

Universiteit
Leiden
The Netherlands

Exploring the molecular pathogenetic basis of cutaneous lymphomas

Bastidas Torres, A.N.

Citation
Bastidas Torres, A. N. (2020, November 11). *Exploring the molecular pathogenetic basis of cutaneous lymphomas*. Retrieved from https://hdl.handle.net/1887/138190

Version:	Publisher's Version
License:	Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from:	https://hdl.handle.net/1887/138190

Note: To cite this publication please use the final published version (if applicable).

Cover Page

Universiteit Leiden

Leiden University Repository

The handle http://hdl.handle.net/1887/138190 holds various files of this Leiden University dissertation.

Author: Bastidas Torres, A.N.
Title: Exploring the molecular pathogenetic basis of cutaneous lymphomas
Issue Date: 2020-11-11

Exploring the molecular pathogenetic basis of cutaneous lymphomas

Armando N. Bastidas Torres



Exploring the molecular pathogenetic basis of cutaneous lymphomas

Armando N. Bastidas Torres

Exploring the molecular pathogenetic basis of cutaneous lymphomas

© Armando N. Bastidas Torres, 2020

No part of this thesis may be reproduced, stored or transmitted in any form or by any means without prior permission of the author.

The research described in this thesis was carried out at the Departments of Dermatology and Cell and Chemical Biology of Leiden University Medical Center (LUMC), Leiden, The Netherlands. This work was financially supported by the Dutch Cancer Society under grant number UL2013-6104.

ISBN: 978-90-7401-315-4

Cover Design: Vera van Ommeren based on ideas by Armando N. Bastidas Torres
Lay-out and design: Vera van Ommeren, persoonlijkproefschrift.nl
Printed by: Ridderprint|www.ridderprint.nl

To my parents Rosa and Clímaco

Exploring the molecular pathogenetic basis of cutaneous lymphomas

Proefschift

Ter verkrijging van
De graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof. mr. C.J.J.M. Stolker
volgens besluit van het College voor Promoties
te verdedigen op woensdag 11 november 2020
klokke 15:00 uur

door

Armando Nicolas Bastidas Torres
geboren te Quito, Ecuador
in 1985

Promotor:	Prof. dr. M.H. Vermeer
Co-promotor:	Dr. C.P. Tensen
Leden promotiecommissie:	Prof. dr. M. Smit (Vrije Universiteit Amsterdam) Prof. dr. N. Ødum (University of Copenhagen, Denmark) Prof. dr. P. Devilee Dr. K. Szuhai

CONTENTS

Chapter 1	General introduction	9
Chapter 2	Molecular advances in cutaneous T-cell lymphoma *Seminars in Cutaneous Medicine and Surgery 37: 81–86 (2018)*	37
Chapter 3	Genomic analysis reveals recurrent deletion of JAK-STAT signaling inhibitors *HNRNPK* and *SOCS1* in mycosis fungoides *Genes, Chromosomes and Cancer 57: 653–664 (2018)*	51
Chapter 4	The genetic landscape of primary cutaneous anaplastic large cell lymphoma *Submitted*	83
Chapter 5	Deregulation of JAK2 signaling underlies primary cutaneous CD8+ aggressive epidermotropic cytotoxic T-cell lymphoma *Submitted*	115
Chapter 6	Whole-genome analysis uncovers recurrent *IKZF1* inactivation and aberrant cell adhesion in blastic plasmacytoid dendritic cell neoplasm *Genes, Chromosomes and Cancer 59: 295–308 (2020)*	155
Chapter 7	General discussion and conclusions	187
Appendix	Nederlandse samenvatting List of abbreviations Curriculum vitae List of publications Acknowledgements	199

1

General introduction



1. CUTANEOUS LYMPHOMAS

The term primary cutaneous lymphoma refers to a group of hematological malignancies that presents in the skin without evidence of extracutaneous disease at the time of diagnosis. After the group of gastrointestinal tract lymphomas, primary cutaneous lymphomas are the second most common group of extra-nodal lymphomas with an estimated annual incidence of 1:100.000 individuals.[1] The World Health Organization – European Organization for Research and Treatment of Cancer (WHO-EORTC) classification for primary cutaneous lymphomas recognizes 21 entities which can be divided into three broad categories: cutaneous T-cell lymphomas (CTCL), cutaneous B-cell lymphomas (CBCL) and precursor neoplasms (PN) (Table 1).[2,3] The malignancies included in these categories vary widely in their clinical presentation, prognosis, histopathology, immunohistochemistry (IHC) and molecular biology.

Table 1. The 2018 updated WHO-EORTC classification of primary cutaneous lymphomas.

Lymphoma	Abbreviation	Incidence (%)
Cutaneous T-cell lymphomas (* = provisional):		75
Mycosis fungoides	MF	39
Mycosis fungoides variants:		
Folliculotropic mycosis fungoides	FMF	5
Pagetoid reticulosis	PR	<1
Granulomatous slack skin	GSS	<1
Sézary syndrome	SS	2
Adult T-cell leukemia/lymphoma	ATLL	<1
Subcutaneous panniculitis-like T-cell lymphoma	SPTCL	1
Extranodal NK/T-cell lymphoma, nasal type	eNK/TCL	<1
Chronic active EBV infection:		
Hydroa vacciniforme-like lymphoproliferative disorders	HV-like LPD	<1
Primary cutaneous CD30+ lymphoproliferative disorders:		
Primary cutaneous anaplastic large cell lymphoma	pcALCL	8
Lymphomatoid papulosis	LyP	12

Table 1. Continued.

Lymphoma	Abbreviation	Incidence (%)
Primary cutaneous peripheral T-cell lymphoma, rare subtypes:		
Primary cutaneous gamma-delta T-cell lymphoma	pcGDTCL	<1
Primary cutaneous CD8+ aggressive epidermotropic cytotoxic T-cell lymphoma*	pcAECyTCL	<1
Primary cutaneous CD4+ small/medium T-cell lymphoproliferative disorder*	pcSMLD	6
Primary cutaneous acral CD8+ T-cell lymphoma*	PCATCL	<1
Primary cutaneous peripheral T-cell lymphoma, not otherwise specified	PTCL-NOS	2
Cutaneous B-cell lymphomas (* = provisional):		25
Primary cutaneous marginal zone lymphoma	pcMZL	9
Primary cutaneous follicle center lymphoma	pcFCL	12
Primary cutaneous diffuse large B-cell lymphoma, leg-type	pcDLBCL, LT	4
Intravascular large B-cell lymphoma	ivLBCL	<1
EBV+ mucocutaneous ulcer*	EMU	<1
Precursor neoplasms:		
Blastic plasmacytoid dendritic cell neoplasm	BPDCN	<1

1.1 Cutaneous T-cell lymphoma

CTCL is a group of malignancies derived from skin-homing T cells. The most recent WHO-EORTC classification recognizes fifteen CTCL variants, three of which are currently considered provisional entities (Table 1). CTCL accounts for 75% of cutaneous lymphoma cases in the Western world, with mycosis fungoides (MF) and primary cutaneous CD30+ lymphoproliferative disorders (LPDs) being the most common.[2,3] In the following paragraphs the CTCL variants relevant for this thesis are discussed in more detail.

1.1.1 Mycosis fungoides

MF, the most common variant of CTCL, is caused by malignant T cells with hyperchromatic-cerebriform nuclei that migrate to the skin.[2] MF cells usually express CD3, CD4 and CD45RO, which corresponds to a memory T-cell phenotype.[4] The disease has an indolent clinical course with slow progression

from patches to more infiltrated plaques and eventually to tumors. In a minority of patients, extracutaneous localization develops in later stages of the disease. The prognosis of MF patients depends on stage, in particular, the type and extent of skin lesions and presence of extracutaneous disease.[2] While patients with early stages of disease have an excellent prognosis, the 10-year disease-specific survival of patients with tumor stage disease and patients with histologically documented lymph node involvement is 42% and 20%, respectively.[5] When MF is limited to the skin, skin-directed therapies (e.g. topical steroids, PUVA phototherapy, UV-B phototherapy, nitrogen mustard, radiotherapy) are indicated.[6-8] However, patients with refractory tumor-stage disease or extracutaneous dissemination usually require multi-agent chemotherapy.[9]

1.1.2 Primary cutaneous anaplastic large cell lymphoma

Primary cutaneous anaplastic large cell lymphoma (pcALCL) is a CD30+ T-cell neoplasm composed of large cells with an anaplastic, pleomorphic or immunoblastic morphology.[10] pcALCL is characterized by the expression of the surface receptor CD30 in more than 75% of tumor cells.[2] pcALCL cells display an activated T-cell phenotype with expression of CD3, CD4 and CD45RO and variable loss of CD2 and CD5.[11] pcALCL typically present in the skin with solitary or localized tumors, which often show ulceration.[10-12] Although some phenotypic and molecular features are shared between pcALCL and systemic anaplastic large cell lymphoma (sALCL), their clinical course differs. pcALCL has significantly better prognosis than sALCL, with a 5-year survival of 90%.[13] Yet, about 10% of pcALCL cases develop extracutaneous involvement and worse prognosis.[12] Treatment of choice for patients with pcALCL depends on the type and number of lesions. Patients with solitary or few lesions are treated with radiotherapy or surgical excision, whereas patients with multifocal lesions are often treated with low-dose methotrexate.[14,15] Patients with extracutaneous dissemination beyond locoregional lymph nodes receive multi-agent chemotherapy.[16]

1.1.3 Primary cutaneous CD8+ aggressive epidermotropic cytotoxic T-cell lymphoma

Primary cutaneous CD8+ aggressive epidermotropic cytotoxic T-cell lymphoma (pcAECyTCL) is an aggressive variant of CTCL still considered a provisional entity in the latest 2018 WHO-EORTC classification due to its rarity.[2,17,18] The most common presentation of pcAECyTCL consists of widely disseminated plaques and tumors with hemorrhagic ulcerations and necrosis in the skin. The disease frequently disseminates to viscera, including the lung, testes, central nervous system, and oral mucosa.[18,19] Neoplastic T cells in pcAECyTCL appear to derive from CD8+ T cells as their immunophenotype usually shows positivity for CD7, CD8, CD45RA, T-BET, TCR-βF1, and at least one cytotoxic marker (e.g. granzyme

B, perforin, TIA-1).[2,17] At present there are no effective therapeutic options for patients with pcAECyTCL.

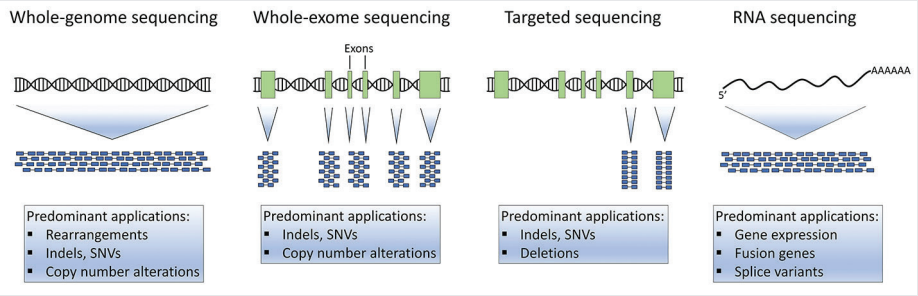
1.2 Precursor neoplasms

1.2.1 Blastic plasmacytoid dendritic cell neoplasm

Blastic plasmacytoid dendritic cell neoplasm (BPDCN) is a rare and clinically aggressive hematological malignancy derived from the precursors of plasmacytoid dendritic cells.[2,20] Patients with BPDCN typically present with solitary or widespread brown to violaceous bruise-like lesions, plaques, or tumors, which are accompanied by simultaneous or subsequent systemic dissemination.[21] Common extracutaneous sites of dissemination are the bone marrow, lymph nodes and peripheral blood. Neoplastic cells are characterized by the expression of CD4, CD56 (NCAM1), CD123 (IL3RA), CD303 (BDCA-2) and TCL1.[20,21] The median overall survival of patients with BPDCN is 12 months from diagnosis as most therapeutic regimes (e.g. ALL-like, CHOP-like, etc.) are not capable of achieving long-term remission, with hematopoietic stem cell transplantation (HSCT) being the only exception in a few cases.[20]

Box 1: Next-generation sequencing

Next-generation sequencing (NGS) refers to a group of high-throughput technologies used to determine the order of nucleotides within nucleic acid molecules (i.e. DNA and RNA). In cancer research, NGS is particularly suitable for the discovery of different types of genetic alterations ranging from pathogenic single nucleotide substitutions to large chromosomal aberrations in cancer genomes.[22] The terms whole-genome sequencing (WGS), whole-exome sequencing (WES) and targeted sequencing (TS) refer to NGS approaches that characterize the entire genome, the complete set of coding genes (known as exome,<2% of the genome) or a defined group of individual genes, respectively.[22] Another NGS application called RNA sequencing (RNA-seq) is used to characterize the collection of expressed genes in the cell known as the transcriptome, which makes possible the identification of abnormal gene expression patterns in malignancies when compared to healthy controls, and the detection of expressed gene fusions derived from chromosomal rearrangements.[22]



Box 1 Figure. NGS approaches in cancer research. Whole-genome sequencing (WGS) evaluates the entire genome, including gene-coding (exons) and non-coding regions (intergenic sequences, introns). Whole-exome sequencing (WES) examines all exons of the entire set of known genes (<2% of the genome). Targeted sequencing (TS) evaluates either all exons (not shown) or just specific spots within exons (shown) of a panel of genes. RNA sequencing (RNA-seq) assesses expressed RNA (protein and non-protein coding) to measure gene expression, discover abnormal transcripts and detect splice variants.[23] Adapted from Simon R, Roychowdhury S. Implementing personalized cancer genomics in clinical trials. Nat Rev Drug Discov. 2013;12(5):358-369.

2. ONCOGENESIS

Cancer is a group of diseases characterized by uncontrolled cell proliferation and the spread of anomalous cells from the site of origin to other body sites. The abnormal proliferation and behavior of cancer cells is the result of defects in regulatory circuits that govern normal cell physiology and homeostasis. Even though the specific molecular defects may vary between cancer types, their functional outcome lead to six core physiological alterations essential for cancer development.[24,25] This group of physiological alterations (known as the hallmarks of cancer) are the acquisition of the capability for autonomous growth signals, evasion of growth inhibitory signals, evasion of apoptotic cell death, unlimited replicative potential, angiogenesis, and tissue invasion and metastasis. In recent years, the reprogramming of energy metabolism and the avoidance of immune destruction have emerged as two additional hallmarks of cancer (Figure 1).[24,26]

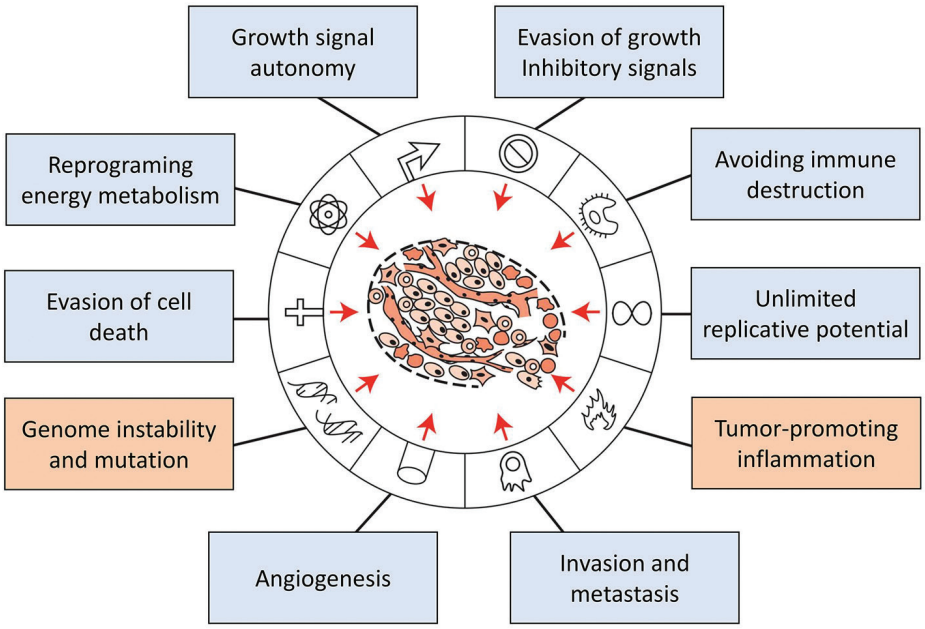


Figure 1. The hallmarks of cancer. Eight core principles underlie the complexity of cancer (light blue). Two special characteristics enable the transformation of normal cells into neoplastic cells and the acquisition of the other eight hallmarks of cancer (light red). Modified from Pecorino L. Molecular Biology of Cancer: Mechanisms, Targets and Therapeutics (ed 3rd). Oxford: Oxford University Press; 2012.

The acquisition of these hallmarks is enabled by two characteristics inherent to cancer: genome instability and tumor-promoting inflammation. Genome instability refers to the high frequency of mutations within the genome of

cancer cells. Mutations, which accumulate in the genome over time, lead to an increasing deregulation of the normal cellular functions.[25,26] Tumor-promoting inflammation refers to the pathogenic inflammatory response resulting from the infiltration of immune cells in the tumor. Tumor-infiltrating immune cells can promote the acquisition of core hallmarks of cancer by providing growth factors and enzymes required to maintain a pro-oncogenic microenvironment in the tumor. Immune cells in the tumor can also contribute to mutagenesis by releasing DNA-damaging reactive oxygen species (ROS).[25,26]

Given that the underlying etiology of cancer is essentially genetic, genome instability has a fundamental role in oncogenesis. Mutations leading to malignant transformation affect the function of genes involved in the regulation of cell proliferation, apoptosis and differentiation. Defective genes driving cancer can be divided into two categories: oncogenes and tumor suppressors.[25] Oncogenes are proliferation-promoting genes whose protein products (e.g. cytokines, cytokine receptors, transcription factors, signal transducers, etc.) are produced in higher quantities or have increased activity as a result of gain-of-function mutations (Figure 2).[27] Tumor suppressors are genes that encode proteins with the ability to inhibit cell proliferation (e.g. cell cycle inhibitors, DNA repair proteins, apoptosis regulators, etc.) whose expression or functionality become impaired in cancer cells due to loss-of-function mutations (Figure 2).[27]

3. MUTATIONS

A mutation is a change in the nucleotide sequence of the genome of an organism, virus or extrachromosomal DNA. Errors in DNA replication, DNA repair, mitosis and meiosis can produce mutations. The source of these errors can be endogenous (e.g. ROS) or exogenous (e.g. ultraviolet (UV) radiation, environmental carcinogens, etc.).[27]

Even though mutations can be classified in many different ways, the classification by effect on DNA structure is the most appropriate in the context of NGS studies.

3.1 Mutations by effect on DNA structure

When considering the effect on DNA structure, mutations are divided into small-scale mutations and large-scale mutations. Small-scale mutations are genetic changes involving 1 to 50 nucleotides, which include small insertions/deletions (indels) and single nucleotide variants (SNVs) (Figure 3).[27] On the contrary, large-scale mutations are changes in chromosome structure which can affect the functioning of numerous genes. These structural chromosomal changes

include deletion or insertion of one or several contiguous genes (known as copy number alterations), inversion of genes on a chromosome, or the exchange of large segments of DNA between nonhomologous chromosomes (known as chromosomal rearrangements or translocations) (Figure 3).[27]

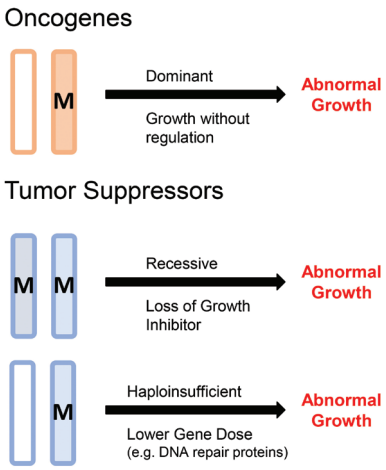


Figure 2. Oncogenes and tumor suppressors. Mutated genes in cancer can be divided into oncogenes and tumor suppressors. As a result of mutations, oncogenes acquire an enhanced ability to promote cell proliferation whereas tumor suppressors lose their natural ability to suppress cell proliferation. Oncogenes are 'dominant' in nature as a mutation in only one allele is sufficient to produce a proliferation-promoting effect. Tumor suppressors are 'recessive' or 'haploinsufficient' in nature, meaning that mutations in either both alleles or only one, respectively, may be needed to produce a proliferation-promoting effect. Adapted from Pecorino L. Molecular Biology of Cancer: Mechanisms, Targets and Therapeutics (ed 3rd). Oxford: Oxford University Press; 2012.

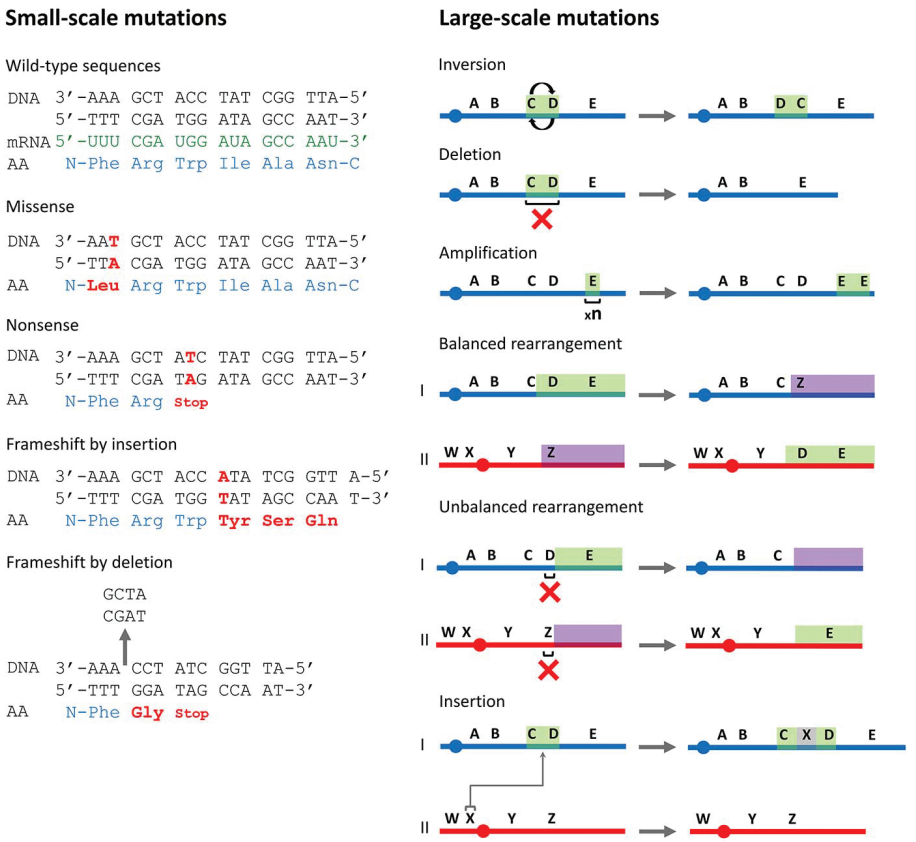


Figure 3. Small-scale and large-scale mutations. Small-scale mutations include single nucleotide variants (SNVs) and small insertions/deletions (indels). SNVs can cause missense and nonsense mutations. Missense mutations lead to a change in a single amino acid in the encoded protein. Nonsense mutations lead to the formation of a stop codon. Indels lead to frameshift mutations causing a change in the reading frame. Large-scale mutations involve alterations in large segments of DNA which often result from errors in DNA double-strand breaks. Adapted from Lodish H, Berk A, Kaiser C, et al. Molecular Cell Biology (ed 8th). New York: W. H. Freeman 2016.

4. MOLECULAR ALTERATIONS IN CUTANEOUS LYMPHOMAS

Prior studies using cytogenetic techniques, array-based methods and NGS have sought to identify genetic alterations underlying various cutaneous lymphomas, including MF, pcALCL, pcAECyTCL and BPDCN. An overview of the most relevant findings in these four cutaneous lymphomas is given below.

4.1 Mycosis fungoides

Despite being the most common CTCL, only few molecular studies have investigated MF separately as an independent entity[22] whereas several others combined MF samples with Sézary syndrome (SS) samples. However, besides the differences in histopathological characteristics, clinical course and prognosis between these two CTCL variants, array-based molecular studies have shown that MF and SS display different copy number alteration (CNA) and microRNA (miRNA) profiles.[28,29] This molecular evidence strongly supports the idea that MF and SS are related but distinct entities.

4.1.1 Studies with mixed SS/MF cohorts

Genetic studies performed on mixed SS/MF cohorts range from cytological to NGS analyses. Overall, these analyses failed to establish informative commonalities between MF patients. Mao et al. studied CNAs in a SS/MF cohort by using comparative genomic hybridization (CGH) and reported loss of 1p36 as the only recurrent event in MF patients.[30] Similarly, Fischer et al. used CGH to assess CNAs in a mixed CTCL group and found that chromosomal aberrations were more abundant in patients with advanced stages of SS/MF. However, no CNA analysis per variant was carried out.[31] Subsequent WES/TS studies with mixed SS/MF cohorts reported gain-of-function SNVs in solitary or few MF cases. These SNVs included *JAK3* (p.A573V), *MAPK1* (p.E322K), *STAT3* (p.Y640F), *PLCG1* (p.S345F, p.S520F), and *TNFRSF1B* (p.T377I).[32-36]

4.1.2 Studies with MF-only cohorts

Several genome-wide studies that analyzed CNAs, small-scale mutations and gene expression in MF separately from SS have defined molecular features of this lymphoma.

4.1.2.1 CNAs

CNA profiles of MF cohorts reported in different studies have shown limited overlap between them, probably due to the use of different platforms and/or biological material (disease stages/subtypes). Prochazkova et al. used CGH to study CNAs in large-cell transformed MF, and described losses within chromosomes 2, 9 and 17 as well as gains within chromosomes 1 7, 9 and 17.[37] Carbone et al. analyzed CNAs in early-stage MF by using array-based CGH (a-CGH) and reported monosomy 19, recurrent losses within chromosomes 7, 9, 12 and 16, and recurrent gains within chromosomes 8 and 21.[8] van Doorn et al. showed that the genome of tumor-stage MF (T-MF) is characterized by losses within chromosomes 1, 5, 9, and 13, and gains within chromosomes 7, 8 and 17 by using higher resolution a-CGH.[28] Notably, patients with loss of 9p21, gain of 8q24.3 or gain of 1q21-1q22, regions where *CDKN2A/B*, *MYC* and *MCL1* reside,

respectively, had lower survival rates than patients without these chromosomal imbalances.[28]

4.1.2.2 microRNA and gene expression

van Kester et al. studied the miRNome of T-MF for the first time by using miRNA arrays and detected up-regulation of oncomirs miR-93, miR-155 and several members of the miR-17-92 cluster. These miRNAs have relevant functions in cancer and the immune response.[29] The up-regulation of miR-155 in T-MF was subsequently confirmed by Ralfkiaer et al. and several other research groups.[38-46] Ralfkiaer et al. also showed that up-regulation of miR-155 together with down-regulation of miR-203 and miR-205 can effectively be used as a classifier to distinguish between CTCL and benign inflammatory dermatoses (BID).[40] The effectiveness of these miRNAs as CTCL classifiers was later confirmed by Marstran et al., Dusilkova et al. and Shen et al.[42,45,46] Additionally, Lindahl et al. developed a miRNA classifier involving miR-106b-5p, miR-148a-3p and miR-338-3p that predicted progression from early- to advanced-stage MF and survival at time of diagnosis.[47]

Moyal et al. reported an apparent correlation between MF progression and increased levels of miR-155.[43] These authors also showed that inhibition of miR-155 in MF cell lines (i.e. MyLa, MJ) leads to cell cycle arrest at G2/M and apoptosis *in vitro* as well as the development of smaller tumors in a xerograph mouse model of MF.[44] McGirt et al. reported down-regulation of miR-223 in early- and advanced-stage MF, showed that elevated miR-233 expression reduces proliferation of MF cell lines, identified putative miR-223 targets (e.g. *E2F1*, *MEF2C*, *TOX*) and found evidence that increase of miR-223 expression post-therapy could be predictive of clinical response to extracorporeal photopheresis in MF patients.[48,49]

Regarding gene expression, van Kester et al. performed an array-based meta-analysis of gene expression where T-MF was compared to normal skin, inflamed skin, psoriasis and normal T cells.[50] T-MF was found to express low levels of NF-kB inhibitor NFKBIZ and high levels of cytokine IL32 and several potential markers (e.g. *GTSF1*, *TRIP13*).[50]

4.1.2.3 Small-scale mutations

McGirt et al. performed the first MF-only WGS/TS study, which mainly focused on the assessment of small-scale mutations. A set of 42 genes were found to be mutated in 5 patients, but 37/42 mutated genes were patient-specific. The somatic mutation rate in MF (average, 12 mutations/Mb) was found to be similar to mutation rates observed in melanoma and lung cancer.[34] The majority of somatic single nucleotide substitutions observed in MF were C>T transitions (40%-65%),

which are much less common in other hematological cancers.[34] This mutational signature was consistent with ultraviolet (UV) light exposure. However, at present it is still unclear whether exposure to UV light plays a causative role in the onset of the disease.[34]

4.2 Primary cutaneous anaplastic large cell lymphoma

The majority of genetic alterations reported in pcALCL derive from comparative studies involving pcALCL and other lymphoma entities. These studies include cytogenetic and molecular analyses focused on CD30+ LPDs, anaplastic large cell lymphoma (systemic and cutaneous) and CTCL variants (e.g. CD30+ MF, SS).

4.2.1 Chromosomal rearrangements

Numerous studies have shown that chromosomal rearrangements typically found in systemic anaplastic large cell lymphoma (sALCL) can also be found in pcALCL, though, at much lower frequencies.[51] In pcALCL, these chromosomal rearrangements involve, in descending order of frequency, *DUSP22/IRF4*, *ALK*, *TP63* and *TYK2*.[51]

4.2.1.1 *DUSP22/IRF4* rearrangements

Wada et al. performed the most comprehensive evaluation of *DUSP22/IRF4* rearrangements in pcALCL to date (182 patients from 3 institutions) and reported a frequency of 20%.[52] Feldman et al. showed that the presence of *DUSP22/IRF4* rearrangements lead to DUSP22 down-regulation while several other studies demonstrated that these events are not correlated with changes in IRF4 expression.[52-55] Unlike sALCL, where patients with *DUSP22/IRF4* rearrangements have a better prognosis than patients with TP63 rearrangements or triple negative patients (no *DUSP22/IRF4*, *ALK* or *TP63* rearrangements), the prognosis of pcALCL patients with and without *DUSP22/IRF4* rearrangements is comparable.[56]

Different lines of evidence suggest that *DUSP22/IRF4* rearrangements, which primarily inactivate *DUSP22*, exert pathogenic effects on T cell physiology. Mélard et al. showed that re-expression of *DUSP22* in DUSP22-deficient malignant T cells impaired proliferation and induced apoptosis, supporting a tumor suppressor role of this gene.[57] Sekine et al. found that DUSP22 is an inhibitor of the IL-6/LIF/STAT3 signaling axis.[58] Li et al. proved that DUPS22 attenuates TCR signaling via Lck dephosphorylation, and consequently, DUSP22-knockout T cells displayed enhanced cell proliferation and cytokine production.[59] Finally, Xing et al. reported an apparent correlation between the presence of *DUSP22/IRF4* rearrangements and higher expression of chemokine receptor CCR8 in pcALCL.[60] CCR8 is a skin-homing receptor preferentially expressed

by skin resident T cells, whose overexpression in pcALCL has been proposed to explain the tendency of this lymphoma to remain restricted to the skin.[61]

4.2.1.2 *ALK* rearrangements

The most extensive characterization of ALK+ pcALCL cases was performed by Melchers et al. Out of 309 pcALCL cases evaluated by the authors, only 6 (2%) adult cases carried *ALK* rearrangements.[16] Three of six affected patients bore fusion gene *NPM1-ALK*, while the rest had other fusion partners (i.e. *ATIC-ALK*, *TPM3-ALK*, *TRAF1-ALK*).[16] All patients with *NPM1-ALK* fusions showed cytoplasmic and nuclear ALK localization, while patients carrying the other fusion partners exhibited cytoplasmic ALK localization only.[16] *ALK* rearrangements have also been found in children with pcALCL. Oschlies et al. reported *ALK* rearrangements in 6 of 23 pediatric cases of pcALCL, all of them accompanied by ALK expression.[62] Although *ALK* rearrangements are indicative of bad prognosis in sALCL, both studies informed that pcALCL patients with and without *ALK* rearrangements had a similar prognosis.

