Whole-genome profiling of primary cutaneous anaplastic large cell lymphoma


1Department of Dermatology, Leiden University Medical Center and 2Sequencing Analysis Support Core, Leiden University Medical Center, Leiden, the Netherlands

*ANBT and RCM contributed equally as co-first authors.

Abstract

Primary cutaneous anaplastic large cell lymphoma (pcALCL), a hematologic neoplasm caused by skin-homing CD30+ malignant T cells, is part of the spectrum of primary cutaneous CD30+ lymphoproliferative disorders. To date, only a small number of molecular alterations have been described in pcALCL and, so far, no clear unifying theme that could explain the pathogenetic origin of the disease has emerged among patients. In order to clarify the pathogenetic basis of pcALCL, we performed high-resolution genetic profiling (genome/transcriptome) of this lymphoma (n=12) by using whole-genome sequencing, whole-exome sequencing and RNA sequencing. Our study, which uncovered novel genomic rearrangements, copy number alterations and small-scale mutations underlying this malignancy, revealed that the cell cycle, T-cell physiology regulation, transcription and signaling via the PI-3-K, MAPK and G-protein pathways are cellular processes commonly impacted by molecular alterations in patients with pcALCL. Recurrent events affecting cancer-associated genes included deletion of PRDM1 and TNFRSF14, gain of EZH2 and TNFRSF8, small-scale mutations in LRP1B, PDPK1 and PIK3R1 and rearrangements involving GPS2, LINC-PINT and TNK1. Consistent with the genomic data, transcriptome analysis uncovered upregulation of signal transduction routes associated with the PI-3-K, MAPK and G-protein pathways (e.g., ERK, phospholipase C, AKT). Our molecular findings suggest that inhibition of proliferation-promoting pathways altered in pcALCL (particularly PI-3-K/AKT signaling) should be explored as potential alternative therapy for patients with this lymphoma, especially, for cases that do not respond to first-line skin-directed therapies or with extracutaneous disease.

Introduction

Primary cutaneous anaplastic large cell lymphoma (pcALCL) is a hematologic neoplasm produced by malignant CD30+ large T cells that infiltrate the skin forming solitary or localized tumors. pcALCL is largely an anaplastic lymphoma kinase (ALK)-negative lymphoma with only ~2% of ALK positive (ALK+) cases. Patients with pcALCL have a favorable prognosis with a 5-year survival of 90%; yet, a minority of patients (~10% of cases) develop extracutaneous dissemination with a more aggressive clinical course.

Unlike systemic anaplastic large cell lymphoma (sALCL), which appears to be driven by a variety of genetic alterations that overactivate STAT3 signaling, the pathogenetic basis of pcALCL has remained elusive. Previous studies using cytogenetic techniques, array-based platforms (DNA, RNA and miRNA) and whole-exome sequencing (WES) have only identified a few recurrent molecular alterations in this lymphoma; thus, genetic abnormalities underlying pcALCL are still largely unknown. Chromosomal rearrangements involving ALK, DUSP22, TP63 and TYK2, common in sALCL, have also been detected in pcALCL, albeit at much lower frequencies. Frequent copy number alterations (CNA) include losses within chromosomes 6 and 13 and gains within chromosomes 7 and 17. Overexpressed genes with potential pathogenic and/or diagnostic relevance include EZH2 and IRF4, Oncomir mirR-155, which is associated with other cutaneous T-cell lymphoma (CTCL) variants, is overexpressed in pcALCL too. Finally, single nucleotide variants (SNV) presumed to be pathogenic have been reported in JAK1, MSC and STAT3 in a few cases.

Given the lack of a clear unifying theme between the few
reported molecular alterations in the disease, we performed a genome-wide analysis (genome/transcriptome) of_pcALCL using whole-genome sequencing (WGS), WES and RNA sequencing (RNA-seq) to investigate genetic defects underlying this neoplasm. Our data show that molecular alterations in genes involved in the regulation of the cell cycle, T-cell physiology, transcription and signaling via the PI-3-K, MAPK and G-protein pathways underlie pcALCL. Hence, the inhibition of proliferation-promoting pathways operative in pcALCL (especially PI-3-K/AKT signaling) could be explored as an alternative therapy for pcALCL, in particular, for patients who do not respond to skin-directed therapies or with extracutaneous dissemination.

Methods

Patient selection and sequencing

Frozen tumor biopsies (≥75% tumor cells) from 12 patients with (ALK–) pcALCL (Online Supplementary Figure S1; Online Supplementary Table S1) were subjected to WGS and RNA-seq. Seven matched tumor/germline (granulocytes) samples from this cohort (i.e., cAL2-5/9-11) were additionally subjected to WES. Details of sequencing, data processing and DNA/RNA analyses are provided in the Online Supplementary Appendix (Online Supplementary Figures S2 to S4; Online Supplementary Tables S2 to S18). Diagnosis was performed by an expert panel of pathologists and dermatologists according to the criteria of the World Health Organization-European Organisation for Research and Treatment of Cancer (WHO-EORTC) classification for primary cutaneous lymphomas. Patient material was approved by the Institutional Review Board of Leiden University Medical Center. Informed consent was obtained from patients in accordance with the Declaration of Helsinki.

