.

18. de Leval L, Rickman DS, Thielen C, et al. The gene expression profile of nodal peripheral T-cell lymphoma demonstrates a molecular link between angioimmunoblastic T-cell lymphoma (AITL) and follicular helper T (TFH) cells. Blood. 2007;109(11):4952–4963.

19. Dobay MP, Lemonnier F, Missiaglia E, et al. Integrative clinicopathological and molecular analyses of angioimmunoblastic T-cell lymphoma and other nodal lymphomas of follicular helper T-cell origin. Haematologica. 2017;102(4):e148–151.

20. Alirkilicarslan AL, Dupuy A, Pujals A, et al. Expression of TFH markers and detection of RHOA p. G17V and IDH2 p. R172K/S mutations in cutaneous localizations of angioimmunoblastic T-cell lymphomas. Am J Surg Pathol. 2017;41(12):1581–1592.

21. Dupuy A, Lemonnier F, Fataccioli V, et al. Multiple ways to detect IDH2 mutations in angioimmunoblastic T cell lymphoma: from immunohistochemistry to Next Generation Sequencing. J Mol Diagn. 2018;20(5):677-685.

22. Hsi ED, Horwitz SM, Carson KR, et al. Analysis of Peripheral T-cell Lymphoma Diagnostic Workup in the United States. Clin Lymphoma Myeloma Leuk. 2017;17(4):193-200.

23. Mareschal S, Ruminy P, Bagacean C, et al. Accurate Classification of Germinal Center B-Cell–Like/Activated B-Cell–Like Diffuse Large B-Cell Lymphoma Using a Simple and Rapid Reverse Transcriptase–Multiplex Ligation-Dependent Probe Amplification Assay: A CALYM Study. J Mol Diagn. 2015;17(3):273–283.

24. Bobée V, Ruminy P, Marchand V, et al. Determination of Molecular Subtypes of Diffuse Large B-Cell Lymphoma Using a Reverse TranscriptaseMultiplex Ligation-Dependant Probe Amplification Classifier: a CALYM Study. J Mol Diagn. 2017;19(6):892–904.

25. Nagata Y, Kontani K, Enami T, et al. Variegated RHOA mutations in adult T-cell leukemia/lymphoma. Blood. 2016;127(5):596-604.

26. Pedersen MB, Hamilton-Dutoit SJ, Bendix K, et al. DUSP22 and TP63 rearrangements predict outcome of ALK-negative anaplastic large cell lymphoma: a Danish cohort study. Blood. 2017;130(4):554–557.

27. Maura F, Agnelli L, Leongamornlert D, et al. Integration of transcriptional and mutational data simplifies the stratification of peripheral T-cell lymphoma. Am J Hematol. 2019;94(6):628-634.

28. Matutes E. Adult T-cell leukaemia/lymphoma. J Clin Pathol. 2006;60(12):1373–1377.

29. Licata MJ, Janakiram M, Tan S, et al. Diagnostic challenges of adult T-cell leukemia/lymphoma in North America – a clinical, histological, and immunophenotypic correlation with a workflow proposal. Leuk Lymphoma. 2018;59(5):1188–1194.

30. Gaulard P, de Leval L. The microenvironment in T-cell lymphomas: Emerging themes. Semin Cancer Biol. 2014;24:49–60.

Table 1. Gene panel designed for the RT-MLPA assay.

Family genes and other targets	Genes	Detection method in the routine practice
Main T-cell subsets	CD4	Immunohistochemistry
	CD8	Immunohistochemistry
	TCRα	Not applicable
T_{FH}	CXCL13	Immunohistochemistry
	CXCR5	Not applicable
	BCL6	Immunohistochemistry
	ICOS	Immunohistochemistry
Th1	TBX21	Immunohistochemistry
	IFNγ	Not applicable
Th2	GATA3	Immunohistochemistry
	CCR4	Not applicable
Treg	FOXP3	Immunohistochemistry
NK-associated and cytotoxic	CD56	Immunohistochemistry
	PRF	Immunohistochemistry
	GZB	Immunohistochemistry
Activation	CD30	Immunohistochemistry
Virus	EBER	In situ hybridization
Mutations	RHOAm G17V	AS-PCR*, other sequencing methods
	IDH2m R172K/T	AS-PCR*, other sequencing methods, immunohistochemistry
Other	ALK	Immunohistochemistry, FISH

*AS-PCR= allele-specific PCR

Table 2. Reproducibility of the RT-MLPA assay among 3 laboratories (Center 1, Center 2 and Center3), evaluated on 40 FFPE samples (6 ALK-positive anaplastic large cell lymphomas = ALCL ALK+, 4 ALCL ALK-, 13 angioimmunoblastic T-cell lymphomas = AITL, 9 NK/T cell lymphomas = NKTCL, 6 adult T cell lymphomas = ATLL, and 2 peripheral T-cell lymphomas not otherwise specified = PTCL-NOS).

	Pathology	RT-MLPA Classification			RHOAG17V/ IDH2 R172 status by RTMLPA (m=mutated, wt= wild-type)	Mutations detected by targeted next generation sequencing (NGS)
		Center1	Center2	Center3		
Concordant samples						
Case1	AITL	AITL	AITL	AITL	m/wt	ND
Case2	AITL	AITL	AITL	AITL	m/wt	RHOA, IDH2 (R712S), TET2,
Case3	AITL	AITL	AITL	AITL	wt/wt	ND
Case4	AITL	AITL	AITL	AITL	m/wt	RHOA, IDH2 (R712S), TET2,
Case5	AITL	AITL	AITL	AITL	wt/wt	ND
Case6	AITL	AITL*	AITL*	AITL*	m/wt	RHOA, IDH2 (R712T), TET2
Case7	AITL	AITL*	AITL*	AITL*	wt/wt	ND
Case8	AITL †	TH2	TH2	TH2	wt/wt	IDH2(R172S), TET2, CD28
Case9	AITL†	TH2	TH2	TH2	wt/wt	TET2, FYN
Case10	AITL□	Cytotoxic/Th1	Cytotoxic/Th1	Cytotoxic/Th1	m/wt	RHOA, TET2
Case11	AITL□	Cytotoxic/Th1	Cytotoxic/Th1	AITL	m/wt	ND
Case12	AITL□	AITL*	TH2	AITL	m/wt	ND
Case13	AITL	Cytotoxic/Th1*	Cytotoxic/Th1	Cytotoxic/Th1	wt/wt	ND
Case14	ALCL ALK-	ALCL ALK-	ALCL ALK-	ALCL ALK-	wt/wt	ND
Case15	ALCL ALK-	ALCL ALK-	ALCL ALK-	ALCL ALK-	wt/wt	ND
Case16	ALCL ALK-	ALCL ALK-	ALCL ALK-	ALCL ALK-	wt/wt	ND
Case17	ALCL ALK-	CD30TH2*	CD30TH2*	CD30TH2*	wt/wt	TET2, TP53
Case18	ALCL ALK+	ALCL ALK+	ALCL ALK+	ALCL ALK+	wt/wt	ND
Case19	ALCL ALK+	ALCL ALK+	ALCL ALK+	ALCL ALK+	wt/wt	ND
Case20	ALCL ALK+	ALCL ALK+	ALCL ALK+	ALCL ALK+	wt/wt	ND
Case21	ALCL ALK+	ALCL ALK+	ALCL ALK+	ALCL ALK+	wt/wt	ND
Case22	ALCL ALK+	ALCL ALK+	ALCL ALK+	ALCL ALK+	wt/wt	ND
Case23	ALCL ALK+	ALCL ALK+	ALCL ALK+	ALCL ALK+	wt/wt	ND
Case24	ATLL	TH2	TH2	TH2	wt/wt	ND
Case25	ATLL	TH2	TH2	TH2	wt/wt	ND
Case26	ATLL	TH2	TH2	TH2	wt/wt	ND
Case27	ATLL	TH2	TH2	TH2	wt/wt	ND
Case28	ATLL	unclassified	unclassified	unclassified	wt/wt	ND
Case29	ATLL□	Failure	Failure	TH2	wt/wt	ND
Case30	NKTCL	NKTCL	NKTCL	NKTCL	wt/wt	ND
Case31	NKTCL	NKTCL	NKTCL	NKTCL	wt/wt	ND
Case32	NKTCL	NKTCL	NKTCL	NKTCL	wt/wt	ND
Case33	NKTCL	NKTCL	NKTCL	NKTCL	wt/wt	ND
Case34	NKTCL	NKTCL	NKTCL	NKTCL	wt/wt	ND
Case35	NKTCL	NKTCL	NKTCL	NKTCL	wt/wt	ND
Case36	NKTCL	NKTCL	NKTCL	NKTCL	wt/wt	ND
Case37	NKTCL	NKTCL	NKTCL	NKTCL	wt/wt	ND
Case38	NKTCL	NKTCL	NKTCL	NKTCL	wt/wt	ND
Case39	PTCL-NOS	Cytotoxic/Th1	Cytotoxic/Th1	Cytotoxic/Th1	wt/wt	ND
Case40	PTCL-NOS (EBV+)	NKTCL	NKTCL	NKTCL	wt/wt	ND

† AITL tumor cell rich * distant of the samples of the predicted SVM class□Among the 4 discrepant samples, the SVM resulted in concordance between two centers for two cases, one case showed discordant result between the three centers, and one ATLL sample had no interpretable profile in two centers whereas the other determined a Th2 profile concordant with the diagnosis

Table 3. Clinical, Pathological and Molecular features of the two subgroups of T_{FH}-PTCL

	TFH signature n=50	TFH and TH2 signatures n=13	p
Clinical data			
Age median (range)	64.4 (54-74.6)	67.4 (56.7-74.7)	0.79
IPI>=3	71.8% (28/39)	75% (9/12)	1
PIT>=2	71% (27/38)	63.6% (7/11)	0.72
Extranodal site>=2	50% (22/44)	58.3% (7/12)	0.75
Stade>=3	100% (44/44)	91.7% (11/12)	0.21
PS>=2	40% (16/40)	8.3% (1/12)	0.076
LDH>1N	64% (25/39)	50% (6/12)	0.5
B signs	70.7% (29/41)	45.5% (5/11)	0.16
Hypergammaglobulinemia	37.5% (12/32)	12.5% (1/8)	0.24
Coombs	46% (12/26)	0% (0/2)	0.49
Anemia	61.5% (24/39)	41.6% (5/12)	0.32
Cutaneous rash	23.8% (10/42)	33% (4/12)	0.48
BM	48.8% (21/43)	33.3% (4/12)	0.51
Complete response	50% (20/40)	41.7% (5/12)	0.75
OS median (range)	22 (5.5-77)	30.5 (6-50.5)	0.91
PFS median (range)	10 (3-39)	12 (5.5-38)	0.42
Pathological data			
Clear cells	65.2% (30/46)	36% (4/11)	0.1
B Blasts	90% (44/49)	66% (8/12)	0.07
EBV positivity	70.2% (33/47)	58.3% (7/12)	0.5
EBV extent >5 large blast-cells/high power field	29.5% (13/44)	0% (0/12)	0.049
Molecular data			
TET2 mutation	60.6%% (20/33)	14.3% (1/7)	0.039
DNMT3a mutation	25% (8/32)	0% (0/6)	0.31
RHOA mutation	52% (26/50)	15.4% (2/13)	0.027
IDH2 mutation	20% (10/50)	0% (0/13)	0.1

Figure Legends

Figure 1. Unsupervised hierarchical clustering of PTCL entities other than PTCL-NOS using RT-MLPA (n = 153). The assay was used to classify AITL (n = 30), PTCL-TFH (n=33), ALCL (n = 55), ATLL (n = 13), HSTL (n = 6), and NKTCL (n = 16). Differential gene expression is depicted according to a red (positive) to blue (negative) color scale, and concordance with histopathological diagnosis (Pathology). Two main branches were observed: the left branch divided in 6 HTSL (C1), 50 T_{FH}-PTCL/AITL (C2), 12 ATLL with 13 T_{FH}-PTCL (C3) and 24 ALK-negative ALCLs (C4) and the right branch contained two clusters of 16 NKTCL (C5) and 31 cytotoxic ALCLs (C6).

