Gene-expression profiling of systemic anaplastic large-cell lymphoma reveals differences based on ALK status and two distinct morphologic ALK+ subtypes

Laurence Lamant,1,2 Aurélien de Reyniès,3 Marie-Michèle Duplantier,1,2 David S. Rickman,3 Frédérique Sabourdy,1,2 Sylvie Giuriato,1,2 Laurence Bruguères,4 Philippe Gaulard,5 Estelle Espinos,1,2 and Georges Delsol1,2

1Institut National de la Santé et de la Recherche Médicale (INSERM) U563 Centre de physiopathologie Toulouse Purpan, Toulouse, France; 2 Université Paul-Sabatier, Toulouse, France; 3Programme Cartes d’Identité des Tumeurs (CIT), Ligue Nationale Contre le Cancer, Paris, France; 4 Department of Pediatric Oncology, Institut Gustave Roussy, Paris, France; 5Department of Pathology, Inserm U 617, Hôpital Henri Mondor, Créteil, France

With the use of microarray gene-expression profiling, we analyzed a homogeneous series of 32 patients with systemic anaplastic large-cell lymphoma (ALCL) and 5 ALCL cell lines. Unsupervised analysis classified ALCL in 2 clusters, corresponding essentially to morphologic subgroups (ie, common type vs small cell and “mixed” variants) and clinical variables. Patients with a morphologic variant of ALCL had advanced-stage disease. This group included a significant number of patients who experienced early relapse. Supervised analysis showed that ALK+ ALCL and ALK− ALCL have different gene-expression profiles, further confirming that they are different entities. Among the most significantly differentially expressed genes between ALK+ and ALK− samples, we found BCL6, PTPN12, CEBPB, and SERPINA1 genes to be overexpressed in ALK+ ALCL. This result was confirmed at the protein level for BCL-6, CEBPβ and serpinA1 through tissue microarrays. The molecular signature of ALK− ALCL included overexpression of CCRT, CNTRF, IL22, and IL21 genes but did not provide any obvious clues to the molecular mechanism underlying this tumor subtype. Once confirmed on a larger number of patients, the results of the present study could be used for clinical and therapeutic management of patients at the time of diagnosis.

© 2007 by The American Society of Hematology

Introduction

Anaplastic large-cell lymphoma (ALCL) is considered a distinct entity in the World Health Organization (WHO) classification of lymphoid tumors.1 Characteristically, 60% to 85% of these tumors express a chimeric protein, with transforming properties, in which the cytoplasmic portion of the anaplastic lymphoma kinase (ALK) tyrosine kinase is linked to a partner protein.1,2 In most patients, ALK protein expression is attributed to the (2;5)(p23;q35) chromosome translocation, which fuses the ALK gene at 2p23 to the nucleophosmin (NPM) gene at 5q35.3 However, several studies have shown that approximately one fifth of ALK+ ALCLs are associated with other cytogenetic abnormalities in which the ALK gene at 2p23 is known to be fused to the TPM3, TPM4, TFG, ATIC, CLTC, MSN, MYH9, and AL017 genes.4 All these translocations result in the expression of chimeric ALK transcripts that are translated into X-ALK fusion proteins, detectable by immunohistochemistry, with tyrosine kinase activity and oncogenic properties.

Although ALK+ ALCL is considered a distinct entity, it shows a broad spectrum of morphologic features, some recognized in the WHO classification as common type, lymphohistiocytic, and smallcell variants.1 In addition, in some patients, different variants coexist in a single biopsy, and these tumors have been designated as mixed ALCL variants.5 Another characteristic of these tumors is the perivascular pattern of neoplastic cells seen in a significant number of patients. The underlying biologic mechanism responsible for these strikingly different aspects is still unknown. Another controversy stems from the fact that some tumors (15%) with worse prognosis are morphologically and phenotypically similar to ALK+ ALCL (ie, CD30+, epithelial membrane antigen-positive [EMA+], T-/null-cell phenotype) but are negative for ALK protein.6 Finally, a significant proportion of patients with ALCL experience early relapses that are unpredictable and considered adverse prognostic factors.7

We undertook gene-expression profiling to identify whether distinct molecular signatures are associated with the presence and absence of ALK protein and with the various morphologic features. In this report, we show that ALK+ tumors and ALK− ALCLs have different gene-expression profiles suggesting different oncogenic mechanisms associated with each type. In addition, we show that morphologic variants of ALCL—for instance, common type versus small cells and mixed variants—have distinct gene-expression signatures.