Validation of genetic alterations and differentially expressed genes

Select rearrangements and SNV were validated by Sanger sequencing. Target sequences were polymerase chain reaction (PCR)-amplified, run on a 1% agarose gel, column-purified and sequenced on the Applied Biosystems ABI3730xl platform (Applied Biosystems) (Online Supplementary Figures S5 and S6). Select CNA were validated by droplet digital PCR (ddPCR) using QX200 ddPCR system (Bio-Rad) following the manufacturer’s guidelines. In short, genomic DNA (20-40 ng) was mixed with a frequent-cutting restriction enzyme, ddPCR supermix and probes against the target/reference genes (FAM/HEX-labeled, respectively). The reaction mix was partitioned into 20,000 nanodroplets and PCR-amplified. FAM/HEX fluorescence was measured with Bio-Rad QX200 droplet reader. Copy number values were determined using Bio-Rad Quantasoft software v1.7.4. (Online Supplementary Table S19). Select differentially expressed (DE) genes were validated by reverse transcriptase quantitative PCR (RT-qPCR) as described by van Doorn et al.17

Immunohistochemistry

Formalin-fixed paraffin-embedded (FFPE) tumor sections from sequenced patients with available tissue material (i.e., cAL3-4/7-12) were immunohistochemically stained with primary antibodies against phospho-AKT (Abcam, Cat. No. ab38449) and phospho-STAT3 (Cell Signaling Technology, Cat. No. 9145) following the manufacturer’s instructions, counterstained in Mayer’s Hematoxylin solution and coverslipped using DPX-mountant (Sigma-Aldrich, Cat. No. 06522).

Results

Landscape of genomic rearrangements

The landscape of rearrangements of pcALCL showed considerable inter-patient heterogeneity. The number of rearrangements varied from one to 51 per patient (205 total events; mean/patient ± standard deviation [SD], 17 ±17) (Figures 1 and 2A). Sixty-one percent of events were intrachromosomal (range/patient, 0–100%) (Figure 2B). We identified events fusing two annotated genes (30%), a gene with a non-genic region (42%), two non-genic regions (25%) or reordering sequences within a single gene (3%). Ten percent of rearrangements led to the expression of fusion transcripts (mean fusions/patient, 1.6; range, 0–4 fusions/patient) (Figure 2C). Complex rearrangements involving chromosome 3 and 7 were observed in patients cAL8 and cAL1, respectively (Online Supplementary Figure S7).

We detected a total of 166 rearranged genes in our cohort (Online Supplementary Table S8), including a subset of 44 genes implicated in cancer (Figure 1; Online Supplementary Table S8). The vast majority of events were patient-specific. DUSP22 and TP63 rearrangements, reported in pcALCL before,26 were found in single cases (Online Supplementary Figure S6). Gene ontology annotation with DAVID revealed that rearranged genes were involved in multiple cellular processes/pathways, several of which have relevant roles in cancer onset and progression (e.g., transcription, cell cycle, chromatin modification, MAPK pathway) (Figure 2D; Online Supplementary Table S7). Analysis of rearranged genes with NCG 6.0 uncovered cancer genes principally associated with three cellular processes/pathways: the cell cycle (i.e., FNN2, FOXM1, HORMAD1, KMT2E, LINC-PINT, NUMA1, TP53, TP63, USP7), signaling via the PI-3-K and MAPK pathways (i.e., DGKI, ERBB2IP, GPS2, GRM8, MAP3K20, MET, PIK3C2B, PPML,
PTPRK, TNK1) and the regulation of RNA expression/pro-cessing (i.e., ETS1, RBFOX1, TRERF1, ZNF208, ZNF28) (Online Supplementary Figure S6).

A group of five genes were found to be recurrently re-arranged in pcALCL: GPS2 (2 patients), LINC-PINT (3 pa-tients), RBFOX1 (2 patients), TNK1 (2 patients) and VPS13D (2 patients). Notably, four of five of these genes encode products with roles in the cell cycle, transcription or sig-naling. LINC-PINT is a long non-coding RNA that acts as a TP53-dependent regulator of the cell cycle and a PRC2-dependent silencer of gene expression.18 RBFOX1 is an RNA-binding protein that regulates alternative splicing.19 GPS2 and TNK1 are negative regulators of RAS/JNK1 and RAS/RAF1 signaling, respectively.20,21 Since deregulated RAS-MAPK signaling is a frequent driver of human cancers, we looked at potential detrimental ef-fects of the rearrangements on the expression of GPS2 and TNK1. The expression data confirmed that both genes were functionally compromised in the affected patients. For instance, patient cAL4 expressed a truncated GPS2-CHD3 transcript which rendered both fusion partners dys-functional (Figure 2E, F). Similarly, patient cAL3 expressed a transcript encoding a truncated form of TNK1 which was structurally analogous to a variant with proven oncogenic activity reported in Hodgkin lymphoma (Figure 2G, H).22