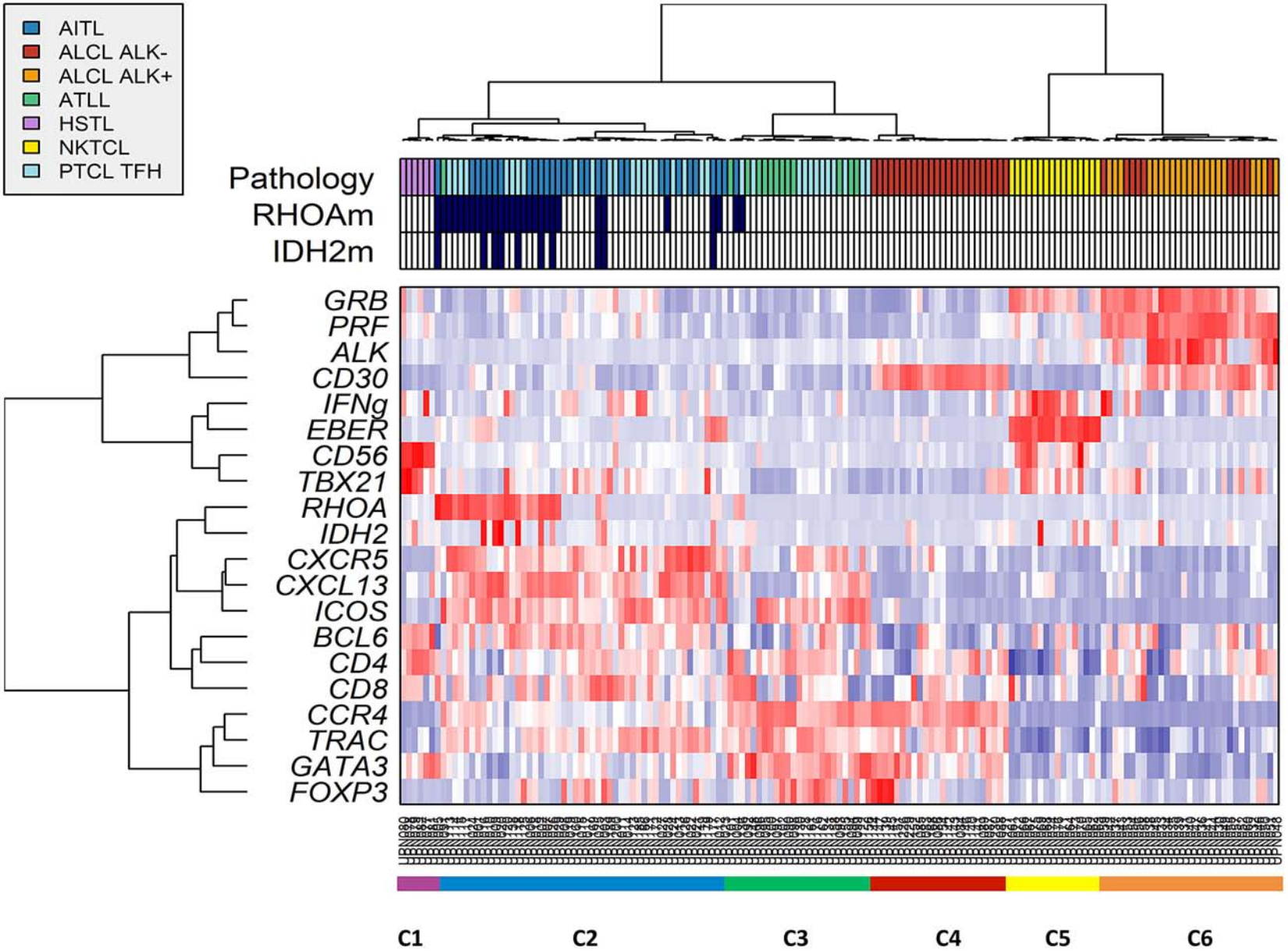
Figure 2. Nodal PTCL with a double T_{FH}/Th2 phenotype and a molecular Th2 signature. A) Diffuse proliferation of large pleomorphic cells; this case would be classified as T_{FH} PTCL according to the WHO2017, based on the expression of 2 T_{FH} markers, ie PD1 (b) and BCL6 (c), but disclosed strong nuclear staining for GATA3 in virtually all tumor cells (d) and, although less uniform, FOXP3 (e). Few tumor cells also expressed CD30 (f). RT-MLPA profile showed a Th2 signature and classified in the Th2 class by the SVM.

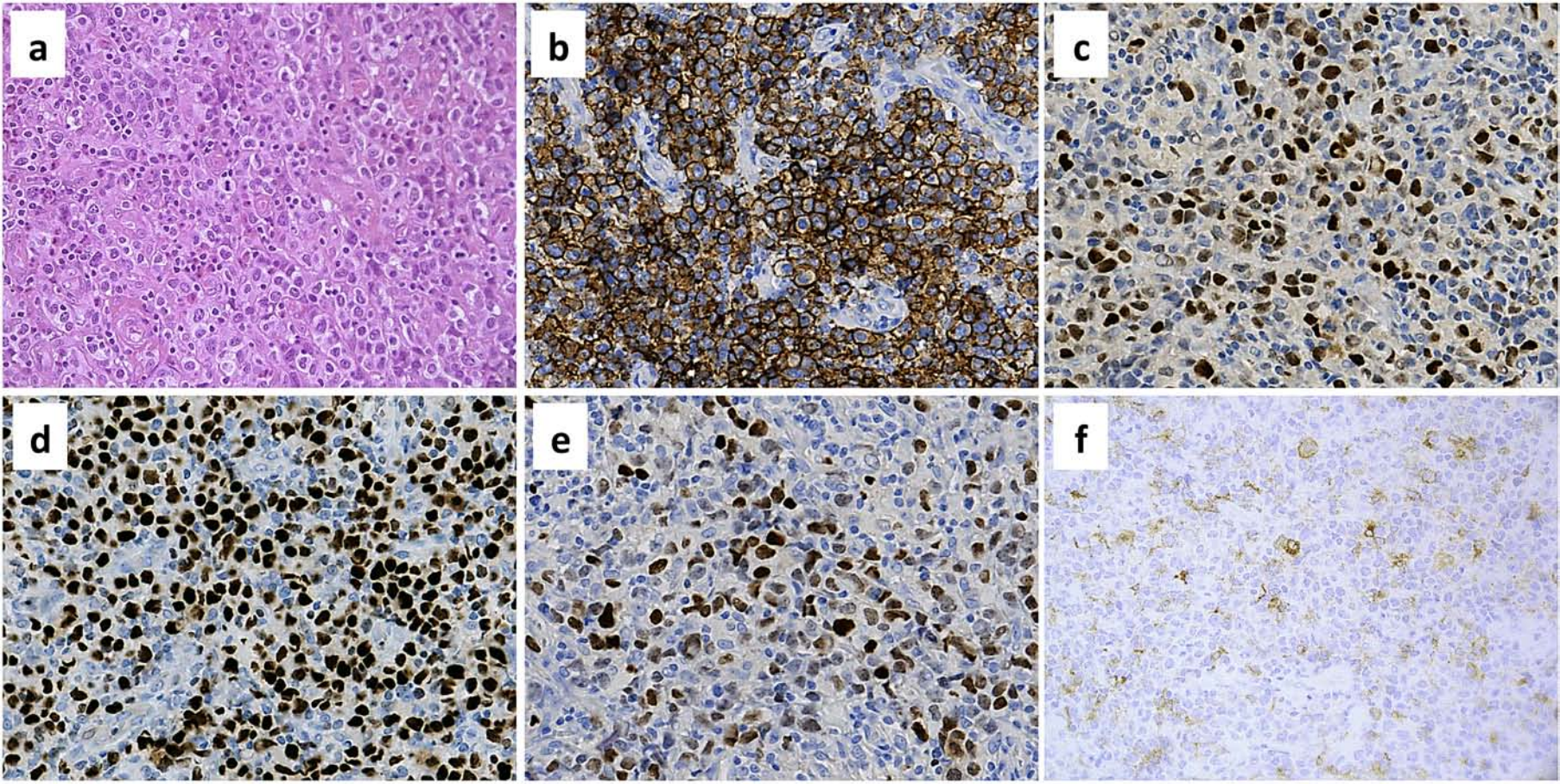
Figure 3. Unsupervised hierarchical clustering of PTCLs including PTCL-NOS (n = 230). PTCL-NOS distributed among 6 defined clusters represented by colored bars under the heat map (from left to right): C1 (red), C2 (green), C3 HSTL (purple), C4 AITL/T_{FH}-derived PTCLs (blue and light blue), C5 NKTCL (yellow), and "cytotoxic ALCL" (orange and red). The 77 PTCL-NOS (grey in Pathology) are distributed among the subgroups