4.2.1.3 *TP63* and *TYK2* rearrangements

Chromosomal rearrangements involving *TP63* and *TYK2* have exceptionally been found in pcALCL. Vasmatzis et al. searched for rearrangements in 19 cases of pcALCL by employing mate-pair genome-wide NGS and detected *TP63* rearrangements in 2 patients with an unusual aggressive clinical course, suggesting a potential correlation between these events and bad prognosis.[63] Both patients with *TP63* rearrangements had strong TP63 protein expression; however, strong TP63 expression was also observed in cases without these alterations. Later, Schrader et al. looked for *TP63* rearrangements in 17 cases of aggressive LyP and pcALCL by using fluorescence *in situ* hybridization (FISH) and no *TP63* rearrangements were found in these cases.[64] The authors also reported variable levels of TP63 expression in this series, suggesting that expression of this protein is not correlated with aggressiveness of the disease.[64]

Contrary to sALCL, genetic alterations affecting genes of the JAK-STAT pathway are uncommon in pcALCL. Velusamy et al. detected *TYK2* rearrangements in 7 of 47 cases of CD30+ LPDs (3 pcALCL, 4 LyP) by using FISH.[65] Two cases (1 pcALCL, 1 LyP) harbored fusion gene *NPM1-TYK2* and exhibited nuclear STAT5 activation. Subsequent functional studies with HEK-293T cells confirmed that *NPM1-TYK2* encoded a self-activating chimera capable of activating STAT1/2/5.[65]

4.2.2 CNAs

The CNA profile of pcALCL has been investigated using a-CGH by three independent research groups. Two genome-wide studies carried out by van

Kester et al. and Laharanne et al. described similar findings, with losses within chromosomes 1, 6 and 13 and gains within chromosomes 1, 7 and 16, being common to both reports.[61,66] Frequent CNAs with potential oncogenic relevance included gain of 7q31 as well as loss of 13q34, 6q16-21 and 16p11.1-11.2, which contained (putative) cancer genes *MET*, *CDC16/CUL4A*, *PRMD1*, *TP53TG3* and *TNFRSF14*, respectively.[61,66] Mao et al., on the contrary, performed targeted a-CGH analysis for 57 common oncogenes together with metaphase-CGH chromosome analysis. These authors reported gain of cancer genes *CBFA2*, *CTSB*, *FES*, *NRAS*, *MYCN* and *RAF1* as well as broad gains within chromosomes 1, 5, 6, 7, 8 and 19 in pcALCL.[67]

4.2.3 Small-scale mutations

Up to now no comprehensive mutational analysis has been performed on cohorts exclusively formed by pcALCL cases. Reports of pathogenic mutations in pcALCL come from studies performed on mixed sALCL/pcALCL cohorts. Crescenzo et al. performed WES on 155 patients with ALK– ALCL (systemic and cutaneous) and identified 6 and 3 cases of pcALCL with gain-of-function SNVs in *JAK1* and *STAT3*, respectively.[68] Similarly, Luchtel et al. performed WES/TS on 119 cases of ALK– ALCL (systemic and cutaneous) and reported a pathogenic MSC (p.E116K) mutation in 2 patients with pcALCL.[69]

4.2.4 Gene and miRNA expression

van Kester et al. performed the only genome-wide gene expression analysis of pcALCL to date by using array technology. Highlights of deregulated expression were the overexpression of *CCR8*, *CCR10*, *EZH2*, *IL23R*, *IRF4*, *TNFRSF8* (CD30), genes with relevant functions in T-cell biology.[61] These results matched prior reports of CD30 and IRF4 overexpression in pcALCL by IHC. Subsequent research by Yi et al. confirmed high expression of the EZH2 protein in pcALCL by IHC, and these authors showed through functional studies that EZH2 promotes cell cycle progression and prevents the migration of anti-cancer T cells into the tumor in pcALCL.[70] Also, Kamstrup et al. showed using IHC that receptor NOTCH1 and its ligand JAG1 (Jagged 1) are overexpressed in tumors from pcALCL patients.[71] Later, the same research group used pcALCL cell lines to evaluate the blockade of Notch signaling as a potential therapeutic target.[72] The study confirmed that pcALCL cell lines overexpress various components of the Notch pathway (i.e. receptors NOTCH1-4, ligand Delta, transcription factor HES1), and these cells underwent marked apoptosis upon treatment with γ-secretase inhibitors.[72]

Benner et al. carried out a comparative analysis between the miRNomes of pcALCL, BID and T-MF by using miRNA arrays. Overexpressed miRNAs in pcALCL included miR-27b, miR-29, miR-30c and miR-155.[73] Later, Sandoval et

al. examined the same disease entities evaluated by Benner et al. using higher resolution arrays and confirmed the overexpression of miR-155 in pcALCL.[74] These authors also reported up-regulation of miR-21, miR-142-3p/5p, let-7i, miR-424, miR-431, miR-542-5p, miR-29b-1, miR-342-p and miR-484, all of which are regarded as oncogenic miRNAs. Down-regulated tumor suppressor miRNAs in pcALCL included miR-23b/miR-27b, miR-203, miR-205 and miR-125b.[74] Notably, the miRNA signatures uncovered in these studies differ from miRNA signatures reported in sALCL.[75,76]

4.3 Primary cutaneous CD8+ aggressive epidermotropic cytotoxic T-cell lymphoma

Molecular research on pcAECyTCL is currently lacking, with the evaluation of chromosomal imbalances being the only molecular alterations investigated thus far.

4.3.1 CNAs

Kato et al. studied CNAs in one patient with pcAECyTCL using spectral karyotyping (SKY) and a single nucleotide polymorphism (SNP) array.[77] The authors reported lack of consistency between the results obtained with both techniques, which were attributed to the lower resolution of SKY compared to the SNP array. Highlights of this analysis was loss of the regions within chromosomes 12, 13 and 17, which contain cancer genes *DUSP6/E2F7/PTPRR*, *E2F7/RB1* and *TP53*, respectively. Also, gain of the region within chromosome 2 where *STAT1* resides.[77] Similarly, Tomasini et al. subjected one pcAECyTCL patient to low-resolution a-CGH to study CNAs. These authors reported broad gains within chromosomes 7, 9, 14, 17 and 20, and broad losses within chromosomes 8, 13 and 18.[78] Finally, Fanoni et al. characterized the CNA profile of 20 pcAECyTCL cases by using a-CGH. Recurrent CNAs between patients included losses within 1p, 9p, 13q and 16p as well as gains within 7q, 8q and 17q. The most frequent CNA in pcAECyTCL was found to be loss of *CDKN2A/B*, which was validated by FISH, and shown not to correlate with worse overall survival.[79]

4.4 Blastic plasmacytoid dendritic cell neoplasm

4.4.1 Chromosomal rearrangements

Several chromosomal rearrangements have been detected in BPDCN primarily through FISH. The majority of these findings were part of single patient case reports, with the studies describing recurrent *MYB* and *MYC* rearrangements by Suzuki et al. and Sakamoto et al., respectively, being the only cohort analyses. Suzuki et al. detected *MYB* rearrangements in 9 of 14 BPDCN patients (children and adults) by RNA-seq.[80] These events mediated the loss of miRNA target

sequences located in the 3'-end of *MYB*, leading to its up-regulation. Sakamoto et al. reported rearrangements at 8q24 that promoted *MYC* overexpression in 47 of 118 (39%) patients.[81] Rearrangement t(6;8)(p21;q24), originally discovered in BPDCN by Nakamura et al., was the commonest.[82] Notably, almost all patients carrying 8q24 rearrangements had an immunoblast-like appearance and more aggressive disease than patients without them.[81]

Additionally, chromosomal rearrangements involving cancer genes *ALK*, *ETV6*, *EWSR1* and *KMT2A* (*MLL*) have been described in clinical case reports. Tokuda et al. reported fusion gene *CLTC-ALK* in tumor cells and normal blood leukocytes (myeloid and lymphoid) of an infant with BPDCN, suggesting that the rearrangement occurred in a multipotent hematopoietic progenitor in utero.[83] Gao et al. informed of a BPDCN case with a rearrangement that disrupted *ETV6*, a tumor suppressor implicated in various hematological cancers.[84] Cao et al. informed of a child with BPDCN that carried an *EWSR1* rearrangement. *EWSR1* is typically rearranged in Ewing's sarcoma, and less frequently, in acute myeloid and lymphoid leukemias.[85] Toya et al. and Yang et al. reported fusion gene *MLL-MLLT1* in and an adult and a child, respectively.[86,87] Fusion gene *MLL-MLLT1* has been shown to immortalize various normal hematopoietic cell types.

4.4.2 CNAs

The CNA profile of BPDCN has been studied by six different research groups using a-CGH.[88-93] These studies showed extensive agreement between them and revealed that BPDCN is mainly characterized by recurrent losses within chromosomes 9, 12 and 13, in particular the regions containing tumor suppressors *CDKN2A/B*, *CDKN1B* and *RB1/LATS2*, respectively. Notably, all of these genes are involved in the regulation of cell cycle progression at G1/S. Lucioni et al. showed that the biallelic loss of *CDKN2A/B* was associated with worse prognosis.[90] Wiesner et al. demonstrated that loss of *CDKN1B* and *CDKN2A* correlated with weak expression of p27 and no expression of p16, respectively.[88] In addition to loss of *CDKN2A/B* and *CDKN1B*, Oiso et al. reported recurrent loss of *CDKN2C*, another G1/S regulator.[89] Dijkman et al. reported correlation between loss of *RB1/LATS2* and low RNA expression of both genes.[92] Other recurrent CNAs reported only in single studies include losses within chromosomes 1, 4, 5, 7, and 15.[88-90,92,93] Emadali et al. reported recurrent loss of the region in chromosome 5 containing *NR3C1* in a subset of patients with highly aggressive BPDCN.[93]

4.4.3 Small-scale mutations

Indels and SNVs underlying BPDCN have been reported in ten studies.[93-102] Taken together these analyses showed that genes involved in chromatin regulation (e.g. *ASXL1*, *TET2*) are the most frequently mutated in BPDCN, followed in no

particular order by genes with roles in the cell cycle (e.g. *TP53*), RAS signaling (e.g. *KRAS*, *NRAS*), splicing (e.g. *ZRSR2*) and transcription (e.g. *IKZF1/2/3*). The predominance of mutations in epigenetic regulators (esp. *TET2*) observed in BPDCN is a common feature of various myeloid malignancies and supports a myeloid origin of the disease.

Jardin et al. were the first to report recurrence of *TET2* and *TP53* mutations in BPDCN.[101] Subsequent studies by Alayed et al. and several others confirmed the high frequency of *TET2* mutations in BPDCN.[93,96,100,102] Taylor et al. performed a TS study with a panel of 219 genes associated with myeloid lineages and found high frequency of *TET2* and *ZRSR2* mutations. This study reported for the first time mutations in *ABL1*, *ARID1A*, *ASXL1*, *GNA13*, *IDH2*, *IRF8*, *KRAS*, *SRSF2* and *U2AF1* in this neoplasm.[95] Stenzinger et al. assessed the mutational status of 50 common cancer genes in BPDCN by using TS. These authors found that mutations in *ATM*, *NRAS*, *KRAS* were recurrent and appeared to be mutually exclusive. Novel mutated genes described in this study included *APC*, *BRAF*, *IDH2*, *KIT*, *MET*, *MLH1*, *RET*, *RB1* and *VHL*, most of them in solitary or few cases.[97]

Menezes et al. performed the first WES/TS study of BPDCN. A group of 38 genes found to be mutated in three patients by WES were included in a TS panel and screened for mutations in 25 patients. *ASXL1* and *TET2* were found to be the most frequently mutated genes, followed by *NRAS*, *NPM1* and members of the Ikaros family.[99] Notably, the overall survival of patients carrying mutations in epigenetic regulators or Ikaros family genes was significantly lower than patients without these mutations.[99] Togami et al. analyzed 12 patients by WES and 12 additional patients by TS. Recurrent mutations were not only found in genes already implicated in BPDCN in previous studies (e.g. *ASXL1*, *ETV6*, *IDH2*, *NRAS*, *TET2*, *TP53*), but also in novel genes (e.g. *CRIPAK*, *HNF1A*, *NEFH*, *PAX3*, *SSC5D*).[94] Finally, Sapienza et al. subjected 14 patients to WES and reported recurrent mutations in *ASXL1* and *TET2* as well as novel mutations in numerous additional epigenetic regulators (e.g. *ARID1A*, *CHD8*, *PHF2*, *SUZ12*, etc.). Predominance of mutations in epigenetic regulators was supported by enrichment analysis of the mutated genes.[98]

4.4.4 Gene expression

Hallmarks of gene expression in BPDCN have been uncovered by using array-based methods and RNA-seq in five studies.[92,98,103-105] Dijkman et al. performed the first array-based gene expression profiling of BPDCN. Given that BPDCN and cutaneous myelomonocytic leukemia (cAML) can exhibit a similar clinical presentation, these authors sought to identify differentially expressed genes between these two entities that could help clarify their pathobiological basis and

cellular origin. BPDCN was found to express oncogenes *HES6*, *FLT3* and *RUNX2* at much higher levels than cAML. BPDCN also displayed robust expression of numerous genes typically expressed by plasmacytoid dendritic cells (pDCs), including *IL3RA* (CD123), *SPIB*, *TCL1A* and Toll-like receptors *TLR9/10*.[92] These findings strongly substantiated prior suggestions of a potential pDC origin of BPDCN cells based on immunophenotype profiling by IHC.

Subsequently, Sapienza et al. compared the transcriptomes of 27 BPDCN cases with the transcriptomes of 8 pDC controls by using arrays. BPDCN displayed signatures consistent with deregulation of cell adhesion, cell-cell communication and vascular development.[103] Overexpressed genes with pathogenic relevance included *CCDN1* (Cyclin D1) and NF-kB target genes *BCL2* and *IRF4*. Activation of NF-kB signaling was confirmed by IHC in cell lines and BPDCN cells from patients.[103] Years later, Sapienza et al. performed a new comparison of gene expression between 5 BPDCN cases and 4 pDC controls by using RNA-seq. The authors looked for expression signatures associated with epigenetic deregulation in BPDCN, and found deregulation of genes modulated by histone demethylase KDM5B34 and methyltransferase PRMT535.[98] In addition, PAT-ChiP analysis of two BPDCN cases was carried out to correlate epigenetically active (H3K27ac) regions of the genome with gene expression. Genes displaying promoter acetylation (H3K27ac) and high expression were all involved in cell cycle progression.[98]

Ceribelli et al. performed an RNAi screening study on CAL-1 cells (BPDCN cell line) where TCF4 (E2-2) was identified as an essential transcription factor for the survival of BPDCN cells.[105] Since TCF4 had been shown before to be crucial for the development and maintenance of pDCs, this finding strengthened the idea that BPDCN cells derive from pDC precursors. Further characterization of TCF4-knockdown BPDCN cell lines by RNA-seq and ChiP-seq showed that depletion of TCF4 caused the down-regulation of genes that define the normal phenotype of mature pDCs (e.g. *BCL11A*, *CLEC4C*, *IRF8*, *IL3RA*, *SPIB*, *TLR9*). The authors also confirmed TCF4 expression in 28/28 cases of BPDCN by IHC and proposed it as a reliable BPDCN marker.[105]

Ceroi et al. performed a comparative analysis of gene expression between BPDCN, acute myeloid leukemia (AML), T-acute lymphoblastic leukemia (T-ALL) and normal pDCs by using arrays.[104] The authors reported a signature associated with deregulation of cholesterol homeostasis in BPDCN, especially involving down-regulation of liver X receptor (LXR) target genes. LXR activation with an agonist induced cholesterol efflux, inhibition of IL-3 signaling (via AKT, NF-kB and STAT5) and apoptosis in BPDCN cell lines (CAL-1, GEN2.2) and BPDCN cells from patients.[104]

5. AIM AND OUTLINE OF THIS THESIS

Despite numerous molecular studies attempted to identify genetic alterations driving the entities included in this thesis, the pathogenetic basis of these cutaneous lymphomas remains poorly understood. As described in this chapter, most prior genome-wide studies made use of array-based techniques whose resolution is significantly lower than current NGS approaches. Moreover, NGS studies performed on these malignancies up to now have mainly been limited to WES studies which focused almost exclusively on small-scale mutations, disregarding structural genomic alterations. Hence, we performed whole-genome and whole-transcriptome analyses of well-characterized cutaneous lymphoma cohorts (i.e. MF, pcALCL, pcAECyTCL, BPDCN) using NGS. These studies had a strong emphasis on structural genomic alterations.

In **chapter 2**, we provide an overview of the genetics of the best studied CTCL variants (MF and SS) derived from NGS studies. This chapter also examines the therapeutic implications of these molecular findings.

In **chapter 3**, we present the first integrated (DNA/RNA) genome-wide analysis of T-MF. To search for structural genomic defects with pathogenic potential, a group of 9 MF patients were analyzed with WGS and RNA-seq. This study aimed to characterize for the first time the landscape of rearrangements (incl. fusion transcripts) and CNAs in this lymphoma using NGS. A comparative analysis of gene expression in MF tumors and normal CD4+ T cells was carried out to identify deregulated cellular processes/pathways in the disease. The copy number status of two novel tumor suppressors found to be recurrently deleted in the sequenced cohort was evaluated in an extension cohort (17 patients) by ddPCR.

Given the limited molecular insight available on pcALCL, in **chapter 4** we performed the most comprehensive genetic study in pcALCL to date. Using a multi-platform approach (WGS, WES, RNA-seq), we investigated genetic alterations and deregulated gene expression in tumors from 12 patients. This study uncovered multiple pathogenic small-scale mutations, CNAs and rearrangements in this lymphoma. Deregulated canonical signaling pathways were detected by comparing gene expression in pcALCL with gene expression in normal CD4+ T cells. Validation of genetic alterations and deregulated genes in pcALCL was performed using PCR-based techniques (Sanger sequencing, ddPCR, RT-qPCR).

In **chapter 5**, we describe the first NGS study ever performed on pcAECyTCL, one of the rarest and most aggressive variants of CTCL. By subjecting 12 patients

to WGS and 6 patients to RNA-seq, we identified for the first time a group of oncogenes and tumor suppressors altered in the disease as well as deregulated cellular processes/pathways. Prominent genetic alterations were validated by FISH, ddPCR and Sanger sequencing. Gene expression in pcAECyTCL was compared to gene expression in normal CD8+ T cells to determine deregulated expression signatures in this lymphoma. The oncogenicity of fusion genes identified in pcAECyTCL was functionally validated and their sensitivity to inhibitors was assessed.

Despite BPDCN has been subjected to multiple molecular studies, the landscape of genomic rearrangements and CNAs had never been examined using NGS. In **chapter 6** we present the first genome-wide analysis of BPDCN focused on structural genomic defects. A group of 10 and 4 patients were subjected to WGS and RNA-seq, respectively. A comparative analysis of gene expression in BPDCN tumors and resting pDCs was performed to detect deregulated cellular processes/pathways in this neoplasm. An extension cohort (15 patients) was used to validate the most relevant finding of the study.

Chapter 7 summarizes and discusses the main findings of this thesis. The chapter ends with conclusions and future perspectives.

REFERENCE LIST

1. Groves FD, Linet MS, Travis LB, Devesa SS. Cancer surveillance series: non-Hodgkin's lymphoma incidence by histologic subtype in the United States from 1978 through 1995. *J Natl Cancer Inst*. 2000;92(15):1240-1251.
2. Willemze R, Jaffe ES, Burg G, et al. WHO-EORTC classification for cutaneous lymphomas. *Blood*. 2005;105(10):3768-3785.
3. Willemze R, Cerroni L, Kempf W, et al. The 2018 update of the WHO-EORTC classification for primary cutaneous lymphomas. *Blood*. 2019;133(16):1703-1714.
4. Kim YH, Liu HL, Mraz-Gernhard S, Varghese A, Hoppe RT. Long-term outcome of 525 patients with mycosis fungoides and Sezary syndrome: clinical prognostic factors and risk for disease progression. *Arch Dermatol*. 2003;139(7):857-866.
5. Agar NS, Wedgeworth E, Crichton S, et al. Survival outcomes and prognostic factors in mycosis fungoides/Sezary syndrome: validation of the revised International Society for Cutaneous Lymphomas/European Organisation for Research and Treatment of Cancer staging proposal. *J Clin Oncol*. 2010;28(31):4730-4739.
6. Whittaker SJ, Marsden JR, Spittle M, Russell Jones R. Joint British Association of Dermatologists and U.K. Cutaneous Lymphoma Group guidelines for the management of primary cutaneous T-cell lymphomas. *Br J Dermatol*. 2003;149(6):1095-1107.
7. Jones GW, Kacinski BM, Wilson LD, et al. Total skin electron radiation in the management of mycosis fungoides: Consensus of the European Organization for Research and Treatment of Cancer (EORTC) Cutaneous Lymphoma Project Group. *J Am Acad Dermatol*. 2002;47(3):364-370.
8. Dummer R, Kempf W, Hess Schmid M, Häffner A, Burg G. Therapy of cutaneous lymphoma--current practice and future developments. *Onkologie*. 2003;26(4):366-372.
9. Kaye FJ, Bunn PA, Jr., Steinberg SM, et al. A randomized trial comparing combination electron-beam radiation and chemotherapy with topical therapy in the initial treatment of mycosis fungoides. *N Engl J Med*. 1989;321(26):1784-1790.
10. Kaudewitz P, Stein H, Dallenbach F, et al. Primary and secondary cutaneous Ki-1+ (CD30+) anaplastic large cell lymphomas. Morphologic, immunohistologic, and clinical-characteristics. *Am J Pathol*. 1989;135(2):359-367.
11. Kummer JA, Vermeer MH, Dukers D, Meijer CJ, Willemze R. Most primary cutaneous CD30-positive lymphoproliferative disorders have a CD4-positive cytotoxic T-cell phenotype. *J Invest Dermatol*. 1997;109(5):636-640.
12. Melchers RC, Willemze R, Vermaat JSP, et al. Outcomes of rare patients with a primary cutaneous CD30+ lymphoproliferative disorder developing extracutaneous disease. *Blood*. 2020;135(10):769-773.
13. Fornari A, Piva R, Chiarle R, Novero D, Inghirami G. Anaplastic large cell lymphoma: one or more entities among T-cell lymphoma? *Hematol Oncol*. 2009;27(4):161-170.
14. Vonderheid EC, Sajjadian A, Kadin ME. Methotrexate is effective therapy for lymphomatoid papulosis and other primary cutaneous CD30-positive lymphoproliferative disorders. *J Am Acad Dermatol*. 1996;34(3):470-481.
15. Bekkenk MW, Geelen FA, van Voorst Vader PC, et al. Primary and secondary cutaneous CD30(+) lymphoproliferative disorders: a report from the Dutch Cutaneous Lymphoma Group on the long-term follow-up data of 219 patients and guidelines for diagnosis and treatment. *Blood*. 2000;95(12):3653-3661.
16. Melchers RC, Willemze R, van de Loo M, et al. Clinical, Histologic, and Molecular Characteristics of Anaplastic Lymphoma Kinase-positive Primary Cutaneous Anaplastic Large Cell Lymphoma. *Am J Surg Pathol*. 2020;44(6):776-781.
17. Berti E, Tomasini D, Vermeer MH, Meijer CJLM, Alessi E, Willemze R. Primary Cutaneous CD8-Positive Epidermotropic Cytotoxic T Cell Lymphomas. *The American Journal of Pathology*. 1999;155(2):483-492.

18. Berti E, Gaulard P, Willemze R, Petrella T, Jaffe E. Primary cutaneous CD8+ aggressive epidermotropic cytotoxic T-cell lymphoma. In: Swerldow SH, Campo E, Harris N, et al., eds. WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues 4th ed. Lyon: IARC; 2017.
19. Guitart J, Martinez-Escala ME, Subtil A, et al. Primary cutaneous aggressive epidermotropic cytotoxic T-cell lymphomas: reappraisal of a provisional entity in the 2016 WHO classification of cutaneous lymphomas. *Mod Pathol*. 2017;30(5):761-772.
20. Pagano L, Valentini CG, Grammatico S, Pulsoni A. Blastic plasmacytoid dendritic cell neoplasm: diagnostic criteria and therapeutical approaches. *Br J Haematol*. 2016;174(2):188-202.
21. Petrella T. FF, Pileri S.A. Blastic plasmacytoid dendritic cell neoplasm. In: Elder D.E. MD, Scolyer R.A., Willemze R., ed. WHO Classification of Skin Tumours Lyon: IARC; 2018.
22. Bastidas Torres AN, Najidh S, Tensen CP, Vermeer MH. Molecular advances in cutaneous T-cell lymphoma. *Semin Cutan Med Surg*. 2018;37(1):81-86.
23. Simon R, Roychowdhury S. Implementing personalized cancer genomics in clinical trials. *Nat Rev Drug Discov*. 2013;12(5):358-369.
24. Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell*. 2000;100(1):57-70.
25. Pecorino L. Molecular Biology of Cancer: Mechanisms, Targets and Therapeutics (ed 3rd). Oxford: Oxford University Press; 2012.
26. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell*. 2011;144(5):646-674.
27. Lodish H, Berk A, Kaiser C, et al. Molecular Cell Biology (ed 8th). New York: W. H. Freeman 2016.
28. van Doorn R, van Kester MS, Dijkman R, et al. Oncogenomic analysis of mycosis fungoides reveals major differences with Sezary syndrome. *Blood*. 2009;113(1):127-136.
29. van Kester MS, Ballabio E, Benner MF, et al. miRNA expression profiling of mycosis fungoides. *Mol Oncol*. 2011;5(3):273-280.
30. Mao X, Lillington D, Scarisbrick JJ, et al. Molecular cytogenetic analysis of cutaneous T-cell lymphomas: identification of common genetic alterations in Sézary syndrome and mycosis fungoides. *Br J Dermatol*. 2002;147(3):464-475.
31. Fischer TC, Gellrich S, Muche JM, et al. Genomic aberrations and survival in cutaneous T cell lymphomas. *J Invest Dermatol*. 2004;122(3):579-586.
32. Vaque JP, Gomez-Lopez G, Monsalvez V, et al. PLCG1 mutations in cutaneous T-cell lymphomas. *Blood*. 2014;123(13):2034-2043.
33. Ungewickell A, Bhaduri A, Rios E, et al. Genomic analysis of mycosis fungoides and Sezary syndrome identifies recurrent alterations in TNFR2. *Nat Genet*. 2015;47(9):1056-1060.
34. McGirt LY, Jia P, Baerenwald DA, et al. Whole-genome sequencing reveals oncogenic mutations in mycosis fungoides. *Blood*. 2015;126(4):508-519.
35. da Silva Almeida AC, Abate F, Khiabanian H, et al. The mutational landscape of cutaneous T cell lymphoma and Sezary syndrome. *Nat Genet*. 2015;47(12):1465-1470.
36. Caumont C, Gros A, Boucher C, et al. PLCG1 Gene Mutations Are Uncommon in Cutaneous T-Cell Lymphomas. *J Invest Dermatol*. 2015;135(9):2334-2337.
37. Prochazkova M, Chevret E, Mainhaguiet G, et al. Common chromosomal abnormalities in mycosis fungoides transformation. *Genes Chromosomes Cancer*. 2007;46(9):828-838.
38. Marosvári D, Téglási V, Csala I, et al. Altered microRNA expression in folliculotropic and transformed mycosis fungoides. *Pathol Oncol Res*. 2015;21(3):821-825.
39. Maj J, Jankowska-Konsur A, Sadakierska-Chudy A, Noga L, Reich A. Altered microRNA expression in mycosis fungoides. *Br J Dermatol*. 2012;166(2):331-336.
40. Ralfkiaer U, Hagedorn PH, Bangsgaard N, et al. Diagnostic microRNA profiling in cutaneous T-cell lymphoma (CTCL). *Blood*. 2011;118(22):5891-5900.
41. Garaicoa FH, Roisman A, Arias M, et al. Genomic imbalances and microRNA transcriptional profiles in patients with mycosis fungoides. *Tumour Biol*. 2016;37(10):13637-13647.
42. Shen X, Wang B, Li K, et al. MicroRNA Signatures in Diagnosis and Prognosis of Cutaneous T-Cell Lymphoma. *J Invest Dermatol*. 2018;138(9):2024-2032.

43. Moyal L, Barzilai A, Gorovitz B, et al. miR-155 is involved in tumor progression of mycosis fungoides. *Exp Dermatol*. 2013;22(6):431-433.
44. Moyal L, Yehezkel S, Gorovitz B, et al. Oncogenic role of microRNA-155 in mycosis fungoides: an in vitro and xenograft mouse model study. *Br J Dermatol*. 2017;177(3):791-800.
45. Dusílková N, Bašová P, Polívka J, et al. Plasma miR-155, miR-203, and miR-205 are Biomarkers for Monitoring of Primary Cutaneous T-Cell Lymphomas. *Int J Mol Sci*. 2017;18(10).
46. Marstrand T, Ahler CB, Ralfkiaer U, et al. Validation of a diagnostic microRNA classifier in cutaneous T-cell lymphomas. *Leuk Lymphoma*. 2014;55(4):957-958.
47. Lindahl LM, Besenbacher S, Rittig AH, et al. Prognostic miRNA classifier in early-stage mycosis fungoides: development and validation in a Danish nationwide study. *Blood*. 2018;131(7):759-770.
48. McGirt LY, Baerenwald DA, Vonderheid EC, Eischen CM. Early changes in miRNA expression are predictive of response to extracorporeal photopheresis in cutaneous T-cell lymphoma. *J Eur Acad Dermatol Venereol*. 2015;29(11):2269-2271.
49. McGirt LY, Adams CM, Baerenwald DA, Zwerner JP, Zic JA, Eischen CM. miR-223 regulates cell growth and targets proto-oncogenes in mycosis fungoides/cutaneous T-cell lymphoma. *J Invest Dermatol*. 2014;134(4):1101-1107.
50. van Kester MS, Borg MK, Zoutman WH, et al. A meta-analysis of gene expression data identifies a molecular signature characteristic for tumor-stage mycosis fungoides. *J Invest Dermatol*. 2012;132(8):2050-2059.
51. Prieto-Torres L, Rodriguez-Pinilla SM, Onaindia A, Ara M, Requena L, Piris M. CD30-positive primary cutaneous lymphoproliferative disorders: molecular alterations and targeted therapies. *Haematologica*. 2019;104(2):226-235.
52. Wada DA, Law ME, Hsi ED, et al. Specificity of IRF4 translocations for primary cutaneous anaplastic large cell lymphoma: a multicenter study of 204 skin biopsies. *Mod Pathol*. 2011;24(4):596-605.
53. Feldman AL, Dogan A, Smith DI, et al. Discovery of recurrent t(6;7)(p25.3;q32.3) translocations in ALK-negative anaplastic large cell lymphomas by massively parallel genomic sequencing. *Blood*. 2011;117(3):915-919.
54. Feldman AL, Law M, Remstein ED, et al. Recurrent translocations involving the IRF4 oncogene locus in peripheral T-cell lymphomas. *Leukemia*. 2009;23(3):574-580.
55. Kiran T, Demirkesen C, Eker C, Kumusoglu H, Tuzuner N. The significance of MUM1/IRF4 protein expression and IRF4 translocation of CD30(+) cutaneous T-cell lymphoproliferative disorders: a study of 53 cases. *Leuk Res*. 2013;37(4):396-400.
56. Fauconneau A, Pham-Ledard A, Cappellen D, et al. Assessment of diagnostic criteria between primary cutaneous anaplastic large-cell lymphoma and CD30-rich transformed mycosis fungoides; a study of 66 cases. *Br J Dermatol*. 2015;172(6):1547-1554.
57. Mélard P, Idrissi Y, Andrique L, et al. Molecular alterations and tumor suppressive function of the DUSP22 (Dual Specificity Phosphatase 22) gene in peripheral T-cell lymphoma subtypes. *Oncotarget*. 2016;7(42):68734-68748.
58. Sekine Y, Tsuji S, Ikeda O, et al. Regulation of STAT3-mediated signaling by LMW-DSP2. *Oncogene*. 2006;25(42):5801-5806.
59. Li JP, Yang CY, Chuang HC, et al. The phosphatase JKAP/DUSP22 inhibits T-cell receptor signalling and autoimmunity by inactivating Lck. *Nat Commun*. 2014;5:3618.
60. Xing X, Feldman AL. Anaplastic large cell lymphomas: ALK positive, ALK negative, and primary cutaneous. *Adv Anat Pathol*. 2015;22(1):29-49.
61. van Kester MS, Tensen CP, Vermeer MH, et al. Cutaneous anaplastic large cell lymphoma and peripheral T-cell lymphoma NOS show distinct chromosomal alterations and differential expression of chemokine receptors and apoptosis regulators. *J Invest Dermatol*. 2010;130(2):563-575.
62. Oschlies I, Lisfeld J, Lamant L, et al. ALK-positive anaplastic large cell lymphoma limited to the skin: clinical, histopathological and molecular analysis of 6 pediatric cases. A report from the ALCL99 study. *Haematologica*. 2013;98(1):50-56.

63. Vasmatzis G, Johnson SH, Knudson RA, et al. Genome-wide analysis reveals recurrent structural abnormalities of TP63 and other p53-related genes in peripheral T-cell lymphomas. *Blood*. 2012;120(11):2280-2289.
64. Schrader AM, Chung YY, Jansen PM, et al. No TP63 rearrangements in a selected group of primary cutaneous CD30+ lymphoproliferative disorders with aggressive clinical course. *Blood*. 2016;128(1):141-143.
65. Velusamy T, Kiel MJ, Sahasrabuddhe AA, et al. A novel recurrent NPM1-TYK2 gene fusion in cutaneous CD30-positive lymphoproliferative disorders. *Blood*. 2014;124(25):3768-3771.
66. Laharanne E, Oumouhou N, Bonnet F, et al. Genome-wide analysis of cutaneous T-cell lymphomas identifies three clinically relevant classes. *J Invest Dermatol*. 2010;130(6):1707-1718.
67. Mao X, Orchard G, Lillington DM, Russell-Jones R, Young BD, Whittaker S. Genetic alterations in primary cutaneous CD30+ anaplastic large cell lymphoma. *Genes Chromosomes Cancer*. 2003;37(2):176-185.
68. Crescenzo R, Abate F, Lasorsa E, et al. Convergent mutations and kinase fusions lead to oncogenic STAT3 activation in anaplastic large cell lymphoma. *Cancer Cell*. 2015;27(4):516-532.
69. Luchtel RA, Zimmermann MT, Hu G, et al. Recurrent MSC (E116K) mutations in ALK-negative anaplastic large cell lymphoma. *Blood*. 2019;133(26):2776-2789.
70. Yi S, Sun J, Qiu L, et al. Dual Role of EZH2 in Cutaneous Anaplastic Large Cell Lymphoma: Promoting Tumor Cell Survival and Regulating Tumor Microenvironment. *J Invest Dermatol*. 2018;138(5):1126-1136.
71. Kamstrup MR, Ralfkiaer E, Skovgaard GL, Gniadecki R. Potential involvement of Notch1 signalling in the pathogenesis of primary cutaneous CD30-positive lymphoproliferative disorders. *Br J Dermatol*. 2008;158(4):747-753.
72. Kamstrup MR, Biskup E, Gniadecki R. Notch signalling in primary cutaneous CD30+ lymphoproliferative disorders: a new therapeutic approach? *Br J Dermatol*. 2010;163(4):781-788.
73. Benner MF, Ballabio E, van Kester MS, et al. Primary cutaneous anaplastic large cell lymphoma shows a distinct miRNA expression profile and reveals differences from tumor-stage mycosis fungoides. *Exp Dermatol*. 2012;21(8):632-634.
74. Sandoval J, Diaz-Lagares A, Salgado R, et al. MicroRNA expression profiling and DNA methylation signature for deregulated microRNA in cutaneous T-cell lymphoma. *J Invest Dermatol*. 2015;135(4):1128-1137.
75. Liu C, Iqbal J, Teruya-Feldstein J, et al. MicroRNA expression profiling identifies molecular signatures associated with anaplastic large cell lymphoma. *Blood*. 2013;122(12):2083-2092.
76. Lawrie CH. MicroRNA expression in lymphoid malignancies: new hope for diagnosis and therapy? *J Cell Mol Med*. 2008;12(5a):1432-1444.
77. Kato K, Oh Y, Takita J, et al. Molecular genetic and cytogenetic analysis of a primary cutaneous CD8-positive aggressive epidermotropic cytotoxic T-cell lymphoma. *Int J Hematol*. 2016;103(2):196-201.
78. Tomasini C, Novelli M, Fanoni D, Berti EF. Erythema multiforme-like lesions in primary cutaneous aggressive cytotoxic epidermotropic CD8+ T-cell lymphoma: A diagnostic and therapeutic challenge. *J Cutan Pathol*. 2017;44(10):867-873.
79. Fanoni D, Corti L, Alberti-Violetti S, et al. Array-based CGH of primary cutaneous CD8+ aggressive EPIDERMO-tropic cytotoxic T-cell lymphoma. *Genes Chromosomes Cancer*. 2018;57(12):622-629.
80. Suzuki K, Suzuki Y, Hama A, et al. Recurrent MYB rearrangement in blastic plasmacytoid dendritic cell neoplasm. *Leukemia*. 2017;31(7):1629-1633.
81. Sakamoto K, Katayama R, Asaka R, et al. Recurrent 8q24 rearrangement in blastic plasmacytoid dendritic cell neoplasm: association with immunoblastoid cytomorphology, MYC expression, and drug response. *Leukemia*. 2018;32(12):2590-2603.
82. Nakamura Y, Kayano H, Kakegawa E, et al. Identification of SUPT3H as a novel 8q24/MYC partner in blastic plasmacytoid dendritic cell neoplasm with t(6;8)(p21;q24) translocation. *Blood Cancer J*. 2015;5:e301.