Patients, materials, and methods

Patients

The following frozen diagnostic tumor specimens from 32 patients with systemic ALCL were studied with patient informed consent: lymph node (29 patients), skin (2 patients), and soft tumor mass (1 patient). Histopathology and immunostaining results were reviewed independently by 3 hematopathologists (L.L., P.G., and G.D.). The percentage of malignant cells in each biopsy specimen was assessed using CD30 staining. Except for 4 patients (indicated by an asterisk in Figure 1), malignant cells accounted for at least 40% of the lymph node involvement. Diagnoses of ALCL were based on morphologic and immunophenotypic criteria, as described in the
WHO classification. All patients coexpressed CD30 and EMA, and most (29 of 32) had the T-cell phenotype, based on the expression of at least one T-cell–associated antigen (CD2, CD3, CD4, CD5, CD8, CD43). Most patients (n = 23) were also positive for cytotoxic associated antigen granzyme B and/or perforin. Twenty-three patients were classified with common type and 9 with morphologic (small cell or mixed) variants. In 11 tumors (4 common type and 7 morphologic variants), neoplastic cells showed a perivascular pattern. ALK expression was detectable in 25 patients using the ALK1 antibody. Staining was nuclear and cytoplasmic in 20 patients, suggesting a t(2;5)(p23;q35)/NPM-ALK translocation, and only cytoplasmic in the 5 remaining patients, with a membrane reinforcement in 3 patients, suggestive of a t(1;2)(q25;p23)/TPM3-ALK translocation. RT-PCR results were in perfect agreement with these staining patterns and demonstrated the expression of TPM3-ALK (3 patients), MYH9-ALK (1 patient), and ATIC-ALK (1 patient) transcripts in the 5 patients with restricted cytoplasmic ALK staining. Because of the controversy regarding ALK-ALCL, only patients with tumors fulfilling strict morphologic and phenotypic criteria were included in the present study (ie, morphologic features consistent with common variant ALCL, coexpression of CD30 and EMA, and T-cell phenotype with the expression of perforin). Patients with primary cutaneous CD30+ T-cell lymphoma and secondary ALK-ALCL were excluded. Among the 32 patients, 19 were male and 13 were female (sex ratio, 1.46), median age at diagnosis was 16 years (range, 4-78 years), and median follow-up was 38.5 months (3.2 years). Eighteen patients had advanced stage III or IV disease at diagnosis, according to the Ann Arbor classification, and 12 patients had localized stage I or II disease. Disease stage in 2 patients, who died immediately after diagnosis, could not be determined. All patients but 2, who died before treatment, received multiagent chemotherapy, based on anthracycline and alkylant agents. Among these patients, 29 achieved complete remission and only one
achieved partial remission. Twelve patients experienced relapse/disease progression, and 9 of these relapses occurred within 12 months of diagnosis (5 children, 4 adults).

Twelve additional cases of systemic ALCL were retrieved from our tumor bank and used only for validation of Affymetrix results by quantitative RT-PCR (qRT-PCR). Approval for this retrospective study protocol was obtained from the CCPPRB (Comité Consultatif de Protection des Personnes dans la Recherche Biomédicale, Hôpitaux de Toulouse, France) Institutional Review Boards. Informed consent was provided in accordance with the Declaration of Helsinki.

Cell lines

Five ALCL cell lines were investigated, of which 4 were t(2;5)(p23;q35)-positive (Karpas 299,11 SU-DHL-1,12 COST,13 PJO [our own laboratory unpublished cell line], and one ALK–[FE-PD]).14 All of them were derived from common type ALCL,11,12 except for the COST cell line, established from a small cell variant ALCL.13

Microarray procedures

Samples were analyzed using Affymetrix GeneChip Human Genome U133A oligonucleotide arrays (http://www.affymetrix.com/support/index.affx). Total RNA was extracted from 40 frozen sections of tumor biopsies (each 5 μm thick) or from 1.0 × 10^6 cells, using the total SV RNA isolation kit (Promega, Charbonnières, France), which included a DNase treatment to minimize genomic DNA contamination. RNA integrity and quantity were assessed on an Agilent 2100 Bioanalyzer (Agilent, Massy, France).