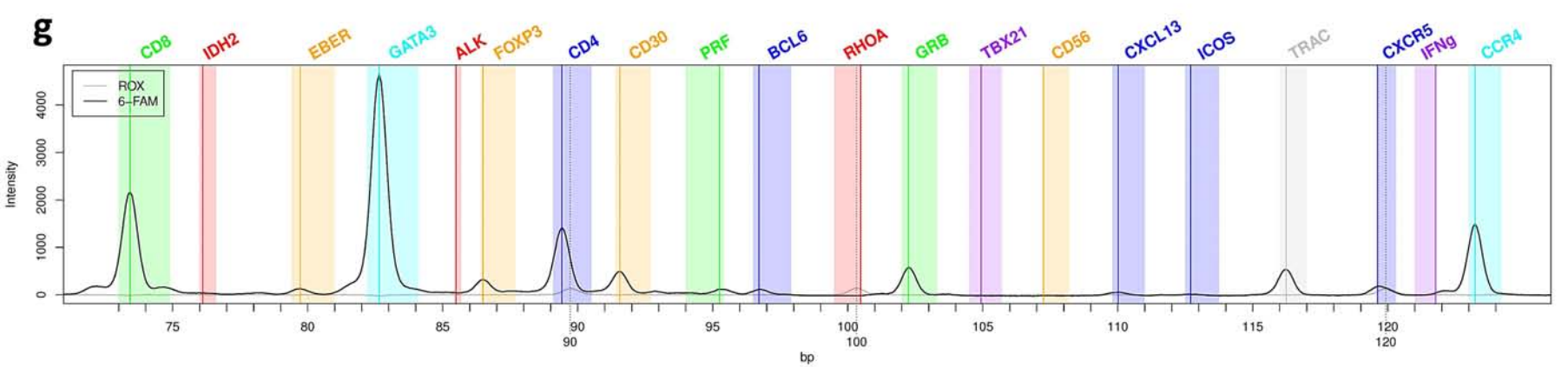
Figure 4. A) Bioinformatic model for the analysis of RT-MLPA data. The SVM model attributes a predicted (rectangle) or provisional (oval) class for each PTCL sample. Post-tests based on ALK and CD30 were designed to distinguish between the subgroups in the heterogeneous "cytotoxic/ALCL-like" category. **B) PCA plot of the SVM classification for PTCL-NOS (n=77)** showed 3 main molecular categories among: T_{FH}/AITL (blue), Th2 (light blue), and cytotoxic/Th1 (green). The latter also comprised NKTCL-like (yellow) and ACL-like (orange) subgroups. Seven cases were unclassified (red).

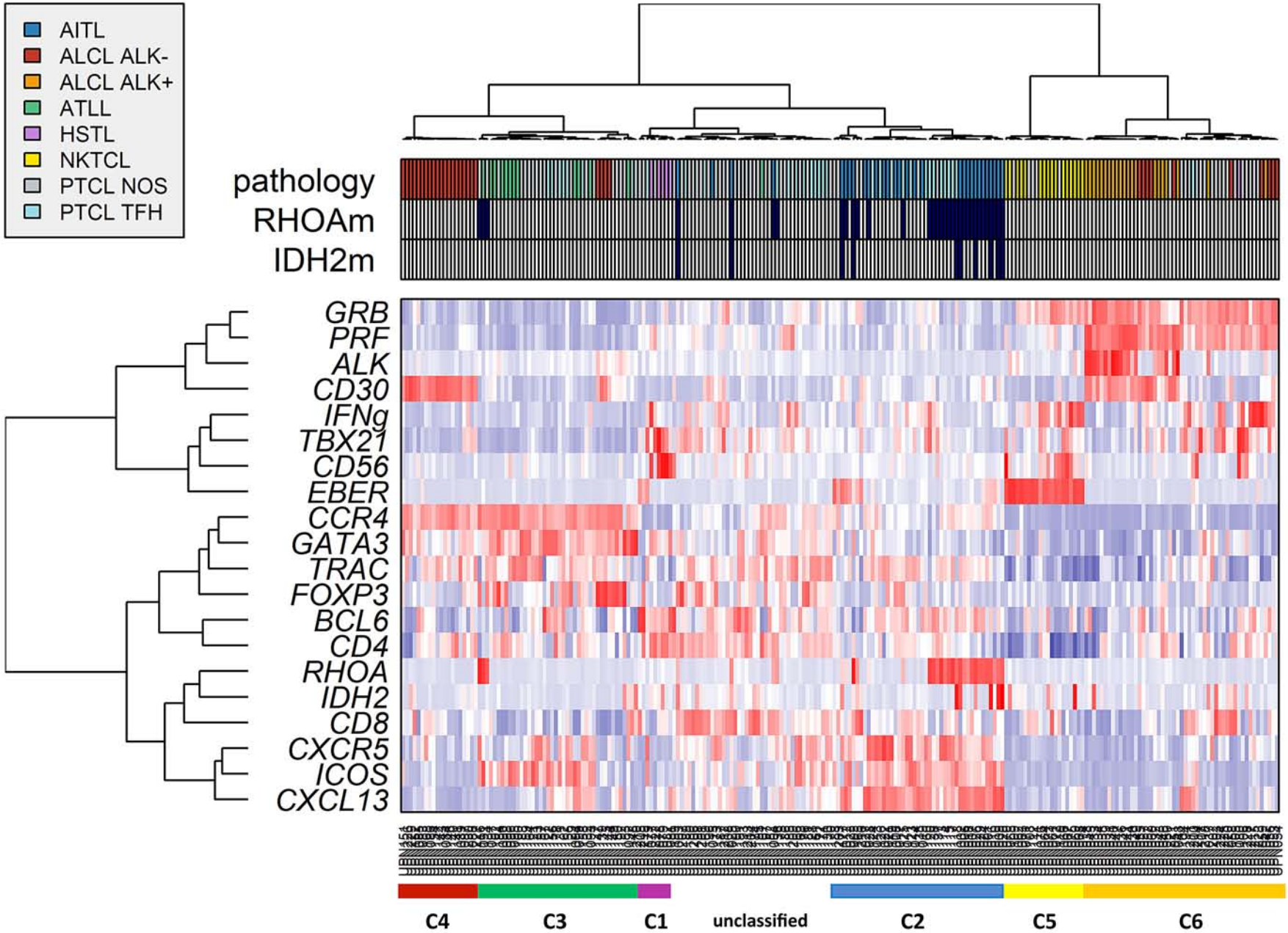
Figure 5. Kaplan Meier survival analysis

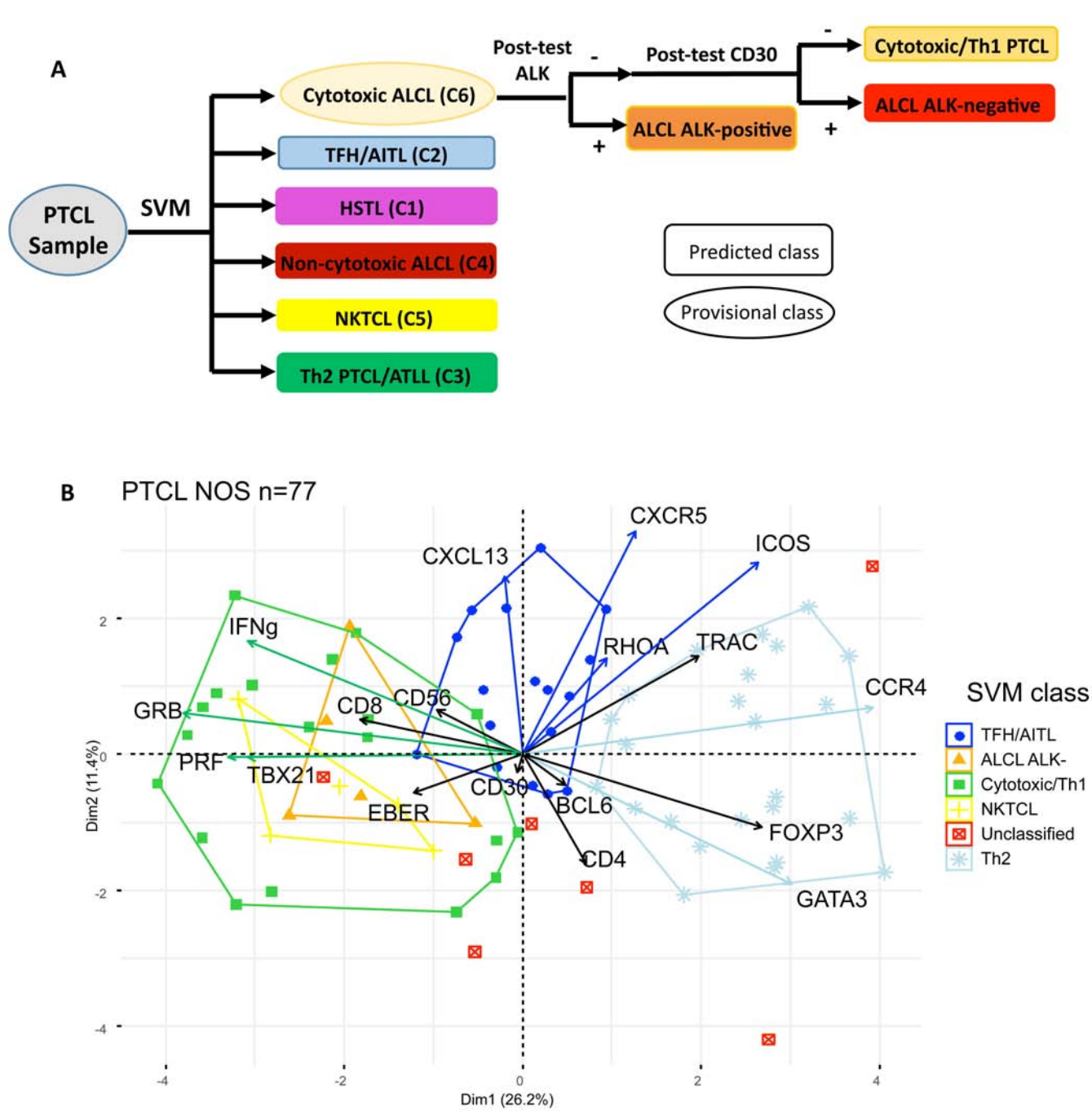
A) Overall survival of the 108 patients corresponding of the main RT-MLPA subgroups (11 ALCL ALK+, 10 cytotoxic ALCL ALK-, 24 non-cytotoxic ALCL ALK-, and 63 T_{FH}/AITL) **B)** Overall survival of the main RT-MLPA subgroups according to the *DUSP22* status within the non-cytotoxic ALCL ALK-negative category **C)** Overall survival of 43 PTCL,NOS according to the molecular status by RT-MLPA (19 cytotoxic/TH1 and 24 TH2) **D)** Overall survival of 30 PTCL-NOS according to the immunohistochemistry data (19 GATA3, 11 TBX21).