**Landscape of copy number alterations**

pcALCL displayed a predominance of large numerical im-balances which included losses within 3q, 6q, 7p, 7q, 8p, 13q, and 16q as well as loss of chromosome Y. Also, gains within 1q, 2p, 2q, 7q and 12q (Figures 3A and 4). Our analy-sis detected 15 focal (≤3Mb) minimal common regions (MCR) between CNA found in our patients (9 deletions, 6 gains), six of which enclosed (putative) cancer genes (Figure 4; Online Supplementary Table S6). Half of these genes participate in the regulation of T-cell physiology. Deletion at 6q21, observed in six of 12 patients, was the most recurrent focal MCR in pcALCL. This region contained tumor suppressor PRDM1 (BLIMP1) (Figure 3B, C), which en-codes a transcription factor that attenuates the prolifer-a tion and survival of T cells through suppression of IL-2 signaling.23 Three of 12 patients had deletions at 13q32.2 which enclosed STK24 (Figure 3D). This gene encodes a pro-apoptotic serine/threonine kinase that modulates phos-phorylation of JNK and p38 kinases in response to stress stimuli.24 Two of 12 patients had deletions at 1p36.32 and
Figure 2. Rearrangements disrupt RAS-MAPK signaling inhibitors in primary cutaneous anaplastic large cell lymphoma. (A) Number of genomic rearrangements per patient. The distribution of inter- and intrachromosomal rearrangements per patient is shown. (B) Distribution of inter- and intrachromosomal rearrangements (cohort). (C) Distribution of genomic rearrangements based on the type of DNA sequences involved in the event (genic or nongenic) and the expression of fusion sequences determined through integration of whole-genome sequencing and RNA sequencing data (cohort). (D) Gene ontology annotation of rearranged genes. (E) Validation of GPS2 rearrangements in patients cAL3 and cAL4 by Sanger sequencing. (F) Diagram of fusion gene GPS2-CHD3 in patient cAL4 and Sashimi plots showing the expression of CHD3 in patients cAL4 (disrupted by rearrangement, purple) and cAL6 (with intact alleles, green). Besides disrupting the signaling-suppressing domain of GPS2, the fusion event also inactivated CHD3, the helicase subunit of chromatin remodeling complex Mi-2/NuRD. In patient cAL4, the GPS2-fused CHD3 transcript (5' side of the gene) ends prematurely. A nonsense-mediated decay (NMD) transcript (3' side of the gene) is expressed too. (G) Validation of TNK1 rearrangements in patients cAL3 and cAL4 by Sanger sequencing. (H) Sashimi plots showing the expression of TNK1 in patients cAL3/4 (disrupted by rearrangements, purple) and cAL7 (with intact alleles, green). Patient cAL3 expressed a transcript encoding a C-terminus truncated TNK1 protein which was structurally analogous to a known oncogenic variant. The expression of TNK1 was abolished in patient cAL4.
16p13.3, which encompassed TNFRSF14 (HVEM) and RBFOX1, respectively. TNFRSF14 is a multifaceted receptor capable of sending stimulatory or inhibitory signals to T cells depending on the ligand it binds.\textsuperscript{25} Of note, loss of TNFRSF14 has been reported in pcALCL before.\textsuperscript{13} In addition, focal (≤3Mb) gains at 1p36.22 and 2p15 were observed in two and three of 12 patients, respectively. The former contained TNFRSF8 which encodes receptor CD30 (Figure 3D), whose overexpression is a hallmark of pcALCL. The latter enclosed XPO1, a gene that encodes a protein involved in the export of a plethora of proteins and RNA from the nucleus to the cytoplasm.\textsuperscript{26} Furthermore, six of 12 patients had gains at 7q36.1 (MCR ~ 6 Mb), which contained EZH2, the catalytic subunit of gene silencing complex PRC2 (Figure 3D).

**Landscape of small-scale mutations and mutational signatures**

In order to detect somatic small-scale mutations (indels, SNV) in pcALCL, we performed WES on matched tumor/germline samples from seven patients. The number of somatic indels/SNV ranged from 44 to 246 (653 total events; mean/patient ± SD, 93±70). The average somatic mutation rate was 15.8 mutations/Mb (range/patient, 9.2–38.6 mutations/Mb). The majority of small-scale mutations were missense SNV (Figure 5A; Online Supplementary Table S9). The most frequent base substitution was C>T (mean/cohort, 56.2%), (Figure 5B), which was attributable to ultraviolet (UV) light exposure and aging-associated spontaneous deamination of 5-methylcytosine. DNA double-strand break (DSB) repair deficiency and activation-induced cytidine deamination were also observed in a subset of patients.