With the 1-cycle eukaryotic target labeling assay (Affymetrix, Santa Clara, CA), total RNA (5 μg) was first reverse transcribed using a T7-oligo(dt) promoter primer in the first-strand cdNA synthesis reaction. After RNase H-mediated second-strand cdNA synthesis, the double-stranded cdNA was purified and served as a template in the subsequent in vitro transcription (IVT) reaction. The IVT reaction was carried out in the presence of T7 RNA polymerase and a biotinylated nucleotide analog/ribonucleotide mix for complementary RNA (cRNA) amplification and biotin labeling. The biotinylated cRNA target was then purified, fragmented, and hybridized to an Affymetrix array. After overnight hybridization, arrays were incubated with phycoerythrin-conjugated streptavidin (SAPE) and biotinylated antibody against streptavidin and were scanned to obtain quantitative gene-expression levels.

Pattern discovery

Except when indicated, transcriptome analysis was carried out using either an assortment of R system software (version 1.9.0) packages including those of Bioconductor (version 1.1.1)15 or original R code. Raw feature data from Affymetrix HG-U133A GeneChip microarrays were normalized using the robust multichip average (RMA) method16 (package affy, version 1.4.32), which yielded log2 intensity expression summary values for each of the 22 823 probe sets. Probe sets corresponding to control genes or having an “_x_” annotation were masked, leaving 19 987 probe sets for further analysis.

Hierarchical clustering analysis of the 32 ALCL samples was derived from a series of 24 hierarchical cluster analyses, obtained from 8 data subsets and 3 different linkage methods (average, complete, and Ward), using 1-Pearson correlation as a distance metric (package class, version 7.2.2). Subsets of data corresponded to supervised selections of probe sets based on 2 criteria: minimal robust coefficient of variation (rCV) and maximal P value of a variance test (see “Variance test”). Between 98 and 5905 probe sets were selected (rCV thresholds spanning the 99.5th to the 60th percentiles and a P value lower than 0.01 for the test of variance). The intrinsic stability of initial dendrograms was assessed at a global level by comparing them with cluster results derived from perturbation/resampling tests (200 iterations for each). Perturbation of data sets was performed by the addition of random Gaussian noise (μ = 0, σ = 1.5 × median variance calculated from the data set). We used the symmetric difference distance17 as a stability measure. This distance is designed to compare 2 partitions and gives the proportion of retention of the pairs of samples that are in the same group (score is between 0 and 1: a score of 1 corresponds to equal partitions). Each dendrogram was cut to yield a partition of k groups (k = 2-8), and we compared partitions at each level of k. We also assessed the stability of the cluster groups obtained for each linkage from the 8 data subsets using the same stability measure. Moreover, for k = 2 groups, we identified consensus clusters, which are groups of samples that clustered together at least 95% of the time in the 24 dendrograms, each cut to yield 2 groups. One of the 24 dendrograms was selected for Figure 1A and the subsequent analyses because we found it to be representative of the segregation between common type ALCL and morphologic variant ALCL globally observed in the 24 dendrograms (see Table S5 [available at the Blood website; see the Supplemental Tables link at the top of the online article]). This dendrogram was obtained with the use of complete linkage on a subset of selected probe sets (n = 495), based on an rCV superior to the 97.5th rCV percentile and less than 10, and a P value of a variance test lower than 0.01.

Robust coefficient of variation. rCV for each probe set was calculated by dividing the standard deviation by the mean, eliminating the highest and lowest expression value across the samples for each probe set.

Variance test. We tested, for each probe set P, whether its variance was different from the median of the variances of the 19 987 probe sets. The statistic used was [(n − 1) × Var(P)/Var(med)], where n refers to the number of samples. This statistic was compared to a percentile of the χ^2 distribution with n − 1 degrees of freedom and yielded a P value for each probe set. This criterion is used in the filtering tool of BRB ArrayTools software (version 3.2.0 beta 5; http://linus.nci.nih.gov/BRB-ArrayTools.html) and is described in the user’s manual.

Association of ALCL subgroups and clinical/genetic variables

Fisher exact tests were carried out to determine the significance of the association between different clinical and genetic variables and the 2 ALCL tumor subgroups derived from each of the cluster analyses. For variables with multiple modalities (eg, stage), we also performed Fisher exact tests after a recoding into binomial values. Each combination was tested (eg, stage I vs II/III/IV; stages I/II vs III/IV). To adjust the association of a given variable to the 2 ALCL clusters for the effect of the other variables, we performed Mantel-Haenzel tests.

t Tests

We performed all univariate t tests using BRB ArrayTools software on the RNA normalized data. We selected probe sets having a P value below a given threshold (.001, except when indicated). To evaluate the number of false positives in this selection (because of multiple testing), we used a Monte-Carlo approach (implemented in BRB comparison tool19) based on 1000 random permutations of the label of the samples. Having ordered the list of probe sets by P value, this method evaluates the rank n for which the false discovery rate is lower than a chosen percentage (with a probability of 90%), where false discovery is relative to the chosen level of significance for the univariate test. This method also evaluates the rank n′, for which the number of false discoveries is less than a chosen number (with a probability of 90%).