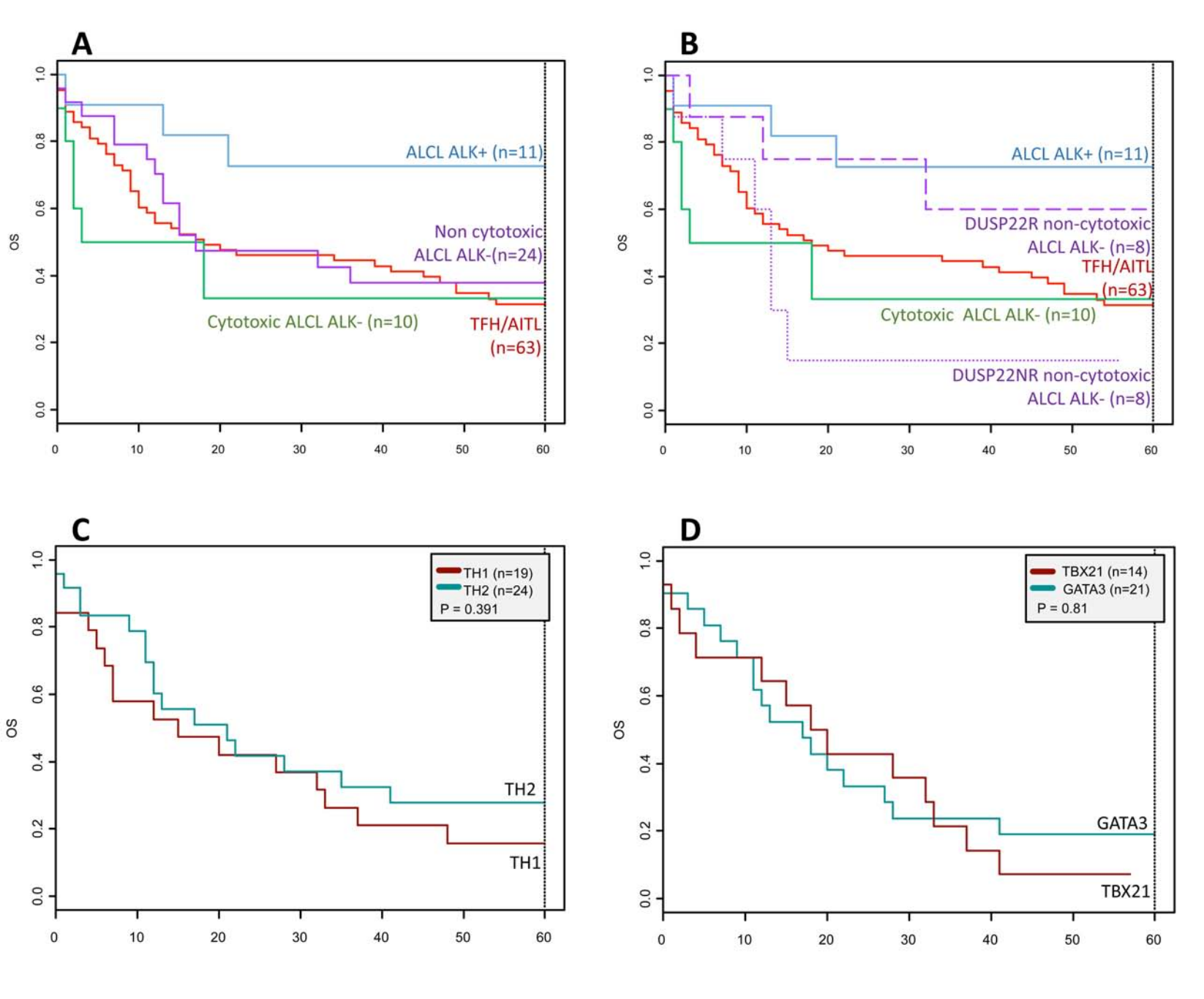

AITL
ALCL ALK-
ALCL ALK+
ATLL
HSTL
NKTCL
PTCL TFH
Pathology
RHOAm
IDH2m
GRB
PRF
ALK
CD30
IFNg
EBER
CD56
TBX21
RHOA
IDH2
CXCR5
CXCL13
ICOS
BCL6
CD4
CD8
CCR4
TRAC
GATA3
FOXP3
C1
C2
C3
C4
C5
C6


a
b
c
d
e
f


g
CD8
IDH2
EBER
GATA3
ALK
FOXP3
CD4
CD30
PRF
BCL6
RHOA
GRB
TBX21
CD56
CXCL13
ICOS
TRAC
CXCR5
IFNg
CCR4
ROX
6-FAM
Intensity
0
1000
2000
3000
4000
75
80
85
90
95
100
105
110
115
120
bp


AITL
ALCL ALK-
ALCL ALK+
ATLL
HSTL
NKTCL
PTCL NOS
PTCL TFH
pathology
RHOAm
IDH2m
GRB
PRF
ALK
CD30
IFNg
TBX21
CD56
EBER
CCR4
GATA3
TRAC
FOXP3
BCL6
CD4
RHOA
IDH2
CD8
CXCR5
ICOS
CXCL13
C4
C3
C1
unclassified
C2
C5
C6


A
PTCL Sample
SVM
Cytotoxic ALCL (C6)
TFH/AITL (C2)
HSTL (C1)
Non-cytotoxic ALCL (C4)
NKTCL (C5)
Th2 PTCL/ATLL (C3)
Post-test ALK
-
+
Post-test CD30
-
+
Cytotoxic/Th1 PTCL
ALCL ALK-negative
ALCL ALK-positive
Predicted class
Provisional class
B
PTCL NOS n=77
CXCR5
ICOS
CXCL13
IFNg
RHOA
TRAC
CCR4
GRB
CD8
CD56
PRF
TBX21
CD30
BCL6
EBER
FOXP3
CD4
GATA3
Dim2 (11.4%)
Dim1 (26.2%)
2
0
-2
-4
-4
-2
0
2
4
SVM class
TFH/AITL
ALCL ALK-
Cytotoxic/Th1
NKTCL
Unclassified
Th2


A
OS
ALCL ALK+ (n=11)
Non cytotoxic
ALCL ALK-(n=24)
Cytotoxic ALCL ALK- (n=10)
TFH/AITL
(n=63)
B
OS
ALCL ALK+ (n=11)
DUSP22R non-cytotoxic
ALCL ALK- (n=8)
TFH/AITL
(n=63)
Cytotoxic ALCL ALK- (n=10)
DUSP22NR non-cytotoxic
ALCL ALK- (n=8)
C
OS
TH1 (n=19)
TH2 (n=24)
P = 0.391
TH2
TH1
D
OS
TBX21 (n=14)
GATA3 (n=21)
P = 0.81
GATA3
TBX21

Supplemental Material

Patients and tumor samples

As shown in Figure S1, the cases were divided into two cohorts. The first cohort (classification cohort, n = 230) used to build and train a predictive classification model, was composed of 230 cases, including 30 AITLs, 33 T_{FH}-PTCL, 21 ALK-positive anaplastic large cell lymphomas (ALCLs), 34 ALK-negative ALCLs, 16 extranodal NK/T-cell lymphomas, nasal-type (NKTCLs), 6 hepatosplenic T-cell lymphomas (HSTLs), and 13 adult T-cell leukemia/lymphomas (ATLLs), and 77 PTCL-NOS cases, according to the WHO 2017 classification. The second cohort (diagnostic cohort) of 40 FFPE PTCL samples (6 ALK-positive ALCLs, 4 ALK-negative ALCLs, 13 AITLs, 9 NKTCLs, 6 ATLLs, and 2 PTCL-NOS) was used to validate the robustness of the assay and its reproducibility between three independent centers.

Immunohistochemistry and EBV *in situ* hybridization

Deparaffinized tissue sections were stained for a panel of T-cell (CD3, CD2, CD7, CD5, CD8, CD4), cytotoxic (TIA-1, granzyme B, perforin), T_{FH} (PD1, CXCL13, BCL6, ICOS, CD10), follicular dendritic-cell (CD21, CD23), B-cell (CD20, CD79a, PAX5), and other (CD30, CD56, ALK) antigens. Immunostains for GATA3 (clone HG3-31, Santa Cruz Biotechnology, Santa Cruz, CA) and TBX21 (clone 4B10, BD Biosciences, San Jose, CA) were performed on a subset of cases. The scoring system was evaluated as follows: score 0: negative (below threshold of 10% positive tumor cells), score 1: 10-50% positive tumor cells, score 2: >50% positive tumor cells as previously described (Dobay, Haematologica, 2017;102(4): e148-e151). The detection of EBV was performed by *in situ* hybridization using EBER probes and was scored as previously described (de Leval, Haemalogica 2015;100(9):e361-364): score 0: absence of large EBV-positive cells, score 1: up to 5 large EBV-positive cells per high power fields (hpf), score 2 : 5 to 50 per hpf and score 3 : > 50 per hpf , or sheets or aggregates of large EBV-positive cells.

FISH for *DUSP22/IRF4* rearrangement

Laboratory-developed fluorescence *in situ* hybridization FISH-probes using bacterial artificial chromosomes (BACs) were used to explore *DUSP22/IRF4* rearrangements by interphase FISH in FFPE tissue sections of ALK-negative ALCL. Break-apart probes consisted of telomeric

RP3-416J7 labelled with SpectrumOrange™ and centromeric RP11-615C17 and CTD-3139L20 labelled with SpectrumGreen™ (Institute of Pathology, Lausanne).

Allele-specific PCR and targeted deep sequencing

RHOA G17V and *IDH2* R172K/T mutational status were also determined by allele-specific qPCR (AS-qPCR) or next generation sequencing (NGS) on a MiSeq instrument with a mean coverage depth of 1200X. (Dupuy et al. J Mol Diag)

RNA extraction and microarray procedure

Total RNA was extracted from frozen and FFPE tissue samples with Trizol reagent and the Maxwell 16 LEV RNA kit (Promega), respectively. High-throughput gene-expression data (HG-U133 plus 2.0 chips, Affymetrix, Santa Clara, CA) from 72 cases (23 AITL and 49 PTCL-NOS) were previously reported (de Leval et al, Blood. 2007;109(11):4952–4963).

RT-MLPA procedure

All probes consisted of a gene-specific region complementary to the cDNA target linked to a tail. The 5' and 3' tails correspond to the primers U1 (TCCAACCCTTAGGGAACCC) and U2 (GTGCCAGCAAGATCCAATCTAGA) used for the final PCR amplification step. Spacers were added between these tails and the gene-specific regions to allow the separation and identification of the PCR products based on size. The 3' probes were phosphorylated at the 5' ends. For four genes, specific probes without the PCR tails were used as competitors to normalize the amplification signal. The competitors were added to the corresponding probes at a ratio of 1 (MLPA) to 4 (competitor) for GZB, CXCL13, TCRα1, and 1 (MLPA) to 3 (competitor) for PRF. For the identification of *RHOAG17V* and *IDH2R172K/T* hotspot mutations, 5' probes with the last nucleotide corresponding to the mutated nucleotide were designed.