### Table 1. Fusion transcripts detected by RNA sequencing in primary cutaneous anaplastic large cell lymphoma.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Fusion transcript</th>
<th>Breakpoints (DNA)</th>
<th>Breakpoint type</th>
<th>Event class</th>
<th>WGS confirmed</th>
</tr>
</thead>
<tbody>
<tr>
<td>cAL1</td>
<td>NLRP1–DERL2</td>
<td>chr17:5554176 - chr17:5477401</td>
<td>Genic - Genic</td>
<td>iDel</td>
<td>Yes</td>
</tr>
<tr>
<td>cAL1</td>
<td>ANKRD11–VPS9D1-AS1</td>
<td>chr16:89306818 - chr16:89713470</td>
<td>Genic - Genic</td>
<td>ITX</td>
<td>Yes</td>
</tr>
<tr>
<td>cAL1</td>
<td>INPP5B–GPN2</td>
<td>chr1:37926300 - chr1:37586211</td>
<td>Genic - Genic</td>
<td>iDel</td>
<td>Yes</td>
</tr>
<tr>
<td>cAL1</td>
<td>IGSF3–CD58</td>
<td>chr1:116578013 - chr1:116525260</td>
<td>Genic - Genic</td>
<td>iDel</td>
<td>Yes</td>
</tr>
<tr>
<td>cAL1</td>
<td>TNK1–GPS2</td>
<td>chr1:37926300 - chr1:37586211</td>
<td>Genic - Genic</td>
<td>iDel</td>
<td>Yes</td>
</tr>
<tr>
<td>cAL1</td>
<td>TP53–TNK1</td>
<td>chr1:116578013 - chr1:116525260</td>
<td>Genic - Genic</td>
<td>iDel</td>
<td>Yes</td>
</tr>
<tr>
<td>cAL1</td>
<td>GPS2–CHD3</td>
<td>chr1:37926300 - chr1:37586211</td>
<td>Genic - Genic</td>
<td>iDel</td>
<td>Yes</td>
</tr>
<tr>
<td>cAL1</td>
<td>LINC-PINT–RNASEH2B-AS1</td>
<td>chr1:116578013 - chr1:116525260</td>
<td>Genic - Genic</td>
<td>iDel</td>
<td>Yes</td>
</tr>
<tr>
<td>cAL1</td>
<td>SMARCAD1–TP63</td>
<td>chr1:37926300 - chr1:37586211</td>
<td>Genic - Genic</td>
<td>iDel</td>
<td>Yes</td>
</tr>
<tr>
<td>cAL1</td>
<td>TP63–SMARCAD1</td>
<td>chr1:37926300 - chr1:37586211</td>
<td>Genic - Genic</td>
<td>iDel</td>
<td>Yes</td>
</tr>
<tr>
<td>cAL1</td>
<td>ETS1–SLC37A4</td>
<td>chr11:128520859 - chr11:119027083</td>
<td>Genic - Genic</td>
<td>iDel</td>
<td>Yes</td>
</tr>
<tr>
<td>cAL1</td>
<td>NFKB2–SIRT3</td>
<td>chr1:102401366 - chr1:102401164</td>
<td>Genic - Genic</td>
<td>iDel</td>
<td>Yes</td>
</tr>
<tr>
<td>cAL1</td>
<td>SIRT3–NFKB2</td>
<td>chr1:128520859 - chr1:119027083</td>
<td>Genic - Genic</td>
<td>iDel</td>
<td>Yes</td>
</tr>
<tr>
<td>cAL1</td>
<td>GTPBP2–TRERF1</td>
<td>chr1:128520859 - chr1:119027083</td>
<td>Genic - Genic</td>
<td>iDel</td>
<td>Yes</td>
</tr>
<tr>
<td>cAL1</td>
<td>USP7–SAMD11</td>
<td>chr16:89306818 - chr16:89713470</td>
<td>Genic - Genic</td>
<td>iDel</td>
<td>Yes</td>
</tr>
<tr>
<td>cAL1</td>
<td>MMS19–RRP12</td>
<td>chr10:97461940 - chr10:97410043</td>
<td>Genic - Genic</td>
<td>iDel</td>
<td>Yes</td>
</tr>
<tr>
<td>cAL1</td>
<td>NFAT5–WWP2</td>
<td>chr16:69595202 - chr16:69801863</td>
<td>Genic - Genic</td>
<td>iDel</td>
<td>Yes</td>
</tr>
<tr>
<td>cAL1</td>
<td>RABEPK–UBAC1</td>
<td>chr9:125226314 - chr9:135964125</td>
<td>Genic - Genic</td>
<td>iDel</td>
<td>Yes</td>
</tr>
<tr>
<td>cAL1</td>
<td>HIVEP3–MRPL43</td>
<td>chr1:128520859 - chr1:119027083</td>
<td>Genic - Genic</td>
<td>iDel</td>
<td>Yes</td>
</tr>
<tr>
<td>cAL1</td>
<td>AGTRAP–LST1</td>
<td>chr1:128520859 - chr1:119027083</td>
<td>Genic - Genic</td>
<td>iDel</td>
<td>Yes</td>
</tr>
<tr>
<td>cAL1</td>
<td>CLK2–NUMA1</td>
<td>chr1:128520859 - chr1:119027083</td>
<td>Genic - Genic</td>
<td>iDel</td>
<td>Yes</td>
</tr>
<tr>
<td>cAL1</td>
<td>KANSL1–ARL17B</td>
<td>chr1:128520859 - chr1:119027083</td>
<td>Genic - Genic</td>
<td>iDel</td>
<td>Yes</td>
</tr>
<tr>
<td>cAL1</td>
<td>ERBB2IP–SQRDL</td>
<td>chr15:46564289 - chr15:46564289</td>
<td>Genic - Genic</td>
<td>iDel</td>
<td>Yes</td>
</tr>
<tr>
<td>cAL1</td>
<td>RBPC1–TRAPP12</td>
<td>chr2:3436901 - chr2:3436901</td>
<td>Genic - Genic</td>
<td>iDel</td>
<td>Yes</td>
</tr>
</tbody>
</table>

CTX: interchromosomal translocation; ITX: intrachromosomal translocation; iDel: interstitial deletion; WGS: whole-genome sequencing.
deaminase (AID)-mediated somatic hypermutation were smaller contributors to the mutational landscape (Figure 5C). A total of 603 genes were found to carry somatic mutations across seven patients, 158 of which were genes associated with cancer (134 candidate cancer genes, 24 bona fide oncogenes/tumor suppressors) (Online Supplementary Table S11). All small-scale mutations were patient-specific. Gene set analysis of mutated genes with Panther ident-