Functional annotation using gene ontology

To assess the biologic functions and pathways associated with ALCL, we measured the underrepresentation and overrepresentation of gene ontology (GO) terms (association of gene products with regard to their associated biologic processes, cellular components, and molecular functions) in gene clusters derived from expression data with the use of FatGO19 and hypergeometric tests provided in the R package GOstats, version 1.1.1. We performed independent analyses of GO categories for overexpressed genes for C1’ relative to C2’ and C2’ relative to C1’. We performed tests on nonredundant lists of Entrez Gene identifiers (we used hgu133a meta-data library [version 1.7.0] to obtain Entrez Gene identifiers corresponding to a given list of probe sets).

From www.bloodjournal.org by guest on January 10, 2018. For personal use only.
Quantitative reverse transcription-PCR

cDNA were reverse transcribed from 2 μg total RNA, using Superscript II reverse transcriptase (Invitrogen, Cergy Pontoise, France) with oligo-dT as primers. Absence of genomic DNA contamination was checked by running RT-PCR with Raf8/raf9 primers, which hybridized on distinct exons spanning an intronic region. qRT-PCR was then performed on the cDNA diluted at 1:10 in sterile water, on an ABI7000 (Applied Biosystems, Courtaboeuf, France) with SYBR Green PCR reagents in a 25 μL reaction mixture (SYBR Green Jumpstart Taq Ready Mix; Sigma, Lyon, France). All primers were designed with Vector NTI advance 9.0 software (Invitrogen), and all spanned the region on the target gene covered by the corresponding Affymetrix probe set. Primers were used at a final concentration of 300 nM. Their sequences are listed in Table 1. All samples were run in duplicate. PCR was performed for 40 cycles at 95°C for 15 seconds and at 60°C for 1 minute after initial incubation at 50°C for 2 minutes and 95°C for 10 minutes. PCR product specificity was evaluated by generating a dissociation curve. The geometric mean Ct of RPS5A, RPS18, and TBP was used as a normalization factor to calculate ΔCt values [ΔCtgene = Ctgene - mean(CtRPS5A, CtRPS18, CtTBP)].

Immunohistochemistry

Immunostaining was performed on paraffin sections of patients analyzed by Affymetrix arrays and additional patients included in an ALCL tissue microarray (TMA). Three antibodies were used: anti–C/EBPβ mAb (code H7; DakoCytomation, Santa Cruz Biotechnology, Santa Cruz, CA), anti–BCL-6 mAb (code H7; DakoCytomation, Trappes, France) and anti-serpinA1 polyclonal antibody (code A0012; DakoCytomation). Antibody binding was detected using the streptavidin-biotin-peroxidase complex (ABC) method and the Dako StreptABCComplex/HRP Duet (mouse/rabbit) kit (code K492; DakoCytomation).

Results

Pattern discovery

Based on extensive cluster analysis using a series of the most varying genes across 32 samples (25 patients ALK+ and 7 patients ALK−) in combination with different linkage methods, we found a robust partition of the ALCL tumor samples. Figure 1A shows a dendrogram of the 32 ALCL samples using the top 2.5% most varying expression profiles (n = 495 probe sets; Table S1) and complete-linkage hierarchical clustering. Two distinct clusters were found, designated C1 and C2. ALK+ patients clustered into 2 subbranches of the C1 and C2 groups (Figure 1A). Patients in cluster C1 primarily had ALCL with common type morphology (Figure 2A-B), whereas patients in group C2 primarily had morphologic variant ALCL (P = .005) (Figure 2C). In addition, the perivascular pattern of neoplastic cells was mostly found in the latter group (P = .009) (Figure 2D). When the same approach was applied to the 25 ALK+ samples using the top 1% most varying expression profiles (n = 198 probe sets; Table S1), segregation of the samples into the same 2 clusters was strongly maintained because each of ALK+ patients was assigned to the same cluster (now designated C1’ and C2’; Figure 1B). In addition to the differentiating morphologic features, patients in group C1’ had Ann Arbor stage II/III at diagnosis (P = .006), whereas those in group C2’ had advanced-stage disease (Ann Arbor III/IV). We found a significant number of patients in group C2’ who experienced early relapse (at most, 1 year after diagnosis; P < .03). Among the 841 probe sets (746 unique Entrez Gene identifiers; Table S2) differentially expressed between the C1’ and C2’ clusters, we found a statistically significant overrepresentation of genes involved in inflammatory and immune response, response to external stimulus, and stress response in the C1’ group. Moreover we found up-regulated genes involved in the leukocyte transendothelial migration pathway in the C1’ group, which is a key event in host defense, cytokine–cytokine receptor interaction, and natural killer cell–mediated cytotoxicity. On the other hand, genes overrepresented and up-regulated in the C2’ group were implicated in the regulation of the cell cycle and in cell proliferation. We also found the up-regulation of pathways such as purine and pyrimidine metabolism, glycolysis, and gluconeogenesis.