Bioinformatic analysis

Starting from raw (.fsa) files generated by the genetic analyzer, the web interface provides a graphical representation of the gene expression profile and a table of normalized gene expression calculated as a function of the FAM fluorescence intensity normalized to the

median intensity of the 20 genes of the signature. For each sample, this interface also returns a class prediction deduced from a support vector machine classifier (SVM), established using the e1071 R package with default settings. To minimize the risks of misclassification, this algorithm was coupled with a second nearest centroid classifier to minimize the risks of misclassification. For each class *C*, the coordinates of the centroid μc was calculated so that

$$\mu_C = \frac{1}{N}\sum_{i=1}^{N} S_i$$

where N represents the number of cases belonging to the class *C* defined by the SVM classifier in the training series and *Si*, a sample defined by the expression of the 20 RT-MLPA markers. For each sample *Si*, the distance d(*Si*, *C*) to the centroid μc of the class *C* is given by

$$d(S_i, C) = \sqrt{\sum_{i=1}^{N}(S_i - \mu_C)^2}$$

For each sample, the class prediction returned by the web interface consists of the results of the SVM algorithm completed by the arithmetic distance to the centroids of each class.

A stratified cross validation was performed to assess the accuracy of our assay. A bootstrap resampling process was first used to build 100 independent training series, randomly selecting two-thirds of the samples within the six categories of the unsupervised hierarchical analysis. For each iteration, a SVM predictor was trained and tested against all remaining samples. A definitive SVM predictor was thus trained using the 184 cases classified by hierarchical analysis.

Post-tests were then built to distinguish ALCL from cytotoxic/Th1 cases. A specific threshold was determined using ROC curves, based on the expression of the *CD30* gene values by RT-MLPA (CD30 threshold = 0.8). A second post-test was performed to distinguish ALK-positive from ALK-negative ALCLs, based on the expression of the *ALK* gene (threshold = 0.2).

Another CD30 post-test was designed among the Th2 category to catch misclassified CD30-positive Th2 cases (threshold = 1.4) due to overlapping *FOXP3* and *ICOS* expression in both the CD30-positive and CD30-negative Th2 groups.

An algorithm was finally performed to calculate the distance of the sample to the centroid of the predicted class and compare it to the other intraclass distances. The algorithm constructs a boxplot based on these intraclass distances and calculates the first and third quartile values, deducting the IQR (InterQuartile Range) value, which is equal to Q3-Q1.

Finally, if the distance of the tested sample was higher than Q3 + 1.5 IQR, the sample was considered to be a mild outlier, which defined the NOS category.

Tables

Table S1- Sequences of the oligonucleotides used for the RT-MLPA assay. Genes are listed according to the probes size. Specific sequences are represented in pink (5'probe) and red (3'probe). The blue sequence corresponds to the common 5' and 3' tails. Nucleotides spacers are in green.

Gene	Oligo	probe	Sequence
CD8	CD8E5L	5'probe	GTGCCAGCAAGATCCAATCTAGATCGTGCCGGTCTTCCTGCCAG
	CD8E6R	3'probe	CGAAGCCCACCACGACGCCTCCAACCCTTAGGGAACCC
IDH2R172K	IDH2R172KL	5'probe	GTGCCAGCAAGATCCAATCTAGACCAAGCCCATCACCATTGGCAA
IDH2R172T	IDH2R172TL	5'probe	GTGCCAGCAAGATCCAATCTAGACCAAGCCCATCACCATTGGCAC
	IDH2R172KR	3'probe	GCACGCCCATGGCGACCAGTTCCAACCCTTAGGGAACCC
EBER1	EBER1L	5'probe	GTGCCAGCAAGATCCAATCTAGATACGTAGCCACCCGTCCCGGGTA
	EBER1R	3'probe	CAAGTCCCGGGTGGTGAGGATATCCAACCCTTAGGGAACCC
GATA3	GATA3E3L	5'probe	GTGCCAGCAAGATCCAATCTAGACCTCATTAAGCCCAAGCGAAGGCTG
	GATA3E4R	3'probe	TCTGCAGCCAGGAGAGCAGGGACTCCAACCCTTAGGGAACCC
ALK	ALKE23F	5'probe	GTGCCAGCAAGATCCAATCTAGACCTCCGAGAGACCCGCCCTCGCCCG
	ALKE24R	3'probe	AGCCAGCCCTCCTCCCTGGCCATGCTCCAACCCTTAGGGAACCC
FOXP3	FOXP3E3L	5'probe	GTGCCAGCAAGATCCAATCTAGATAGGACAGGCCACATTTCATGCACCAG
	FOXP3E4R	3'probe	CTCTCAACGGTGGATGCCCACGCTTCCAACCCTTAGGGAACCC
CD4	CD4E4L	5'probe	GTGCCAGCAAGATCCAATCTAGAGAGGAGGTGCAATTGCTAGTGTTCGGAT
	CD4E5R	3'probe	TGACTGCCAACTCTGACACCCACCTTCCAACCCTTAGGGAACCC
CD30	CD30E3L	5'probe	GTGCCAGCAAGATCCAATCTAGATTGTACAGCCTGCGTGACTTGTTCTCGAG
	CD30E4R	3'probe	ACGACCTCGTGGAGAAGACGCCGTACTCCAACCCTTAGGGAACCC
PFR	PFRE2L	5'probe	GTGCCAGCAAGATCCAATCTAGATAACACGGTGGAGTGCCGCTTCTACAG
	PFRE3R	3'probe	TTTCCATGTGGTACACACTCCCCCGTACTACTCCAACCCTTAGGGAACCC
BCL6	BCL6E3Lb	5'probe	GTGCCAGCAAGATCCAATCTAGACATAAAACGGTCCTCATGGCCTGCAG
	BCL6E4Rb	3'probe	TGGCCTGTTCTATAGCATCTTTACAGACCAGTTGTCCAACCCTTAGGGAACCC
RHOA mut	RHOmutL	5'probe	GTGCCAGCAAGATCCAATCTAGAGGTGATTGTTGGTGATGGAGCCTGTGT
RHOA	RHOR	3'probe	AAAGACATGCTTGCTCATAGTCTTCAGCAAGGACCTCCAACCCTTAGGGAACCC
GZMB	GRBE3L	5'probe	GTGCCAGCAAGATCCAATCTAGATACTAACTTCTCCAACGACATCATGCTACTGCAG
	GRBE4R	3'probe	CTGGAGAGAAAGGCCAAGCGGACCAGTACTACTCCAACCCTTAGGGAACCC
TBET	TBETE5L	5'probe	GTGCCAGCAAGATCCAATCTAGATACTACCCAAAGGATTCCGGGAGAACTTTGAGTC
	TBETE6R	3'probe	CATGTACACATCTGTTGACACCAGCATCCCTACTTCCAACCCTTAGGGAACCC
CD56	CD56E11L	5'probe	GTGCCAGCAAGATCCAATCTAGATACTACTACTCACCCCCTCTGCCAGCTATCTGGAG
	CD56E12R	3'probe	GTGACCCCAGACTCTGAGAATGATTTTGGTACTACTCCAACCCTTAGGGAACCC
CXCL13	CXCL13E2L	5'probe	GTGCCAGCAAGATCCAATCTAGATACTACTGGTCAGCAGCCTCTCTCCAGTCCAAG
	CXCL13E3R	3'probe	GTGTTCTGGAGGTCTATTACACAAGCTTGAGGTGTTACTTCCAACCCTTAGGGAACCC
ICOS	ICOSE2L	5'probe	GTGCCAGCAAGATCCAATCTAGAAAGTAACTCTTACAGGAGGATATTTGCATATTTATG
	ICOSE3R	3'probe	AATCACAACTTTGTTGCCAGCTGAAGTTCTGTACTACTTCCAACCCTTAGGGAACCC
TRAC	TRACE3L	5'probe	GTGCCAGCAAGATCCAATCTAGATACTACTACTACTACTACCTGCGGCTGTGGTCCAGCTGAG
	TRACE4R	3'probe	ATCTGCAAGATTGTAAGACAGCCTGTGCTCTACTACTATCCAACCCTTAGGGAACCC
CXCR5	CXCR5E1L	5'probe	GTGCCAGCAAGATCCAATCTAGATACTACTACTACTACTGGACCTCGAGAACCTGGAGGACCTG
	CXCR5E2R	3'probe	TTCTGGGAACTGGACAGATTGGACAACTATAACGTACTACTCCAACCCTTAGGGAACCC
INFg	INFgE3L	5'probe	GTGCCAGCAAGATCCAATCTAGATACTAAACGAGATGACTTCGAAAAGCTGACTAATTATTCG
	IFNgE4R	3'probe	GTAACTGACTTGAATGTCCAACGCAAAGCATACTACTACTACTTCCAACCCTTAGGGAACCC
CCR4	CCR4E1L	5'probe	GTGCCAGCAAGATCCAATCTAGATACTACTACTACTCCTCAGAGCCGCTTTCAGAAAAGCAAG
	CCR4E2R	3'probe	CTGCTTCTGGTTGGGCCCAGACCTTACTACTACTACTACTACTACTCCAACCCTTAGGGAACCC

Table S2. Comparison of *RhoA*G17V (n = 33) and *IDH2*R172K/T (n = 8) mutations detected by RT-MLPA and next generation sequencing (NGS) or allele-specific-qPCR (AS-qPCR) analysis.
Concordant results are represented in green, while discordant result in red.