Figure 3. Landscape of copy number alterations reveals recurrent deletion of PRDM1 in primary cutaneous anaplastic large cell lymphoma. (A) Human chromosome ideograms showing regions of gain and loss detected by whole-genome sequencing in 12 primary cutaneous anaplastic large cell lymphoma (pcALCL) genomes. Blue bars to the right of the chromosomes depict regions of loss whereas red bars to the left of the chromosomes depict regions of gain. (B) Diagram representing deletions (light blue bars) in six patients with pcALCL. The minimal common region (MCR) of losses at 6q21 in the analyzed patients maps to PRDM1, a gene encoding a transcription factor that attenuates proliferation and survival of T cells. (C) Sashimi plots showing expression of PRDM1 in patients cAL6 (affected by a focal deletion, purple) and cAL9 (with intact alleles, green). Patient cAL6 expressed a PRDM1 transcript that lacked exons 5, 6 and 7 as a result of a monoallelic focal deletion (blue bar, 4.6 Kb). This structurally defective transcript encodes a protein lacking the Zn finger domains of PRDM1 which are required for its functionality. IDEL: interstitial deletion. (D) Recurrent copy number alterations affecting EZH2, PRDM1, RBFOX1, STK24 and TNFRSF8 in pcALCL were validated by droplet digital polymerase chain reaction. Ctrl: healthy CD4+ T cells.
ified enrichment of genes involved in signaling mediated by G-proteins (β-adrenergic receptors, glutamate receptors, etc.) and PI-3-K signaling (Figure 5D; Online Supplementary Table S12). Analysis of mutated genes with NCG 6.0 database revealed a group of cancer genes with established roles in chromatin modification (i.e., ARID1B, BCOR, KDM5A, PRDM16, TRRAP), hematopoietsis (i.e., FLT3, FSTL3 and GATA2), signaling via the PI-3-K and MAPK pathways (i.e., ERBB4, MAP3K1, PDPK1, PIK3R1, RET and WNK2) and G-protein signaling (i.e., CREB3L1, DRD5, RGS7 and RGS12). Of note, although JAK–STAT mutations leading to STAT3 overactivation are common in sALCL, no mutations in JAK–STAT pathway genes were found in our pcALCL patients.

We identified a group of 30 recurrently mutated genes (n=2–5) in our cohort (Online Supplementary Table S10), which included (putative) oncogenes (i.e., CDK14, CNOT7, LRP2, PDPK1) and tumor suppressors (i.e., CSMD1, LRP1B, PIK3R1) (Online Supplementary Figure S5). Within this group, PDPK1 and PIK3R1 were genes of special interest. PDPK1 encodes a master serine/threonine kinase that activates AKT, a kinase with a central role in the PI-3-K/AKT pathway. Two of 12 patients had mutations in the PH domain of PDPK1 (cAL4: p.T518M; cAL10: p.W535L) (Figure 5E and F). In contrast, PIK3R1 encodes the regulatory (inhibitory) subunit of class IA phosphoinositide 3-kinases (PI-3-K). Class IA PI-3-K mediate the production of PIP3, a membrane phospholipid that induces PI-3-K/AKT pathway activation by recruiting PDPK1 and AKT to the plasma membrane where AKT activation occurs. Activated AKT triggers downstream signaling events leading to cell proliferation and survival. Two of 12 patients had splice site mutations in PIK3R1 (cAL5; G>T at chr5:68,296,342; cAL9: G>A at chr5:68,293,835) (Figure 5E, F).

Figure 4. Distribution of recurrent genetic alterations in primary cutaneous anaplastic large cell lymphoma. First panel: recurrent chromosomal rearrangements impacting cancer genes. Second panel: recurrent large-scale copy number alterations (CNA) (>3 Mb). Third panel: focal minimal common regions (MCR) shared by copy number alterations (CNA). Fourth panel: cancer genes recurrently affected by indels/single nucleotide variants (SNV). CTX: interchromosomal rearrangement; ITX: intrachromosomal rearrangement; Del: deletion.
Figure 5. Landscape of small-scale mutations reveals recurrent events affecting PI-3-K/AKT pathway genes in primary cutaneous anaplastic large cell lymphoma. (A) Number of small-scale mutations per patient. The distribution of missense, nonsense, frameshift and splice site mutations per patient is shown. (B) Distribution of nucleotide substitutions per patient. (C) Median contribution of COSMIC mutational signatures in primary cutaneous anaplastic large cell lymphoma (pcALCL) (n=7) and Sézary syndrome (n=30). AID SHM: activation-induced cytidine deaminase-mediated somatic hypermutation; DSBRD: double-strand break repair deficiency; SS: Sézary syndrome; UV: ultraviolet. (D) Gene set analysis of genes affected by indels and single nucleotide variants (SNV) in pcALCL. Mutated genes were primarily involved in G-protein signaling and the PI-3-K pathway. (E) Diagram showing small-scale mutations affecting \( \text{PIK3R1} \) and \( \text{PDPK1} \) in patients with pcALCL. Splice site mutations in \( \text{PIK3R1} \) are predicted to affect the three known protein isoforms (i.e., p85\( \alpha \), p55\( \alpha \) and p50\( \alpha \)) encoded by this gene. Missense mutations in \( \text{PDPK1} \) impacted the PH domain of its encoded protein. (F) Sanger chromatograms showing validation of splice site and missense mutations in \( \text{PIK3R1} \) and \( \text{PDPK1} \), respectively.
upregulated, 1,446 downregulated, false discovery rate [FDR] >0.01) (Figure 6A; Online Supplementary Table S15). The analysis not only confirmed upregulation of genes found to be overexpressed in pcALCL in prior studies (e.g., CCR8, CCR10, EZH2, IRF4, MIR155HG and TNFRSF8) (Figure 6B), but also revealed upregulation of a group of (proto)-oncogenes (e.g., CSF1R, EGFR, FGFR1/2/3, MET and PDGFRα/B) involved in signaling through the P-1-3K/AKT pathway and/or the MAPK pathway (Figure 6C). Of note, CSF1R overexpression has been reported in sALCL before.27 Downregulated cancer genes included inhibitors of the P-1-3K/AKT pathway (i.e., PIK3IP1 and PIK3R1), inhibitors of the MAPK pathway (i.e., DUSP2, DUSP16, GPS2) and tumor suppressors implicated in other hematological cancers (i.e., FHIT, ROH and TNFAIP3) (Figure 6C).