Molecular signature associated with ALK status in ALCL

To determine the most significant differentially expressed genes between ALK+ ALCL and ALK− ALCL, 4 ALCL ALK+ and 1 ALK− cell lines were included in the t test in addition to the 32 ALCL samples (25 patients with ALK+, 7 patients with ALK−), which gave more weight to genes expressed by the tumor cells. Among the list of 303 genes differentially expressed (Table S3), 117 genes were overexpressed in ALK+ samples, and 186 were overexpressed in ALK− samples. Among the 117 genes overexpressed in patients with ALK+ ALCL, several genes were highly statistically discriminatory between ALK+ and ALK− samples (P < .001) (Table S3). The 4 top genes (excluding ALK) were BCL6 (fold-change, 2.45), PTPN12 (tyrosine phosphatase) (fold-change, 2.28), CEBPB (CCAAT/enhancer-binding protein β) (fold-change, 2.05), and SERPIN1A (alpha-1 antitrypsin: a1-AT or αPI) (fold-change, 5.12).

The most significant overrepresented GO terms (biologic processes, cellular components, and molecular functions associated with gene products) in the ALK+ group were related to immunologic functions (immune response, response to pathogen and parasite) and the I-kB kinase/NF-κB cascade.20 Looking more specifically at the pathways associated with ALK+ patients, we found an overrepresentation of pathways such as the leukocyte transendothelial migration (ITGAM, PXN, JAM3), focal adhesion, and adherens junction (CTNNB1/-catenin, RAC1). In contrast, the functional categories of the genes overexpressed in ALK− ALCL did not provide any obvious clues to the molecular mechanism(s) underlying this tumor subtype because they tended to be associated
with nonspecific biologic functions. We only found a statistically significant overrepresentation of genes up-regulated in these cases such as CCR7, CNTFR, IL22, and IL21, known to be involved in the cytokine–cytokine receptor interaction pathway.

Molecular signatures associated with morphologic features in ALCL

We found 491 probe sets corresponding to 447 genes that statistically differentiated common type ALCL and morphologic variants (only ALK+ patients were considered because ALK− [7 patients] was of common morphology) (Table S4). Of these, 248 genes were overexpressed in morphologic variants compared with common type ALCL, with a clear overrepresentation of 66 genes grouped together in the GO term cell communication. Most of these genes encode proteins involved in adhesion and migration, such as the focal adhesion kinase PYK2 (or PTK2B), a cytoplasmic tyrosine kinase of the FAK family. Genes coding for proteins that belonged to the focal adhesion complex and that played a critical role in the integrin-dependent cell motility—such as PLCG2, the catalytic subunits of the phosphoinositide 3-kinase (PIK3CA, PIK3CD) and several small G-proteins (ARHGDIR, GNB1, GRK5, RALB, RASSF2, RGS19, GNA13, CENPD2, DOK2, ARHGEF2, CNTD1, GPR18, ABR, RHOG)—were also overexpressed. We also found overexpression of the SYK gene encoding a tyrosine kinase protein, which becomes activated upon integrin-mediated adhesion and several genes encoding integrin-binding adhesion coreceptors, known to interact with these intracellular proteins, such as metalloproteases (ADAMTS2, ADAM12, ADAMDEC1) and molecules of the tetraspanin superfamily (CD37, CD53).