Id	Pathology	RhoAG17V status by RT-MLPA	RhoA status by AS-qPCR /NGS (allele frequency)	IDH2 R172K/T status by RT-MLPA	IDH2 status with AS-qPCR/ NGS (allele frequency)
UPN001	AITL	+	+/NA	WT	WT/NA
UPN002	AITL	+	+/NA	+	+/NA
UPN003	AITL	+	+/NA	+	+/NA
UPN004	AITL	+	+/NA	WT	WT/NA
UPN005	AITL	+	+/NA	+	+/NA
UPN006	AITL	+	+/+ (6.12%)	WT	NA/NA
UPN007	AITL	+	+/+ (19.27%)	WT	NA/NA
UPN009	AITL	+	+/+ (12.93%)	+	+/+ (5.08%)
UPN010	AITL	+	+/+ (8.53%)	WT	NA/NA
UPN012	AITL	+	+/+ (23.3%)	WT	NA/NA
UPN016	AITL	+	+/+ (22.38%)	+	+/+ (14.9%)
UPN018	AITL	+	+/+ (9.13%)	WT	NA/NA
UPN020	AITL	+	+/+ (12.24%)	+	+/+ (10.23%)
UPN024	AITL	+	+/+ (18.2%)	WT	NA/NA
UPN025	AITL	+	+/+ (7.44%)	WT	WT/+ (2.83%)
UPN026	AITL	+	+/+ (17.52%)	+	+/NA
UPN028	AITL	+	+/NA	WT	NA/NA
UPN029	AITL	+	+/+ (18.31%)	WT	NA/NA
UPN177	AITL	+	+/NA	+	+/NA
UPN117	PTCL TFH	+	+/+ (38.93%)	WT	NA/NA
UPN136	PTCL TFH	+	+/NA	WT	NA/NA
UPN113	PTCL TFH	+	+/NA	WT	NA/NA
UPN120	PTCL TFH	+	+/+ (23.54%)	WT	NA/NA
UPN118	PTCL TFH	+	+/NA	+	+/+ (2.22%)
UPN115	PTCL TFH	+	+/+ (23.15%)	WT	NA/NA
UPN114	PTCL TFH	+	+/NA	WT	NA/NA
UPN138	PTCL TFH	+	+/+ (21.87%)	WT	NA/NA
UPN134	PTCL TFH	+	+/NA	WT	NA/NA
UPN165	PTCL TFH	+	+/NA	WT	NA/NA
UPN116	PTCLnos	+	+/+ (50%)	WT	NA/NA
UPN125	PTCLnos	+	+/NA	WT	NA/NA
UPN137	PTCLnos	+	+/NA	WT	NA/NA
UPN091	ATLL	+	+/NA	WT	NA/NA

Table S3. Correlation scores of the RT-MLPA values (n = 20) of 40 FFPE samples between three independent centers (very strong correlation (VSC): rho > 0.9, strong correlation (SC) rho > 0.7)

	corr-C2-C3	corr-C1-C3	corr-C1-C2	
CD8	0.9777741	0.9553372	0.9683209	VSC
IDH2	-0.2502684	0.4048631	0.3518407	NA
EBER	0.9951752	0.9954745	0.9877439	VSC
GATA3	0.9771558	0.939488	0.9531494	VSC
ALK	0.9894248	0.9789414	0.9823311	VSC
FOXP3	0.9509514	0.927661	0.9438251	VSC
CD4	0.967501	0.8858032	0.9321464	VSC
CD30	0.9897987	0.9903066	0.9864354	VSC
PRF	0.971159	0.987516	0.9687311	VSC
BCL6	0.9412178	0.8655903	0.8919673	SC
RHOA	0.9324257	0.7869005	0.7019153	SC
GRB	0.9740104	0.9800914	0.9456018	VSC
TBX21	0.9366044	0.9629732	0.9254932	VSC
CD56	0.9362774	0.9749429	0.9381184	VSC
CXCL13	0.9633031	0.9472914	0.9743159	VSC
ICOS	0.9041137	0.8439778	0.8613445	SC
TRAC	0.9728271	0.9564053	0.9828094	VSC
CXCR5	0.9810425	0.9339171	0.9412344	VSC
IFNg	0.8904373	0.8534687	0.8465664	SC
CCR4	0.9921214	0.9822178	0.9819727	VSC

***** There was no correlation for *IDH2 R172m* in the absence of *IDH2* mutation detected in these 40 FFPE samples

Table S4. Clinical parameters of ALK-negative ALCL and comparison of the non-cytotoxic ALCL ALK-negative subgroup according to *DUSP22* status

	ALCL ALK- n=13		Non-cytotoxic ALCL ALK-negative n=24		DUSP22R (n=8)		DUSP22NR (n=8)		p
	median (IQ)	% (n)	median (IQ)	% (n)	median (IQ)	% (n)	median (IQ)	% (n)	
age median (range)	53.8 (40.8-67.2)		58 (52-70)		55 (43-72)		56.45 (53.1-70.3)		0.63
<= 60 years		63.6% (7/11)		54.2% (13/24)		62.5% (5/8)		62.5% (5/8)	1
gender male		53.8% (7/13)		79.2% (19/24)		87.5% (7/8)		75% (6/8)	1
IPI>=3		62.5% (5/8)		33.3% (7/21)		50% (4/8)		16.6% (1/6)	0.3
PIT>=2		85.7% (6/7)		42.8% (9/21)		50% (4/8)		16.6% (1/6)	0.3
extranodal site >=2		85.7% (6/7)		21.7% (5/23)		25% (2/8)		28.6% (2/7)	1
stage >=3		77.8% (7/9)		73.9% (17/23)		75% (6/8)		85.7% (6/7)	1
PS>=2		50% (4/8)		21% (4/19)		28.6% (2/7)		0 (0/5)	0.47
LDH>=1		38.5% (5/13)		50% (11/22)		50% (4/8)		60% (3/5)	1
B symptoms		50% (4/8)		35% (7/20)		57.1% (4/7)		20% (1/5)	0.33

Figures

Figure S1. Study design. 270 cases were divided into two cohorts. The classification cohort (n=230 including 30 AITL, 33 PTCL-TFH, 55 ALCL, 13 ATLL, 6 HSTL, 16 NKTCL and 77 PTCL-NOS and consisting mostly of fresh-frozen (FF) samples) was used to train a SVM classifier. The diagnostic cohort (n=40 including 13 AITL, 10 ALCL, 6 ATLL, 9 NKTCL and 2 PTCL-NOS) was used for the independent validation on FFPE samples and for the inter-laboratory reproducibility study.

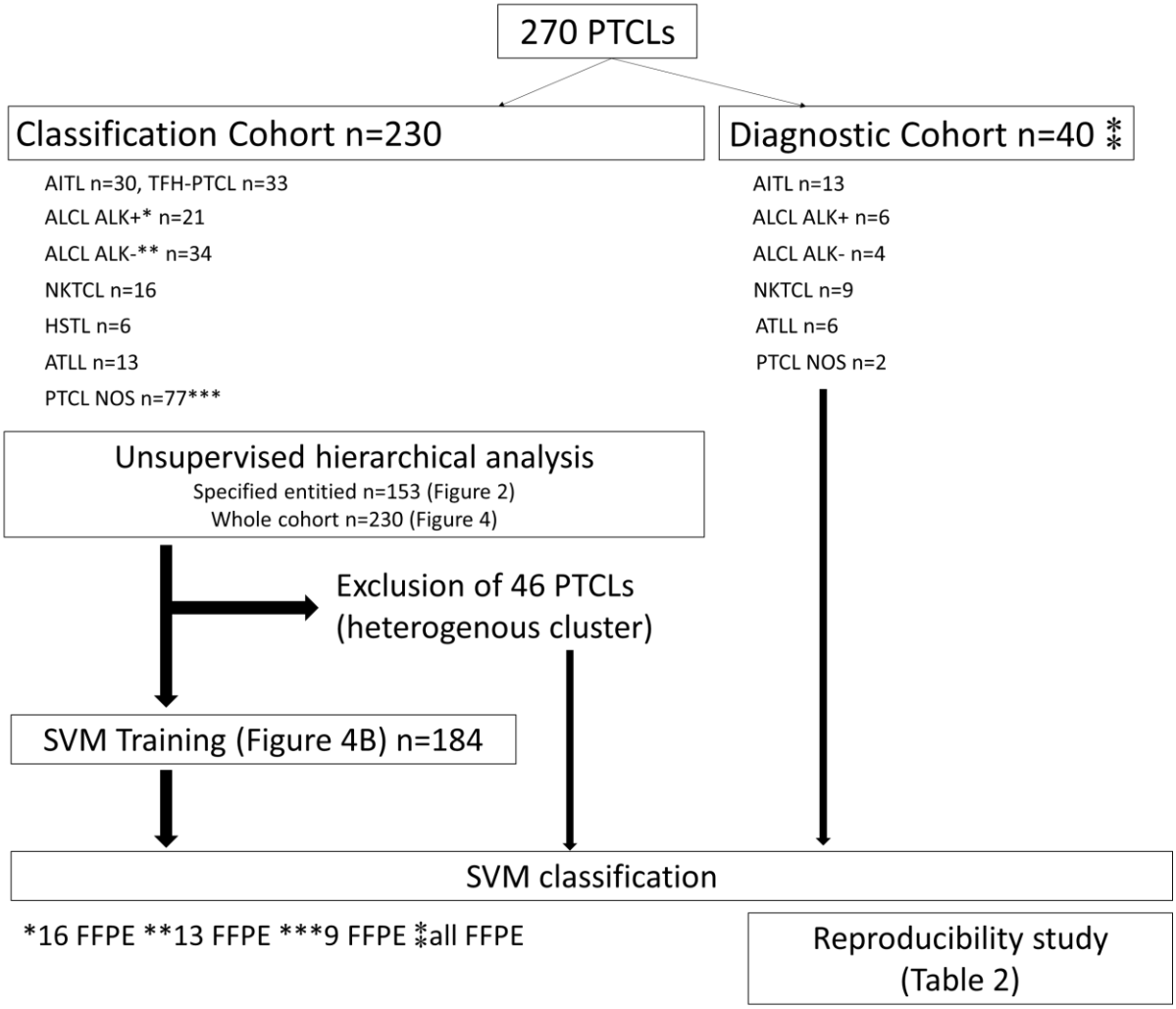


Figure S2. Representation of the RT-MLPA profiles of each PTCL molecular category **A)** AITL/T_{FH} profile, showing the expression of *CXCL13*, *CXCR5*, *ICOS*, and *BCL6*, and in this case the presence of *RHOA* and *IDH2* mutations. **B)** NKTCL signature, characterized by high EBER expression, and that of *CD56* and cytotoxic markers. **C and D)** Cytotoxic ALCL profile, defined by the expression of *CD30* and cytotoxic markers with (C) or without (D) *ALK*. **E)** Non-cytotoxic ALK-negative ALCL signature, characterized by the expression of *CD30* and TH2 markers (*GATA3* and *CCR4*). **F)** Cytotoxic/Th1 signature, defined by the expression of cytotoxic markers with inconsistent expression of Th1 markers. **G)** ATLL/TH2 signature, characterized by the expression of *GATA3*, *CCR4*, and *ICOS*, with inconsistent *FOXP3* expression. **H)** HSTL profile, showing the expression of *CD56*, *TBX21*, *GATA3*, and *BCL6*.