83. Tokuda K, Eguchi-Ishimae M, Yagi C, et al. CLTC-ALK fusion as a primary event in congenital blastic plasmacytoid dendritic cell neoplasm. *Genes Chromosomes Cancer*. 2014;53(1):78-89.
84. Gao NA, Wang XX, Sun JR, Yu WZ, Guo NJ. Blastic plasmacytoid dendritic cell neoplasm with leukemic manifestation and ETV6 gene rearrangement: A case report. *Exp Ther Med*. 2015;9(4):1109-1112.
85. Cao Q, Liu F, Niu G, Xue L, Han A. Blastic plasmacytoid dendritic cell neoplasm with EWSR1 gene rearrangement. *J Clin Pathol*. 2014;67(1):90-92.
86. Toya T, Nishimoto N, Koya J, et al. The first case of blastic plasmacytoid dendritic cell neoplasm with MLL-ENL rearrangement. *Leuk Res*. 2012;36(1):117-118.
87. Yang N, Huh J, Chung WS, Cho MS, Ryu KH, Chung HS. KMT2A (MLL)-MLLT1 rearrangement in blastic plasmacytoid dendritic cell neoplasm. *Cancer Genet*. 2015;208(9):464-467.
88. Wiesner T, Obenauf AC, Cota C, Fried I, Speicher MR, Cerroni L. Alterations of the cell-cycle inhibitors p27(KIP1) and p16(INK4a) are frequent in blastic plasmacytoid dendritic cell neoplasms. *J Invest Dermatol*. 2010;130(4):1152-1157.
89. Oiso N, Tatsumi Y, Arao T, et al. Loss of genomic DNA copy numbers in the p18, p16, p27 and RB loci in blastic plasmacytoid dendritic cell neoplasm. *Eur J Dermatol*. 2012;22(3):393-394.
90. Lucioni M, Novara F, Fiandrino G, et al. Twenty-one cases of blastic plasmacytoid dendritic cell neoplasm: focus on biallelic locus 9p21.3 deletion. *Blood*. 2011;118(17):4591-4594.
91. Jardin F, Callanan M, Penther D, et al. Recurrent genomic aberrations combined with deletions of various tumour suppressor genes may deregulate the G1/S transition in CD4+CD56+ haematodermic neoplasms and contribute to the aggressiveness of the disease. *Leukemia*. 2009;23(4):698-707.
92. Dijkman R, van Doorn R, Szuhai K, Willemze R, Vermeer MH, Tensen CP. Gene-expression profiling and array-based CGH classify CD4+CD56+ hematodermic neoplasm and cutaneous myelomonocytic leukemia as distinct disease entities. *Blood*. 2007;109(4):1720-1727.
93. Emadali A, Hoghoughi N, Duley S, et al. Haploinsufficiency for NR3C1, the gene encoding the glucocorticoid receptor, in blastic plasmacytoid dendritic cell neoplasms. *Blood*. 2016;127(24):3040-3053.
94. Togami K, Madan V, Li J, et al. Blastic Plasmacytoid Dendritic Cell Neoplasm (BPDCN) Harbors Frequent Splicesosome Mutations That Cause Aberrant RNA Splicing Affecting Genes Critical in pDC Differentiation and Function. *Blood*. 2016;128(22):738-738.
95. Taylor J, Kim SS, Stevenson KE, et al. Loss-Of-Function Mutations In The Splicing Factor ZRSR2 Are Common In Blastic Plasmacytoid Dendritic Cell Neoplasm and Have Male Predominance. *Blood*. 2013;122(21):741-741.
96. Szczepaniak A, Machnicki M, Gniot M, et al. Germline missense NF1 mutation in an elderly patient with a blastic plasmacytoid dendritic cell neoplasm. *Int J Hematol*. 2019;110(1):102-106.
97. Stenzinger A, Endris V, Pfarr N, et al. Targeted ultra-deep sequencing reveals recurrent and mutually exclusive mutations of cancer genes in blastic plasmacytoid dendritic cell neoplasm. *Oncotarget*. 2014;5(15):6404-6413.
98. Sapienza MR, Abate F, Melle F, et al. Blastic plasmacytoid dendritic cell neoplasm: genomics mark epigenetic dysregulation as a primary therapeutic target. *Haematologica*. 2019;104(4):729-737.
99. Menezes J, Acquadro F, Wiseman M, et al. Exome sequencing reveals novel and recurrent mutations with clinical impact in blastic plasmacytoid dendritic cell neoplasm. *Leukemia*. 2014;28(4):823-829.
100. Ladikou E, Ottolini B, Nawaz N, et al. Clonal evolution in the transition from cutaneous disease to acute leukemia suggested by liquid biopsy in blastic plasmacytoid dendritic cell neoplasm. *Haematologica*. 2018;103(5):e196-e199.
101. Jardin F, Ruminy P, Parmentier F, et al. TET2 and TP53 mutations are frequently observed in blastic plasmacytoid dendritic cell neoplasm. *Br J Haematol*. 2011;153(3):413-416.
102. Alayed K, Patel KP, Konoplev S, et al. TET2 mutations, myelodysplastic features, and a distinct immunoprofile characterize blastic plasmacytoid dendritic cell neoplasm in the bone marrow. *Am J Hematol*. 2013;88(12):1055-1061.

103. Sapienza MR, Fuligni F, Agostinelli C, et al. Molecular profiling of blastic plasmacytoid dendritic cell neoplasm reveals a unique pattern and suggests selective sensitivity to NF-kB pathway inhibition. *Leukemia*. 2014;28(8):1606-1616.
104. Ceroi A, Masson D, Roggy A, et al. LXR agonist treatment of blastic plasmacytoid dendritic cell neoplasm restores cholesterol efflux and triggers apoptosis. *Blood*. 2016;128(23):2694-2707.
105. Ceribelli M, Hou ZE, Kelly PN, et al. A Druggable TCF4- and BRD4-Dependent Transcriptional Network Sustains Malignancy in Blastic Plasmacytoid Dendritic Cell Neoplasm. *Cancer Cell*. 2016;30(5):764-778.

2

Molecular advances in cutaneous T-cell lymphoma

Armando N. Bastidas Torres
Safa Najidh
Cornelis P. Tensen
Maarten H. Vermeer

Seminars in Cutaneous Medicine and Surgery 37: 81–86 (2018)



ABSTRACT

Cutaneous T-cell lymphoma (CTCL) is a group of malignancies derived from skin-homing T cells. Mycosis fungoides (MF) and Sézary syndrome (SS), are the most common CTCL variants. In recent years, the genetic landscape of SS/MF has been characterized using genome-wide next-generation sequencing (NGS) approaches. These studies have revealed that genes subjected to oncogenic mutations take part in cell cycle regulation, chromatin modification, JAK-STAT signaling, TCR/NF-κB signaling and MAPK signaling, which suggests that deregulation of these cellular processes underlie lymphomagenesis. These studies provide the groundwork for functional and clinical studies that will lead to better risk assessment and more effective therapeutic approach in CTCL patients.

1. INTRODUCTION

Cutaneous T-cell lymphomas (CTCLs) are malignancies derived from skin-homing T cells. Mycosis fungoides (MF), the most common CTCL, typically presents with patches and plaques especially on sun-protected areas of the skin. The disease has an indolent clinical course with slow progression from patches to more infiltrated plaques and eventually to tumors. In a minority of patients, extracutaneous localization develops in later stages of the disease. The prognosis of MF patients depends on stage, in particular, the type and extent of skin lesions and presence of extracutaneous disease.[1] While patients with early stages of disease have an excellent prognosis, the 10-year disease-specific survival of patients with tumor stage disease and patients with histologically documented lymph node involvement is 42% and 20%, respectively.[2]

Sézary syndrome (SS), on the contrary, is a leukemic CTCL variant. SS presents with pruritic erythroderma, lymphadenopathy and presence of tumor cells (known as Sézary cells) in skin, lymph nodes and peripheral blood. Patients with SS have a poor prognosis, with a 5-year survival of 30%.[1,2]

The genetic alterations in CTCL, in particular SS, and to a lesser extent MF, have been studied using cytogenetic and array-based methods. These studies identified extensive genetic instability with complex karyotypes but no highly recurrent translocations. In contrast, recurrent copy number alteration (CNA) were identified including deletion of *CDKN2A* in both SS and MF, deletion of *RB1* in SS, and gain of *MYC* in SS.[3,4]

In recent years, genome-wide studies using next-generation sequencing (NGS) have provided a more comprehensive picture of the genetic landscape of CTCL. Most of these studies have focused exclusively on SS (7 of 12),[5-11] while some others included MF samples together with SS samples (4 of 12),[12-15] and one study focused on MF only (Table 1).[16] Taken together, these studies include ~8 times more SS samples than MF samples, making the reported findings mostly a molecular representation of SS. Recently, all CTCL samples with publicly available sequencing data have been re-analyzed applying uniform methods and metrics to generate a more representative overview of point mutation frequencies in putative driver genes.[17] This article summarizes the most relevant findings in CTCL genetics derived from NGS studies and examines their therapeutic implications.

2. NEXT GENERATION SEQUENCING

NGS refers to a group of high-throughput technologies used to determine the order of nucleotides within nucleic acid molecules (i.e. DNA and RNA). In cancer research, NGS is particularly suitable for the discovery of different types of genetic alterations ranging from pathogenic single nucleotide substitutions to large chromosomal aberrations in cancer genomes. The terms whole-genome sequencing (WGS), whole-exome sequencing (WES) and targeted sequencing (TS) refer to NGS approaches that characterize the entire genome, the complete set of coding genes (known as exome, <2% of the genome) or a defined group of individual genes, respectively.[18,19] Another NGS application called RNA sequencing (RNA-seq) is used to characterize the collection of expressed genes in the cell known as the transcriptome, which makes possible the identification of abnormal gene expression patterns in malignancies when compared to healthy controls, and the detection of expressed gene fusions derived from chromosomal rearrangements.[18] The large majority of studies reviewed in this article have made use of WES and/or TS to investigate SS/MF, whereas a minority of them opted for WGS or RNA-seq.

3. DISCOVERY OF MOLECULAR ALTERATIONS IN SS/MF USING NGS

NGS data derived from the genome or transcriptome are analyzed using a variety of bioinformatic tools that allow the identification of genetic alterations or abnormal gene expression, respectively. In the case of SS/MF, these analyses have uncovered a distinctive mutational landscape and revealed that genes affected by point mutations and CNAs cluster mainly into five cellular processes: cell cycle regulation, chromatin modification, JAK-STAT signaling, TCR/NF-κB signaling and MAPK signaling (Figure 1). An overview of these findings is presented below.

Table 1. NGS studies on CTCL. The majority of published studies used WES to characterize the mutational landscape of SS/MF genomes. CTCL, cutaneous T cell lymphoma. SS, Sézary syndrome. MF, mycosis fungoides. WGS, whole-genome sequencing. WES, whole-exome sequencing. TS, targeted sequencing. RNA-seq, RNA sequencing.

Authors	Journal	Year	CTCL variant (# of samples)	WGS	WES	TS	RNA-seq
Lee et al.	Blood	2012	SS (3), MF (24)				27
Sekulic et al.	Mol Genet Genomic Med	2014	SS (1)	1			1
Vaqué et al.	Blood	2014	SS (4), MF (7)			11	
Choi et al.	Nat Genet	2015	SS (40)	2	40		
McGirt et al.	Blood	2015	MF (30)	5		25	
Ungewickell et al.	Nat Genet	2015	SS (32), MF (41)		11	73	
Kiel et al.	Nat Comm	2015	SS (66)	6	66		
Da Silva Almeida et al.	Nat Genet	2015	SS (25), MF (8), Other CTCLs (9)		42		
Wang et al.	Nat Genet	2015	SS (37)		37	37	32
Prasad et al.	J Invest Dermatol	2016	SS (12)		12		10
Woollard et al.	Blood	2016	SS (101)		10	101	
Izykowska et al.	Oncotarget	2017	SS (9)	9			9

4. THE GENOMIC LANDSCAPE OF SS/MF

The somatic mutation rate in SS (i.e. average, 3.85 mutations/Mb) is similar to mutation rates observed in adult solid tumors.[6] The majority of somatic single nucleotide substitutions observed in SS/MF are C>T transitions (40%-75%),[6,16] which are much less common in other hematological cancers.[6] This mutational signature is regarded as caused by UV light exposure when occurring at NpCpG sites,[6,11,16] or aging when occurring at NpCpC sites. In SS, both aging (43%) and UV radiation (30%) contribute to C>T transitions.[6,20] At present it is unclear whether exposure to UV light plays a causative role in the onset of the disease.

The fact that oncogenic alterations in SS much more frequently result from CNAs rather than single nucleotide substitutions suggests that the former may play an important role in SS lymphomagenesis.[11] Choi and colleagues reported complex genomic rearrangements associated with deletion of tumor suppressors in SS, suggesting that phenomena such as chromothripsis or chromoplexy could be key

events in the development of CTCL.[11] These authors proposed that deregulation of a group of DNA-cutting enzymes encoded by recombination activated genes (RAGs) may be implicated in the characteristic genomic instability of SS.[11]

4.1 Cell cycle and DNA repair

The cell cycle is a series of systematic events that regulate the correct division of a cell into two daughter cells. Although cyclins and cyclin-dependent kinases (CDKs) govern cell cycle progression, multiple other proteins participate indirectly in the regulation of the cell cycle.[21] For instance, ATM, RB1, TP53 and CDK inhibitors (CKIs) act on different signaling levels to ultimately induce cell cycle arrest in response to DNA damage. In SS, CKIs encoded by *CDKN1A* and *CDKN2A*, are deleted in 11% and 40% of cases, respectively.[6,11] Deletions of *ATM* (30%), *RB1* (39%) and *TP53* (93%) are frequent as well.[6,11]

4.2 JAK-STAT signaling

Cytokines are small secreted proteins that mediate communication between cells. Receptors on T cells bind specific cytokines, become activated, and elicit changes in gene expression through intracellular signaling pathways. The JAK-STAT pathway is central to cytokine stimulation in T cells.[22] Cytokine-bound receptors activate JAK kinases, which then activate STAT proteins. Activated STAT proteins dimerize and translocate to the nucleus where they induce expression of genes responsible for T-cell proliferation, differentiation and apoptosis.[22]

Point mutations either predicted or confirmed to be gain-of-function are reported in *JAK1* (0.9% of cases), *JAK3* (2.7%), *STAT3* (0.9%) and *STAT5B* (3.6%).[17] Hyperactive JAK and STAT proteins often result from structural changes in their functional domains (Figure 1). For instance, amino acid substitutions in the pseudokinase domain of JAK proteins enhance their tyrosine kinase activity while changes in the SH2 domain of STAT proteins are predicted to enhance dimerization. In addition, copy number gains of *JAK2* (13% of cases), *STAT3* (60%) and *STAT5B* (60%) are frequent and have been shown to correlate with increased expression (Figure 1).[17]

The existence of FDA-approved JAK inhibitors such as Tofacitinib and Ruxolitinib makes the idea of targeting malignant T cells driven by hyperactive JAK kinases attractive. In two separate NGS studies, the authors demonstrated that CTCL cell lines bearing gain-of-function point mutations in JAK3 are sensitive to these compounds, opening the possibility for their potential use in CTCL patients carrying JAK mutations.[15,16]

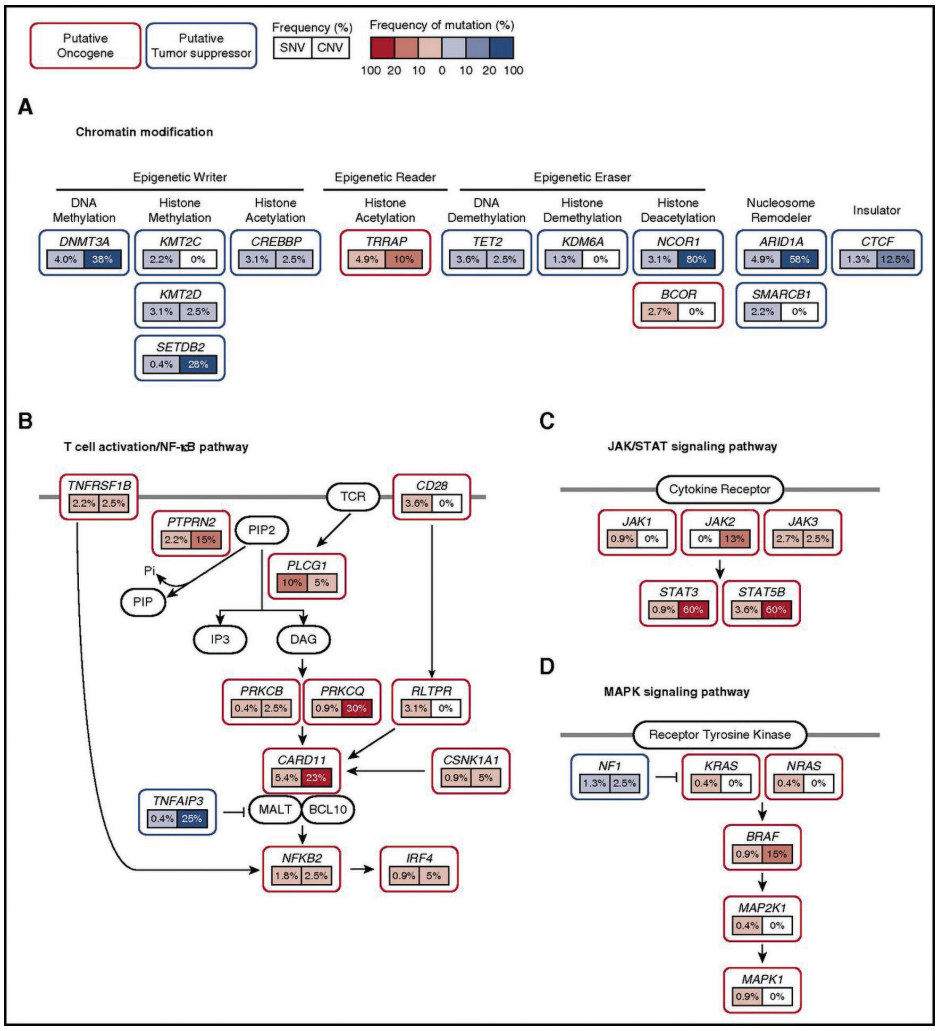


Figure. Schematic of mutations in recurrently mutated signaling pathways in CTCL. CTCL harbors recurrent mutations that are predicted to affect (A) chromatin, (B) T-cell activation/NF-kB signaling, (C) JAK/STAT signaling, and (D) MAPK signaling. From: Park J, Yang J, Wenzel AT, et al. Genomic analysis of 200 CTCLs identifies a novel recurrent gain-of-function alteration in RLTPR (p.Q575E). Page 1433, Blood, 2017;130(12):1433. doi: 10.1182/blood-2017-02-768234. Used with permission from the American Society of Hematology.

4.3 MAPK signaling

Similar to the JAK-STAT pathway, the MAPK pathway propagates towards the nucleus extracellular signals received through cytokine receptors. These signals elicit cellular processes such as growth, proliferation and apoptosis.[23] Subsets

of SS/MF patients carry established gain-of-function mutations (<1%) in several members of this pathway such as *KRAS*, *NRAS*, *BRAF*, *MAP2K1* and *MAPK1* (Figure 1).[17]

4.4 TCR/NF-κB signaling

Initiation of the adaptive immune response against pathogens relies on the activation of T cells capable of recognizing foreign antigens. Upon binding between the T-cell receptor (TCR) and its cognate antigen, a series of signaling events lead to the expression of genes that promote T-cell survival, proliferation and differentiation.[24] Activated T cells then orchestrate multiple immunological events directed at clearing the pathogen from the body.

In CTCLs, hyperactive TCR signaling might drive uncontrolled proliferation of malignant T cells. Genetic alterations reported in SS/MF target a variety of components of the TCR pathway.[6,11,13,15] CD28 is a co-stimulatory molecule of the TCR complex, which by engaging ligand B7 (CD80 and CD86) on antigen presenting cells (APCs), co-signals for T-cell activation.[24] Increased affinity of CD28 for B7 due to amino acid substitutions or gene fusions results in augmented TCR signaling. Point mutations in CD28 occur in 3.6% of cases while in-frame gene fusions (i.e. *CD28-ICOS* and *CD28-CTLA4*) have been observed in a small number of cases (Figure 1).[6,7,13,17]

Downstream to TCR activation, phospholipase PLCγ1 catalyzes the production of second messengers (i.e. DAG and IP3), which in turn, activate downstream kinases.[24] Point mutations in the catalytic core of PLCγ1 (10% of cases) and copy number gains (5%) increase second messenger production which enhance antigen-dependent T-cell activation (Figure 1).[12,17] Furthermore, point mutations in DAG-activated kinases PRKCB and PRKCQ (<1% of cases) and amplification of the latter (30%) may contribute to TCR signaling deregulation (Figure 1).[17] However, the functional relevance of these alterations remains uncertain.

Gene expression changes triggered by TCR activation are effected through canonical NF-κB signaling and other downstream pathways.[24] *CARD11* encodes a scaffold protein needed for the activation of transcription factor complex NF-κB. In SS/MF, point mutations (5.4% of cases) and copy number gains (23%) involving *CARD11* lead to increased NF-κB activation (Figure 1).[13,15,17] In addition, recurrent deletions (2.5% of cases) of the C-terminal auto-inhibitory IκB domain of NFKB2 are expected to render it constitutively active.[13,17]

Besides the TCR complex, other receptors can also elicit canonical NF-κB signaling in CTCL. Membrane receptor TNFR2 (encoded by *TNFRSF1B*) initiates NF-κB signaling upon binding to ligand TNFα. Reported activating point

mutations in *TNFRSF1B* enable constitutive NF-κB signaling without the need of a ligand.[13] Overall, point mutations and amplifications involving *TNFRSF1B* are reported in 2.2% and 2.5% of cases, respectively .[13,17] In addition, *TNFAIP3*, which encodes an inhibitor of TNFα-induced NF-κB signaling, is deleted in 25% of SS cases (Figure 1).

In line with these studies, pharmacological inhibition of complex NF-κB decreased its DNA binding ability and induced cell death of SS cells *in vitro*.[25,26] In a phase II study the safety and efficacy of the NF-κB inhibitor bortezomib were evaluated in 10 MF patients.[27] The overall response rate (ORR) for MF was 70 % and the complete response rate (CRR) was 10 % suggesting the potential usefulness of NF-κB inhibitors in CTCL treatment.[27]

4.5 Chromatin modification

Chromatin consists of genomic DNA and associated proteins. Histone proteins, the main structural components of chromatin, form a protein core around which DNA winds called the nucleosome. Histones not only play roles in the packaging, replication, and repair of DNA, but also regulate gene expression.[28,29] Both DNA and histones are subjected to chemical (known as epigenetic) modifications that direct changes in chromatin architecture. DNA methylation, histone methylation, and histone acetylation are epigenetic modifications with established effects on gene expression.[28,29]

Proteins that chemically alter chromatin, named epigenetic regulators, can be divided into three categories: writers, erasers and readers. Writers are enzymes that add epigenetic marks to either DNA or histones, erasers reverse the action of writers, and readers interpret epigenetic marks to elicit downstream signals.[28,29] Epigenetic regulators are frequently mutated in CTCLs, particularly, SS.

4.5.1 DNA methylation

In normal cells, widespread DNA methylation maintains chromosomal stability whereas low methylation at gene promoters allows transcription. In cancer, this methylation pattern is typically inverted promoting genome instability and transcriptional misregulation. Loss-of-function alterations in DNA methyltransferases (DNMT) can lead to global hypo-methylation.[30] In SS, DNA methyltransferase DNMT3A is frequently deleted (38%) and mutated (4%) (Figure 1). Opposing DNMTs, TET enzymes demethylate DNA by oxidizing methylated cytosine residues in successive enzymatic steps. *TET2* is often inactivated in hematological malignancies, though, its role in cancer development is not clear.[31] In SS, *TET2* mutations (3.6% of cases) and deletions (2.5%) are reported too (Figure 1). In line with these results, it was shown that SS cells are characterized by

widespread changes in DNA methylation suggesting that SS could be amenable to treatment with demethylating agents.[32]

4.5.2 Histone methylation

Histone methyltransferases (HMTs) are enzymes that add one, two or three methyl groups to specific lysine or arginine residues on histones, while histone demethylases (HDMs) remove them. Histone methylation can signal either for gene activation or repression depending on the location of the epigenetic mark.[29] Aberrant methylation patterns on histones are extensively observed in malignancies and are likely to be a consequence of damaging mutations in HMTs and HDMs.[29] In SS, two histone demethylases, KMT2C and KMT2D, are mutated in 2.2% and 3.1% of cases, respectively; while a third enzyme, SETDB2, is deleted in 28% of cases. In addition, mutations in histone demethylase KDM6A are also seen in 1.3% of SS cases (Figure 1).

4.5.3 Histone Acetylation

Acetylation of lysine residues on histones activates gene expression by inducing chromatin relaxation. Acetyl groups neutralize the positive charge of lysine residues which decreases the attraction between the nucleosome and the negatively charged DNA. The addition of acetyl groups is mediated by histone acetyltransferases (HATs). In SS, histone acetyltransferase CREBBP, can be mutated (3.1% of cases) or deleted (2.5%). By contrast, removal of acetyl groups is carried out by histone deacetylases (HDACs). Deletions involving histone deacetylase NCOR1 are frequent in SS/MF (80%) while point mutations are less common (3.1%) (Figure 1). Another histone deacetylase, BCOR, is mutated in 2.7% of cases.[17] Finally, adapter protein TRRAP which is part of various multiprotein complexes with HAT activity, is amplified in 10% of cases and mutated in 4.9% of cases (Figure 1).[17]

Two systemic HDAC inhibitors (HDACis) are FDA-approved for the treatment of CTCL, including romidepsin (class I-specific HDACi) and vorinostat (pan-HDACi). Trials with (novel) HDACis as single agents or in combinations are now underway in diverse cancer types, including hematologic and solid tumors, but have not yet reached the clinic.[33] However, insight in the mechanisms through which different HDACis suppress tumors is limited, and elucidation on how epigenetic drugs exert their anti-cancer effect will be crucial to further develop and optimize these therapies.

4.5.4 Chromatin remodeling

Chromatin remodelers are multiprotein complexes that change the degree of chromatin compaction and nucleosome position in response to upstream epigenetic signals. These changes in chromatin architecture regulate gene

expression. Genetic alterations involving members of the SWI/SNF family of chromatin remodelers are common in human cancers.[34] For instance, *ARID1A* and *SMARCB1* are frequently mutated in gastric and rhabdoid tumors, respectively.[34] In SS, deletion of *ARID1A* (58%), *ARID5B* (29%) and *SMARCC1* (21%) are the most frequent alterations involving chromatin remodeling genes.[9,17] Furthermore, *ARID1A* deletions have been shown to correlate with lower expression levels in SS.[11] Similarly, point mutations in chromatin remodeling genes *ARID2* and *SMARCB1* have been reported in subsets of MF patients.[16]

5. CONCLUSION/OUTLOOK

The combined results from recent NGS studies have uncovered extensive genetic changes in SS/MF, including recurrent CNAs and point mutations that affect key pathways that likely are of critical importance in driving the disease. In the forthcoming years, it will be essential to integrate genetic studies with transcriptional data, and develop pre-clinical models to determine the functional consequences of the identified genetic alterations. These studies will guide the development of new therapeutic strategies including the targeting of frequently affected cellular processes, such as NF-κB and JAK-STAT signaling, and chromatin regulation. Prospective studies on large numbers of CTCL patients, such as the Cutaneous Lymphoma International Consortium, will be essential to correlate specific genetic alterations with clinical course and response to therapy. It is hoped that ultimately these studies will lead to an individualized risk assessment and effective therapeutic approach.

Acknowledgments
The authors are funded by the Dutch Cancer Society (grant UL2013-6104)

REFERENCES

1. Willemze R, Jaffe ES, Burg G, et al. WHO-EORTC classification for cutaneous lymphomas. *Blood*. 2005;105(10):3768-3785.
2. Agar NS, Wedgeworth E, Crichton S, et al. Survival outcomes and prognostic factors in mycosis fungoides/Sezary syndrome: validation of the revised International Society for Cutaneous Lymphomas/European Organisation for Research and Treatment of Cancer staging proposal. *J Clin Oncol*. 2010;28(31):4730-4739.
3. Vermeer MH, van Doorn R, Dijkman R, et al. Novel and highly recurrent chromosomal alterations in Sezary syndrome. *Cancer Res*. 2008;68(8):2689-2698.
4. Mao X, Lillington D, Scarisbrick JJ, et al. Molecular cytogenetic analysis of cutaneous T-cell lymphomas: identification of common genetic alterations in Sezary syndrome and mycosis fungoides. *Br J Dermatol*. 2002;147(3):464-475.
5. Woollard WJ, Pullabhatla V, Lorenc A, et al. Candidate driver genes involved in genome maintenance and DNA repair in Sezary syndrome. *Blood*. 2016;127(26):3387-3397.
6. Wang L, Ni X, Covington KR, et al. Genomic profiling of Sezary syndrome identifies alterations of key T cell signaling and differentiation genes. *Nat Genet*. 2015;47(12):1426-1434.
7. Sekulic A, Liang WS, Tembe W, et al. Personalized treatment of Sezary syndrome by targeting a novel CTLA4:CD28 fusion. *Mol Genet Genomic Med*. 2015;3(2):130-136.
8. Prasad A, Rabionet R, Espinet B, et al. Identification of Gene Mutations and Fusion Genes in Patients with Sezary Syndrome. *J Invest Dermatol*. 2016;136(7):1490-1499.
9. Kiel MJ, Sahasrabuddhe AA, Rolland DC, et al. Genomic analyses reveal recurrent mutations in epigenetic modifiers and the JAK-STAT pathway in Sezary syndrome. *Nat Commun*. 2015;6:8470.
10. Izykowska K, Przybylski GK, Gand C, et al. Genetic rearrangements result in altered gene expression and novel fusion transcripts in Sezary syndrome. *Oncotarget*. 2017;8(24):39627-39639.
11. Choi J, Goh G, Walradt T, et al. Genomic landscape of cutaneous T cell lymphoma. *Nat Genet*. 2015;47(9):1011-1019.
12. Vaque JP, Gomez-Lopez G, Monsalvez V, et al. PLCG1 mutations in cutaneous T-cell lymphomas. *Blood*. 2014;123(13):2034-2043.
13. Ungewickell A, Bhaduri A, Rios E, et al. Genomic analysis of mycosis fungoides and Sezary syndrome identifies recurrent alterations in TNFR2. *Nat Genet*. 2015;47(9):1056-1060.
14. Lee CS, Ungewickell A, Bhaduri A, et al. Transcriptome sequencing in Sezary syndrome identifies Sezary cell and mycosis fungoides-associated lncRNAs and novel transcripts. *Blood*. 2012;120(16):3288-3297.
15. da Silva Almeida AC, Abate F, Khiabanian H, et al. The mutational landscape of cutaneous T cell lymphoma and Sezary syndrome. *Nat Genet*. 2015;47(12):1465-1470.
16. McGirt LY, Jia P, Baerenwald DA, et al. Whole-genome sequencing reveals oncogenic mutations in mycosis fungoides. *Blood*. 2015;126(4):508-519.
17. Park J, Yang J, Wenzel AT, et al. Genomic analysis of 220 CTCLs identifies a novel recurrent gain-of-function alteration in RLTPR (p.Q575E). *Blood*. 2017.
18. Koboldt DC, Steinberg KM, Larson DE, Wilson RK, Mardis ER. The next-generation sequencing revolution and its impact on genomics. *Cell*. 2013;155(1):27-38.
19. Behjati S, Tarpey PS. What is next generation sequencing? *Arch Dis Child Educ Pract Ed*. 2013;98(6):236-238.
20. Alexandrov LB, Nik-Zainal S, Wedge DC, et al. Signatures of mutational processes in human cancer. *Nature*. 2013;500(7463):415-421.
21. Otto T, Sicinski P. Cell cycle proteins as promising targets in cancer therapy. *Nat Rev Cancer*. 2017;17(2):93-115.
22. Waldmann TA, Chen J. Disorders of the JAK/STAT Pathway in T Cell Lymphoma Pathogenesis: Implications for Immunotherapy. *Annu Rev Immunol*. 2017;35:533-550.