Confirmation of differential gene expression by qRT-PCR

To validate microarray experiment data, qRT-PCR was performed on 6 selected genes (ALK, CEBPB, TMD1, MAC30, CCR7, GREM1), in the list of differentially expressed genes between ALK+ and ALK− samples (Table S3). We selected these genes based on a prediction analysis of ALK status (data not shown). Three of them (ALK, CEBPB, TMD1) were highly significantly differentially expressed between ALK+ and ALK− samples. Another 12 ALCL samples not analyzed with Affymetrix arrays were included in the analysis to test the validity of the results. Highly concordant results (median Pearson correlation coefficient between qRT-PCR and Affymetrix data, 0.86) were obtained for all 6 genes with statistically significant differences between ALK+ and ALK− tumors. Concordant results were also obtained by analyzing 10 randomly selected genes (data not shown).

Immunohistochemical validation of overexpressed genes using tissue microarrays

Among genes overexpressed in ALK+ ALCL, we assessed the protein expression of C/EBPB, BCL-6, and serpinA1. These proteins were chosen because corresponding antibodies suitable for use on tissue microarrays were commercially available. C/EBPB expression was examined in 86 patients (65 ALK+, 21 ALK−), including 16 patients investigated with the use of Affymetrix. Strong nuclear staining for C/EBPB was detected in 97% (63 of 65) of ALK+ patients and in only 4.8% (1 of 21) of ALK− patients (Figure 3). Forty (81.6%) of the 49 ALK+ patients tested for BCL-6 showed a nuclear staining of neoplastic cells, whereas only 2 (28%) of the 7 ALK− patients were positive (P = .002). Twenty-nine (78%) of the 37 ALK+ patients were positively stained for serpinA1 whereas only 2 of the 9 ALK− patients showed a cytoplasmic serpinA1 expression in tumor cells (P < .003).

Discussion

The recent development of DNA microarrays has provided new biologic insight into lymphoproliferative disorders. However, this approach has mainly been used for B-cell lymphomas, in particular diffuse large B-cell lymphoma,21-23 primary cutaneous large B-cell lymphoma,25 primary mediastinal B-cell lymphoma,26 mantle-cell lymphoma,27 B-cell chronic lymphocytic leukemia,28 hairy-cell leukemia,29 follicular lymphoma,30 and multiple myeloma.31 Using gene-expression profiling, some of these lymphoid neoplasms display a distinct molecular signature related to the cell of origin (eg, diffuse large B-cell lymphoma),21,23 and in some tumors (eg, diffuse large B-cell lymphoma and follicular lymphoma),23,30 it is even possible to predict outcome on the basis of individual genes expressed by nonneoplastic infiltrating cells. By contrast, few studies have been dedicated to peripheral T-cell lymphomas.
T-cell lymphomas represent a heterogeneous group of tumors that could benefit from DNA microarrays for defining unrecognized subcategories, identifying new oncogenic pathways or pathogenetically relevant genes and biologic predictors for their clinical course. With the exception of previous reports based on cell lines derived from T-cell malignancies,36–38 few studies involve biopsy specimens from patients with T-cell lymphomas.32–35 In the study by Martínez-Delgado et al,32 gene-expression profiles have been analyzed in 34 peripheral T-cell lymphoma pertaining to different categories. They found only 35 genes distinguishing all peripheral T-cell lymphomas from reactive lymph node, but they were unable to establish the molecular signature of each category. The latter series included only 5 patients with systemic ALCL.

Our study is the first report using oligonucleotide microarray gene-expression profiling in a homogeneous series of 32 patients (25 ALK+, 7 ALK–) with systemic ALCL. In contrast to Thompson et al,33 only systemic ALK– ALCL fulfilling strict criteria were included in the present study (see “Patients, materials, and methods”). The main goal of the present study was to address 2 major open questions regarding ALCL. First, do ALK+ and ALK– ALCL tumors have different gene-expression profiles, suggesting that they are different entities involving distinct oncogenic mechanisms? Second, are the various morphologic features of ALCL (eg, common type vs morphologic variants) merely morphologic variations of the same disease entity with a single molecular signature or are they associated with different gene profiling expression?

Although primary systemic ALK+ ALCL is a well-defined entity in the WHO classification, so-called ALK– ALCL is still the subject of controversy. In our unsupervised analysis (Figure 1A), ALK– patients clustered into 2 subbranches in C1 and C2, further suggesting that ALK– ALCL is heterogeneous and belongs to entities different from those for ALK+ ALCL. These results are in agreement with the recent study by Ballester et al35 showing a clear segregation between ALK+ and ALK– patients by using hierarchical clustering analysis of T-cell non-Hodgkin lymphomas. In the present study, the supervised analysis by class comparison between ALK+ and ALK– ALCL tumors provided distinct molecular signatures. The molecular signature of ALK– ALCL includes overexpression of 4 genes with the most significant P value—CCR7, CNTFR, IL22, and IL21—but did not allow identification of the underlying oncogenic mechanism associated with these tumors. Among the genes overexpressed in ALK+ ALCL, BCL6, PTPN12, SERPINA1, and CEBPB are the 4 top genes overexpressed with the most significant P value.