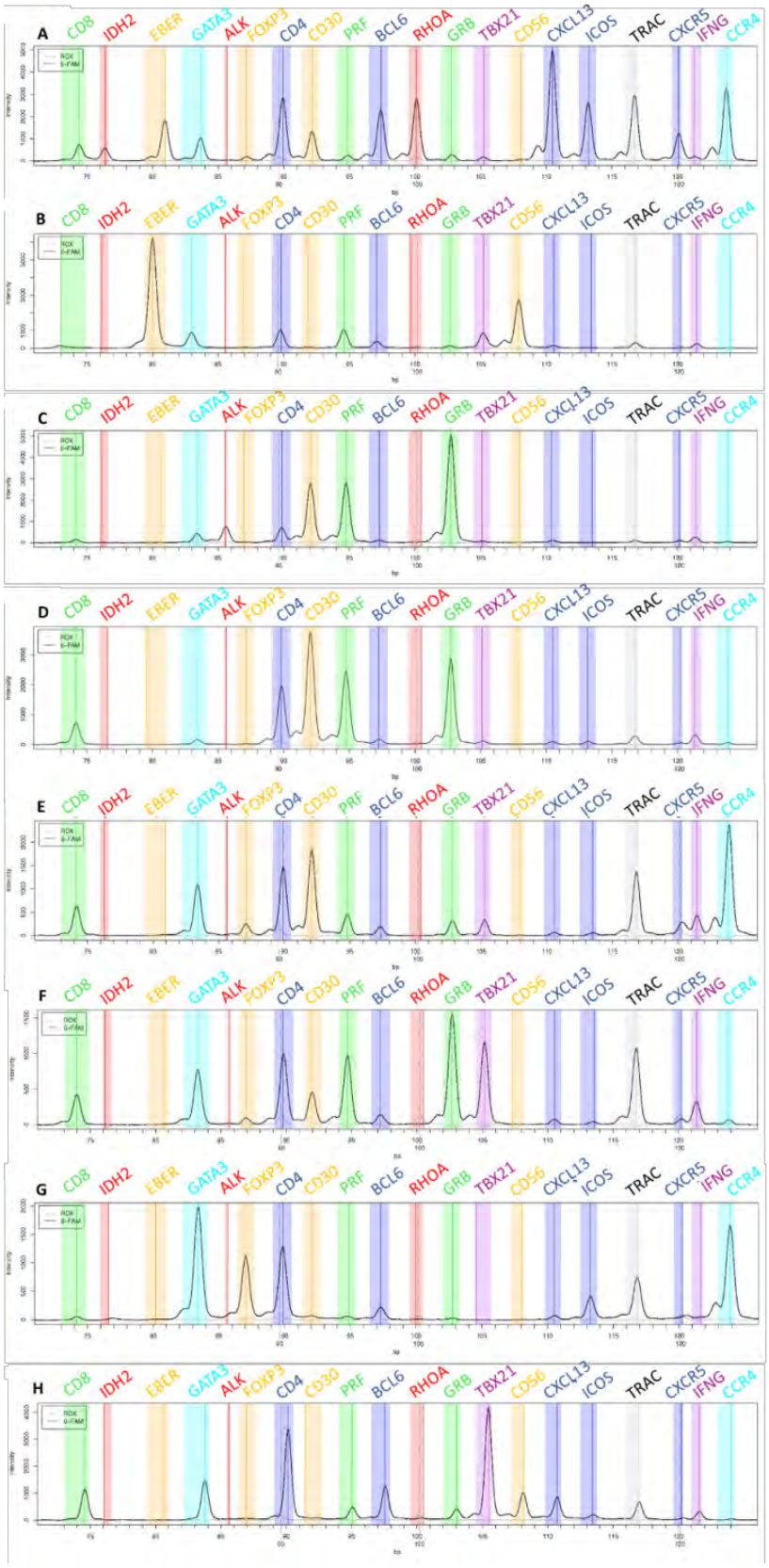

A
B
C
D
E
F
G
H
CD8
IDH2
EBER
GATA3
ALK
FOXP3
CD4
CD30
PRF
BCL6
RHOA
GRB
TBX21
CD56
CXCL13
ICOS
TRAC
CXCR5
IFNG
CCR4

Figure S3. ALCL case reclassified from ALK-negative to ALK-positive based on RT-MLPA assay (UPN051) a) Histopathology of the misdiagnosed ALK- ALCL case, based on the negative immunohistochemistry with ALK1 (performed twice). Strong cytoplasmic staining with the D5F3 clone was obtained, retrospectively. **b)** RT-MLPA profile of the case showing ALK expression (red arrow). **c)** Sequencing of the specific RT-PCR products confirmed an ATIC-ALK fusion transcript. No mutation in the region coding for the epitope of ALK1 was found (data not shown).

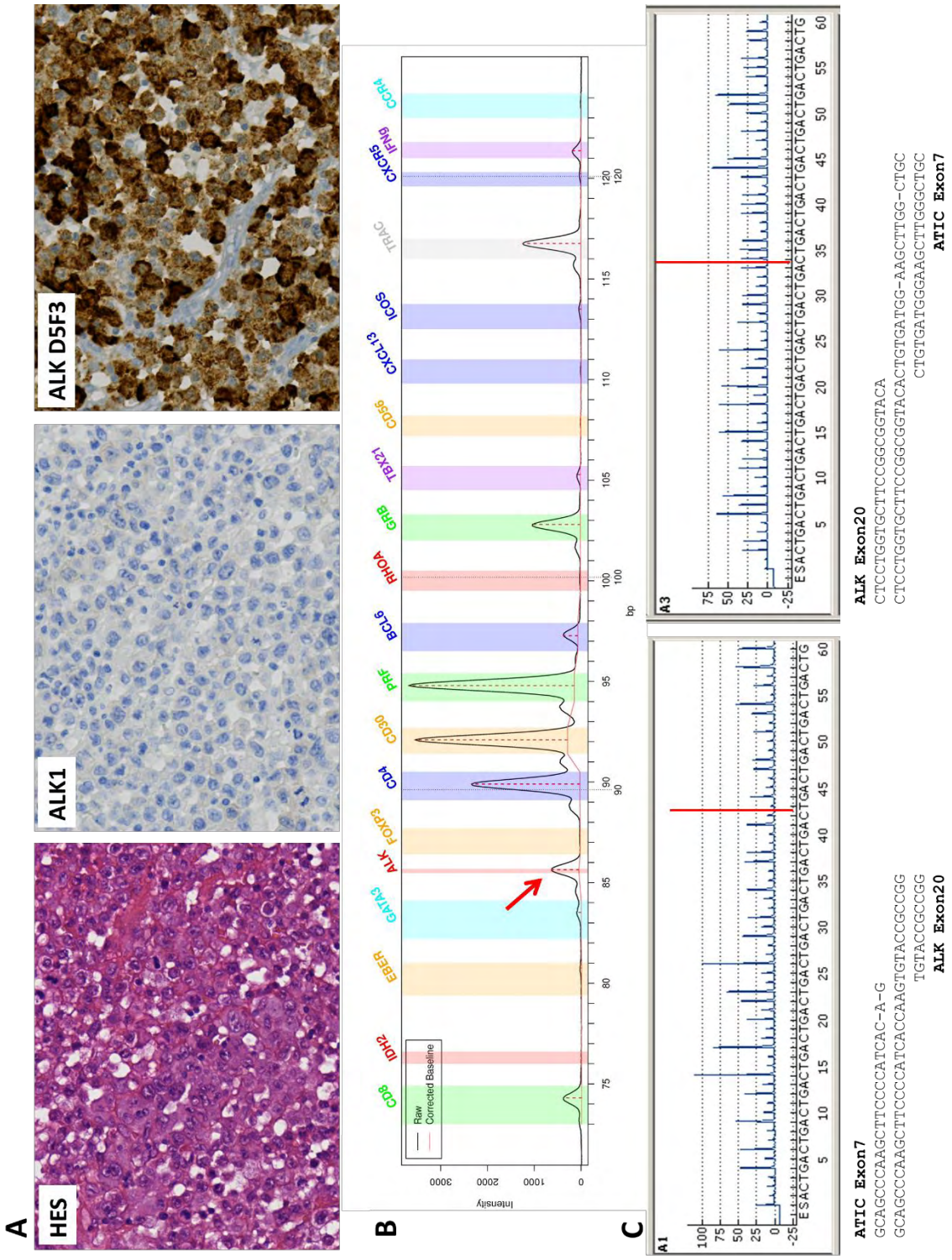


Figure S4. Scatterplot representation of the correlation between RT-MLPA and Affymetrix gene expression values (n = 71 cases, 23 AITL, and 49 PTCL NOS). There were significant correlations (Spearman test, $p < 0.05$) for each gene, with rho > 0.9 for *CXCL13* and *CCR4,* rho > 0.7 for *PRF, GZMB, GATA3, ICOS, CXCR5, BCL6, TNFRSF8/CD30, CD8,* and *TBX21,* and rho > 0.5 for *CD4* and *FOXP3* (x = Affymetrix values log, y = RT-MLPA values). The correlation for *ALK* and *CD56* was not determined because these genes were not evaluated in the Affymetrix series in the absence of ALCL, HSTL, and NKTCL. *RHOA*m, *IDH2*m, TCR, and EBER expression were not studied by the Affymetrix chip.