Twenty-five fusion transcripts were identified in our cohort (Table 1), nine of which were formed by genes with established roles in cancer (i.e., ERBB2IP–SQRDL, ETS1–SLC37A4, GPS2–CHD3, NFAT5–WWP2, SIRT3–NFKB2, SMARCAD1–TP63, TNK1–GPS2, TP53–TNK1 and TP63–SMARCAD1).

Deregulated cellular processes/pathways

In order to investigate deregulated cellular processes/pathways in pcALCL, we analyzed up- and downregulated genes separately with GeneAnalytics. Uregulated canonical signal transduction routes were ERK signaling (enrichment score: 166.10), phospholipase C signaling (enrichment score: 86.47), PAK signaling (enrichment score: 59.21), AKT signaling (enrichment score: 49.09) and G-protein couple receptor (GPCR) signaling (enrichment score: 48.91). In addition, the analysis revealed upregulation of adhesion-related cellular processes (e.g., focal adhesion, integrin pathway) (Figure 6D). In contrast, all downregulated profiles were signaling processes leading to T-cell activation (T-cell receptor (TCR) signaling, Icos–IcosL pathway, etc.) (Figure 6E).

Transcriptome analysis with Enrichr using the ARCS4 database revealed that genes known to be co-expressed with receptor tyrosine kinasenses DDR2 and PDGFR (especially PDGFRβ) were the most overrepresented in the group of upregulated genes (Figure 6F), suggesting that these receptors and/or their associated signal transduction networks might play relevant roles in the disease. DDR2 binds to collagen and activates signaling pathways (e.g., MAPK pathway) implicated in cell adhesion, proliferation and extracellular matrix (ECM) remodeling. Similarly, PDGFR binds to its cognate ligands (PDGFs) and activates the PI-3-K/AKT and MAPK pathways, promoting cell proliferation and survival.

In order to validate the upregulation of DDR2 and PDGFRB in pcALCL, transcript abundance of these genes was quantified by RT-qPCR in pcALCL tumors and a control group formed by benign inflammatory dermatoses (BID) and healthy CD4+ T cells (HC). Both genes were confirmed to be consistently upregulated in pcALCL in comparison to the control group (P<0.0001, Mann-Whitney U test) (Figure 6G). Additionally, we used RT-qPCR to assess in these two groups the expression of tumor suppressors GPS2 and PIK3R1, which displayed recurrent genetic defects and were downregulated in pcALCL according to the sequencing data. The analysis confirmed the downregulation of both tumor suppressors in pcALCL as well (P<0.05, Mann-Whitney U test) (Figure 6H).

Since the PI-3-K/AKT pathway was the most consistently affected canonical signal transduction route in pcALCL (i.e., mutations in PDPK1/PIK3R1, generalized PIK3R1 downregulation, AKT signature), we investigated the presence of activated AKT (pAKT) by immunohistochemistry (IHC) in eight sequenced patients with available tumor tissue (i.e., cAL3-4/7-12). Robust AKT activation was observed in seven of the eight evaluated patients (Figure 7A; Online Supplementary Figure S8). Surprisingly, although our pcALCL patients carried no mutations in JAK-STAT pathway genes, we detected STAT3 activation (pSTAT3) by IHC in three of eight evaluated patients. This suggests that STAT3 signaling might be induced in some cases by alternative non-JAK-STAT molecular alterations or via mutation-independent mechanisms (Figure 7B; Online Supplementary Figures S4 and S8).

Discussion

In order to gain insight into the pathogenetic basis of pcALCL, we investigated rearrangements, CNA, small-scale mutations and gene expression in this lymphoma by analyzing genome (WGS, WES) and transcriptome (RNA-seq) data. In agreement with previous studies, we observed solitary cases with DUSP22 and TP63 rearrangements in our cohort, which are events known to occur at low frequencies in pcALCL.10,28 We also confirmed that losses within 6q and gains within 7q were the most recurrent large-scale chromosomal imbalances. Moreover, similar to other CTCL variants (e.g., mycosis fungoides, Sézary syndrome, cutaneous γ/δ T-cell lymphomas), we found that pcALCL displayed a predominant C>T mutational signature which was mainly attributable to UV light exposure and aging. Although UV-associated mutations appear to be a common feature among distinct CTCL variants, their significance in the lymphomagenesis of this group of neoplasms remains unclear.