23. Dhillon AS, Hagan S, Rath O, Kolch W. MAP kinase signalling pathways in cancer. *Oncogene*. 2007;26(22):3279-3290.
24. Huse M. The T-cell-receptor signaling network. *J Cell Sci*. 2009;122(Pt 9):1269-1273.
25. Sors A, Jean-Louis F, Pellet C, et al. Down-regulating constitutive activation of the NF-kappaB canonical pathway overcomes the resistance of cutaneous T-cell lymphoma to apoptosis. *Blood*. 2006;107(6):2354-2363.
26. Nicolay JP, Muller-Decker K, Schroeder A, et al. Dimethyl fumarate restores apoptosis sensitivity and inhibits tumor growth and metastasis in CTCL by targeting NF-kappaB. *Blood*. 2016;128(6):805-815.
27. Zinzani PL, Musuraca G, Tani M, et al. Phase II trial of proteasome inhibitor bortezomib in patients with relapsed or refractory cutaneous T-cell lymphoma. *J Clin Oncol*. 2007;25(27):4293-4297.
28. Arrowsmith CH, Bountra C, Fish PV, Lee K, Schapira M. Epigenetic protein families: a new frontier for drug discovery. *Nat Rev Drug Discov*. 2012;11(5):384-400.
29. Plass C, Pfister SM, Lindroth AM, Bogatyrova O, Claus R, Lichter P. Mutations in regulators of the epigenome and their connections to global chromatin patterns in cancer. *Nat Rev Genet*. 2013;14(11):765-780.
30. Shen H, Laird PW. Interplay between the cancer genome and epigenome. *Cell*. 2013;153(1):38-55.
31. Huang Y, Rao A. Connections between TET proteins and aberrant DNA modification in cancer. *Trends Genet*. 2014;30(10):464-474.
32. van Doorn R, Slieker RC, Boonk SE, et al. Epigenomic Analysis of Sezary Syndrome Defines Patterns of Aberrant DNA Methylation and Identifies Diagnostic Markers. *J Invest Dermatol*. 2016;136(9):1876-1884.
33. Olsen EA, Kim YH, Kuzel TM, et al. Phase IIb multicenter trial of vorinostat in patients with persistent, progressive, or treatment refractory cutaneous T-cell lymphoma. *J Clin Oncol*. 2007;25(21):3109-3115.
34. Lu C, Allis CD. SWI/SNF complex in cancer. *Nat Genet*. 2017;49(2):178-179.

3

Genomic analysis reveals recurrent deletion of JAK-STAT signaling inhibitors *HNRNPK* and *SOCS1* in mycosis fungoides

Armando N. Bastidas Torres
Davy Cats
Hailiang Mei
Karoly Szuhai
Rein Willemze
Maarten H. Vermeer
Cornelis P. Tensen

Genes, Chromosomes and Cancer 57: 653–664 (2018)



ABSTRACT

Mycosis fungoides (MF) is the most common cutaneous T-cell lymphoma (CTCL). Causative genetic alterations in MF are unknown. The low recurrence of pathogenic small-scale mutations (i.e. nucleotide substitutions, indels) in the disease, calls for the study of additional aspects of MF genetics. Here, we investigated structural genomic alterations in T-MF by integrating whole-genome sequencing and RNA-sequencing. Multiple genes with roles in cell physiology (n=113) and metabolism (n=92) were found to be impacted by genomic rearrangements, including 47 genes currently implicated in cancer. Fusion transcripts involving genes of interest such as *DOT1L*, *KDM6A*, *LIFR*, *TP53* and *TP63* were also observed. Additionally, we identified recurrent deletions of genes involved in cell cycle control, chromatin regulation, the JAK-STAT pathway and the PI-3-K pathway. Remarkably, many of these deletions result from genomic rearrangements. Deletion of tumor suppressors *HNRNPK* and *SOCS1* were the most frequent genetic alterations in MF after deletion of *CDKN2A*. Notably, *SOCS1* deletion could be detected in early-stage MF. In agreement with the observed genomic alterations, transcriptome analysis revealed up-regulation of the cell cycle, JAK-STAT, PI-3-K and developmental pathways. Our results position inactivation of *HNRNPK* and *SOCS1* as potential driver events in MF development.

1. INTRODUCTION

Mycosis fungoides (MF) is the most common type of cutaneous T-cell lymphoma (CTCL), a heterogeneous group of neoplasms derived from malignant skin-homing T cells. MF typically evolves from erythematous cutaneous patches and/or plaques to tumors. Patients with T-MF have a 10-year survival of 42%, which shows the need of a better understanding of the disease and more effective treatments.[1]

Inactivation of tumor suppressors *CDKN2A* and *CDKN2B* are established genetic alterations in MF, whereas mutations in *FAS* have been reported in subsets of patients.[2-4] In recent years, the copy number alteration (CNA), micro-RNA (miRNA) and mutational profiles of MF have been investigated using genome-wide array technologies and next generation sequencing (NGS).

Common CNAs include losses within chromosomes 1, 5, 9 and 13, and gains within chromosomes 7 and 17.[5] Highlights of miRNA expression are up-regulation of oncomirs miR-93 and miR-155.[6] Gain-of-function single nucleotide variants (SNVs) found in solitary or few cases include *JAK3* (p.A573V), *MAPK1* (p.E322K), *STAT3* (p.Y640F), *PLCG1* (p.S345F, p.S520F) and *TNFRSF1B* (p.T377I).[7-11]

Even though this body of information has shed some light on the pathogenetics of MF, driver genetic alterations remain unknown. Particularly, the low recurrence of pathogenic small-scale mutations (i.e. SNVs, indels) manifest the need of research on additional facets of MF genetics.

To date no study has provided insight into the landscape of genomic rearrangements underlying MF. Consequently, we performed an integrated Whole Genome Sequencing (WGS)/RNA-sequencing (RNA-seq) analysis of T-MF to investigate structural aberrations at base-level resolution.

Our results reveal numerous rearrangements associated with CNAs, and affecting genes involved in signal transduction and transcriptional regulation. Moreover, we identify two recurrently deleted tumor suppressors, *HNRNPK* and *SOCS1*, that are novel to MF genetics. These findings broaden our understanding of MF and provide new potential targets for treatment.

2. MATERIALS AND METHODS

2.1 Patient selection

Frozen skin biopsies from nine patients with T-MF (IIB-IVA2, Supplementary Table 1) were selected for this study. Diagnosis was based on the criteria of the WHO-EORTC classification for primary cutaneous lymphomas[1] and confirmed by an expert panel of dermatologist and pathologists. We subjected eight tumor biopsies to WGS and RNA-seq, and one biopsy to WGS only. Eighteen additional frozen tumor biopsies (IIB-IVA2, Supplementary Table 1) were used as a validation cohort. Whenever available, formalin-fixed paraffin-embedded (FFPE) tumor biopsies were used for validation experiments (sequenced and extension cohorts) by fluorescence *in situ* hybridization (FISH). Frozen and FFPE tumor biopsies contained ≥70% malignant T cells. Patient material was approved by the Leiden University Medical Center institutional review board and informed consent was obtained in accordance with the declaration of Helsinki.

2.2 DNA and RNA isolation

Genomic DNA was isolated using Genomic-tip 20/G kit (Qiagen) following the manufacturer's protocol. DNA purity (A260/280 and A260/230 ratios) was evaluated using a Nanodrop 1000 system (Nanodrop Technologies, Wilmington, CA). DNA integrity was verified by gel electrophoresis (0.7% agarose, ethidium bromide). Total RNA was isolated using RNeasy mini kit (Qiagen). RNA integrity was verified with an Agilent 2100 Bioanalyzer.

2.3 Sequencing

DNA and RNA were sequenced by the Beijing Genomics Institute (BGI). For whole-genome sequencing, DNA libraries were prepared using TruSeq Nano DNA HT sample prep kit (Illumina), which involves DNA fragmentation by Covaris sonication, end-repair, A-tailing, adapter ligation and fragment enrichment. Purified DNA fragments were subjected to paired-end sequencing (2 x 150 bp) on the Illumina HiSeq X-Ten platform. For RNA sequencing, total RNA was depleted from rRNA using Ribo-Zero Magnetic kit (Epicentre Biotechnologies, Madison, WI, USA), fragmented, and followed by first-strand cDNA synthesis, second-strand cDNA synthesis (with dUTP instead of dTTP), end repair, A-tailing, adapter ligation, Uracil-N-glycosylase treatment, and cDNA library enrichment. Purified cDNA libraries were subjected to paired-end sequencing (2 x 100 bp) on the Illumina HiSeq 4000 platform. All NGS data have been deposited in the European Genome-Phenome Archive (EGA) under study number EGAS00001002860.

2.4 Sequence data processing

For whole-genome sequencing, raw reads were processed using FastQC (v0.11.2), SeqTK (v1.0-r63), Cutadapt and Sickle (v1.33). Clean reads were aligned to the human reference genome Hg19 using BWA-mem (v0.7.10) (Supplementary Table 2). For RNA sequencing, raw reads were processed with FastQC (v0.10.1), Cutadapt (v1.5), and Sickle (v1.33). Clean reads were aligned to human reference genome Hg19 using GSNAP (release 2014-12-23). SAM alignments were compressed and indexed with Picard (v1.120), and fragment counts were obtained with HTseq (v0.6.1p1) using UCSC RefSeq annotations (downloaded 2015-07-01).

2.5 Discovery of DNA rearrangements and fusion transcripts

Genome structural variation (SV) analysis was performed using a set of tools that included Pindel (v0.2.5b8), CleverSV (v2.0rc3), Breakdancer-max (v1.4.4) and Delly (v0.6.7). Post processing of the SV calls included sorting and merging of the calls using a local script and pySVTools (v0.1.3). Each of the structural variant callers were used with default settings and following best practices. SV calls were manually verified and curated using the Integrative Genomic Viewer[12] (IGV, v2.3.78). Select events were validated by PCR or FISH. Star Fusion[13] (v0.8.0) and FusionCatcher[14] (v0.99.6a) were used to detect fusion transcripts in RNA-seq data. Fusion transcript calls were contrasted with DNA SV data and visually verified on DNA level using IGV. Rearranged genes implicated in cancer were identified using the Network of Cancer Genes 5.0 (NCG 5.0) and literature search.

2.6 Detection of CNAs

Copy number alterations (CNAs) were identified by Control-FREEC[15] using a window size of 50 Kb. The output was then subjected to a Wilcoxon rank test and a Kolmogorov-Smirnov test to generate a list of genomic regions with statistically supported copy number alterations (CNAs). CNA regions were visually verified using IGV and select CNA events were validated by ddPCR. Subclonal CNA events were visually detected (coverage changes + supporting reads) using IGV. Select subclonal events were validated by sanger sequencing.

2.7 Discovery of pathogenic SNVs

SNVs were detected using GATK (v3.5). SNVs present in the dbSNP database were filtered out. We searched for pathogenic SNVs in 1461 genes involved in signaling pathways and cellular processes previously reported as affected in CTCL.[7-9,16-22] Gene lists were retrieved from the PathCards database (http://pathcards.genecards.org/). Only SNVs predicted to produce highly deleterious amino acid substitutions by both SIFT and PolyPhen-2 were further investigated on ClinVar, COSMIC and literature.

2.8 Differential expression analysis

EdgeR (v3.14.0) was used to normalize fragments per gene counts and determine differentially expressed genes (DEG). Each MF sample was compared individually to a control group formed by seven CD4+ T-cell subsets ($T_{naïve}$, T_H1, T_H2, T_H17, T_{reg}, T_{EM}, T_{CM}; 5 samples per subset). RNA-seq data of T-cell controls were generated by Ranzani et al. and downloaded from EBI (https://www.ebi.ac.uk/).[23] Testing was performed using negative binomial generalized log-linear models. Only genes found to be commonly up-regulated or down-regulated (FDR < 0.01) in all MF samples were regarded as DEG. DEG implicated in cancer were identified using NCG 5.0 and literature search.

2.9 Functional annotation, GSEA and pathway analysis

Functional annotation of rearranged genes was performed using Panther[24] (v11.1). Gene set enrichment analysis[25] (GSEA, v2.2.4) was run as a pre-ranked analysis with 1000 permutations using the hallmarks gene set from the Molecular Signatures Database (MSigDB). Normalized enrichment scores (NES) were calculated to determine expression signatures. FDR *q* values were obtained. Pathway analysis with DEG was performed with DAVID[26] (v6.8) using default settings.

2.10 FISH

Fluorescence *in situ* hybridization (FISH) for CLEC16A was performed on formalin-fixed paraffin-embedded (FFPE) tumor and plaque biopsies using bacterial artificial chromosome (BAC) probes. For all tumor samples except one (MF5), probe mix A (RP11-727C18, RP11-916G12 and RP11-959J23; telomeric/5′) and probe mix B (RP11-722I5, RP11-829F21 and RP11-936M1; centromeric/3′) were employed (Supplementary Figure 1; Supplementary Table 3). For sample MF5, probe mix A (telomeric/5′) and RP11-396B14 (centromeric/3′) were used. For the two plaque samples included in this study, we used a combination of break apart and fusion FISH. Break apart probe mixes A and B were used together with patient-specific fusion probes (MF3: RP11-107L10, RP11-1097K16 and RP11-421N18; MF4: RP11-625M5 and RP11-1083M15) (Supplementary Figure 1; Supplementary Table 3). Probes were purchased from BACPAC resources at Children's Hospital Oakland (CHORI) and their identity confirmed by FISH on metaphase controls and sanger sequencing. DNA was isolated from BAC clones by alkaline extraction, and labelled with haptens (digoxigenin (DIG)-, biotin (BIO)- or dinitrophenyl (DNP)-coupled dideoxynucleotides) by nick translation. FFPE tissue sections were subjected to deparaffinization in xylene, pre-treatment in 10mM citrate buffer, digestion in 0.4% pepsin, co-denaturation and hybridization with the probes. Hybridized sections were then incubated with fluorescein isothiocyanate (FITC)-conjugated mouse anti-DIG antibodies, Cy3-conjugated streptavidin and rabbit anti-DNP antibodies; and subsequently,

incubated with Alexa 488-conjugated goat anti-mouse antibodies and Cy5-conjugated goat anti-rabbit antibodies. Finally, sections were counterstained with 4,6 diamidino-2-phenylindole (DAPI). Analysis was performed by manually scoring 100 tumor cells per section.

2.11 Digital droplet PCR

Select CNAs were validated by using Bio-Rad QX200 ddPCR system (Supplementary Figure 2) following the manufacturer's guidelines. In short, 20-40 ng of genomic DNA was mixed with a frequent-cutting restriction enzyme, ddPCR supermix, FAM-labeled probes against the gene of interest and HEX-labeled probes against the reference gene in a 96-well PCR plate. Each 20 µL reaction was then transferred to a droplet generation cartridge, partitioned into nano droplets, and pipetted back to a fresh 96-well PCR plate by using Bio-Rad QX200 automated droplet generator. The plate containing the emulsified samples was sealed with foil and amplified on a Bio-Rad T100 thermocycler. PCR program was the following: 95 °C for 10 min, 39 cycles of 94 °C for 30 seconds and 60 °C for 1 min, and 98 °C for 10 min. The plate containing the post-PCR nano droplets was then placed into Bio-Rad QX200 droplet reader, which aspirates droplets and measures FAM/HEX fluorescence one droplet at a time. Copy number values were determined with Bio-Rad Quantasoft software v1.7.4. Reported copy numbers of *HNRNPK* and *SOCS1* in samples from the validation cohort are the average of 3 independent measurements using different reference genes.

3. RESULTS

3.1 Landscape of genomic rearrangements

The number of rearrangements ranged from 13 to 62 per patient (352 total events; mean/patient ± standard deviation, 39±18) (Figure 1; Figure 2A; Supplementary Table 4). Fifty-two percent of events were interchromosomal (range/patient, 35%-85%) (Figure 2B). Thirty-two percent of events fused 2 annotated genes, while the rest joined either a gene with a non-genic region or 2 non-genic regions, or reshuffled sequences within a single gene. Seven percent of rearrangements resulted in fusion transcripts (mean fusions/patient, 3; range, 1-5 fusions/patient) (Figure 2C; Supplementary Table 5). We also observed chromothripsis-like events in three patients (i.e. MF1, MF6, MF9) who carry numerous complex rearrangements in chromosomes 1 and X, 6 and 10, and 1 and 5, respectively (Figure 1; Figure 2D).

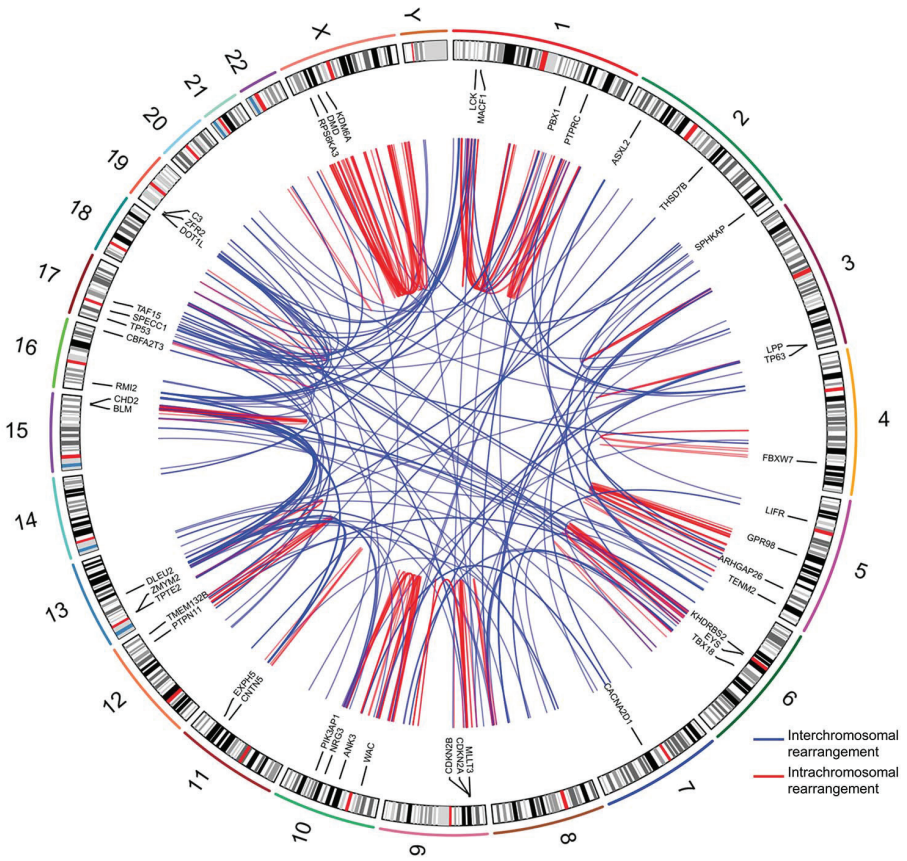


Figure 1. Landscape of genomic rearrangements in mycosis fungoides. Circos plot depicting 352 genomic rearrangements identified in nine MF genomes by WGS. The outer ring consists of chromosome ideograms arranged circularly end to end. The inner area in the plot shows arcs that represent interchromosomal (blue) and intrachromosomal (red) rearrangements. The ring between the chromosome ideograms and the arcs contains labels indicating rearranged genes implicated in cancer.

A total of 270 genes were found to be rearranged (Supplementary Table 6), 47 of which are currently implicated in cancer (Supplementary Table 7). This group includes genes previously associated with MF or Sézary Syndrome (SS) (i.e. *CDKN2A*, *CDKN2B*, *DLEU2*, *KDM6A*, *TP53*, *TP63* and *VAV1*)[5,19] and genes implicated in other hematological malignancies (e.g. *ARHGAP26*, *CBFA2T3*, *CHD2*, *DOT1L*, *LCK*, *LPP*, *PBX1*, *PTPN11*, *MLLT3*, *TAF15*, *SPECC1*, *ZMYM2*) (Figure 1).[27-38]

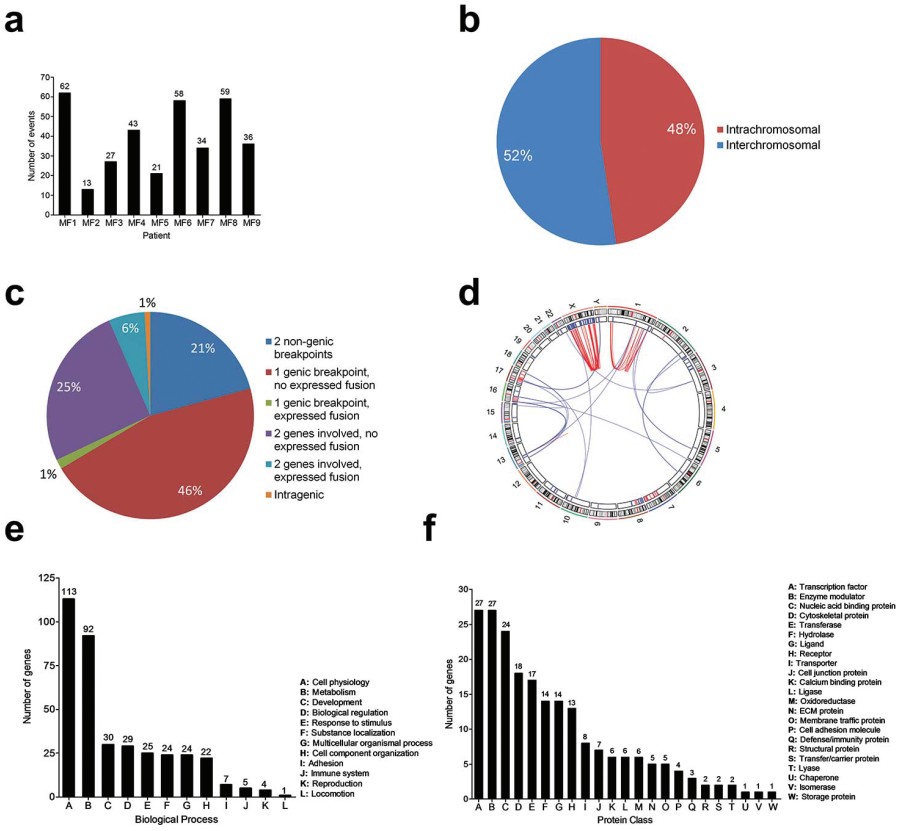


Figure 2. Distribution and functional annotation of genomic rearrangements in mycosis fungoides. A, Number of genomic rearrangements per sample. B, Distribution of inter- and intrachromosomal rearrangements. C, Distribution of genomic rearrangements based on the type of DNA sequences involved in the event (genic or nongenic) and the expression of fusion sequences determined through integration of WGS and RNA-seq data. D, Circos plot illustrating chromothripsis-like events in chromosomes 1 and X of sample MF1. The plot shows that complex rearrangements are associated with deletion of genomic regions. E, Distribution of rearranged genes according to the biological process their protein products take part in. F, Distribution of rearranged genes according to the protein class they encode (143 of 270 rearranged genes were assigned to a protein class by Panther).

Functional annotation of rearranged genes reveals a diverse set of biological processes (Figure 2E; Supplementary Table 8), being cell physiology (n_{genes}=113) and metabolism (n_{genes}=92) the highest ranking categories. Breakdown of these two categories shows that cell communication (n_{genes}= 34) and cell cycle (n_{genes}=16) are the most affected physiological processes while nucleic acid metabolism

(n_{genes} = 48, esp. transcription) leads the group of impacted biochemical processes (Supplementary Table 8).

In support of the postulated deleterious effects of genomic rearrangements on normal cellular functions, at least 100 rearranged genes (127/270 genes could not be assigned to a protein class by Panther) encode proteins with relevant roles in signal transduction (e.g. ligands, receptors, enzyme modulators) and transcriptional regulation (e.g. transcription factors, chromatin regulators) (Figure 2F; Supplementary Table 9). Nonetheless, rearranged genes do not group into a single or a few signaling pathways, but take part in numerous different pathways/processes.

We found nine recurrently rearranged genes, *ARHGAP26* (two of nine patients), *ATXN1* (two of nine), *CLEC16A* (four of nine), *ELF1* (two of nine), *EYS* (two of nine), *RBPJ* (two of nine), *RPS6KA3* (two of nine), *SLC24A2* (two of nine) and *SSH2* (two of nine) in our sequenced cohort. However, in all cases fusion partners differ between patients and the resulting chimeric sequences are expressed only in single patients (five of nine rearranged genes) or not expressed (four of nine). Interestingly, 50% of recurrent CNAs containing cancer genes are associated with inter- or intrachromosomal rearrangements. For instance, *ARID1A*, *CDKN2A/B*, *PTPRC*, *SOCS1* and *STK11* are deleted as a consequence of chromosomal rearrangements in two or more patients. This fact, together with the small percentage of expressed gene fusions, suggests that in MF, rearrangements more often mediate inactivation of tumor suppressors, rather than generate oncogenic fusions.

3.2 Aneuploidies and CNAs

Large-scale numerical abnormalities (> 3 Mb) included trisomy 4 (1/9 patients), 7 (two of nine) and 18 (two of nine), as well as deletions within 9q (four of nine patients), 10q (two of nine) and 16q (three of nine), and gains within 3q (three of nine), 5p (three of nine), 7q (two of nine), 8q (two of nine) and 17q (four of nine) (Figure 3; Figure 4).

We identified 18 focal (≤ 3 Mb) minimal common regions (MCRs) shared by CNAs, with 15/18 MCRs containing cancer genes (Figure 3; Figure 4; Supplementary Table 10). These focal MCRs affect genes primarily involved in cell cycle control, chromatin regulation, the JAK-STAT pathway and the PI-3-K pathway.

The most frequent focal MCR was 9p21.3 deletion, found in 7/9 patients. This region encloses exon 2 and 3 of *CDKN2A*. We found deletions at 16p13.13 and

9q21.32 in five of nine patients, which include JAK-STAT pathway regulator *SOCS1* and *TP53*-dependent p21 co-activator *HNRNPK*, respectively.

Deletions at 1q31.3 and 13q14.3 were observed in four of nine patients. The former involves JAK-STAT inhibitor *PTPRC* while the latter contains the *DLEU2/Mir-15a/16-1* locus, which is frequently deleted in chronic lymphocytic leukemia (CLL).[39]

Additionally, three of nine patients had deletions at 1p36.11, 9q21.31, 10q23.31, 13q14.11, 19p13.3 and 20q13.13, which include tumor suppressors *ARID1A*, *TLE4*, *PTEN*, *FOXO1*, *STK11* (alongside *TCF3*) and *PTPN1*, respectively. Lastly, two of nine patients presented deletions at 5q15-21.1, 6p22.3 and 16q24.3 which involve (putative) tumor suppressors *CHD1*, *JARID2* and *CBFA2T3*, respectively.

In contrast, focal MCRs within gain areas were rare (n=3), with gain at 8q24.21 (involving *MYC*) found in two of nine patients, being the only event containing a cancer gene.

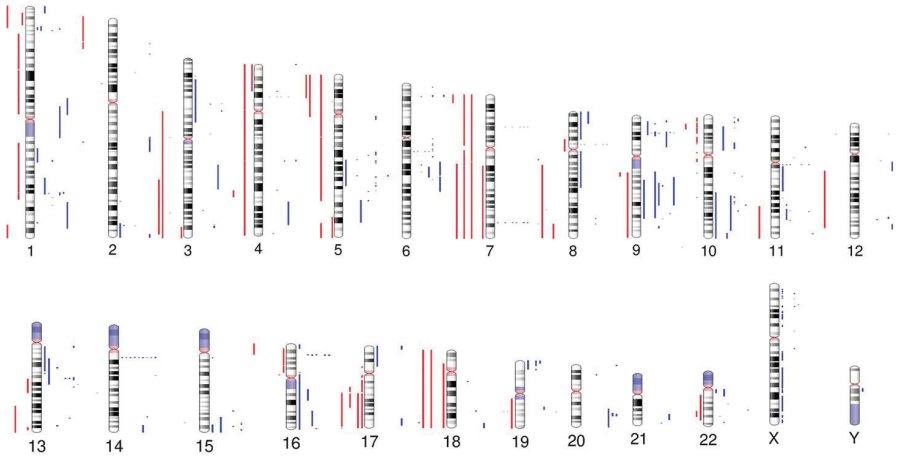


Figure 3. Overview of CNAs in mycosis fungoides. Human chromosome ideograms showing regions of gain and loss identified by WGS in nine MF genomes. Red bars to the left of each chromosome represent regions of gain while blue bars to the right of each chromosome represent regions of loss

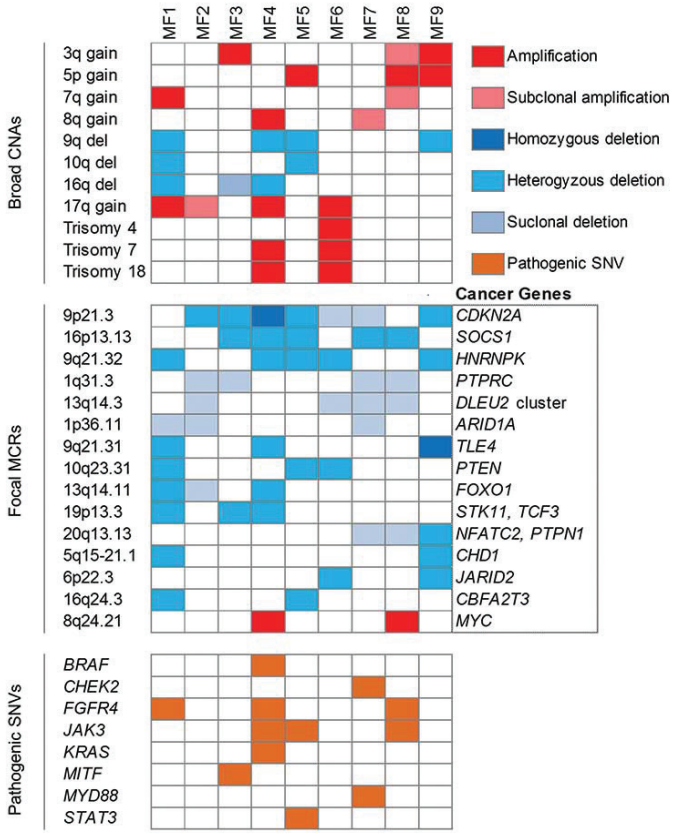


Figure 4. Distribution of recurrent CNAs and pathogenic SNVs in mycosis fungoides. Upper panel, broad CNAs (>3 Mb); middle panel, focal MCRs (≤3 Mb) shared by CNAs. Bona fide cancer genes contained within each focal MCR are indicated; bottom panel, pathogenic SNVs. Only SNVs for which functional validation has been reported in literature are shown.

3.3 Pathogenic SNVs

Prior studies showed that recurrent pathogenic SNVs in MF are rare.[7-9,19] Nevertheless, we looked for pathogenic SNVs in exonic sequences of genes involved in the JAK-STAT pathway, the MAPK pathway, the NF-κB pathway, the PI-3-K pathway, the T-cell receptor (TCR) pathway, cell cycle control, chromatin organization and genes that are presumed drivers[19] in CTCL (Supplementary Table 11).

Recurrent SNVs were found in two genes, *FGFR4* (p.G388R[40], three of nine patients; expressed: 1/3) and *JAK3* (p.A573V[41], two of nine patients), both of which are known gain-of-function mutations. The remaining SNVs predicted

as pathogenic occur in single patients only. These include gain-of-function SNVs in *BRAF* (p.G466E[42]), *JAK3* (p.V722I[43]), *KRAS* (p.G13C[44]), *MYD88* (p.L273P[45]), and *STAT3* (p.Y640F[46]), which have been reported either in CTCL or other malignancies and functionally validated elsewhere (Figure 4; Supplementary Figure 3); also, SNVs in *CHEK2* (p.I200T[47]) and *MITF* (p.E419K[48]), which are germline risk factors for breast cancer and melanoma, respectively (Figure 4); and 47 other patient-specific SNVs (predicted as highly deleterious by PolyPhen-2 and SIFT) (Supplementary Table 12) located in genes with relevant roles in the aforesaid pathways. Importantly, pathogenic SNVs in genes from the JAK-STAT and MAPK pathway are not mutually exclusive.

3.4 Differentially expressed genes and fusion transcripts

We identified differentially expressed (DE) genes by comparing expression in our MF cohort with expression in normal $CD4^+$ T cells. Since the cell of origin of MF remains unidentified, transcriptome analysis was performed using a control group formed by several CD4+ T-cell subsets (see Materials and Methods) with the aim of detecting aberrant expression patterns that are absent in a range of normal CD4+ phenotypes. A total of 733 genes (450 up-regulated, 283 down-regulated, FDR <0.01) were found to be differentially expressed (Figure 5A, Supplementary Table 13). We next used NCG 5.0 to pinpoint DE genes implicated in cancer. Eighty-one cancer genes (51 up-regulated, 30 down-regulated) were identified (Supplementary Table 14). Up-regulated genes include oncogenes *MALAT1*, *MECOM*, *PBX1*, *TTK* and *WWTR1*, whereas down-regulated genes include tumor suppressors *BRD7*, *CDKN1B*, *CYLD*, *HNRNPK*, *TSC1* and *XPA*. The expression profile also comprises up-regulation of developmental genes *GLI3*, *JAG1* and *NOTCH3*, and down-regulation of transcriptional repressor *FOXP1* and cell proliferation inhibitors *GPS2* and *RHOH* (Figure 5B, 5C and 5D).