Overexpression of the BCL6 proto-oncogene in ALK+ ALCL was confirmed by immunohistochemistry. These results are in agreement with previous reports, although the percentage of positive cases (approximately 40% in previous studies, including ALK+ and ALK– patients) was lower than in our study (overall, 75% of all patients).39,40 Such a discrepancy could be explained by the overrepresentation of ALK+ patients in our series. BCL-6 is known to be predominantly expressed in germinal center B lymphocytes and their neoplastic counterparts,41 and there is no clear explanation concerning its expression in T-/null-cell ALCL.39 However, BCL-6 is also expressed by a subpopulation of germinal center and perifollicular CD4+ T lymphocytes and by a small population of CD30+ cells localized in the perifollicular area and at the edge of germinal centers.39,42 These CD4+/BCL-6+ T cells might correspond to the recently described follicular B-helper T cells (Tfh),43 but the latter cells are characteristically positive for CXCL13, a chemokine that is exceptionally expressed in ALCL and, in contrast to CXCL12, not found in the present study. Moreover, BCL-6 overexpression in ALK+ ALCL is in agreement with the down-regulation of the cyclin D2 gene, also found in these tumors (data not shown) because the latter gene is known to be negatively regulated by BCL-6.45

PTPN12 (also named PTP-PEST) was the second statistically discriminatory gene overexpressed in ALK+ samples. The expression of this tyrosine phosphatase, originally cloned from a human colon cancer cDNA library,46 has never been reported in lymphoma. This phosphatase has been described as a negative regulator of c- abl kinase activity.57 We should note that we have recently shown that SHP1, another protein tyrosine phosphatase, is also expressed in ALK+ ALCL and involved in NPM-ALK dephosphorylation.48

A potentially interesting gene overexpressed in ALK+ ALCL is the SERPINA1 gene. SerpinA1 is a major plasma serine proteinase inhibitor that is mostly synthesized by the liver but is also
synthesized by a number of other cells, including monocytes/macrophages. On the basis of morphologic ground and serpinA1 expression, some hematopoietic neoplasms were originally diagnosed as malignant histiocytosis or true histiocytic lymphoma and were later reclassified as ALCL. The overexpression of serpinA1 in ALK+ ALCL is in agreement with the results reported by Villalva et al using subtractive suppressive hybridization technique between an ALK+ ALCL cell line and an ALK− ALCL tumor. Of note, increased serpinA1 levels in sera or tumor tissues from various cancer patients have been reported. Interestingly, in these studies, patients with tumors positive for serpinA1 had worse prognoses than those with negative ones. The adverse prognostic effect of serpinA1 expression could be related to its inhibition of natural killer (NK)−cell activity.

C/EBPB is the fourth gene that discriminated ALK+ and ALK− samples and was highly correlated with ALK expression. Overexpression of C/EBPB in ALK+ tumors was confirmed by immunohistochemistry. Recently, Jundt et al demonstrated in ALCL cell lines that the overexpression of C/EBPB, via the Akt/mTOR (mammalian target of rapamycin) pathway activation, led to increased cell proliferation. A body of evidence suggests that C/EBPB proteins play a central role in the balance between cell differentiation and proliferation. This dual activity is based on the ratio between the main isoforms (LAP and LIP) transcribed from a single mRNA. Full-length LAP (liver-enriched activating protein) drives the cell to the differentiation pathway and cycle arrest, whereas the truncated isoform LIP (liver-enriched inhibitory protein) promotes cell proliferation. Because of the lack of specific anti-LIP antibody, we could not determine the expression of the C/EBPB LIP isoform through immunohistochemistry in biopsy specimens. Given that expression of the LIP isoform of C/EBPB was based on activation of the Akt/mTOR pathway, in the absence of available specific ALK inhibitors, therapeutic drugs such as rapamycin targeting this pathway could be envisaged.