Even though pcALCL exhibited a manifest inter-patient heterogeneity in terms of number of genetic abnormalities and affected genes, our analysis showed that the cell cycle, T-cell physiology regulation, transcription and particularly signaling through the PI-3-K and MAPK pathways, were commonly affected cellular processes/pathways in our cohort. Recurrent genetic alterations involving cancer-associated genes included rearrangements involving GPS2, LINC–PINT and TNK1, deletion of PRDM1 and TNFRSF14, amplification of TNFRSF8 (CD30)
and EZH2, and small-scale mutations in CDK14, LRP1B, PDK1 and PIK3R1.
Deletion of PRDM1 and gain of EZH2, each observed in six of 12 patients, were the two most frequent genetic alter-
ations in our cohort. PRDM1, which is recurrently deleted in sALCL too, acts as a tumor suppressor. Prior research
showed that reintroduction of PRDM1 in a sALCL cell line triggered cell proliferation impairment, cell cycle block

Continued on following page.
Figure 6. RNA sequencing supports upregulation of signaling via the PI-3-K/AKT pathway, the MAPK pathway and G-proteins in primary cutaneous anaplastic large cell lymphoma. (A) Heatmap showing 3,162 differentially expressed genes (1,716 upregulated, 1,446 downregulated, false discovery rate [FDR] <0.01) in primary cutaneous anaplastic large cell lymphoma (pcALCL) when compared to CD4+ T cells. (B) Heatmap showing expression of genes with diagnostic and/or pathogenic relevance in pcALCL identified in prior studies. RNA sequencing confirmed the upregulation of several markers and oncogenes that characterize this lymphoma. (C) Heatmap showing (proto-)oncogenes (Proto-OG) and tumor suppressors (TSG) deregulated in pcALCL which are involved in the PI-3-K/AKT pathway and the MAPK pathway. (D) Upregulated cellular processes/pathways in pcALCL as determined by gene set pathway analysis with GeneAnalytics. (E) Downregulated cellular processes/pathways in pcALCL as determined by gene set pathway analysis with GeneAnalytics. (F) Top three human kinases for which known co-expressed genes were overrepresented in the transcriptome of pcALCL. (G) Validation of DDR2 and PDGFRB upregulation in pcALCL by reverse transcriptase quantitative polymerase chain reaction (RT-qPCR). (H) Validation of GPS2 and PIK3R1 downregulation in pcALCL by RT-qPCR. (*P<0.05, ****P<0.0001); pcALCL (n=12), BID (n=8), HC (n=4). BID: benign inflammatory dermatoses; HC: healthy CD4+ T cells.

and apoptosis.28 Notably, PRDM1 reintroduction also downregulated in these cells the expression of MIR155HG (miR-155 precursor) and IRF4, two genes that are consistently overexpressed in sALCL and pcALCL.29 On the other hand, previous studies have shown that EZH2 is overexpressed in pcALCL.114 EZH2 appears to operate as an oncogene in the disease by epigenetically silencing TXNIP and CXCL10, two genes with anti-tumor effects in pcALCL.14 Although the molecular mechanisms leading to EZH2 over-expression in pcALCL require further study, our data suggest that gene amplification may be one of them. Four additional genes implicated in cancer underwent recurrent alterations in pcALCL (LINC-PINT, LRPIB, RBFOX1 and USH2A). LINC-PINT and RBFOX1 behave as tumor suppressors in gastric cancer and glioblastoma.30-33 LRPIB, an LDL family receptor involved in endocytosis, is inactivated by genetic and epigenetic mechanisms in several cancers (gastric, renal, thyroid, etc.).34-36 USH2A, one of the two SNV-mutated genes in our cohort, is the 21st most mutated gene in breast cancer according to The Cancer Genome Atlas (TCGA)37 yet its function remains unclear. Genetic alterations impacting genes associated with the P-I-3K/AKT pathway, the MAPK family of pathways (via ERK, JNK and p38 kinases), and to a lesser extent G-protein signaling, were the most notable common denominator in patients with pcALCL. Nine of 12 patients had abnormalities in one or more genes linked to the aforementioned pathways (Online Supplementary Table S20). Patient-specific genetic alterations, which accounted for the vast majority of events, included rearrangements and small-scale mutations in CREB3L1, DGKI, DARD5, ERBB2IP, ERBB4, GRM8, MAP3K1, MAP3K20, MET, PIK3C2B, PPM1L, PTPRK, RET, RGS7, RGS12 and WNK2. In addition, recurrent mutations in the catalytic subunit of PI-3-K. We observed small-scale mutations affecting PDPK1 or PIK3R1 in four of 12 patients. PDPK1 is frequently amplified or overexpressed in human cancers, but uncommonly affected by pathogenic SNV.40 Surprisingly, we found two patients with SNV (p.T518M; p.W535L) affecting the PH domain of this kinase. PDPK1 becomes constitutively active as a monomer upon dissociation from its inactive homodimeric form.41 As activation of PDPK1 requires autophosphorylation of residues in the PH domain to destabilize the inactive homodimeric state, these mutations might impact the stability of the latter favoring an active monomeric state. This mechanism has been demonstrated in the PH domain mutation p.T513E.42 In contrast, loss-of-function mutations in PIK3R1 occur in glioblastoma, endometrial carcinoma and immune disorders.43 One of the two splice site mutations found in our cohort is causative of PASLI-like disease, a disorder characterized by immunodeficiency, lymphoproliferation, antibody defects and senescence of CD8+ T cells. This splice site mutation (G>A at chr5:68,293,835), which results in exon 11 skipping during mRNA maturation, leads to overactivation of the PI-3-K/AKT pathway in immune cells from these patients.44 The other observed splice site mutation (G>T at chr5:68,296,342) is predicted to result in exon 15 skipping, and consequently, deletion of the C-terminal SH2 domain of the protein which is essential for inhibiting the catalytic subunit of PI-3-K.
In line with the genomic data, transcriptome analysis revealed upregulation of several signal transduction routes associated with the PI-3-K/AKT pathway, the MAPK pathway and G-protein signaling (e.g., ERK, PLC and AKT). The PI-3-K/AKT pathway, the signaling route most prominently altered in our cohort, was confirmed to be activated in our patients through IHC. Thus, suppression of this pathway using PI-3-K inhibitors (e.g., duvelisib) could be explored as an alternative therapy for patients who do not respond to skin-directed therapies or with extracutaneous involvement. Upregulated genes in pcALCL matched expression patterns primarily associated with overexpression of receptor kinases DDR2 and PDGFRβ, suggesting a potential role of these receptors in the development of the disease. The upregulation of DDR2 and PDGFβ in pcALCL was confirmed by RT-qPCR. DDR2 is overexpressed in subsets of patients with sALCL and most Hodgkin lymphoma cases.\textsuperscript{43,44} Similarly, PDGF is commonly overexpressed in sALCL and various other human cancers.\textsuperscript{45,46} Of note, PDGFR blockade with imatinib has been shown to be therapeutically effective in murine models of sALCL and a patient with refractory sALCL.\textsuperscript{45} Since our data suggest that PDGFR is overexpressed in pcALCL too, blockade of this receptor might also have a positive therapeutic effect on patients with pcALCL. However, validation of our molecular findings in large series of patients will be essential before the eventual clinical exploration of these potential therapies.