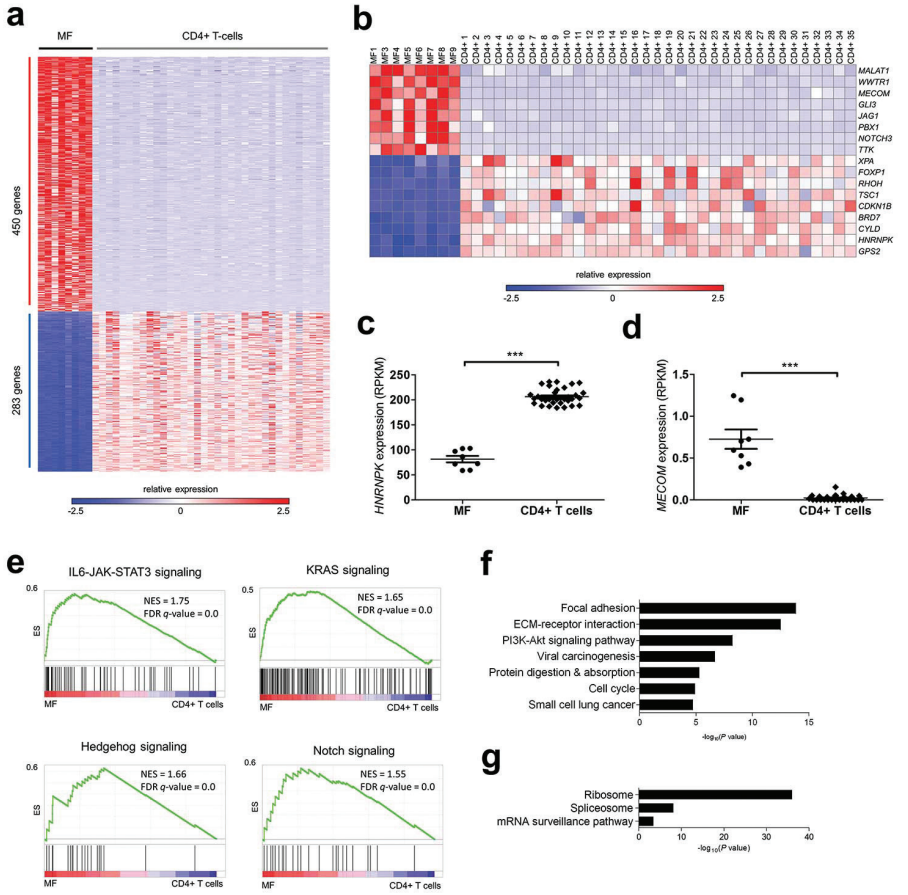


Figure 5. RNA-seq identifies differentially expressed cancer genes and deregulated signaling pathways in mycosis fungoides. A, Heat map of differentially expressed (DE) genes (FDR < 0.01) in MF when compared to CD4+ T cells. Of 733 DE genes, 450 were commonly up-regulated and 283 were commonly down-regulated. B, Oncogenes and tumor suppressors with roles in cell cycle control and development are among the group of DE genes. C and D, *HNRNPK* and *MECOM*, whose deregulation (down- and up-, respectively) are reported to enhance the JAK-STAT pathway, are differentially expressed in MF (*HNRNPK* : −2.5-fold average, ***P < 1 × 10−4; *MECOM* : 31-fold average, ***P < 1 × 10−4, Mann-Whitney test). E, Gene set enrichment analysis. Select GSEA plots showing up-regulation of STAT3 signaling (upper left), KRAS signaling (upper right), Hedgehog signaling (lower left) and Notch signaling (lower right) in MF compared to CD4+ T cells (see Supporting Information Table 15 for a complete list of GSEA signatures). NES, normalized enrichment score; FDR q -value, false discovery rate q -value. F and G, Pathway analysis by DAVID reveals up-regulation of the PI-3-K/Akt pathway, the cell cycle and cancer signatures, and down-regulation of ribosome, spliceosome and mRNA surveillance (see Supporting Information Table 16 for a complete list of enriched terms/processes).

We detected 24 patient-specific fusion transcripts (Table 1), including 6 (i.e. *ATXN1-TP63*, *CCR7-DOT1L*, *KDM6A-IL1RAPL1*, *LMF1-TAF15*, *TP53-GPR3* and *YTHDF3-LIFR*) that contain genes implicated in cancer. To our knowledge, with the exception of *ATXN1-TP63*,[49] all these chimeric transcripts are reported for the first time.

Table 1. Fusion transcripts detected by RNA-seq in T-MF. CTX, interchromosomal translocation. ITX, intrachromosomal translocation. iDel, interstitial deletion.

Sample	Fusion transcript	Breakpoints (DNA)	Breakpoint Type	Event Class	WGS con-firmed
MF1	*KDM6A–IL1RAPL1*	chrX:44746566 - chrX:29451290	Genic - Genic	ITX	Yes
MF1	*CHIC1–RP2*	chrX:72844450 - chrX:46680435	Genic - Nongenic	ITX	Yes
MF3	*ANKRD13A–CUL9*	chr12:110448655 - chr6:43160142	Genic - Genic	CTX	Yes
MF3	*CLEC16A–SCARB1*	chr16:11067010 - chr12:125350896	Genic - Nongenic	CTX	Yes
MF3	*SSH2–GRAP2*	chr17:28059210 - chr22:40314573	Genic - Genic	CTX	Yes
MF3	*LMF1–TAF15*	chr16:986148 - chr17:34145925	Genic - Genic	CTX	Yes
MF3	*ATXN1–TP63*	chr6:16307814 - chr3:189470345	Genic - Genic	CTX	Yes
MF4	*CCR7–DOT1L*	chr17:38718403 - chr19:2181252	Genic - Genic	CTX	Yes
MF5	*PHACTR4–EPB41*	chr1:28755797 - chr1:29246304	Genic - Genic	iDel	Yes
MF5	*ADAM12–MMRN2*	chr10:127935628 - chr10:88698606	Genic - Genic	iDel	Yes
MF5	*TRAPPC10–TRPM2*	chr21:45487448 - chr21:45795335	Genic - Genic	iDel	Yes
MF5	*ARHGAP26–TENM2*	chr5:142272242 - chr5:167448836	Genic - Genic	ITX	Yes
MF6	*ANK3–RNLS*	chr10:62168031 - chr10:90101461	Genic - Genic	ITX	Yes
MF6	*ELF1–SATB2*	chr13:41540637 - chr2:200369768	Genic - Nongenic	CTX	Yes
MF7	*TP53–GPR3*	chr17:7579754 - chr1:27718138	Genic - Nongenic	CTX	Yes
MF7	*CLPP–NR3C1*	chr19:6365544 - chr5:142800539	Genic - Genic	CTX	Yes
MF7	*SARNP–WRAP53*	chr12:56161974 - chr17:7593927	Genic - Genic	CTX	Yes
MF8	*SETD5–RNF19A*	chr3:9497374 - chr8:101391443	Genic - Nongenic	CTX	Yes
MF8	*SUDS3–TMEM132B*	chr12:118847216 - chr12:125987360	Genic - Genic	ITX	Yes
MF8	*AACS–STAB2*	chr12:125625503 - chr12:104094623	Genic - Genic	ITX	Yes
MF8	*RPUSD3–RNF19A*	chr3:9882804 - chr8:101303862	Genic - Genic	CTX	Yes
MF8	*YTHDF3–LIFR*	chr8:64081882 - chr5:38586949	Genic - Genic	CTX	Yes
MF9	*DPM1–UBE2V1*	chr20:49574368 - chr20:48703893	Genic - Genic	iDel	Yes
MF9	*KCNAB2–ESPN*	chr1:6071941 - chr1:6493075	Genic - Genic	iDel	Yes

3.5 Deregulated signaling pathways

To look for evidence of deregulated pathways in MF, we performed GSEA using annotated gene sets from MSigDB to look for expression signatures. The analysis revealed up-regulation of IL6-JAK-STAT3 signaling (NES = 1.75, FDR *q*-value = 2.62×10^{-4}), KRAS signaling (NES = 1.65, FDR *q*-value = 1.8×10^{-3}), Hedgehog signaling (NES = 1.66, FDR *q*-value = 1.8×10^{-3}) and Notch signaling (NES = 1.55, FDR *q*-value = 0.01) (Figure 4E; Supplementary Table 15)

In addition, we performed DAVID pathway analysis with up- and down-regulated genes separately. Cellular processes associated with integrin-mediated signaling (i.e. ECM-receptor interaction, focal adhesion), PI-3-K/Akt signaling ($P = 4.6x10^{-10}$), cancer signatures (i.e. viral carcinogenesis, small cell lung cancer) and cell cycle ($P = 1.1x10^{-5}$) are prominent up-regulated profiles in MF (Figure 5F; Supplementary Table 16). Down-regulated profiles include ribosome ($P = 1.0x10^{-36}$) and spliceosome ($P = 8.7x10^{-9}$) activity, and mRNA surveillance ($P = 4.1x10^{-4}$) (Figure 5G; Supplementary Table 16).

3.6 *SOCS1* and *HNRNPK* are recurrently deleted

From all structural alterations revealed by our analysis, deletion of *HNRNPK* and *SOCS1* stand out because of their novelty and recurrence. Notably, apart from being deleted in five of nine sequenced patients, *HNRNPK* is down-regulated in eight of eight transcriptomes (2.5-fold average, $P < 1x10^{-4}$) (Figure 5C; Figure 7A) whereas *SOCS1* deletions are invariably focal (≤ 3 Mb) in 5/5 affected patients (MCR: 305kb, 3 genes) (Figure 6; Figure 7B; Supplementary Figure 4). Consequently, we evaluated copy number of *HNRNPK* and *SOCS1* in 18 additional tumor biopsies by ddPCR. In this validation cohort we found *HNRNPK* deletion in 5 patients and *SOCS1* deletion in 4 patients (Figure 7C and 7D). Taking together the sequenced and validation cohorts, *HNRNPK* was deleted in 10 of 27 (37%) patients and *SOCS1* in 9 of 27 (33%) patients.

3.7 Deletion of *SOCS1* can be found at early stage

We seized upon the fact that *SOCS1* deletions mostly result from translocations in our sequenced cohort to investigate their occurrence at early stage. We used a combination of break-apart and fusion FISH to search for *SOCS1*-deleting translocations in available plaque-stage tissue from two sequenced patients (MF3 and MF4) with confirmed *SOCS1*-deleting translocations in tumor-stage tissue (Figure 6). Plaque biopsies from patients MF3 and MF4 were procured, respectively, 3 years and 8 months prior to tumor development. We found that patient MF4 bears the translocation at plaque-stage too (Figure 7E), suggesting that *SOCS1* deletion is an early event in this individual.

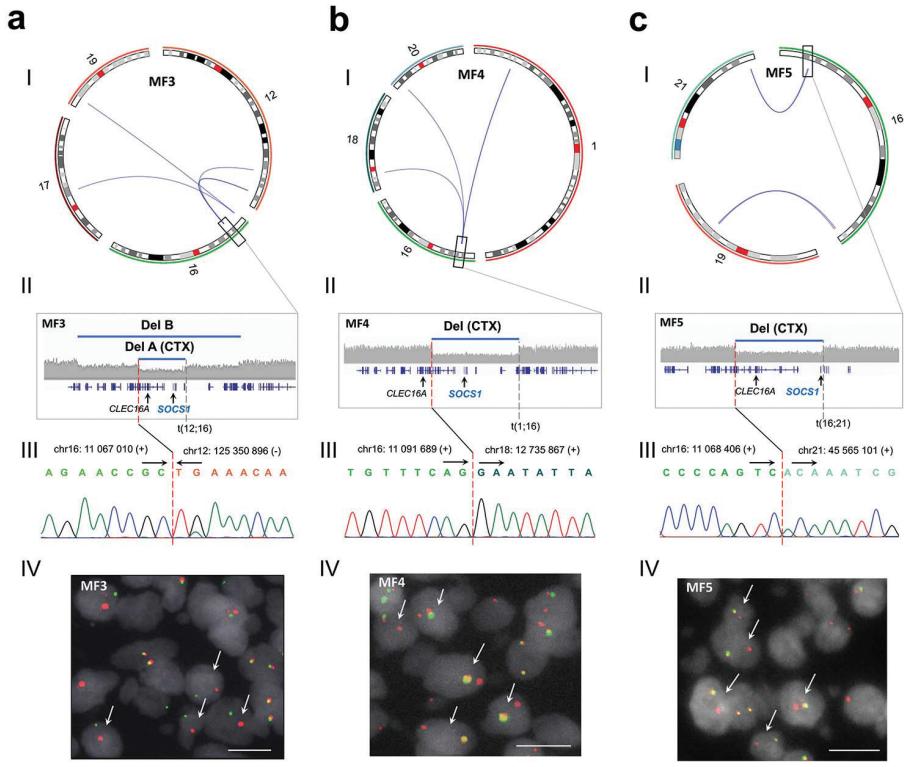


Figure 6. Genomic rearrangements at 16q13.13 are associated with focal *SOCS1* deletions in mycosis fungoides. (I) Circos plots displaying genomic rearrangements at 16q13.13. (II) Magnified views of deletions at 16q13.13 resulting from structural alterations. Genomic rearrangements at 16q13.13 validated by (III) Sanger sequencing and (IV) break apart FISH in (A) MF3, (B) MF4, and (C) MF5. Del, deletion. CTX, interchromosomal translocation. Scale bar, 10 µm.

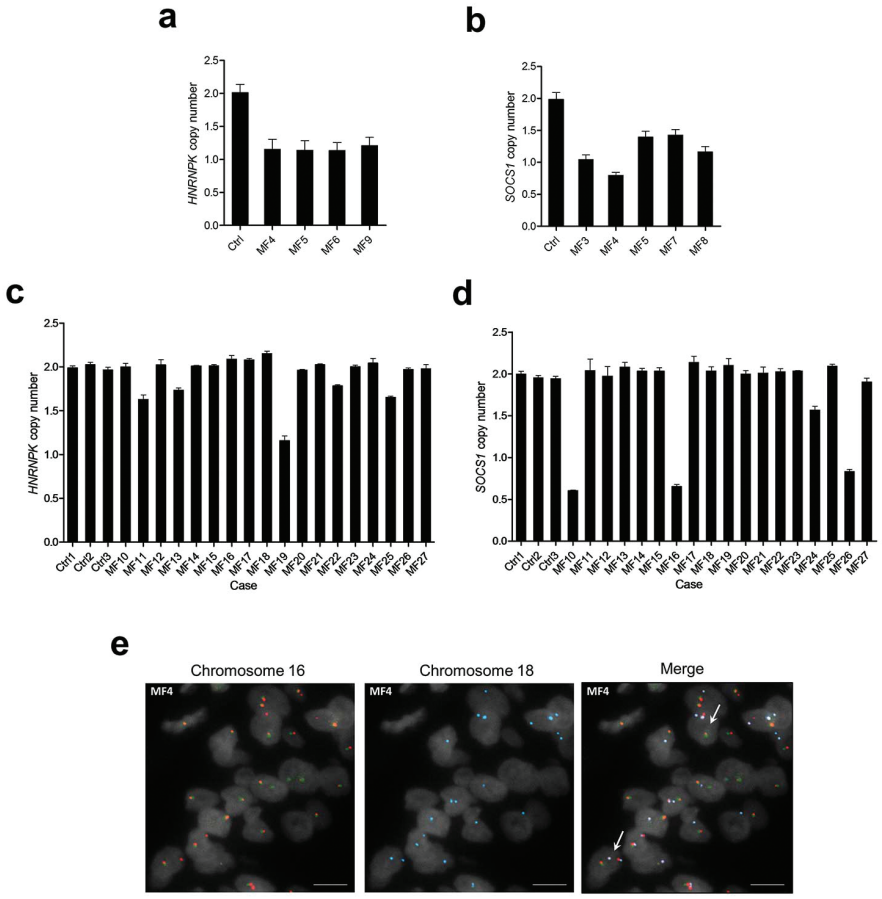


Figure 7. *HNRNPK* and *SOCS1* are recurrently deleted in mycosis fungoides. Deletion of (A) *HNRNPK* and (B) *SOCS1* in sequenced tumor samples was confirmed by ddPCR. Deletion of (C) *HNRNPK* and (D) *SOCS1* was also identified in samples from the extension cohort by ddPCR. E, The translocation responsible for *SOCS1* deletion in sample MF4 was found in (early-stage) plaque tissue by FISH. Ctrl, CD4+ T cells. Scale bar, 10 µm.

4. DISCUSSION

This study represents the first integrated analysis (DNA/RNA) of genomic rearrangements in MF. The analysis reveals that MF displays a complex and heterogeneous landscape of inter- and intrachromosomal rearrangements. We observed, among others, translocations leading to deletion of *ARID1A*, *CDKN2A/B*, *PTPRC*, *SOCS1* and *STK11*. This suggests that rearrangements mediate the deletion of tumor suppressors involved in pathways that are commonly deregulated in MF patients. We detected 270 rearranged genes, of which at least 100 play diverse roles in signal transduction and transcriptional regulation, and 47 are currently implicated in neoplasms. Our analysis identified 24 fusion transcripts, including 6 containing *bona fide* cancer genes, which though not recurrent, may contribute to MF development in individual patients.

All potentially deleterious SNVs we observed are patient-specific, with the exception of *FGFR4* (p.G388R) and *JAK3* (p.A573V). Still, pathogenic SNVs may play relevant roles in signaling deregulation in some individuals. Importantly, we observed numerous SNVs in two genes reported as putative drivers in CTCL (i.e. *KMT2C, NCOR1*), which albeit predicted as highly deleterious, were not found to be expressed in the RNA-seq data. This highlights the importance of integrating DNA and RNA analyses to evaluate mutational data.

Although phenotypic resemblance between MF cells and several CD4+ T-cell subsets (i.e. T_H2[50], T_H17[51], T_{RM}[52]) has been documented in previous years, there is no definitive proof of any of these potential origins. Despite the cell of origin of MF remains undetermined, aberrant expression detected in our MF cohort using a 'pan'- CD4+ control group matches earlier observations made by other groups. For instance, overactivation of JAK-STAT[53,54] and NOTCH[55] signaling, and mutations that enhance RAS-mediated signaling[56] have been previously described in MF. Yet, our transcriptome data should be interpreted with caution as further confirmation is required once the exact CD4+ T-cell subset giving rise to MF is identified.

Interestingly, transcriptome analysis reveals a subset of DE cancer genes that play roles in cell cycle regulation and development. Tumor suppressors *BRD7*, *CDKN1B*, *GPS2* and *HNRNPK*, which are down-regulated in MF, are known to prompt cell cycle arrest at G1/S.[57-59] *TSC1*, down-regulated too, sustains quiescence in naïve T cells and its abrogation results in rapid cycling behavior.[60] *PBX1*, a direct transcriptional repressor of *CDKN2B*, is consistently up-regulated in MF.[61] Additionally, up-regulation of mitotic checkpoint kinase *TTK* might contribute to genomic instability in MF, since its expression has been shown to prevent aneuploidy-induced cell death.[62]

On the other hand, up-regulation of developmental genes *GLI3*, *JAG1* and *NOTCH3* might contribute to boost cell proliferation. Notably, NOTCH3 over-activation has been shown to induce an auto-sustaining JAG1 expression loop in T-cell acute lymphoblastic leukemia (T-ALL), which in turn, enhances expression of Notch target genes responsible for the progression of the disease.[63]

Moreover, transcriptome analysis also shows that processes related to transcription (i.e. spliceosome activity, mRNA surveillance) are flawed in MF, which might be linked to the considerable number of transcription-related genes affected by genomic rearrangements. Taken together, the structural and expression analyses show that the cell cycle, the JAK-STAT pathway, the PI-3-K pathway and developmental pathways are deregulated in MF.

We report for the first time recurrent deletion of *HNRNPK* and *SOCS1* not only in MF, but any CTCL. Furthermore, we found evidence that *SOCS1* deletion is an early event in 1 of 2 patients with available plaque-stage material by FISH. Importantly, while the incidence of deletion of both genes in the extension cohort was lower compared to the sequenced cohort, this difference was not statistically significant ($P > 0.05$, Fisher exact test). Moreover, we can rule out the existence of clinical differences between the tumors from the two cohorts.

hnRNP-K is a nuclear ribonucleoprotein implicated in leukemogenesis of acute myeloid leukemia (AML).[59] Interestingly, studies have shown that haploinsufficiency of *HNRNPK* not only downregulates p21, but also up-regulates STAT3 signaling and give rise to B- and T-cell lymphomas in a mouse model.[59] On the other hand, *SOCS1*, which is silenced in several cancers including multiple myeloma (MM)[64], inhibits JAK-STAT signaling by suppressing the tyrosine kinase activity of JAK proteins.[65]

A noteworthy fact is that miR-155, which is often up-regulated in MF[6], has been found to target *SOCS1* in breast cancer and laryngeal carcinoma, leading to constitutive STAT3 activation in both cancers.[66,67] We observed that 2 of 3 patients without *SOCS1* deletions express much higher levels (7-fold average) of miR-155 precursor *MIR155HG* than patients with *SOCS1* deletions (Supplementary Figure 5), which suggests that miR-155 levels rise to inhibit *SOCS1* in patients with functional copies of the gene. Moreover, *SOCS1* might be suppressed in MF in one additional way. *MECOM*, which is consistently up-regulated in our sequenced cohort, has been found to inhibit the expression of several regulators of the JAK-STAT pathway in AML, particularly *SOCS1*.[68] This evidence suggests that deregulation of STAT3 signaling via inactivation of *HNRNPK* and *SOCS1* might be important events in the pathogenesis of MF. Future studies with cells and animal models will be essential to functionally confirm the association

between these tumor suppressors, their regulators, and STAT3 signaling in MF. In this scenario, targeting miR-155 and/or *MECOM* to treat patients with functional *SOCS1* alleles constitute potential novel therapeutic strategies.

Overall, the findings in this study reveal that genomic rearrangements and CNAs play relevant roles in the pathogenetics of MF and position *HRNRPK* and *SOCS1* as putative drivers of MF development.

Acknowledgements
The authors thank D. de Jong, from the Department of Cell and Chemical Biology, Leiden University Medical Center, Leiden, The Netherlands, for providing assistance in performing FISH. This study was funded by the Dutch Cancer Society (grant UL2013-6104).

Authorship
A.N.B.T. and C.P.T. designed the project and wrote the manuscript. A.N.B.T. and D.C. performed the bioinformatic analyses and analyzed the data. A.N.B.T. performed all the experiments. A.B, D.C., H.M., K.S., M.V. and C.T reviewed the data and the manuscript.

Conflict of interest disclosure
The authors declare no conflicts of interest.

REFERENCES

1. Willemze R, Jaffe ES, Burg G, et al. WHO-EORTC classification for cutaneous lymphomas. *Blood*. 2005;105:3768-3785.
2. Nicolae-Cristea AR, Benner MF, Zoutman WH, et al. Diagnostic and prognostic significance of CDKN2A/CDKN2B deletions in patients with transformed mycosis fungoides and primary cutaneous CD30-positive lymphoproliferative disease. *Br J Dermatol*. 2015;172:784-788.
3. Laharanne E, Chevret E, Idrissi Y, et al. CDKN2A-CDKN2B deletion defines an aggressive subset of cutaneous T-cell lymphoma. *Mod Pathol*. 2010;23:547-558.
4. van Doorn R, Dijkman R, Vermeer MH, Starink TM, Willemze R, Tensen CP. A novel splice variant of the Fas gene in patients with cutaneous T-cell lymphoma. *Cancer Res*. 2002;62:5389-5392.
5. van Doorn R, van Kester MS, Dijkman R, et al. Oncogenomic analysis of mycosis fungoides reveals major differences with Sezary syndrome. *Blood*. 2009;113:127-136.
6. van Kester MS, Ballabio E, Benner MF, et al. miRNA expression profiling of mycosis fungoides. *Mol Oncol*. 2011;5:273-280.
7. Ungewickell A, Bhaduri A, Rios E, et al. Genomic analysis of mycosis fungoides and Sezary syndrome identifies recurrent alterations in TNFR2. *Nat Genet*. 2015;47:1056-1060.
8. McGirt LY, Jia P, Baerenwald DA, et al. Whole-genome sequencing reveals oncogenic mutations in mycosis fungoides. *Blood*. 2015;126:508-519.
9. da Silva Almeida AC, Abate F, Khiabanian H, et al. The mutational landscape of cutaneous T cell lymphoma and Sezary syndrome. *Nat Genet*. 2015;47:1465-1470.
10. Vaque JP, Gomez-Lopez G, Monsalvez V, et al. PLCG1 mutations in cutaneous T-cell lymphomas. *Blood*. 2014;123:2034-2043.
11. Caumont C, Gros A, Boucher C, et al. PLCG1 Gene Mutations Are Uncommon in Cutaneous T-Cell Lymphomas. *J Invest Dermatol*. 2015;135:2334-2337.
12. Robinson JT, Thorvaldsdottir H, Winckler W, et al. Integrative genomics viewer. *Nat Biotechnol*. 2011;29:24-26.
13. Haas B, Dobin A, Stransky N, et al. STAR-Fusion: Fast and Accurate Fusion Transcript Detection from RNA-Seq. *bioRxiv*. 2017.
14. Nicorici D, Satalan M, Edgren H, et al. FusionCatcher - a tool for finding somatic fusion genes in paired-end RNA-sequencing data. *bioRxiv*. 2014.
15. Boeva V, Popova T, Bleakley K, et al. Control-FREEC: a tool for assessing copy number and allelic content using next-generation sequencing data. *Bioinformatics*. 2012;28:423-425.
16. Choi J, Goh G, Walradt T, et al. Genomic landscape of cutaneous T cell lymphoma. *Nat Genet*. 2015;47:1011-1019.
17. Kiel MJ, Sahasrabuddhe AA, Rolland DC, et al. Genomic analyses reveal recurrent mutations in epigenetic modifiers and the JAK-STAT pathway in Sezary syndrome. *Nat Commun*. 2015;6:8470.
18. Lee CS, Ungewickell A, Bhaduri A, et al. Transcriptome sequencing in Sezary syndrome identifies Sezary cell and mycosis fungoides-associated lncRNAs and novel transcripts. *Blood*. 2012;120:3288-3297.
19. Park J, Yang J, Wenzel AT, et al. Genomic analysis of 220 CTCLs identifies a novel recurrent gain-of-function alteration in RLTPR (p.Q575E). *Blood*. 2017.
20. Prasad A, Rabionet R, Espinet B, et al. Identification of Gene Mutations and Fusion Genes in Patients with Sezary Syndrome. *J Invest Dermatol*. 2016;136:1490-1499.
21. Wang L, Ni X, Covington KR, et al. Genomic profiling of Sezary syndrome identifies alterations of key T cell signaling and differentiation genes. *Nat Genet*. 2015;47:1426-1434.
22. Woollard WJ, Pullabhatla V, Lorenc A, et al. Candidate driver genes involved in genome maintenance and DNA repair in Sezary syndrome. *Blood*. 2016;127:3387-3397.

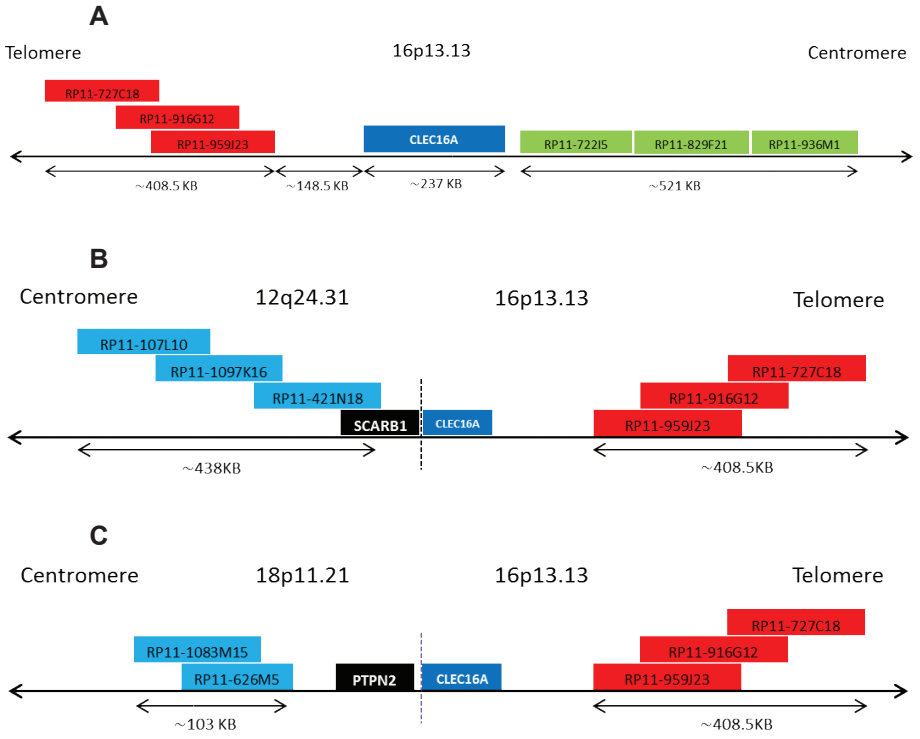
23. Ranzani V, Rossetti G, Panzeri I, et al. The long intergenic noncoding RNA landscape of human lymphocytes highlights the regulation of T cell differentiation by linc-MAF-4. *Nat Immunol*. 2015;16:318-325.
24. Mi H, Muruganujan A, Thomas PD. PANTHER in 2013: modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees. *Nucleic Acids Res*. 2013;41:D377-386.
25. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A*. 2005;102:15545-15550.
26. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc*. 2009;4:44-57.
27. Gruber TA, Larson Gedman A, Zhang J, et al. An Inv(16)(p13.3q24.3)-encoded CBFA2T3-GLIS2 fusion protein defines an aggressive subtype of pediatric acute megakaryoblastic leukemia. *Cancer Cell*. 2012;22:683-697.
28. Borkhardt A, Bojesen S, Haas OA, et al. The human GRAF gene is fused to MLL in a unique t(5;11)(q31;q23) and both alleles are disrupted in three cases of myelodysplastic syndrome/acute myeloid leukemia with a deletion 5q. *Proc Natl Acad Sci U S A*. 2000;97:9168-9173.
29. Vieira L, Sousa AC, Matos P, et al. Three-way translocation involves MLL, MLLT3, and a novel cell cycle control gene, FLJ10374, in the pathogenesis of acute myeloid leukemia with t(9;11;19)(p22;q23;p13.3). *Genes Chromosomes Cancer*. 2006;45:455-469.
30. Tartaglia M, Martinelli S, Cazzaniga G, et al. Genetic evidence for lineage-related and differentiation stage-related contribution of somatic PTPN11 mutations to leukemogenesis in childhood acute leukemia. *Blood*. 2004;104:307-313.
31. Ren M, Cowell JK. Constitutive Notch pathway activation in murine ZMYM2-FGFR1-induced T-cell lymphomas associated with atypical myeloproliferative disease. *Blood*. 2011;117:6837-6847.
32. Morerio C, Acquila M, Rosanda C, et al. HCMOGT-1 is a novel fusion partner to PDGFRB in juvenile myelomonocytic leukemia with t(5;17)(q33;p11.2). *Cancer Res*. 2004;64:2649-2651.
33. Martini A, La Starza R, Janssen H, et al. Recurrent rearrangement of the Ewing's sarcoma gene, EWSR1, or its homologue, TAF15, with the transcription factor CIZ/NMP4 in acute leukemia. *Cancer Res*. 2002;62:5408-5412.
34. Diakos C, Xiao Y, Zheng S, Kager L, Dworzak M, Wiemels JL. Direct and indirect targets of the E2A-PBX1 leukemia-specific fusion protein. *PLoS One*. 2014;9:e87602.
35. Daheron L, Veinstein A, Brizard F, et al. Human LPP gene is fused to MLL in a secondary acute leukemia with a t(3;11) (q28;q23). *Genes Chromosomes Cancer*. 2001;31:382-389.
36. Bernt KM, Zhu N, Sinha AU, et al. MLL-rearranged leukemia is dependent on aberrant H3K79 methylation by DOT1L. *Cancer Cell*. 2011;20:66-78.
37. Majolini MB, D'Elios MM, Galieni P, et al. Expression of the T-cell-specific tyrosine kinase Lck in normal B-1 cells and in chronic lymphocytic leukemia B cells. *Blood*. 1998;91:3390-3396.
38. Rodriguez D, Bretones G, Quesada V, et al. Mutations in CHD2 cause defective association with active chromatin in chronic lymphocytic leukemia. *Blood*. 2015;126:195-202.
39. Klein U, Lia M, Crespo M, et al. The DLEU2/miR-15a/16-1 cluster controls B cell proliferation and its deletion leads to chronic lymphocytic leukemia. *Cancer Cell*. 2010;17:28-40.
40. Ulaganathan VK, Sperl B, Rapp UR, Ullrich A. Germline variant FGFR4 p.G388R exposes a membrane-proximal STAT3 binding site. *Nature*. 2015;528:570-574.
41. Koo GC, Tan SY, Tang T, et al. Janus kinase 3-activating mutations identified in natural killer/T-cell lymphoma. *Cancer Discov*. 2012;2:591-597.
42. Kamata T, Hussain J, Giblett S, Hayward R, Marais R, Pritchard C. BRAF inactivation drives aneuploidy by deregulating CRAF. *Cancer Res*. 2010;70:8475-8486.
43. Walters DK, Mercher T, Gu TL, et al. Activating alleles of JAK3 in acute megakaryoblastic leukemia. *Cancer Cell*. 2006;10:65-75.

44. Niemela JE, Lu L, Fleisher TA, et al. Somatic KRAS mutations associated with a human nonmalignant syndrome of autoimmunity and abnormal leukocyte homeostasis. *Blood*. 2011;117:2883-2886.
45. Ngo VN, Young RM, Schmitz R, et al. Oncogenically active MYD88 mutations in human lymphoma. *Nature*. 2011;470:115-119.
46. Koskela HL, Eldfors S, Ellonen P, et al. Somatic STAT3 mutations in large granular lymphocytic leukemia. *N Engl J Med*. 2012;366:1905-1913.
47. Roeb W, Higgins J, King MC. Response to DNA damage of CHEK2 missense mutations in familial breast cancer. *Hum Mol Genet*. 2012;21:2738-2744.
48. Yokoyama S, Woods SL, Boyle GM, et al. A novel recurrent mutation in MITF predisposes to familial and sporadic melanoma. *Nature*. 2011;480:99-103.
49. Boddicker RL, Razidlo GL, Dasari S, et al. Integrated mate-pair and RNA sequencing identifies novel, targetable gene fusions in peripheral T-cell lymphoma. *Blood*. 2016;128:1234-1245.
50. Vowels BR, Cassin M, Vonderheid EC, Rook AH. Aberrant cytokine production by Sezary syndrome patients: cytokine secretion pattern resembles murine Th2 cells. *J Invest Dermatol*. 1992;99:90-94.
51. Ciree A, Michel L, Camilleri-Broet S, et al. Expression and activity of IL-17 in cutaneous T-cell lymphomas (mycosis fungoides and Sezary syndrome). *Int J Cancer*. 2004;112:113-120.
52. Campbell JJ, Clark RA, Watanabe R, Kupper TS. Sezary syndrome and mycosis fungoides arise from distinct T-cell subsets: a biologic rationale for their distinct clinical behaviors. *Blood*. 2010;116:767-771.
53. Sommer VH, Clemmensen OJ, Nielsen O, et al. In vivo activation of STAT3 in cutaneous T-cell lymphoma. Evidence for an antiapoptotic function of STAT3. *Leukemia*. 2004;18:1288-1295.
54. Nielsen M, Kaltoft K, Nordahl M, et al. Constitutive activation of a slowly migrating isoform of Stat3 in mycosis fungoides: tyrphostin AG490 inhibits Stat3 activation and growth of mycosis fungoides tumor cell lines. *Proc Natl Acad Sci U S A*. 1997;94:6764-6769.
55. Gallardo F, Sandoval J, Diaz-Lagares A, et al. Notch1 Pathway Activation Results from the Epigenetic Abrogation of Notch-Related MicroRNAs in Mycosis Fungoides. *J Invest Dermatol*. 2015;135:3144-3152.
56. Kiessling MK, Oberholzer PA, Mondal C, et al. High-throughput mutation profiling of CTCL samples reveals KRAS and NRAS mutations sensitizing tumors toward inhibition of the RAS/RAF/MEK signaling cascade. *Blood*. 2011;117:2433-2440.
57. Zhou M, Liu H, Xu X, et al. Identification of nuclear localization signal that governs nuclear import of BRD7 and its essential roles in inhibiting cell cycle progression. *J Cell Biochem*. 2006;98:920-930.
58. Peng YC, Kuo F, Breiding DE, Wang YF, Mansur CP, Androphy EJ. AMF1 (GPS2) modulates p53 transactivation. *Mol Cell Biol*. 2001;21:5913-5924.
59. Gallardo M, Lee HJ, Zhang X, et al. hnRNP K Is a Haploinsufficient Tumor Suppressor that Regulates Proliferation and Differentiation Programs in Hematologic Malignancies. *Cancer Cell*. 2015;28:486-499.
60. Yang K, Neale G, Green DR, He W, Chi H. The tumor suppressor Tsc1 enforces quiescence of naive T cells to promote immune homeostasis and function. *Nat Immunol*. 2011;12:888-897.
61. Koss M, Bolze A, Brendolan A, et al. Congenital asplenia in mice and humans with mutations in a Pbx/Nkx2-5/p15 module. *Dev Cell*. 2012;22:913-926.
62. Stratford JK, Yan F, Hill RA, et al. Genetic and pharmacological inhibition of TTK impairs pancreatic cancer cell line growth by inducing lethal chromosomal instability. *PLoS One*. 2017;12:e0174863.
63. Pelullo M, Quaranta R, Talora C, et al. Notch3/Jagged1 circuitry reinforces notch signaling and sustains T-ALL. *Neoplasia*. 2014;16:1007-1017.
64. Galm O, Yoshikawa H, Esteller M, Osieka R, Herman JG. SOCS-1, a negative regulator of cytokine signaling, is frequently silenced by methylation in multiple myeloma. *Blood*. 2003;101:2784-2788.