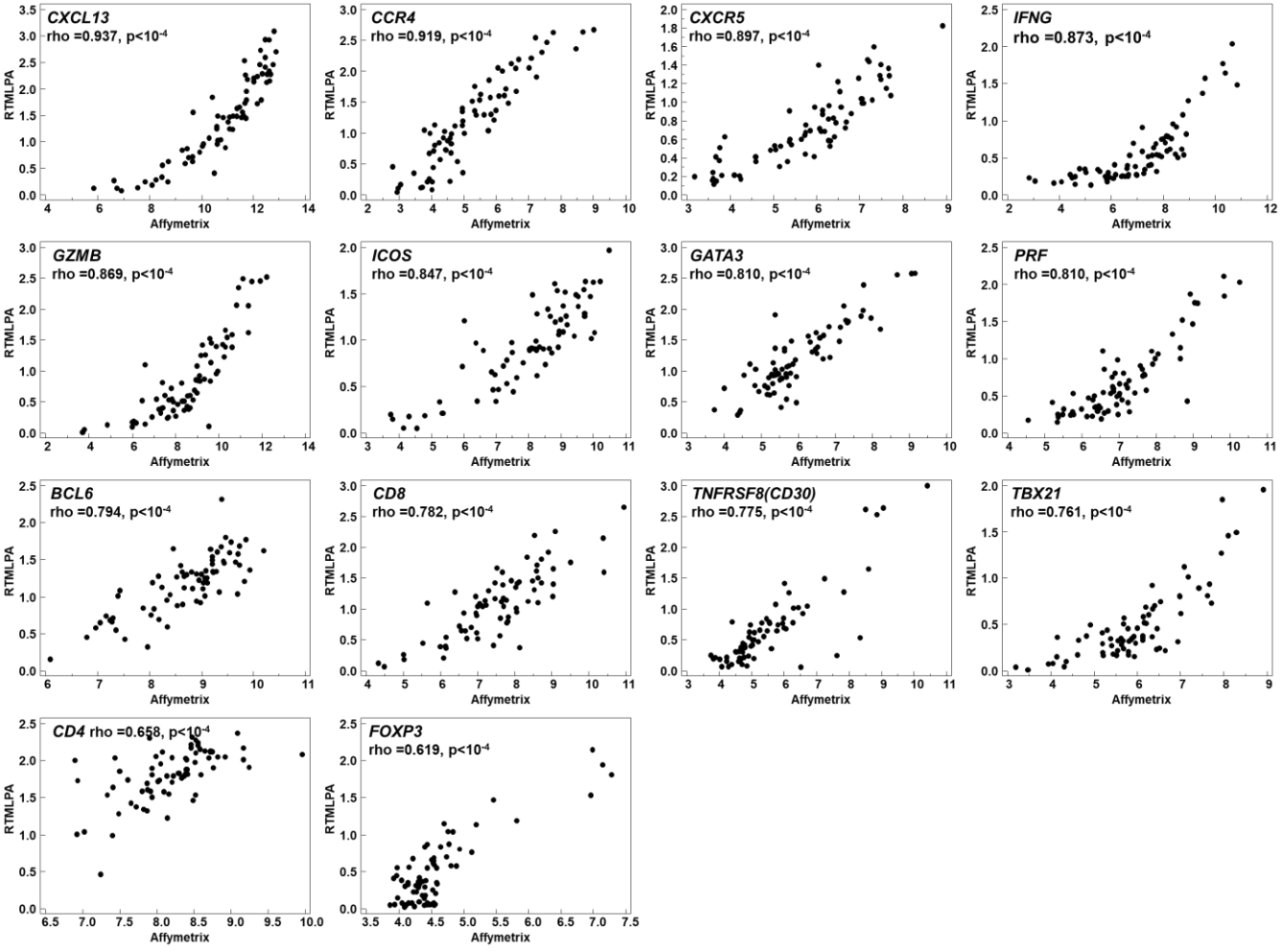


Figure S5. Comparison of individual RT-MLPA gene expression values and immunohistochemical results in 224 PTCLs, including 20 ALK+ ALCLs, 34 ALK- ALCLs, 29 AITLs, 36 PTCL-TFH, 15 NKTCLs, 13 ATLLs, 6 HSTLs, and 70 PTCL-NOS (6 cases with no IHC data were not considered). There was a significant correlation by Wilcoxon's rank- sum test between the RT-MLPA expression level of each gene and negative (-) or positive (+) staining by immunohistochemistry (CD30, TBX21, PRF, GZMB, GATA3, ALK, CXCL13, CD56, ICOS, CD8, CD4, and BCL6) or *in situ* hybridization (EBER).

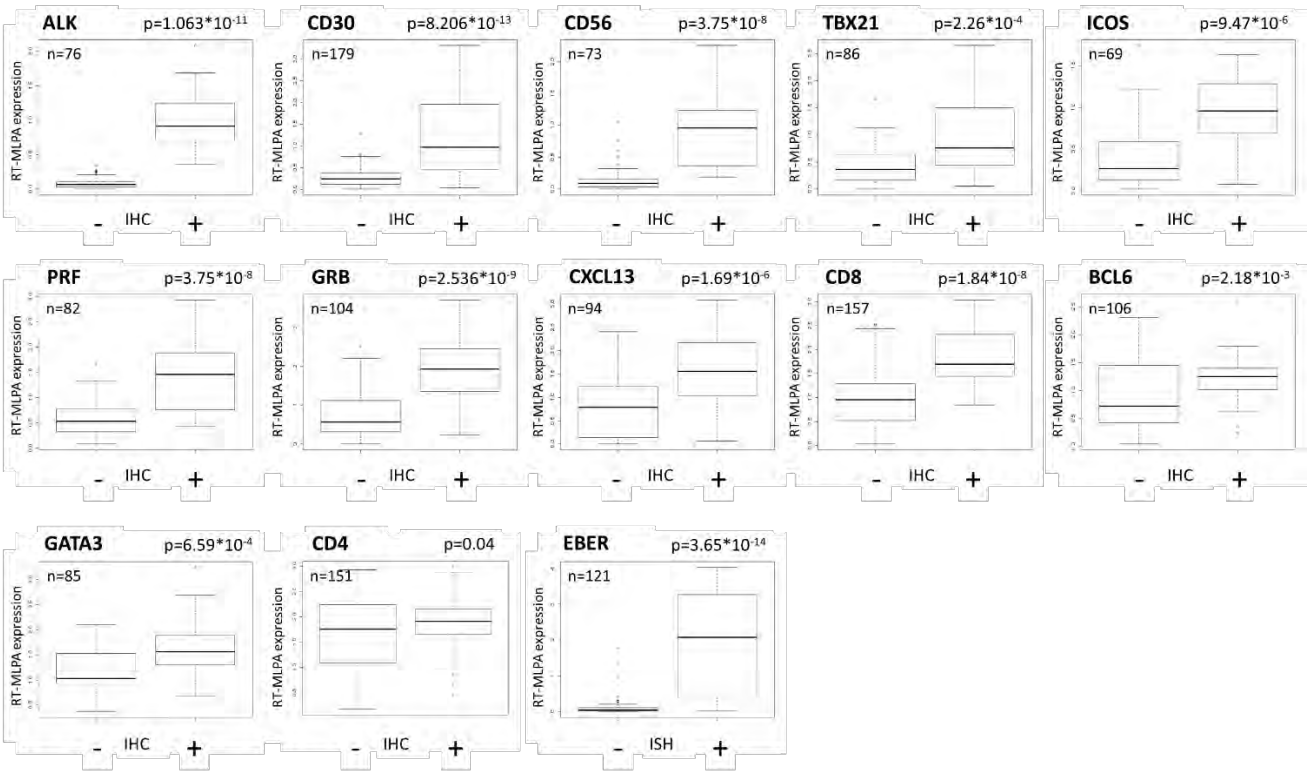


Figure S6. Examples of RT-MLPA profiles for paired FFPE and frozen samples. A) Superimposed profiles showed similar peaks for each 3 paired cases (NKTCL, AITL and non-cytotoxic ALK-negative ALCL). **B**) There was a strong correlation (rho>0.7) of RTMLPA normalized data between frozen (blue) and FFPE (orange) samples.

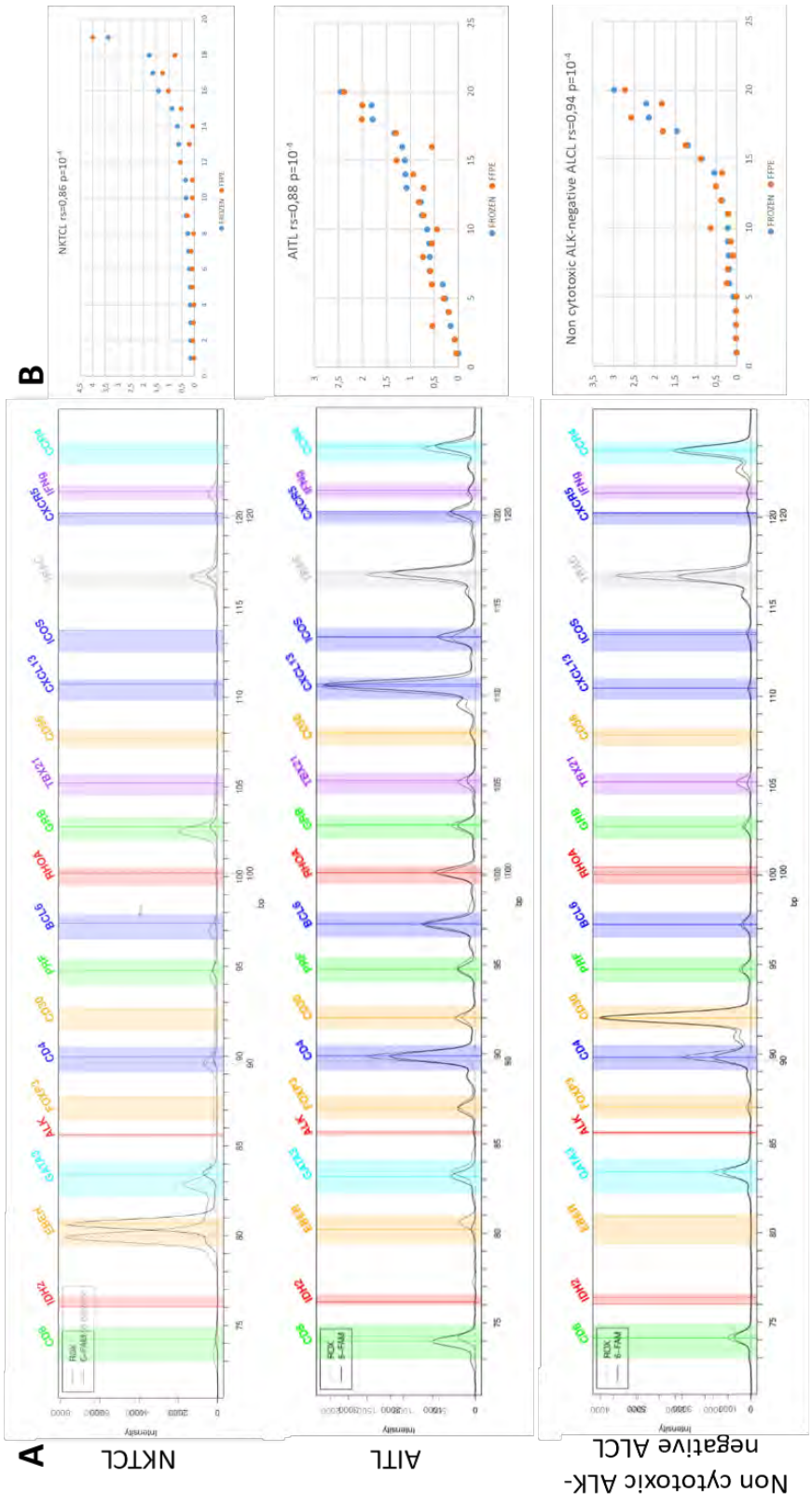


Figure S7. Molecular prediction of PTCL-NOS (n = 77) by the SVM model and correlation with immunohistochemical data. The SVM classification is presented in the top line, the pathological diagnosis in the second line, and immunohistochemical markers in the map). The SVM proposed a molecular class for 92% (69/75) of PTCL-NOS: 17 with a TFH signature, 28 with a cytotoxic/Th1 signature (5 ALK-negative ALCL, 19 cytotoxic/Th1, 5 NKTCL) and 24 with a Th2 signature. Among the 29 cases with a cytotoxic molecular signature, 23 were characterized as cytotoxic by immunochemistry and 6 were undetermined. Among the 24 molecular Th2 PTCL-NOS, 14/18 tested cases had a positive immunostaining for GATA3. Only four discrepancies (5%) were noted: 4 cases with a cytotoxic phenotype were classified in the T_{FH}/ AITL group.