In agreement with prior reports, our data showed that pcALCL shares some pathogenic features with sALCL such as the overexpression of several markers and oncogenes...
(e.g., CSF1R, IRF4, MIR155HG, PDGFRB and TNFRSF8), the deletion of PRDM1, and the occasional presence of specific rearrangements (e.g., DUSP22 and TP63). However, our cohort displayed recurrent molecular alterations in cancer genes thus far not associated with sALCL (e.g., GPS2, LINC-PINT, PDKPI, PIK3R1 and TNK1). In addition, our patients were devoid of alterations in JAK-STAT pathway genes which are common in sALCL. Still, our results should be interpreted with caution as future whole-genome studies involving larger cohorts are necessary to obtain a more precise estimate of the frequency of molecular alterations underlying pcALCL, including possible JAK-STAT mutations. Despite the absence of JAK-STAT alterations, some of our pcALCL patients exhibited STAT3 activation, suggesting that STAT3 signaling was either induced by non-JAK-STAT molecular alterations (e.g., oncogenic TNK1 variant in patient cAL3) or mediated by mutation-independent mechanisms in these individuals. In line with the latter, ALK- sALCL cell lines lacking JAK-STAT alterations have been found to induce STAT3 signaling through aberrant autocrine cytokine receptor stimulation. Yet, alternative activation mechanisms of JAK-STAT signaling (e.g., signaling crosstalk, epigenetics-related) cannot be excluded.

In summary, our study has uncovered novel genetic alterations with pathogenic significance in pcALCL and established relevant molecular characteristics of this lymphoma. Firstly, pcALCL was found to be genetically heterogeneous. Secondly, pcALCL displayed a mutational signature mainly attributable to UV light and aging with smaller contributions from DSB repair deficiency and AID-mediated somatic hypermutation. Thirdly, some pcALCL cases displayed STAT3 activation even in the absence of canonical JAK-STAT alterations which generally underlie aberrant STAT3 signaling in other lymphomas. Finally, pcALCL appears to develop as a result of defects in genes with roles in the cell cycle, T-cell physiology, transcription, and especially signaling via the PI3-K/AKT pathway, the MAPK pathway and G-proteins.

Disclosures
No conflicts of interest to disclose.

Contributions
ANBT, RM, KQ, RW, MV and CPT conceptualized and designed the project; ANBT, RM and CPT wrote the manuscript; ANBT, LG, HM, CA and TK performed the bioinformatic analyses; ANBT and JO performed the experiments; ANBT and RM analyzed the results, interpreted the data and produced figures and tables; KQ, RW and MV provided valuable biological specimens; ANBT, RM, LG, JO, HM, CA, TK, KQ, RW, MV and CPT revised and approved the final manuscript.

Acknowledgments
The authors thank Julian van Toledo for valuable bioinformatic support.

Funding
This study was funded by the Dutch Cancer Society (KWF, grant UL2013-6104), the Netherlands Organization for Health Research and Development (ZonMw, grant 40-43500-98-4027/435004503), Takeda Nederland B.V. and the Foundation for Pathological Research and Development (S.P.O.O., grant SPOO-2016003).

Data sharing statement
The sequencing data that support the findings of this study are openly available in the European Genome-Phenome Archive (EGA) at https://ega-archive.org, reference number [EGAS00001004429]. Additional data and protocols can be made available to other investigators upon request.

References