65. Yoshimura A, Suzuki M, Sakaguchi R, Hanada T, Yasukawa H. SOCS, Inflammation, and Autoimmunity. *Front Immunol*. 2012;3:20.
66. Jiang S, Zhang HW, Lu MH, et al. MicroRNA-155 functions as an OncomiR in breast cancer by targeting the suppressor of cytokine signaling 1 gene. *Cancer Res.* 2010;70:3119-3127.
67. Zhao XD, Zhang W, Liang HJ, Ji WY. Overexpression of miR -155 promotes proliferation and invasion of human laryngeal squamous cell carcinoma via targeting SOCS1 and STAT3. *PLoS One.* 2013;8:e56395.
68. Glass C, Wuertzer C, Cui X, et al. Global Identification of EVI1 Target Genes in Acute Myeloid Leukemia. *PLoS One.* 2013;8:e67134.

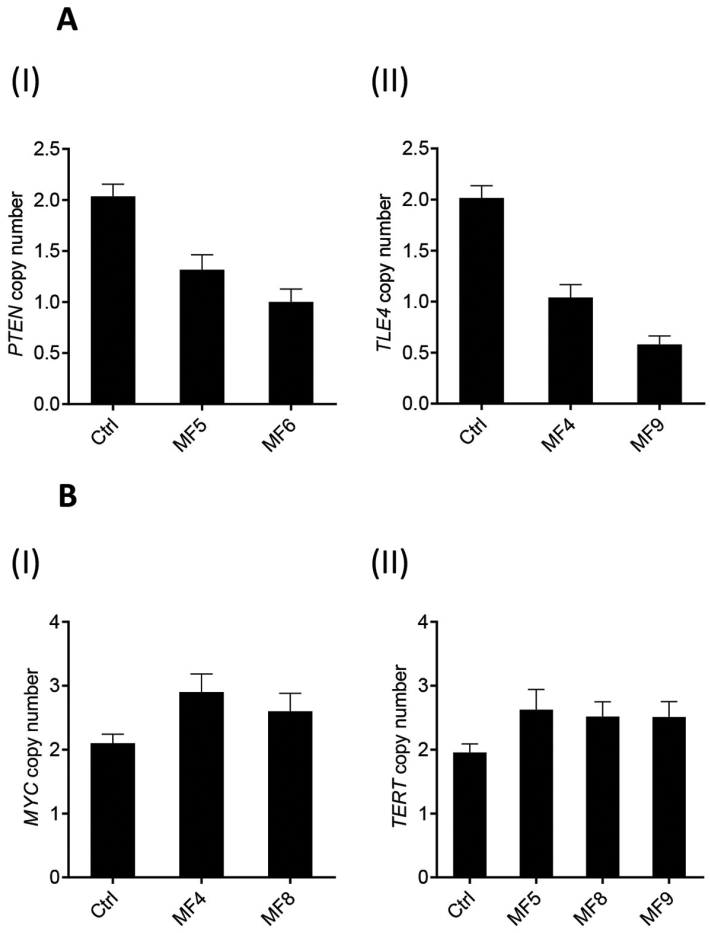
SUPPLEMENTARY FIGURES

Supplementary Figure 1.



Genomic location of BAC probes used in FISH experiments. (A) Break-apart BAC probes for the validation of DNA breaks at 16q13.13 in T-MF. (B) Fusion BAC probes for the detection of t(12;16) in plaque-stage tissue of sample MF3. (C) Fusion BAC probes for the detection of t(16;18) in plaque-stage tissue of sample MF4. See Supplementary Table 3 for exact genomic positions of BAC probes used in this study.

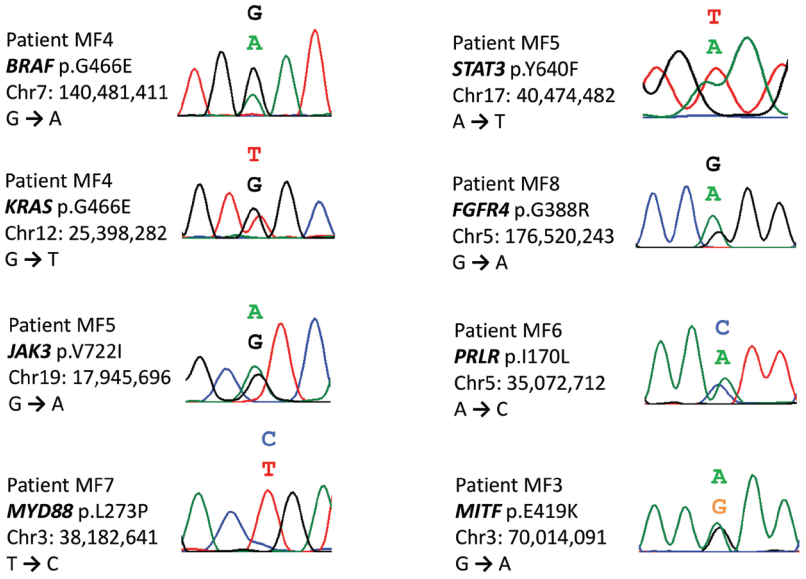
Supplementary Figure 2.



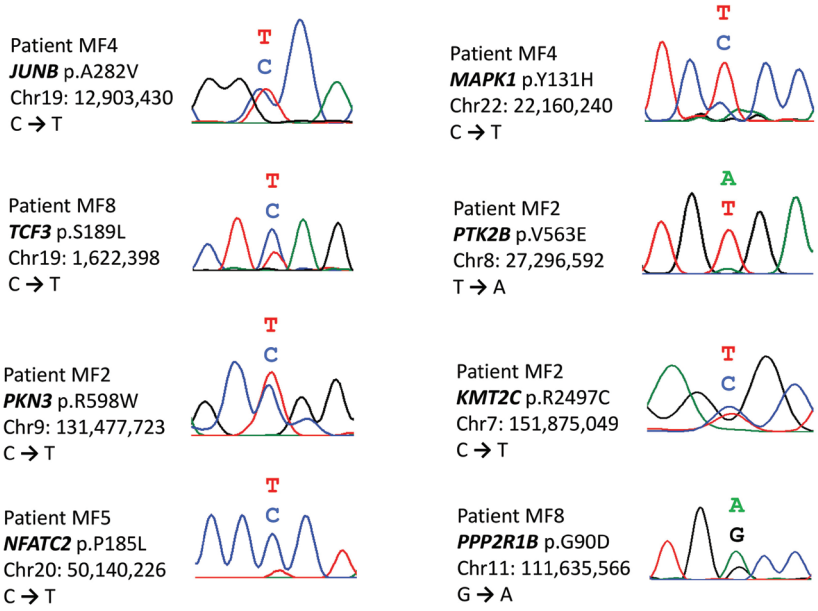
Validation of select CNAs by ddPCR. (A) Deletion of tumor suppressors (I) *PTEN* and (II) *TLE4* in T-MF were confirmed by ddPCR. (B) Gain of oncogenes (I) *MYC* and (II) *TERT* in T-MF were confirmed by ddPCR. Ctrl, CD4+ T cells.

Supplementary Figure 3.

A

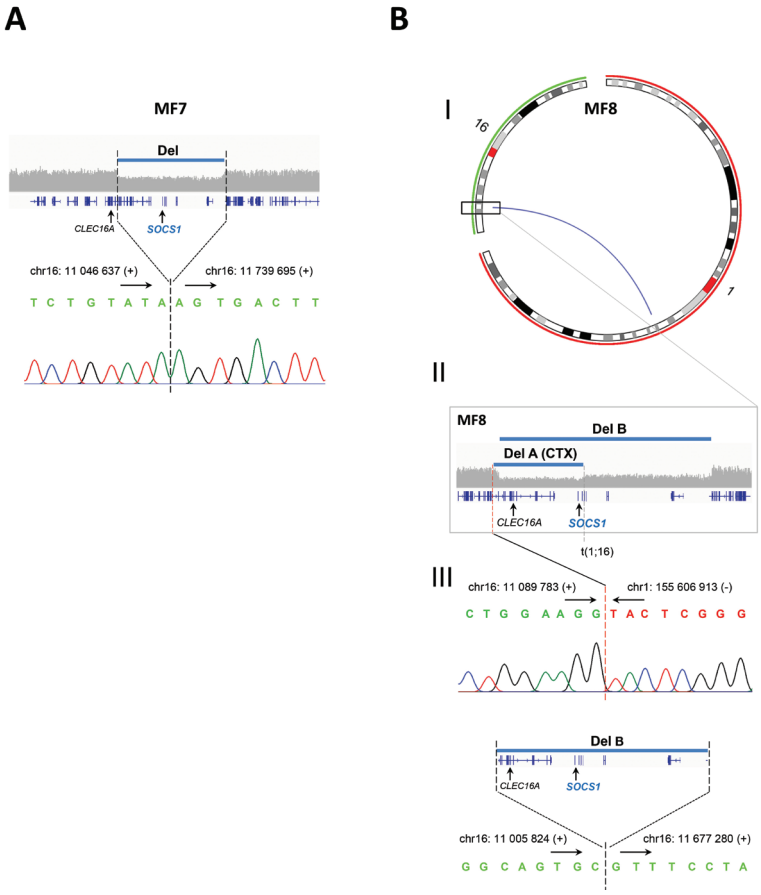


B



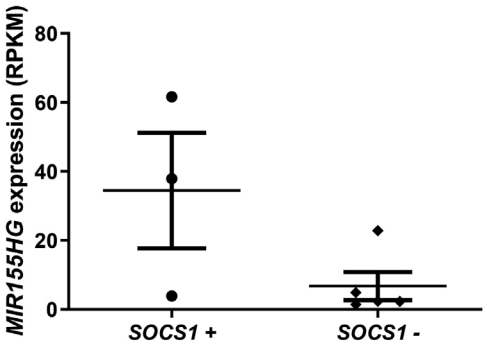
Sanger sequencing validation of select SNVs. (A) Sanger chromatograms of T-MF DNA showing select point mutations with known pathogenic effects (i.e. gain-of-function, susceptibility to disease). (B) Sanger chromatograms of T-MF DNA showing select point mutations with predicted deleterious effects on protein function. Genomic coordinates according to reference genome Hg19.

Supplementary Figure 4.



Validation of genomic alterations at 16q13.13 associated with focal *SOCS1* deletions in mycosis fungoides. (A) Deletion event at 16q13.13 in MF7 confirmed by sanger sequencing. (B) (I) Circos plot displaying a genomic rearrangement at 16q13.13 in MF8. (II) Magnified view of deletions (Del A and Del B) at 16q13.13 resulting from structural alterations in MF8. (III) Genomic events (Del A and Del B) at 16q13.13 validated by sanger sequencing in MF8. Del, deletion. CTX, interchromosomal translocation.

Supplementary Figure 5.



***MIR155HG* expression in T-MF**. Two out of three MF samples with intact copies of *SOCS1* (*SOCS1 +*) express higher levels (7-fold average) of transcript *MIR155HG* compared to MF samples with *SOCS1* deletions (*SOCS1 -*). *MIR155HG* hosts miR-155, a known inhibitor of *SOCS1*.

Note: Supplementary tables available online at https://doi.org/10.1002/gcc.22679.

4

The genetic landscape of primary cutaneous anaplastic large cell lymphoma

Armando N. Bastidas Torres†
Rutger Melchers†
Liana van Grieken
Hailiang Mei
Koen Quint
Rein Willemze
Maarten H. Vermeer
Cornelis P. Tensen
†These authors share first authorship

Manuscript submitted



ABSTRACT

Primary cutaneous anaplastic large cell lymphoma (pcALCL), a hematological neoplasm caused by skin-homing CD30+ malignant T cells, is part of the spectrum of primary cutaneous CD30+ lymphoproliferative diseases. To date, only a small number of genetic alterations have been described in pcALCL, and so far, no unifying theme that could explain the pathogenetic origin of the disease has emerged among patients. To clarify the pathogenetic basis of pcALCL, we performed an integrated (genome/transcriptome) high resolution genetic profiling of this lymphoma by using whole-genome sequencing, whole-exome sequencing and RNA sequencing. Our study, which comprehensively delineated the landscape of genomic rearrangements, copy number alterations and small-scale mutations in this malignancy, revealed that the cell cycle, T-cell physiology regulation, transcription and signaling through the PI-3-K, MAPK and G protein pathways are cellular processes commonly impacted by genetic 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 up-regulation of signal transduction routes associated with the PI-3-K, MAPK and G protein pathways (ERK, phospholipase-C, AKT, etc.). These findings suggest that unlike its systemic counterpart, pcALCL is not driven by deregulated JAK-STAT signaling; thus, therapies targeting PI-3-K/MAPK/G protein signaling might benefit patients who do not respond to first line skin-directed therapies or patients with extracutaneous disease.

1. INTRODUCTION

Primary cutaneous anaplastic large cell lymphoma (pcALCL) is a hematological neoplasm derived from malignant CD30+ large T cells that infiltrate the skin forming solitary or localized tumors.[1,2] 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.[1-3]

Unlike systemic anaplastic large cell lymphoma (sALCL), which appears to be driven by a variety of genetic alterations that overactivate STAT3 signaling,[4-9] 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;[10] thus, genetic alterations underlying pcALCL are still unknown.

Chromosomal rearrangements involving *ALK*, *DUSP22*, *TP63* and *TYK2*, common in sALCL, have also been detected in pcALCL, albeit, at much lower frequencies.[10,11] Frequent copy number alterations (CNAs) include losses within chromosomes 6 and 13 and gains within chromosomes 7 and 17.[12,13] Overexpressed genes with potential pathogenic and/or diagnostic relevance include *EZH2* and *IRF4*.[12,14,15] Oncomir miR-155, which is associated with other cutaneous T-cell lymphoma (CTCL) variants, is overexpressed in pcALCL as well.[16,17] Finally, single nucleotide variants (SNVs) presumed as pathogenic have been reported in *JAK1*, *MSC* and *STAT3* in a few cases.[5,8]

Given the lack of a clear unifying theme between the few reported molecular alterations in the disease, we performed an integrated (genome/transcriptome) analysis 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-mediated pathways underlie pcALCL. These findings suggest that contrary to sALCL, pcALCL is not driven by deregulated JAK-STAT signaling.

2. MATERIALS AND METHODS

2.1 Patient selection

Frozen tumor biopsies (≥ 75% tumor cells) from 12 patients with pcALCL (Table S1) were subjected to WGS and RNA-seq. Seven matched tumor/germline (granulocytes) samples from this cohort (i.e., cAL2, cAL3, cAL4, cAL5, cAL9, cAL10 and cAL11) were additionally subjected to WES. Diagnosis was performed by an expert panel of dermatologist and pathologists according to the criteria of the 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.

2.2 Sequencing

For WGS, 1.2 ng of high-molecular weight was used to generate DNA linked-read libraries using the Chromium Genome Library & Gel Bead kit v2 (10x Genomics) according to the manufacturer's protocol. Libraries were sequenced on the Illumina HiSeq X-Ten platform using paired-end 150bp reads. For WES, genomic DNA was fragmented, 3'-adenylated and subjected to whole-exome capture using SureSelect Exome Target Enrichment (Agilent). Captured exome DNA was converted into single-stranded circle DNA libraries, nanoball amplified and sequenced on the BGISEQ-500 platform. For RNA-seq, transcripts were enriched with Oligo(dT) Dynabeads (Invitrogen), and libraries were prepared using reverse transcribed cDNA using TruSeq library preparation kit (Illumina) according to the manufacturer's protocol. Libraries were sequenced on the Illumina HiSeq4000 platform. All datasets (WGS, WES, RNA-seq) have been deposited in the European Genome-Phenome Archive (EGA) under study number EGAS00001004429.

2.3 Sequence data processing

For WGS, raw data were processed using the Long Ranger pipeline (10X Genomics). Raw files were demultiplexed and converted into fastq files with Long Ranger mkfastq (v2.1.3) function. Fastq files were aligned to reference genome GRCh38 using Long Ranger's 'Lariat' aligner (Table S2). For WES, raw reads were processed using FastQC (v0.11.7), SeqTK (v1.0-r63) and Cutadapt (v2.4). Clean reads were aligned to the human reference genome GRCh38 using BWA-MEM (v0.7.17-r1188) (Table S3). For RNA-seq, raw reads were processed with FastQC (v0.11.7) and Cutadapt (v2.4). Clean reads were aligned to human reference genome GRCh38 using STAR (v2.6.0c). Fragment counts were obtained with HTSeq-count (v0.9.1) using Ensembl gene annotations (v94).

2.4 Discovery of DNA rearrangements and fusion transcripts

Detection of structural genomic variants (SV) was performed using the integrated structural variant caller of 10X Genomics Long Ranger pipeline. Split read calls were manually verified and curated using the Integrative Genomic Viewer (IGV, v2.3.78) (Table S4). Fusion transcripts were investigated using an in-house pipeline that included callers FusionCatcher (v0.99.6a) and Star Fusion (v0.8.0). Fusion transcript calls were contrasted with genomic SV data and visually verified on DNA level using IGV (Table S5). Rearranged genes implicated in cancer were identified using the Network of Cancer Genes 6.0 (NCG 6.0)[18] and literature search.

2.5 Detection of CNAs

Copy number alterations (CNAs) were identified with Control-FREEC (v8.7) and GISTIC2.0 (Figure S1, Table S6). Recurrent CNAs derived from these analyses were visually verified and curated using IGV (Table S7). Events confirmed after inspection are reported in this manuscript.

Detection of indels and SNVs

For patients with matched germline controls, indels/SNVs were called from WES data using MuTect (Table S8). Genes found to be mutated in patients with matched tumor/germline samples were screened for mutations in WGS data from patients without germline DNA.

2.6 Mutational signature analysis

The assessment of mutational signatures in pcALCL was performed with MuSiCa. A group of 30 Sézary syndrome patients with previously characterized signatures were included in the analyses as a technical control group.

2.7 Differential expression analysis

Differential expression (DE) analysis was performed using Limma-Voom (v3.40.6) with TMM-normalized data. pcALCL samples were compared to a control group formed by seven CD4+ T-cell subsets ($T_{naïve}$, T_H1, T_H2, T_H17, T_{reg}, T_{EM}, T_{CM}; 5 samples per subset) generated by Ranzani et al. and downloaded from EBI (https://www.ebi.ac.uk/).[19] Genes with FDR < 0.01 were considered differentially expressed (Table S9).

2.8 Functional annotation and pathway analysis

Gene ontology annotation of rearranged genes was performed with DAVID (v6.8) (Table S10).[20] Enrichment analysis of mutated genes (indels/SNVs) was performed with Panther (v.14.0) (Table S11). Gene set analysis with DE genes was performed with GeneAnalytics (Weizmann Institute of Science) using default

settings (Table S12).[21] Additional characterization of the transcriptome (ARCHS4 database) was performed with Enrichr (Table S13).[22]

2.9 Validation of genetic alterations and DE genes

Select rearrangements and SNVs were validated by Sanger sequencing on the ABI3730xl platform (Applied Biosystems) (Figure S2, Figure S3). Select CNAs were validated by digital droplet PCR (ddPCR) using QX200 ddPCR system (Bio-Rad) as previously described. Select DE genes were validated by RT-qPCR as described by van Doorn et al.

3. RESULTS

3.1 Landscape of genomic rearrangements

The landscape of rearrangements of pcALCL showed considerable inter-patient heterogeneity. The number of rearrangements varied from 1 to 51 per patient (205 total events; mean/patient ± SD, 17 ± 17) (Figure 1, Figure 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 nongenic region (42%), two nongenic regions (25%) or reordering sequences within a unique 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 (Figure S4).

We detected a total of 166 rearranged genes in our cohort (Table S14), including a subset of 44 genes implicated in cancer (Figure 1, Table S15). The majority of events were patient-specific. *DUSP22* and *TP63* rearrangements, reported in pcALCL before,[10] were found in single cases (Figure S3). 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, Table S10). Analysis of rearranged genes with NCG 6.0 uncovered cancer genes principally associated with three cellular processes/pathways: the cell cycle (i.e. *FMN2*, *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*, *PPM1L*, *PTPRK*, *TNK1*) and the regulation of RNA expression/processing (i.e. *ETS1*, *RBFOX1*, *TRERF1*, *ZNF208*, *ZNF28*) (Figure S3).

A group of five genes were found to be recurrently rearranged in pcALCL: *GPS2* (two patients), *LINC-PINT* (three patients), *RBFOX1* (two patients), *TNK1*

(two patients) and *VPS13D* (two patients). Notably, four of five of these genes encode products with roles in the cell cycle, transcription or signaling. *LINC-PINT* is a long noncoding RNA that acts as a TP53-dependent regulator of the cell cycle and a PRC2-dependent silencer of gene expression.[23] RBFOX1 is an RNA-binding protein that regulates alternative splicing.[24] GPS2 and TNK1 are negative regulators of RAS/JNK1 and RAS/RAF1 signaling, respectively.[25,26]

Since deregulated RAS-MAPK signaling is a common driver of human cancers, we looked at potential detrimental effects of the rearrangements on the expression of *GPS2* and *TNK1*. The expression data confirmed that both genes were functionally compromised. For instance, patient cAL4 expressed a truncated *GPS2-CHD3* transcript which rendered both fusion partners dysfunctional (Figure 2E-F). Similarly, patient cAL3 expressed transcripts encoding a truncated form of TNK1 which was structurally analogous to a variant with proven oncogenic activity (Figure 2G-H).[27]

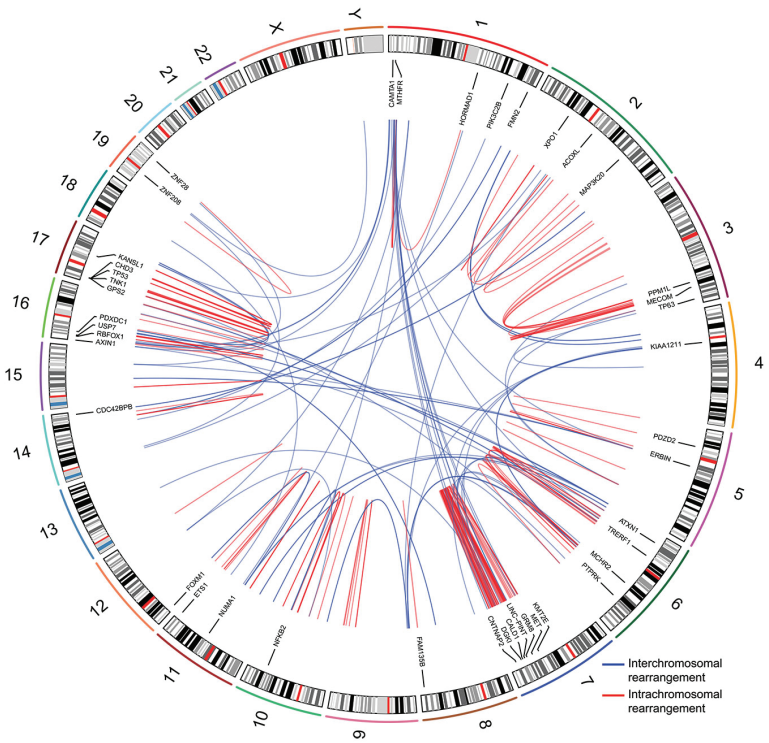


Figure 1. Landscape of genomic rearrangements in pcALCL. Circos plot showing 205 genomic rearrangements detected in twelve pcALCL genomes by WGS. The outer ring is formed by human chromosome ideograms arranged circularly end to end. The area at the center of the plot contains arcs representing interchromosomal (blue) and intra-chromosomal (red) events. The ring between the chromosome ideograms and the arcs contains rearranged genes with established roles in cancer.

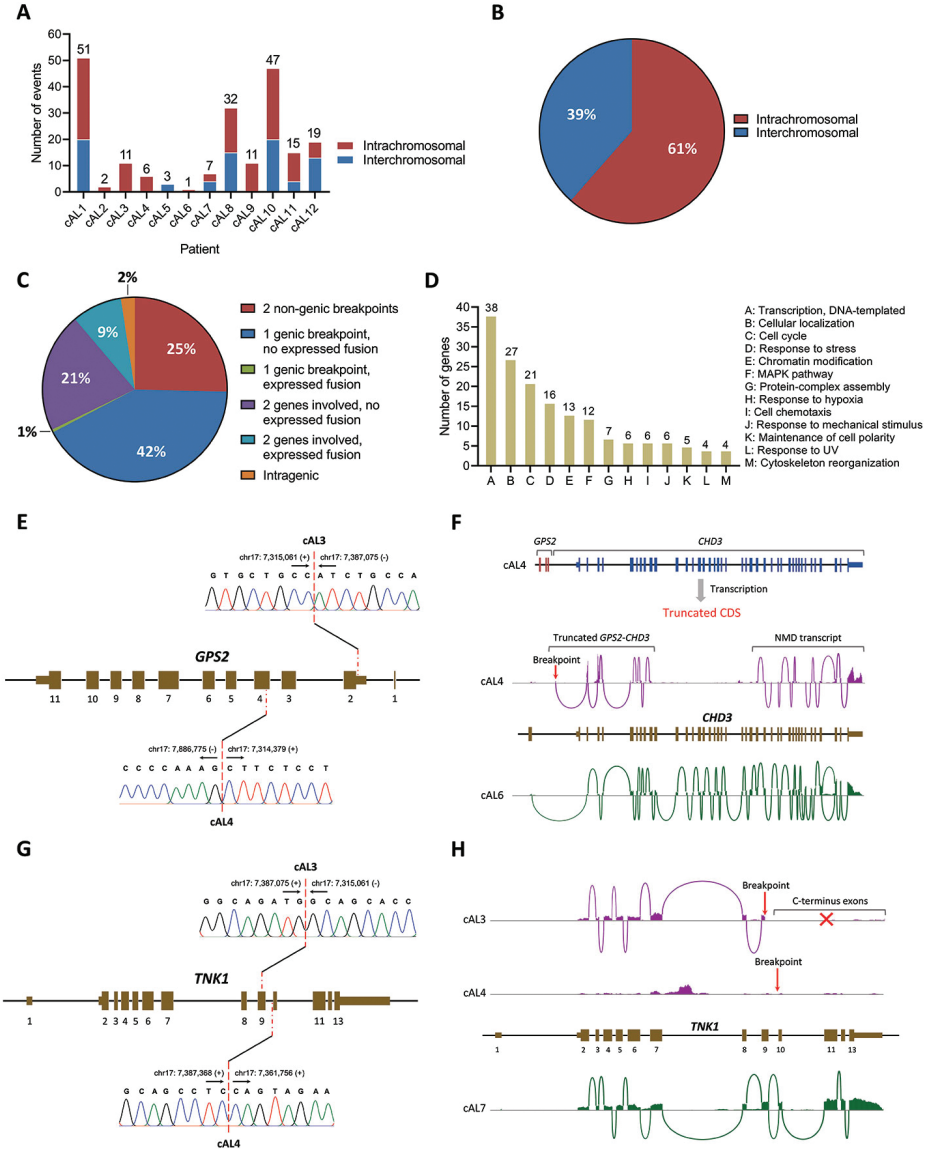


Figure 2. Rearrangements disrupt RAS-MAPK signaling inhibitors in pcALCL. (A) Number of genomic rearrangements per patient. The distribution of inter- and intrachromosomal rearrangements per patient is shown; (B) Distribution of inter- and intrachromosomal rearrangements per cohort; (C) Distribution 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 WGS and RNA-seq data; (D) Gene ontology annotation of rearranged genes; (E) Validation of *GPS2* rearrangements in patients cAL3 and cAL4 by Sanger sequencing; (F) Diagram of *GPS2-CHD3* fusion in patient cAL4 and Sashimi plots showing the expression of *CHD3* in patients cAL4 (disrupted by rearrangement) and cAL6 (with intact alleles). Besides disrupting the sup-

pressor domain of GPS2, the fusion event also inactivated chromatin remodeler *CHD3*. In patient cAL4, the *GPS2*-fused *CHD3* transcript (5'-side of the gene) ends prematurely. An NMD transcript (3'-side of the gene) is expressed too. NMD, nonsense mediated decay; (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) and cAL7 (with intact alleles). 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.

3.2 Landscape of CNAs

pcALCL displayed a predominance of large numerical imbalances 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 (Figure 3A, Figure 4). Our analysis detected 15 focal (≤ 3Mb) minimal common regions (MCRs) between CNAs found in our patients (9 deletions, 6 gains) (Table S7), 6 of which enclosed (putative) cancer genes (Figure 4). Half of these genes participate in the regulation of T-cell physiology.

Deletion at 6q21, observed in six of twelve patients, was the most recurrent focal MCR in pcALCL. This region contained tumor suppressor *PRDM1* (*BLIMP1*) (Figure 3B-C), which encodes a transcription factor that attenuates the proliferation and survival of T cells through suppression of IL-2 signaling.[28] Three of twelve patients had deletions at 13q32.2 which enclosed *STK24* (Figure 3D). This gene encodes a pro-apoptotic serine/threonine kinase that modulates phosphorylation of JNK and p38 kinases in response to stress stimuli.[29] Two of twelve patients had deletions at 1p36.32 and 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.[30] Of note, loss of *TNFRSF14* has been reported in pcALCL before.[13]

In addition, focal (≤ 3Mb) gains at 1p36.22 and 2p15 were observed in two and three out of twelve 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 RNAs from the nucleus to the cytoplasm.[31] Furthermore, 6 of 12 patients had gains at 7q36.1 (MCR ~ 6 Mb), which contained *EZH2*, the catalytic subunit of polycomb repressive complex 2 (PRC2) (Figure 3D).

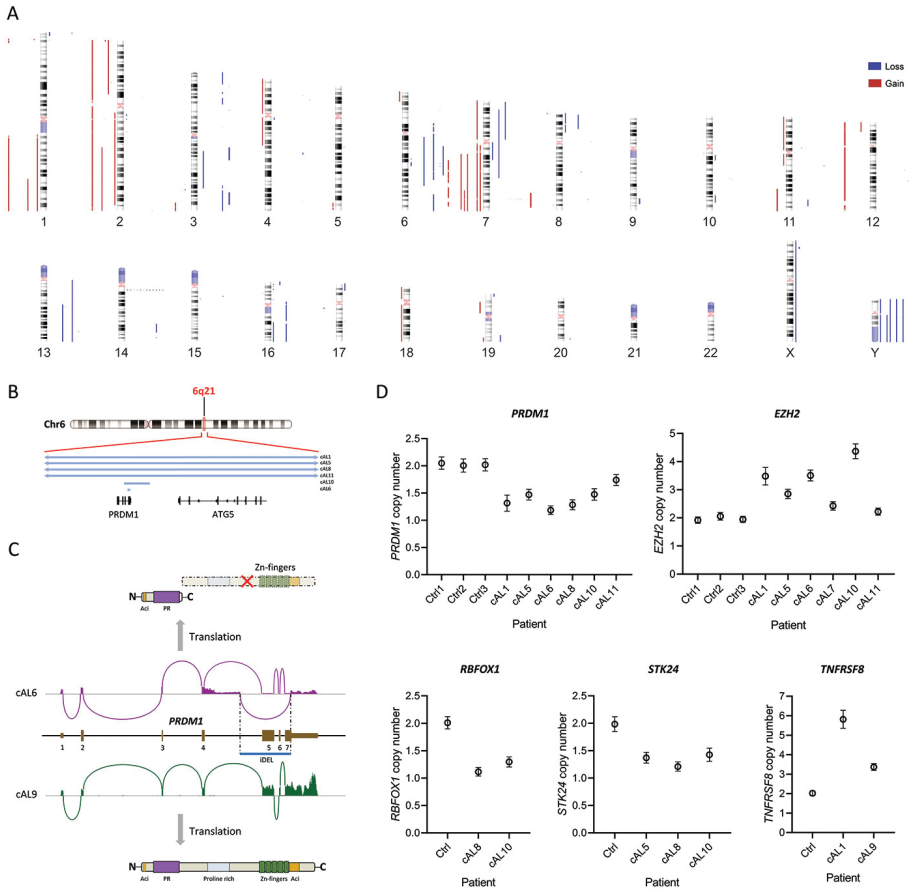


Figure 3. Landscape of CNAs in pcALCL. (A) Human chromosome ideograms showing regions of gain and loss detected through WGS in twelve pcALCL genomes. Blue bars to the right of the chromosomes depict regions of loss whereas red bars to the left of the chromosomes depict regions of gain. (B) Diagram representing deletions (light blue bars) at 6q21 in six patients with pcALCL. The minimum common region between these events 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) and cAL9 (with intact alleles). Patient cAL6 expressed *PRDM1* transcripts that lacked exons 5, 6 and 7 as a result of a monoallelic focal deletion (blue bar, 4.6 Kb). This transcript encodes a PRDM1 protein lacking the Zn finger domains which are required for its functionality. iDEL, interstitial deletion. (D) Recurrent CNAs affecting *PRDM1*, *EZH2*, *RBFOX*, *STK24* and *TNFRSF8* in pcALCL were validated by ddPCR. Ctrl, CD4+ T cells.