The ALK+ signature also included genes implicated in the leukocyte transendothelial migration pathway and adherens junction pathway. Most of the latter genes are involved in the Wnt signaling pathway such as CTNNB1/β-catenin, frequently deregulated in a number of human malignancies and RAC1. It has been demonstrated that in the presence of Wnt, CTNNB1/β-catenin is stabilized and translocated to the nucleus, where it associates with members of the TCF/LEF family of transcription factors, such as TCF4. Interestingly, in ALK+ ALCL, we found an overexpression of TCF4 (3-fold average change), which could account for the overexpression of the Wnt target gene FOSL1 (fos-like antigen or fra-1, 2-fold average change). The deregulation of the β-catenin gene expression could lead to the loss of cell-cell adhesions and might explain the frequent extranodal involvement in ALK+ patients.

Another aim of the present study was to determine whether the various morphologic features of ALCL (common type vs small cell/lymphohistiocytic variants) are merely morphologic variations of the same disease entity and share the same molecular signature or whether they are associated with a distinct gene profile. It could be anticipated that the gene-expression profiles of these 2 groups (common type vs morphologic variants) would be different. Using cluster analysis based on the most varying genes across our series of ALK+ samples, we found a robust bipartition of the ALCL tumor samples. One cluster (C1) consisted mainly of common type ALCL, whereas morphologic variants clustered in the other group (C2), where tumors showing a perivascular disposition of neoplastic cells were also found. Such a clustering may appear to be at odds with previous studies, suggesting that these morphologic variants are variations of a single disease entity. Most important, a significant number of patients in group C2 experienced early relapse (at most, 1 year after diagnosis; P < .03). Among the genes differentially expressed between the C1 and C2 clusters, we found a statistically significant overrepresentation of genes involved in inflammatory and immune response, response to external stimulus, and stress response in the C1 group. Moreover, in the latter group, we found a clear overrepresentation of genes encoding proteins involved in adhesion and migration, particularly in focal adhesion complexes such as PYK2, PLCγ2, PIK3CA, PIK3CD, and several small G-proteins. On the other hand, genes overrepresented and up-regulated in the C2 group were implicated in the regulation of the cell cycle and in cell proliferation. We also found the up-regulation of pathways, such as purine and pyrimidine metabolism, glycolysis, and gluconeogenesis, that could reflect the cellular hyperactive status. To reconcile the concept of a unique entity and the finding of the 2 gene-expression profiles found in clusters C1 and C2, we postulate that in the course of the disease, some patients may switch on or switch off a set of genes, resulting in different morphologic features.

In summary, with the use of gene-expression profiling, we show that ALK+ and ALK− ALCLs have different gene-expression profiles, further confirming that they are different entities. In addition, preliminary results of the present study may have clinical implications because they suggest that tumors in patients who experience relapse within 1 year of diagnosis express a set of genes that, if confirmed on a larger number of patients, could be of clinical and therapeutic value.

Acknowledgments

This work was supported by the Ligue Contre le Cancer, Comités de la Haute-Garonne et de l’Aveyron, and the program Carte d’Identité des Tumeurs, Fondation de France, Cancéroplôle Grand Sud-Ouest: projet cibles thérapeutiques dans les hémopathies malignes, Conseil Régional Midi-Pyrénées and Contrat Interface INSERM-CHU Purpan (Toulouse, France) (L.L.).

We thank Michel March for his excellent technical work performing immunohistochemistry. We also thank Nicole Dastugue for assistance with cytogenetic work and Fabien Petel (Ligue Nationale Contre le Cancer) for his help formatting the data to be MIAME compliant and submitting the data to the EBI. We thank the Société Française des Cancers de l’Enfant and the Groupe d’Etude des Lymphomes de l’Adulte for providing clinical annotations. We thank our fellow pathologists Dr Fanny Gaillard (CHU de Nantes) and Dr Frédérique Djoud (CHU de Lyon) for providing frozen samples.

Authorship

Conflict-of-interest disclosure: The authors declare no competing financial interests.

L.L. and A.D.R. contributed equally to this work. E.E. and G.D. are senior coauthors.

Correspondence: Georges Delsol, Centre de Physiopathologie Toulouse-Purpan, Inserm U-563, Bât. B, CHU-Purpan, Place du Docteur Baylac, BP 3028, 31024 Toulouse Cedex 3, France; e-mail: gdelsol@toulouse.inserm.fr.

References


Gene-expression profiling of systemic anaplastic large-cell lymphoma reveals differences based on ALK status and two distinct morphologic ALK+ subtypes

Laurence Lamant, Aurélien de Reyniès, Marie-Michèle Duplantier, David S. Rickman, Frédérique Sabourdy, Sylvie Giuriato, Laurence Brugières, Philippe Gaulard, Estelle Espinos and Georges Delsol