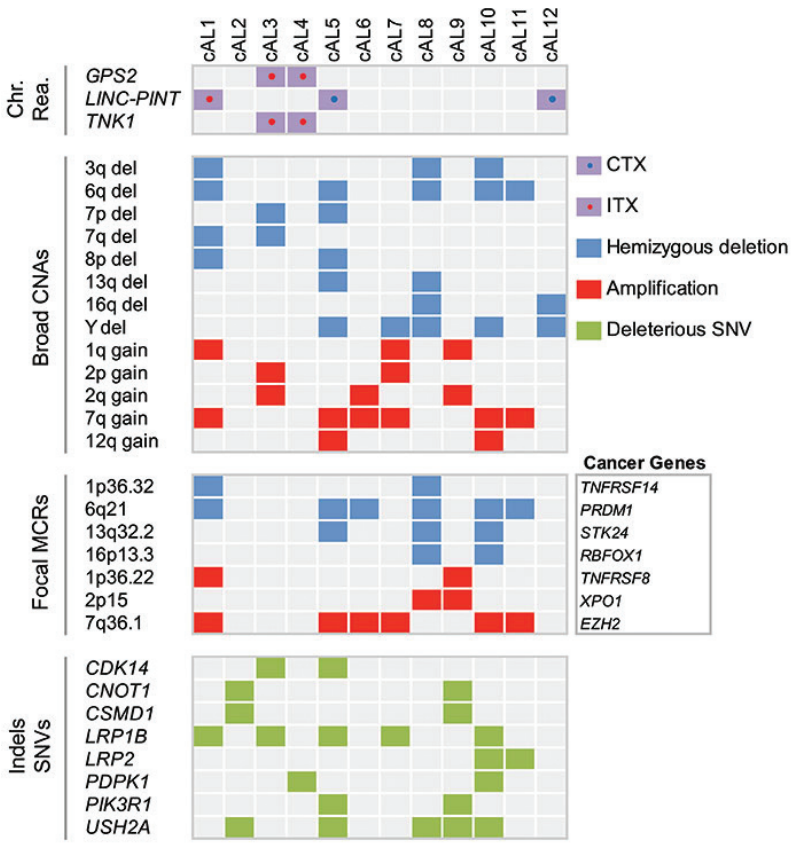


Figure 4. Distribution of recurrent genetic alterations in pcALCL. First panel: recurrent chromosomal rearrangements impacting cancer genes. Second panel: recurrent large-scale CNAs (> 3 Mb). Third panel: focal MCRs shared by CNAs; cancer-associated genes residing within focal MCRs are specified. Fourth panel: Recurrently mutated (indels/SNVs) cancer genes. CTX, interchromosomal rearrangement; ITX, intrachromosomal rearrangement.

3.3 Landscape of small-scale mutations and mutational signatures

To detect somatic small-scale mutations (indels, SNVs) in pcALCL, we performed WES on matched tumor/germline samples from seven patients. The number of somatic indels/SNVs ranged from 48 to 235 (665 total events; mean/patient ± SD, 95 ± 65). The average somatic mutation rate was 15.8 mutations/Mb (range/patient, 9.2–38.6 mutations/Mb). The large majority of small-scale mutations were missense SNVs (Figure 5A, Table S8). 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 AID-mediated somatic hypermutation were smaller contributors to the mutational landscape (Figure 5C).

A total of 601 genes were found to carry somatic mutations across 7 patients, 153 of which were genes associated with cancer (130 candidate cancer genes, 23 *bona fide* oncogenes/tumor suppressors) (Table S16). Most mutations were patient-specific. Gene set analysis of mutated genes with Panther identified enrichment of genes involved in signaling mediated by G-proteins (e.g. beta-adrenergic receptors, glutamate receptors, etc.) and PI-3-K signaling (Figure 5D, Table S11). 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*, *KMT2C*, *PRDM16*, *TRRAP*), hematopoiesis (i.e. *FLT3*, *FSTL3*, *GATA2*), signaling via the P-I-3K and MAPK pathways (i.e. *ERBB4*, *FGFR1*, *MAP3K1*, *PDPK1*, *PIK3R1*, *RASA1*, *RET*, *WNK2*) and G-protein signaling (i.e. *CREB3L1*, *DRD5*, *RGS7*, *RGS12*).

We identified a group of 30 recurrently mutated genes ($n_{patients}$ = 2-4) in our cohort (Figure S2, Table S17), which included (putative) oncogenes (i.e. *CDK14*, *CNOT1*, *LRP2*, *PDPK1*) and tumor suppressors (i.e. *CSMD1*, *LRP1B*, *PIK3R1*) (Figure S2). 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 twelve patients had mutations in the PH domain of PDPK1 (cAL4: p.T518M; cAL10: p.W535L) (Figure 5E-F). In contrast, *PIK3R1* encodes the regulatory (inhibitory) subunit of class IA phosphoinositide 3-kinases (PI3Ks). Class IA PI3Ks mediate the production of PIP_3, 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 twelve 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).

3.4 Differentially expressed genes and fusion transcripts

Given the fact that the CD4+ T cell subset giving rise to pcALCL is unknown at present, we compared gene expression in pcALCL with gene expression in a control group formed by several CD4+ T-cell subsets to find in this neoplasm abnormal expression profiles absent in various normal phenotypes of CD4+ T-cells. This comparison identified 3162 differentially expressed (DE) genes (1716 up-regulated, 1446 down-regulated, FDR < 0.01) (Figure 6A, Table S9).

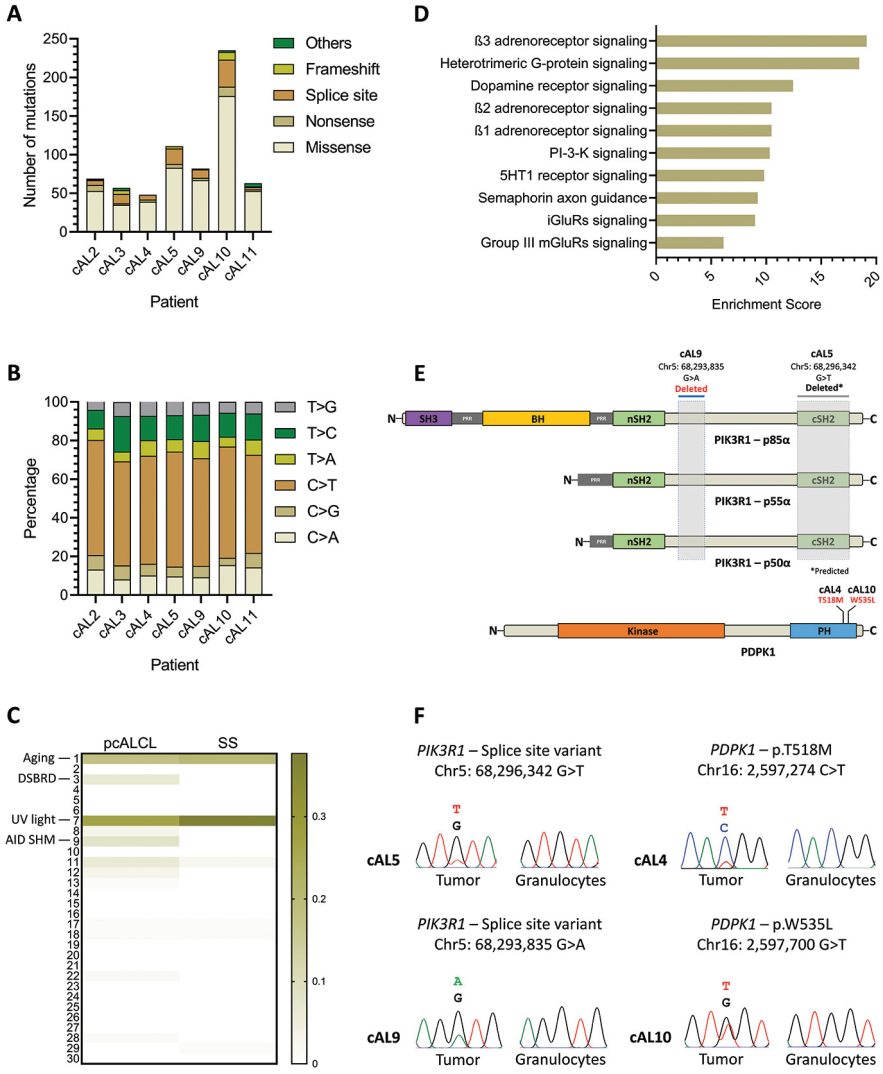


Figure 5. Landscape of small-scale mutations in pcALCL reveals recurrent events in PI-3-K/Akt pathway genes. (A) Number of small-scale mutations per patient. The distribution of missense, nonsense, frameshift and splice site mutations per patient is shown. (B) Distribution of nucleotide substitutions per patient. (C) Median contribution of COSMIC mutational signatures in pcALCL (n=7) and SS (n=30). AID SHM, Activation-induced cytidine deaminase somatic hypermutation; DSBRD, double-strand break repair deficiency; SS, Sézary syndrome; UV, ultraviolet. (D) Gene set analysis of genes affected by indels and SNVs in pcALCL. (E) Diagram showing small-scale mutations affecting PIK3R1 and PDPK1 in patients with pcALCL. Splice site mutations in *PIK3R1* are predicted to affect the three known protein isoforms (i.e. p85α, p55α and p50α) encoded by this gene. Missense mutations in *PDPK1* impacted the PH domain of its protein product. (F) Sanger chromatograms showing validation of splice site and missense mutations in *PIK3R1* and *PDPK1*, respectively.

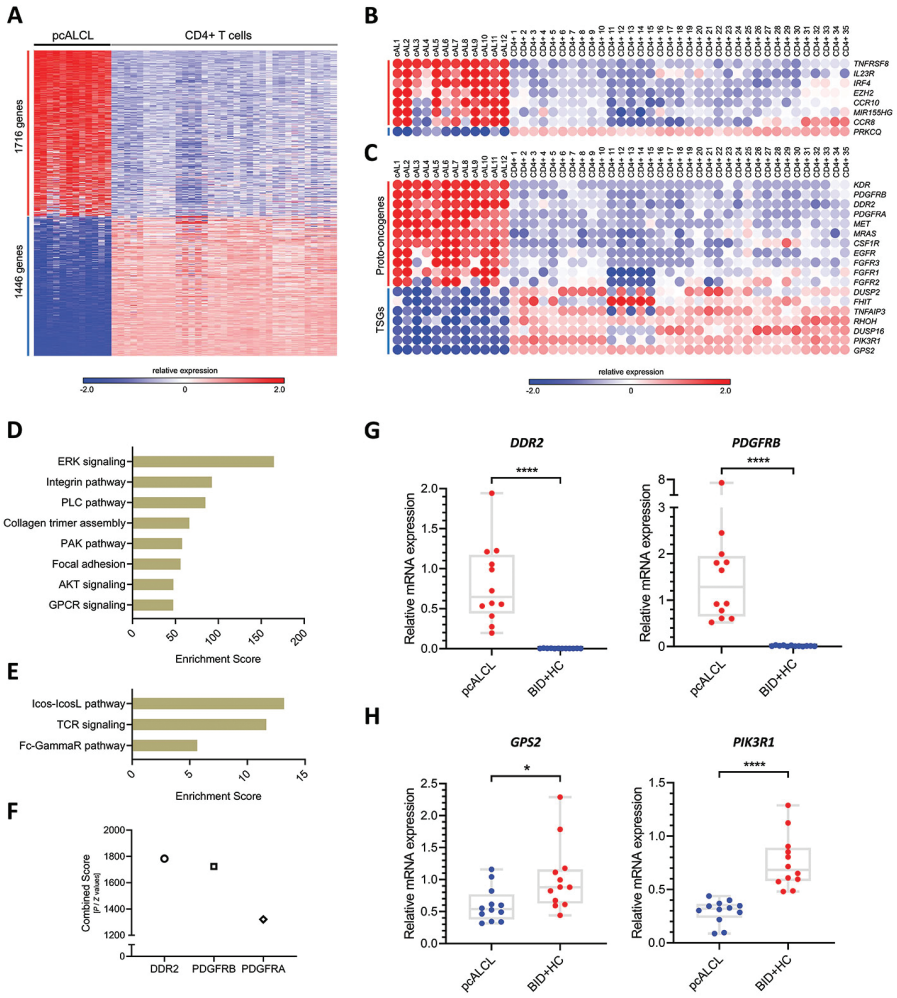


Figure 6. RNA-seq supports up-regulation of signaling via the PI-3-K/Akt pathway, the MAPK pathway and G proteins in pcALCL. (A) Heat map showing 3162 differentially expressed genes (1716 up-regulated, 1446 down-regulated, FDR <0.01) in pcALCL when compared to CD4+ T cells. (B) Heat map showing expression of genes with diagnostic or pathogenic relevance in pcALCL identified in previous studies. RNA-seq confirmed the up-regulation of markers and oncogenes that characterize this lymphoma. (C) Heatmap showing (proto)-oncogenes and tumor suppressors deregulated in pcALCL and involved in the PI-3-K/Akt pathway and the MAPK pathway. (D) Up-regulated cellular processes/pathways in pcALCL as determined by gene set pathway analysis with GeneAnalytics. (E) Down-regulated cellular processes/pathways in pcALCL as determined by gene set pathway analysis with GeneAnalytics. (F) Top three receptor kinases for which known co-expressed genes were overrepresented in the transcriptome of pcALCL. (G) Validation of *DDR2* and *PDGFRB* up-regulation in pcALCL by RT-qPCR. (H) Validation of *GPS2* and *PIK3R1* down-regulation in pcALCL by RT-qPCR. (* $P < 0.05$, **** $P < 0.0001$); pcALCL (n=12), BID (n=8), HC (n=4).

The analysis not only confirmed up-regulation of genes found to be overexpressed in pcALCL in prior studies (e.g. *CCR8*, *CCR10*, *EZH2*, *IRF4*, *MIR155HG*, *TNFRSF8*) (Figure 6B),[12,15-17] but also revealed up-regulation of a group of (proto)-oncogenes (e.g. *CSF1R*, *EGFR*, *FGFR1/2/3*, *MET*, *PDGFRA/B*) involved in signaling through the P-I-3K/Akt pathway and/or the MAPK pathway (Figure 6C). Of note, *CSF1R* overexpression has been reported in sALCL before.[32] Down-regulated genes included inhibitors of the P-I-3K/Akt pathway (i.e. *PIK3IP1*, *PIK3R1*), inhibitors of the MAPK pathway (i.e. *DUSP2*, *DUSP16*, *GPS2*) and tumor suppressors implicated in other hematological cancers (i.e. *FHIT*, *RHOH*, *TNFAIP3*) (Figure 6C).

Table 1. Fusion transcripts detected by RNA-seq in pcALCL. CTX, interchromosomal translocation. ITX, intrachromosomal translocation. iDel, interstitial deletion.

Sample	Fusion transcript	Breakpoints (DNA)	Breakpoint Type	Event Class	WGS con-firmed
cAL1	*NLRP1–DERL2*	chr17:5554176 - chr17:5477401	Genic - Genic	iDel	Yes
cAL1	*ANKRD11–VPS9D1-AS1*	chr16:89306818 - chr16:89713470	Genic - Genic	ITX	Yes
cAL1	*INPP5B–GNL2*	chr1:37926300 - chr1:37586211	Genic - Genic	iDel	Yes
cAL2	*IGSF3–CD58*	chr1:116578013 - chr1:116525260	Genic - Genic	iDel	Yes
cAL3	*TNK1–GPS2*	chr17:7386883 - chr17:7315061	Genic - Genic	ITX	Yes
cAL3	*TP53–TNK1*	chr17:7682311 - chr17:7382806	Genic - Genic	ITX	Yes
cAL4	*GPS2–CHD3*	chr17:7314379 - chr17:7886522	Genic - Genic	ITX	Yes
cAL5	*LINC-PINT–RNASEH2B-AS1*	chr7:130895998 - chr13:50895286	Genic - Genic	CTX	Yes
cAL5	*SMARCAD1–TP63*	chr4:94244727 - chr3:189786186	Genic - Genic	CTX	Yes
cAL5	*TP63–SMARCAD1*	chr3:189786186 - chr4:94244727	Genic - Genic	CTX	Yes
cAL5	*ETS1–SLC37A4*	chr11:128520859 - chr11:119027083	Genic - Genic	ITX	Yes
cAL7	*NFKB2–SIRT3*	chr10:102401366 - chr11:217595	Genic - Genic	CTX	Yes
cAL7	*SIRT3–NFKB2*	chr11:217564 - chr10:102401164	Genic - Genic	ITX	Yes
cAL7	*GTPBP2–TRERF1*	chr6:43623524 - chr6:42232607	Genic - Genic	CTX	Yes
cAL8	*USP7–SAMD11*	chr16:8943214 - chr1:942313	Genic - Genic	CTX	Yes
cAL8	*MMS19–RRP12*	chr10:97461940 - chr10:97410043	Genic - Nongenic	iDel	Yes
cAL9	*NFAT5–WWP2*	chr16:69595202 - chr16:69801863	Genic - Genic	iDel	Yes
cAL10	*RABEPK–UBAC1*	chr9:125226314 - chr9:135960125	Genic - Genic	ITX	Yes
cAL10	*HIVEP3–MRPL43*	chr1:41890425 - chr10:100998633	Genic - Genic	CTX	Yes
cAL10	*DUSP22–UNC93B1*	chr6:314945 - chr11:67992497	Genic - Genic	CTX	Yes
cAL11	*AGTRAP–LST1*	chr1:11738153 - chr6:31588590	Genic - Genic	CTX	Yes
cAL11	*CLK2–NUMA1*	chr1:155267627 - chr11:72044112	Genic - Genic	CTX	Yes
cAL11	*KANSL1–ARL17B*	chr17:46087894 - chr17:46356512	Genic - Genic	ITX	Yes
cAL12	*ERBB2IP–SQRDL*	chr5:66078502 - chr15:45654289	Genic - Genic	CTX	Yes
cAL12	*RBPJ–TRAPPC12*	chr4:26193798 - chr2:3436901	Nongenic - Genic	CTX	Yes

Twenty-five fusion transcripts were identified in our cohort (Table 1), 9 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*, *TP63-SMARCAD1*).

3.5 Deregulated cellular processes/pathways

To investigate deregulated cellular processes/pathways in pcALCL, we analyzed up- and down-regulated genes separately with GeneAnalytics. Up-regulated 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 up-regulation of adhesion-related cellular processes (i.e. focal adhesion, integrin pathway) (Figure 6D). On the contrary, all down-regulated profiles were signaling processes leading to T-cell activation (e.g. T-cell receptor (TCR) signaling, Icos-IcosL pathway, etc.) (Figure 6E).

Transcriptome analysis with Enrichr using ARCHS4 database revealed that genes known to be co-expressed with receptor tyrosine kinases DDR2 and PDGFR (esp. PDGFRβ) were the most overrepresented in the group of up-regulated genes (Figure 6F), suggesting that these receptors and/or their associated signal transduction networks might play relevant roles in the disease. DDR2 binds collagen and activates signaling pathways (e.g. the MAPK pathway) implicated in cell adhesion, proliferation and ECM remodeling. Similarly, PDGFR binds its cognate ligands (PDGFs) and activates the PI-3-K/Akt and MAPK pathways, promoting cell proliferation and survival.

To validate the up-regulation 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 up-regulated in pcALCL in comparison to the control group ($P < 0.0001$, Mann-Whitney 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 down-regulated in pcALCL according to the sequencing data. The analysis confirmed the down-regulation of both tumor suppressors in pcALCL as well ($P < 0.05$, Mann-Whitney test) (Figure 6H).

4. DISCUSSION

To gain insight into the pathogenetic basis of pcALCL, we investigated rearrangements, CNAs, small-scale mutations and gene expression in this lymphoma by integrating 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.[10,33] 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. Sézary syndrome, mycosis fungoides), we found that pcALCL displayed a predominant C>T mutational signature which was mainly attributable to UV light exposure and aging.

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*, *PDPK1* and *PIK3R1*.

Deletion of *PRDM1* and gain of *EZH2*, each observed in six of twelve patients, were the two most frequent genetic alterations in our cohort. *PRDM1*, which is recurrently deleted in sALCL too, acts as a tumor suppressor.[34] Prior research showed that reintroduction of *PRDM1* in a sALCL cell line triggered cell proliferation impairment, cell cycle block and apoptosis.[34] Notably, *PRDM1* reintroduction also down-regulated in these cells the expression of miR-155 precursor *MIR155HG* and *IRF4*, two genes that are consistently overexpressed in sALCL and pcALCL.[34] On the other hand, previous studies have shown that *EZH2* is overexpressed in pcALCL.[12,15] *EZH2* appears to operate as an oncogene in the disease by epigenetically silencing *TXNIP* and *CXCL10*, two genes with anti-tumor effects in pcALCL.[15] Although the molecular mechanisms leading to EZH2 overexpression 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 (i.e. *LINC-PINT*, *LRP1B*, *RBFOX1* and *USH2A*). *LINC-PINT* and *RBFOX1* behave as tumor suppressors in gastric cancer and glioblastoma.[35-38] *LRP1B*, an LDL family receptor involved in endocytosis, is inactivated by genetic and

epigenetic mechanisms in several cancers (e.g. gastric, renal, thyroid, etc.).[39-41] *USH2A*, one of the two most SNV-mutated genes in our cohort, is the 21st most mutated gene in breast cancer according to the Cancer Genome Atlas (TCGA);[42] 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 between patients with pcALCL. Nine of twelve patients had abnormalities in one or more genes linked to the aforementioned pathways (Table S18). Patient-specific genetic alterations, which accounted for the large majority of events, included rearrangements and small-scale mutations in *CREB3L1*, *DGKI*, *DRD5*, *ERBB2IP*, *ERBB4*, *FGFR1*, *GRM8*, *MAP3K1*, *MAP3K20*, *MET*, *PIK3C2B*, *PPM1L*, *PTPRK*, *RASA1*, *RET*, *RGS7*, *RGS12* and *WNK2*. In addition, recurrent genetic defects were found in two inhibitors of the MAPK pathway (*GPS2*, *TNK1*) and two key regulators of the PI-3-K/Akt pathway (*PDPK1*, *PIK3R1*).

GPS2 and *TNK1* had structural defects in two of twelve patients. *GPS2* formed dysfunctional fusions with other cancer genes (i.e. *CHD3*, *TNK1*), inactivating both fusion partners in a single event. *GPS2* is lowly expressed in liposarcoma and rearranged in spindle cell sarcoma.[43] *In vitro* knockdown of *GPS2* has been shown to enhance proliferation, migration and dedifferentiation of liposarcoma cells, supporting its role as a tumor suppressor in this neoplasm.[43] Similarly, rearrangements abolished *TNK1* expression or rendered it dysfunctional in two of twelve patients. *TNK1* is often deleted and lowly expressed in diffuse large B-cell lymphoma.[44] Mouse models carrying monoallelic or biallelic *TNK1* inactivation exhibit tissues with higher levels of RAS activation and develop spontaneous tumors at high rates, including lymphomas and carcinomas.[26] Curiously, albeit wild type TNK1 exhibits anti-proliferative properties, C-terminus truncated forms of the protein like the one observed in patient cAL3 appear to activate proliferation-promoting kinases such as AKT and STATs.[27]

We observed small-scale mutations affecting *PDPK1* or *PIK3R1* in four of twelve patients. *PDPK1* is frequently amplified or overexpressed in human cancers, but uncommonly affected by pathogenic SNVs.[45] Surprinsingly, we found two patients with SNVs (p.T518M; p.W535L) affecting the PH domain of this kinase. PDPK1 becomes contitutively active as a monomer upon dissociation from its inactive homodimeric form.[46] As activation of PDPK1 requires auto-phosphorylation 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 for PH domain mutation p.T513E.[46] In contrast, loss-of-function mutations in

PIK3R1 occur in glioblastoma, endometrial carcinoma and immune disorders.[47] 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.[47] 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 essensial for inhibiting the catalytic subunit of PI3Ks.

In line with the genomic data, transcriptome analysis revealed up-regulation of several signal transduction routes associated with the PI-3-K/Akt pathway, the MAPK pathway and G protein signaling (e.g. ERK, PLC, ATK, etc). Up-regulated genes in pcALCL matched expression patterns primarily associated with overexpression of receptor kinases DDR2 and PDGFR, suggesting a potential role of these receptors in the development of the disease. The up-regulation of *DDR2* and *PDGFRB* in pcALCL was confirmed by RT-qPCR. DDR2 is overexpressed in subsets of patients with sALCL and most Hodgkin lymphoma cases.[48,49] Similarly, PDGFR is commonly overexpressed in sALCL and various other human cancers.[50,51] Of note, PDGFR blockade with imatinib has been shown to be therapeutically effective in murine models of sALCL and a patient with refractory sALCL.[50] 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.

Despite our data showed that pcALCL shares some pathogenetic features with sALCL such as the overexpression of several markers and oncogenes (e.g. *CSF1R*, *IRF4*, *MIR155HG*, *PDGFRB*, *TNFRSF8*), the deletion of *PRDM1*, and the occasional presence of specific rearrangements (e.g. *DUSP22*, *TP63*), pcALCL was devoid of genetic alterations in JAK-STAT pathway genes and showed no expression signatures suggestive of JAK-STAT pathway activation. Given that sALCL is regarded as driven by aberrant STAT3 signaling, our analysis supports the idea that pcALCL and sALCL are related but distinct entities.

In summary, our study has comprehensively delineated the genetic landscape of pcALCL for the first time 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 of DSB repair deficiency and AID hypermutation. Thirdly, athough pathogenetic features of pcALCL and sALCL partially overlapped, pcALCL lacked genetic alterations associated with

deregulation of JAK-STAT signaling which characterize sALCL. Finally, pcALCL appears to develop as a result of defects in the cell cycle, T-cell physiology, transcription, and especially signaling via the PI-3-K/Akt pathway, the MAPK pathway and G proteins.

Acknowledgements
The authors thank Julian van Toledo for valuable bioinformatic support. 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).

Conflict of interest
The authors declare no conflicts of interest.

Author contributions
A.N.B.T., R.M., K.Q. and C.P.T. designed the project. A.N.B.T., R.M. and C.P.T wrote the manuscript. A.N.B.T., L.G., and H.M. performed the bioinformatic analyses. A.N.B.T. performed the experiments. A.N.B.T. and R.M. analyzed the results, interpreted the data and produced figures and tables. K.Q., R.W. and M.V. provided valuable biological specimens. A.N.B.T., R.M., L.G., H.M., K.Q., R.W., M.V. and C.P.T. revised and approved the final manuscript.

REFERENCES

1. Willemze R, Jaffe ES, Burg G, et al. WHO-EORTC classification for cutaneous lymphomas. *Blood*. 2005;105(10):3768-3785.
2. Bekkenk MW, Geelen FA, van Voorst Vader PC, et al. Primary and secondary cutaneous CD30(+) lymphoproliferative disorders: a report from the Dutch Cutaneous Lymphoma Group on the long-term follow-up data of 219 patients and guidelines for diagnosis and treatment. *Blood*. 2000;95(12):3653-3661.
3. Melchers RC, Willemze R, Vermaat JSP, et al. Outcomes of rare patients with a primary cutaneous CD30+ lymphoproliferative disorder developing extracutaneous disease. *Blood*. 2020;135(10):769-773.
4. Velusamy T, Kiel MJ, Sahasrabuddhe AA, et al. A novel recurrent NPM1-TYK2 gene fusion in cutaneous CD30-positive lymphoproliferative disorders. *Blood*. 2014;124(25):3768-3771.
5. Crescenzo R, Abate F, Lasorsa E, et al. Convergent mutations and kinase fusions lead to oncogenic STAT3 activation in anaplastic large cell lymphoma. *Cancer Cell*. 2015;27(4):516-532.
6. Hu G, Dasari S, Asmann YW, et al. Targetable fusions of the FRK tyrosine kinase in ALK-negative anaplastic large cell lymphoma. *Leukemia*. 2017;32(2):565-569.
7. Luchtel RA, Dasari S, Oishi N, et al. Molecular profiling reveals immunogenic cues in anaplastic large cell lymphomas with DUSP22 rearrangements. *Blood*. 2018;132(13):1386-1398.
8. Luchtel RA, Zimmermann MT, Hu G, et al. Recurrent MSC (E116K) mutations in ALK-negative anaplastic large cell lymphoma. *Blood*. 2019;133(26):2776-2789.
9. Prutsch N, Gurnhofer E, Suske T, et al. Dependency on the TYK2/STAT1/MCL1 axis in anaplastic large cell lymphoma. *Leukemia*. 2019;33(3):696-709.
10. Prieto-Torres L, Rodriguez-Pinilla SM, Onaindia A, Ara M, Requena L, Piris MA. CD30-positive primary cutaneous lymphoproliferative disorders: molecular alterations and targeted therapies. *Haematologica*. 2019;104(2):226-235.
11. Melchers RC, Willemze R, van de Loo M, et al. Clinical, Histologic, and Molecular Characteristics of Anaplastic Lymphoma Kinase-positive Primary Cutaneous Anaplastic Large Cell Lymphoma. *The American Journal of Surgical Pathology*. 9000;Publish Ahead of Print.
12. van Kester MS, Tensen CP, Vermeer MH, et al. Cutaneous anaplastic large cell lymphoma and peripheral T-cell lymphoma NOS show distinct chromosomal alterations and differential expression of chemokine receptors and apoptosis regulators. *J Invest Dermatol*. 2010;130(2):563-575.
13. Laharanne E, Oumouhou N, Bonnet F, et al. Genome-wide analysis of cutaneous T-cell lymphomas identifies three clinically relevant classes. *J Invest Dermatol*. 2010;130(6):1707-1718.
14. Wada DA, Law ME, Hsi ED, et al. Specificity of IRF4 translocations for primary cutaneous anaplastic large cell lymphoma: a multicenter study of 204 skin biopsies. *Mod Pathol*. 2011;24(4):596-605.
15. Yi S, Sun J, Qiu L, et al. Dual Role of EZH2 in Cutaneous Anaplastic Large Cell Lymphoma: Promoting Tumor Cell Survival and Regulating Tumor Microenvironment. *J Invest Dermatol*. 2018;138(5):1126-1136.
16. Benner MF, Ballabio E, van Kester MS, et al. Primary cutaneous anaplastic large cell lymphoma shows a distinct miRNA expression profile and reveals differences from tumor-stage mycosis fungoides. *Exp Dermatol*. 2012;21(8):632-634.
17. Sandoval J, Diaz-Lagares A, Salgado R, et al. MicroRNA expression profiling and DNA methylation signature for deregulated microRNA in cutaneous T-cell lymphoma. *J Invest Dermatol*. 2015;135(4):1128-1137.
18. Repana D, Nulsen J, Dressler L, et al. The Network of Cancer Genes (NCG): a comprehensive catalogue of known and candidate cancer genes from cancer sequencing screens. *Genome Biol*. 2019;20(1):1.

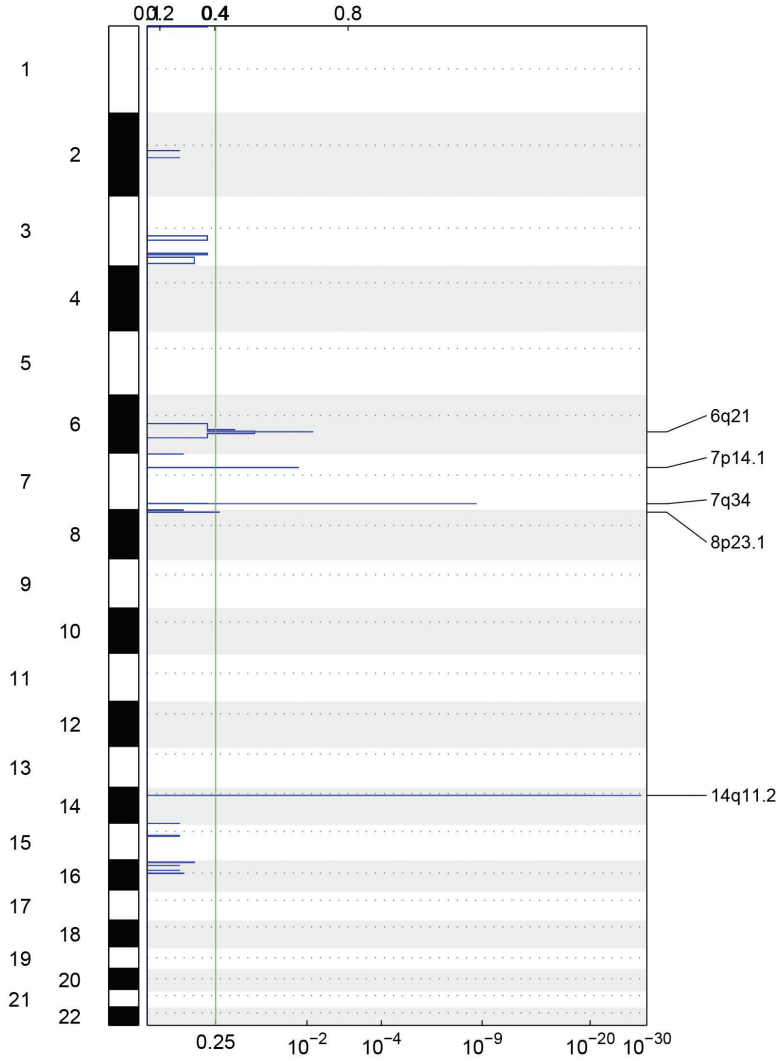
19. Ranzani V, Rossetti G, Panzeri I, et al. The long intergenic noncoding RNA landscape of human lymphocytes highlights the regulation of T cell differentiation by linc-MAF-4. *Nat Immunol*. 2015;16(3):318-325.
20. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc*. 2009;4(1):44-57.
21. Ben-Ari Fuchs S, Lieder I, Stelzer G, et al. GeneAnalytics: An Integrative Gene Set Analysis Tool for Next Generation Sequencing, RNAseq and Microarray Data. *OMICS*. 2016;20(3):139-151.
22. Kuleshov MV, Jones MR, Rouillard AD, et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res*. 2016;44(W1):W90-97.
23. Marin-Bejar O, Mas AM, Gonzalez J, et al. The human lncRNA LINC-PINT inhibits tumor cell invasion through a highly conserved sequence element. *Genome Biol*. 2017;18(1):202.
24. Carreira-Rosario A, Bhargava V, Hillebrand J, Kollipara RK, Ramaswami M, Buszczak M. Repression of Pumilio Protein Expression by Rbfox1 Promotes Germ Cell Differentiation. *Dev Cell*. 2016;36(5):562-571.
25. Spain BH, Bowdish KS, Pacal AR, et al. Two human cDNAs, including a homolog of Arabidopsis FUS6 (COP11), suppress G-protein- and mitogen-activated protein kinase-mediated signal transduction in yeast and mammalian cells. *Mol Cell Biol*. 1996;16(12):6698-6706.
26. Hoare S, Hoare K, Reinhard MK, Lee YJ, Oh SP, May WS, Jr. Tnk1/Kos1 knockout mice develop spontaneous tumors. *Cancer Res*. 2008;68(21):8723-8732.
27. Gu TL, Cherry J, Tucker M, Wu J, Reeves C, Polakiewicz RD. Identification of activated Tnk1 kinase in Hodgkin's lymphoma. *Leukemia*. 2010;24(4):861-865.
28. Martins G, Calame K. Regulation and functions of Blimp-1 in T and B lymphocytes. *Annu Rev Immunol*. 2008;26:133-169.
29. Ling P, Lu TJ, Yuan CJ, Lai MD. Biosignaling of mammalian Ste20-related kinases. *Cell Signal*. 2008;20(7):1237-1247.
30. Steinberg MW, Cheung TC, Ware CF. The signaling networks of the herpesvirus entry mediator (TNFRSF14) in immune regulation. *Immunol Rev*. 2011;244(1):169-187.
31. Kim J, McMillan E, Kim HS, et al. XPO1-dependent nuclear export is a druggable vulnerability in KRAS-mutant lung cancer. *Nature*. 2016;538(7623):114-117.
32. Mathas S, Kreher S, Meaburn KJ, et al. Gene deregulation and spatial genome reorganization near breakpoints prior to formation of translocations in anaplastic large cell lymphoma. *Proc Natl Acad Sci U S A*. 2009;106(14):5831-5836.
33. Schrader AM, Chung YY, Jansen PM, et al. No TP63 rearrangements in a selected group of primary cutaneous CD30+ lymphoproliferative disorders with aggressive clinical course. *Blood*. 2016;128(1):141-143.
34. Boi M, Rinaldi A, Kwee I, et al. PRDM1/BLIMP1 is commonly inactivated in anaplastic large T-cell lymphoma. *Blood*. 2013;122(15):2683-2693.
35. Zhang M, Zhao K, Xu X, et al. A peptide encoded by circular form of LINC-PINT suppresses oncogenic transcriptional elongation in glioblastoma. *Nat Commun*. 2018;9(1):4475.
36. Feng H, Zhang J, Shi Y, Wang L, Zhang C, Wu L. Long noncoding RNA LINC-PINT is inhibited in gastric cancer and predicts poor survival. *J Cell Biochem*. 2019;120(6):9594-9600.
37. Hu J, Ho AL, Yuan L, et al. Neutralization of terminal differentiation in gliomagenesis. *Proceedings of the National Academy of Sciences*. 2013;110(36):14520-14527.
38. Sengupta N, Yau C, Sakthianandeswaren A, et al. Analysis of colorectal cancers in British Bangladeshi identifies early onset, frequent mucinous histotype and a high prevalence of RBFOX1 deletion. *Mol Cancer*. 2013;12:1.
39. Lu YJ, Wu CS, Li HP, et al. Aberrant methylation impairs low density lipoprotein receptor-related protein 1B tumor suppressor function in gastric cancer. *Genes Chromosomes Cancer*. 2010;49(5):412-424.
40. Ni S, Hu J, Duan Y, et al. Down expression of LRP1B promotes cell migration via RhoA/Cdc42 pathway and actin cytoskeleton remodeling in renal cell cancer. *Cancer Sci*. 2013;104(7):817-825.

41. Prazeres H, Torres J, Rodrigues F, et al. Chromosomal, epigenetic and microRNA-mediated inactivation of LRP1B, a modulator of the extracellular environment of thyroid cancer cells. *Oncogene*. 2011;30(11):1302-1317.
42. Cancer Genome Atlas N. Comprehensive molecular portraits of human breast tumours. *Nature*. 2012;490(7418):61-70.
43. Huang XD, Xiao FJ, Wang SX, et al. G protein pathway suppressor 2 (GPS2) acts as a tumor suppressor in liposarcoma. *Tumour Biol*. 2016;37(10):13333-13343.
44. May WS, Hoare K, Hoare S, Reinhard MK, Lee YJ, Oh SP. Tnk1/Kos1: a novel tumor suppressor. *Trans Am Clin Climatol Assoc*. 2010;121:281-292; discussion 292-283.
45. Gagliardi PA, Puliafito A, Primo L. PDK1: At the crossroad of cancer signaling pathways. *Semin Cancer Biol*. 2018;48:27-35.
46. Ziemba BP, Pilling C, Calleja V, Larijani B, Falke JJ. The PH domain of phosphoinositide-dependent kinase-1 exhibits a novel, phospho-regulated monomer-dimer equilibrium with important implications for kinase domain activation: single-molecule and ensemble studies. *Biochemistry*. 2013;52(28):4820-4829.
47. Lucas CL, Zhang Y, Venida A, et al. Heterozygous splice mutation in PIK3R1 causes human immunodeficiency with lymphoproliferation due to dominant activation of PI3K. *J Exp Med*. 2014;211(13):2537-2547.
48. Willenbrock K, Kuppers R, Renne C, et al. Common features and differences in the transcriptome of large cell anaplastic lymphoma and classical Hodgkin's lymphoma. *Haematologica*. 2006;91(5):596-604.
49. Renne C, Willenbrock K, Kuppers R, Hansmann ML, Brauninger A. Autocrine- and paracrine-activated receptor tyrosine kinases in classic Hodgkin lymphoma. *Blood*. 2005;105(10):4051-4059.
50. Laimer D, Dolznig H, Kollmann K, et al. PDGFR blockade is a rational and effective therapy for NPM-ALK-driven lymphomas. *Nat Med*. 2012;18(11):1699-1704.
51. Papadopoulos N, Lennartsson J. The PDGF/PDGFR pathway as a drug target. *Mol Aspects Med*. 2018;62:75-88.

SUPPLEMENTARY FIGURES

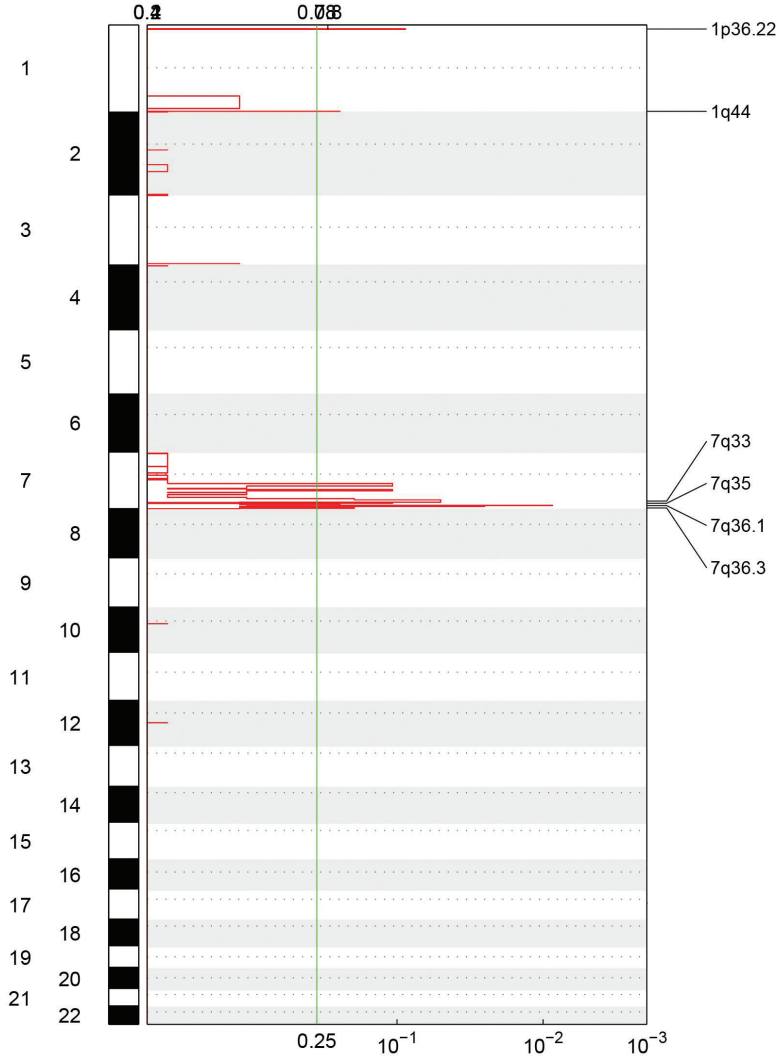
Supplementary Figure S1. Common focal CNAs in pcALCL detected by GISTIC 2.0. analysis.

A)



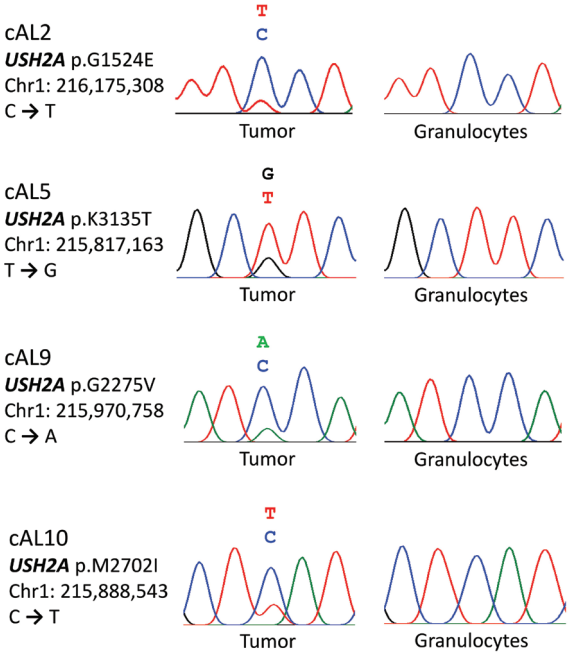
Common focal losses in pcALCL detected by GISTIC 2.0. analysis. GISTIC 2.0 detected common losses shared by pcALCL patients within chromosomes 6, 7, 8 and 14. Peaks at 6q21, 7p14.1, 7q34 and 14q11.2 were all ≤ 3Mb long (as verified with IGV). The peak at 6q21 enclosed tumor suppressor *PRDM1* (*BLIMP1*). Peaks at 7p14.1, 7q34 and 14q11.2 contained loci encoding the α (TRA), β (TRB) and γ (TRG) chains of the T-cell receptor (TCR), which are rearranged during T-cell development. Loss of the exact same region within *TRA*, *TRB* and *TRG* in tumor DNA from patients suggests that pcALCL is a clonal disease.

B)

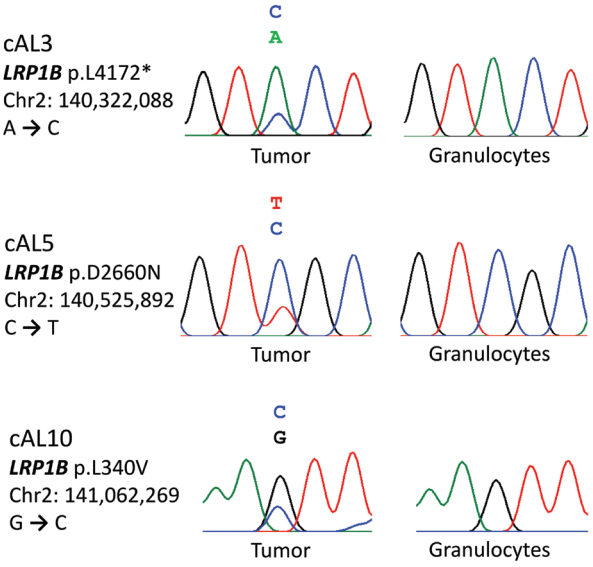


Common focal gains in pcALCL detected by GISTIC 2.0. analysis. GISTIC 2.0 detected common gains between pcAECyTCL patients within chromosomes 1 and 7. The peak at 1p36.22 contained *TNFRSF8* (CD30), whose overexpression is a hallmark of pcALCL. Except for 7q36.1, which enclosed *EZH2*, peaks within chromosome 7 were either > 3Mb long or devoid of (putative) cancer genes, and consequently, excluded.

Supplementary Figure S2. Sanger sequencing validation of select SNVs.
A)

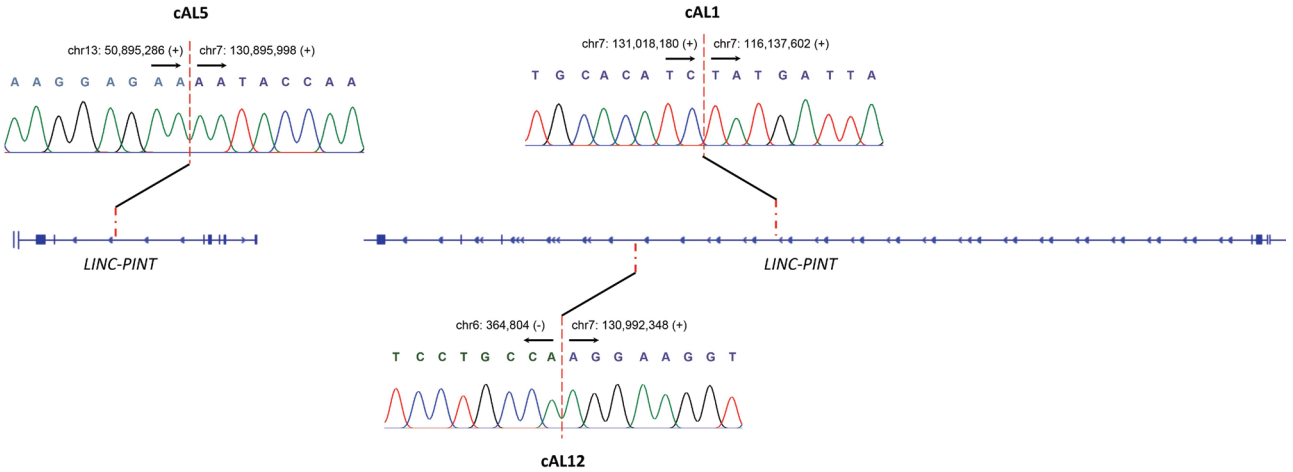


B)

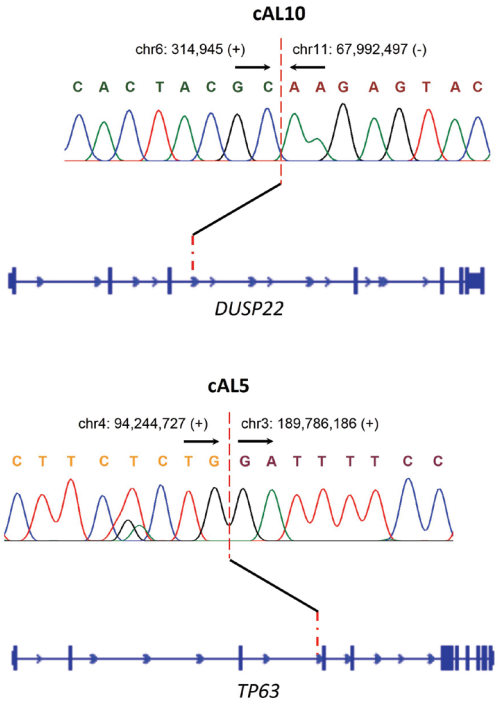


Supplementary Figure S2. Sanger chromatograms showing somatic point mutations in (A) *USH2A* and (B) *LRP1B*, the two most commonly mutated genes in pcALCL tumors, detected by WES. Genomic coordinates according to reference genome GRCh38.

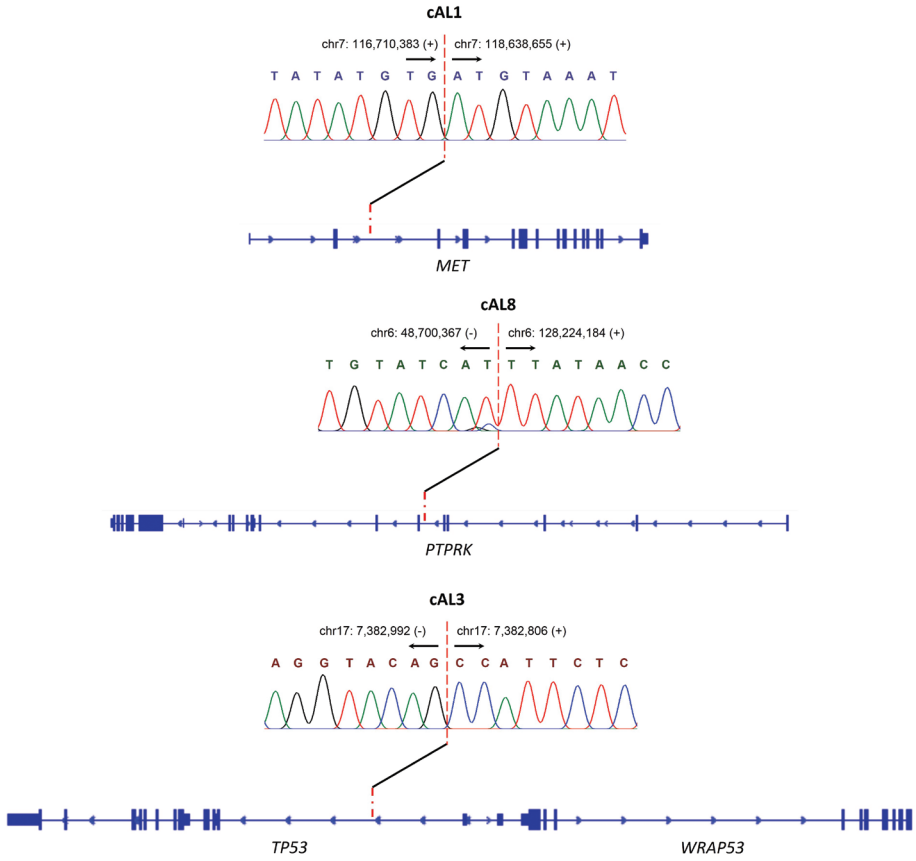
Supplementary Figure S3. Sanger sequencing validation of select genomic rearrangements.
A)



B)

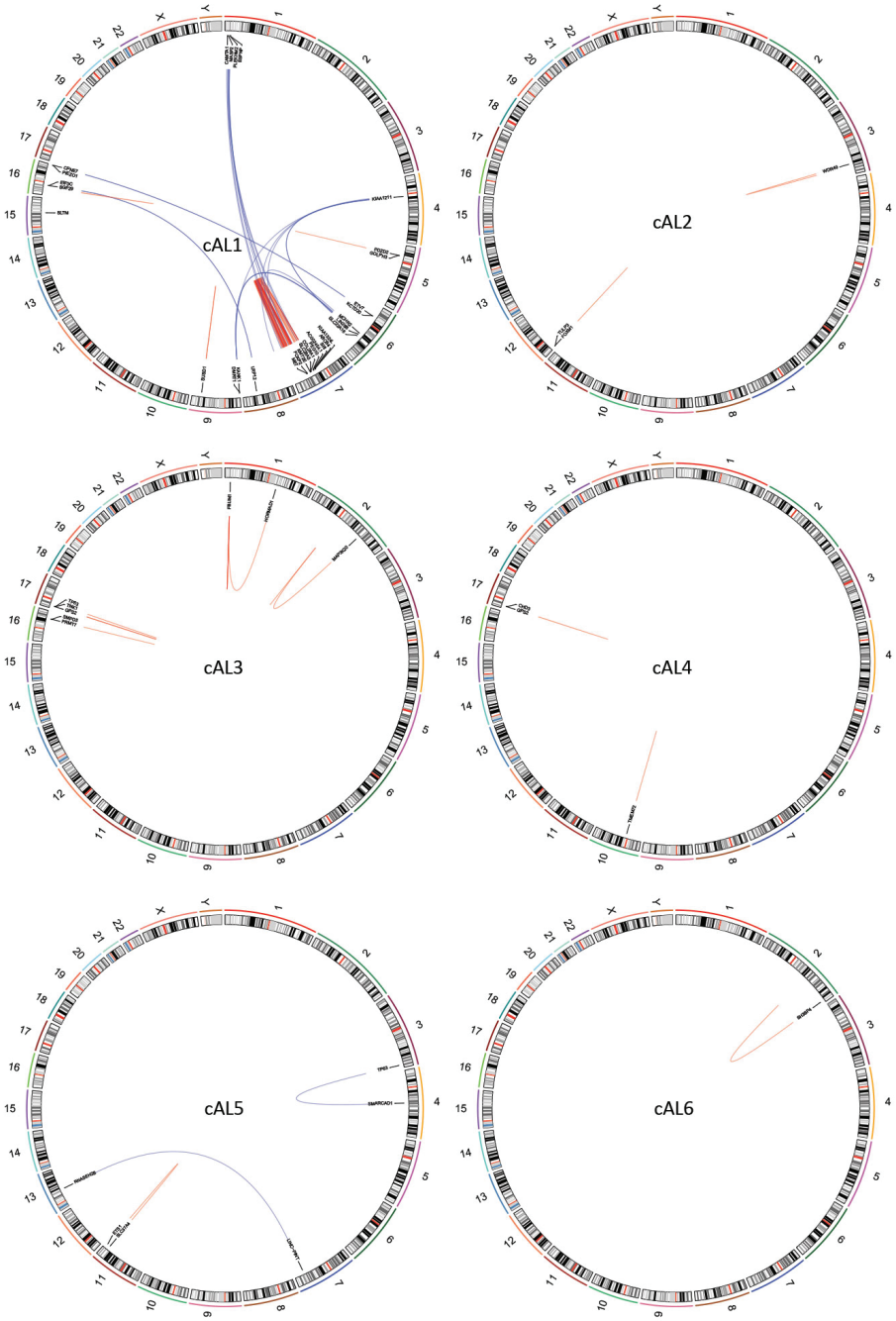

cAL10
chr6: 314,945 (+)
chr11: 67,992,497 (-)
C A C T A C G C A A G A G T A C
DUSP22
cAL5
chr4: 94,244,727 (+)
chr3: 189,786,186 (+)
C T T C T C T G G A T T T T C C
TP63

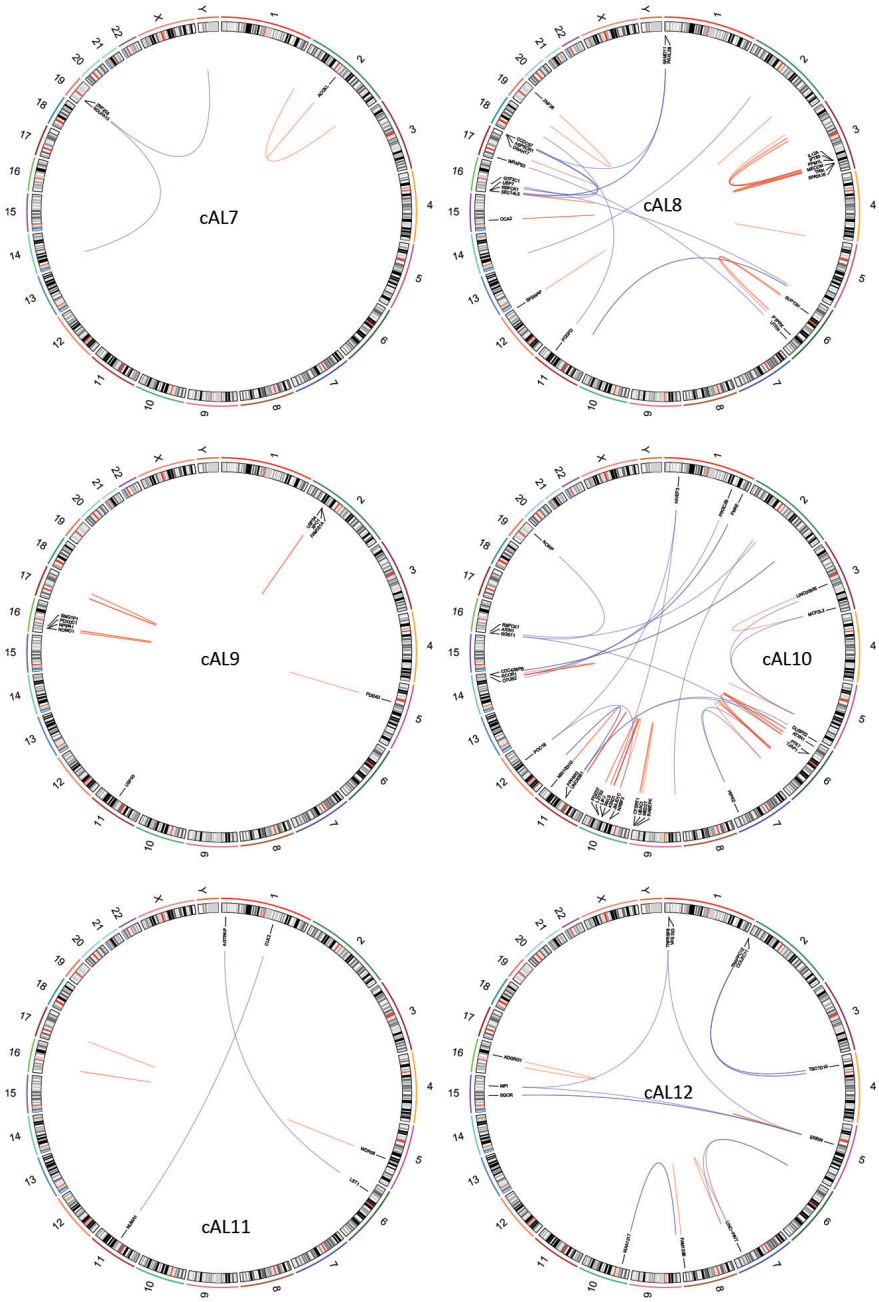
C)



Supplementary Figure S3. Sanger chromatograms showing validation of genomic rearrangements involving cancer genes detected by WGS. (A) Rearrangements at *LINC-PINT,* were the most recurrent in pcALCL. (B) Rearrangements at *DUSP22* and *TP63* were observed in solitary cases. These events are known to occur at low frequencies in pcALCL. C) Rearrangements were found to impact numerous *bona fide* cancer genes in pcALCL (e.g. *MET, PTPRK* and *TP53*). Genomic coordinates according to reference genome GRCh38.

Supplementary Figure S4. Circos plots displaying genomic rearrangements per patient.





Supplementary Figure S4. Tumors from different pcALCL patients contained a variety of inter- and intrachromosomal rearrangements. The number of events ranged from 1 (AEC6) to 51 (AEC1). Rearrangements involving genes involved in RAS signaling: *GPS2* (AEC3 and AEC4) and *TNK1* (AEC3 and AEC4) were the most prominent events in our sequenced cohort.

5

Deregulation of JAK2 signaling underlies primary cutaneous CD8+ aggressive epidermotropic cytotoxic T-cell lymphoma

Armando N. Bastidas Torres
Davy Cats
Jacoba J. Out-Luiting
Daniele Fanoni
Hailiang Mei
Luigia Venegoni
Rein Willemze
Maarten H. Vermeer
Emilio Berti
Cornelis P. Tensen

Manuscript submitted



ABSTRACT

Primary cutaneous CD8+ aggressive epidermotropic cytotoxic T-cell lymphoma (pcAECyTCL) is a rare variant of cutaneous T-cell lymphoma with an aggressive clinical course and a very poor prognosis. Until now, neither a systematic characterization of genetic alterations driving pcAECyTCL has been performed, nor effective therapeutic regimes for patients have been defined. Here, we present the first high resolution genetic characterization of pcAECyTCL by using whole-genome sequencing and RNA sequencing. Our study provides a comprehensive description of genetic alterations (i.e. genomic rearrangements, copy number alterations and small-scale mutations) with pathogenic relevance in this lymphoma, including events that recurrently impact genes with important roles in the cell cycle, chromatin regulation and the JAK-STAT pathway. In particular, we show that mutually exclusive structural alterations involving *JAK2* and *SH2B3* underlie predominantly pcAECyTCL. In line with the genomic data, transcriptome analysis uncovered up-regulation of the cell cycle, JAK2 signaling, NF-κB signaling and high inflammatory response in this cancer. Functional studies confirmed oncogenicity of *JAK2* fusions identified in pcAECyTCL and their sensitivity to JAK inhibitor treatment. Our findings strongly suggest that overactive JAK2 signaling is a central driver of pcAECyTCL, and consequently, patients with this neoplasm would likely benefit from therapy with JAK2 inhibitors such as FDA-approved ruxolitinib.

1. INTRODUCTION

Primary cutaneous CD8+ aggressive epidermotropic cytotoxic T-cell lymphoma (pcAECyTCL) is a rare variant of cutaneous T-cell lymphoma (CTCL) still regarded as a provisional entity by the World Health Organization (WHO) and characterized by an abrupt onset and a highly aggressive clinical course.[1,2] pcAECyTCL presents primarily in the skin with widespread plaques and tumors, often with hemorrhagic ulcerations and necrosis; however, dissemination to extracutaneous sites (esp. central nervous system, lungs, oral cavity and testes) is not uncommon.[3,4] Malignant T cells causing pcAECyTCL typically express CD3, CD7, CD8, CD45RA, TCR-βF1, T-BET and one or more cytotoxic markers (e.g. granzyme B, perforin, TIA-1), which strongly suggests that neoplastic cells in this lymphoma derive from CD8+ T cells.[2,3] Effective therapeutic regimes for pcAECyTCL are currently lacking, and consequently, patients have a poor prognosis with a median overall survival of 12 months.[1]

Thus far the study of the pathogenetic basis of this malignancy has been marginal due to its rarity. Recently, a study performed on tumors from 20 patients defined the copy number alteration (CNA) profile of pcAECyTCL by using array-based comparative genomic hybridization,[5] and before this, two clinical case reports included the evaluation of CNAs in single patients by using array-based methods as well.[6,7] Recurrent CNAs uncovered by these studies include losses within 1p, 9p, 13q and 16p as well as gains within 7q, 8q and 17q, with loss of the region containing *CDKN2A/B* being the most frequent CNA.[5] However, aside from the aforementioned chromosomal imbalances, causative genetic changes in pcAECyTCL remain unknown.

Here, we present the first high resolution genomic analysis of pcAECyTCL using whole-genome sequencing (WGS) and RNA sequencing (RNA-seq). We describe for the first time a number of genomic rearrangements, CNAs and small-scale mutations with pathogenic relevance in this lymphoma. In particular, our results suggest that overactivation of JAK2 signaling due to oncogenic changes in *JAK2* and *SH2B3*, two genes with key roles in this signaling pathway, underlie predominantly pcAECyTCL. These findings have important implications for patient standard of care.

2. MATERIALS AND METHODS

2.1 Patient selection

Frozen tumor biopsies (≥70% tumor cells) from 12 patients with pcAECyTCL (Table S1, Figure S1) were subjected to WGS. Six samples of this cohort (i.e. AEC2-4/6/8/12) were additionally subjected to RNA-seq. WGS and RNA-seq data have been deposited in the European Genome-Phenome Archive (EGA) under study number EGAS00001004332. Sequencing and data processing were performed as previously described[8] and details are included in the Supplement (Table S2). Diagnosis was performed by an expert panel of dermatologists/pathologists according to the criteria of the WHO-EORTC classification for primary cutaneous lymphomas.[1,2] Patient material was approved by the institutional review boards of Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico and Leiden University Medical Center. Informed consent was obtained from patients in accordance with the declaration of Helsinki.

2.2 DNA sequencing analyses

Detection of structural genomic variants (SV) was performed using Breakdancer-max (v1.4.4) and Delly (v0.6.7). Events detected in a healthy control (PBMC) processed with the same pipeline were regarded as mapping artefacts and filtered out accordingly. Rearranged genes implicated in cancer were identified using the Network of Cancer Genes 6.0[9] (NGC 6.0) and literature search. Copy number alterations (CNAs) were identified with Control-FREEC[10] (v8.7) and GISTIC2.0[11] (Figure S2, Table S3). Non-cancer structural variants were identified using the Database of Genomic Variants[12] and filtered out accordingly. SV and CNAs were visually verified/curated using the Integrative Genomic Viewer[13] (IGV, v2.3.78) (Table S4-S5). GATK (v3.5) was used to detect deleterious indels/SNVs (Table S6). Indels/SNVs were investigated in exonic sequences of 1461 genes taking part in cellular processes/pathways previously identified as mutated in CTCL (Table S7).[14,15] Variant allele frequency of ≥0.1 and ≤0.9 and variant read coverage of ≥2 were applied. Non-cancer indels/SNVs were eliminated by filtering out variants with >5% minor allele frequency in ExAC and variants detected in exomes of an in-house cohort formed by eleven healthy volunteers. Indels/SNVs predicted as benign/tolerated (SIFT and PolyPhen-2) and not flagged on COSMIC/ClinVar were excluded. Indels/SNVs predicted as deleterious and/or flagged as pathogenic on COSMIC/ClinVar were further investigated on Varsome and literature.

2.3 RNA sequencing analyses

EdgeR was used to apply TMM normalization on raw data. Limma-Voom was used to perform differential expression analysis. pcAECyTCL samples were compared to a control group formed by skin-resident CCR8+/CD8+ T cells.

RNA-seq data of CCR8+/CD8+ T cell controls[16] were downloaded from SRA (https://www.ncbi.nlm.nih.gov/sra; ERP105873). Genes with FDR <0.01 were considered differentially expressed (DE) (Table S8). Fusion transcripts were investigated in six pcAECyTCL samples with available RNA-seq data using FusionCatcher (v0.99.6a) and Star Fusion (v0.8.0). Fusion transcript calls were manually verified/curated using IGV (Table S9). Gene set enrichment analysis (GSEA, v2.2.4) was run using gene sets from the Molecular Signatures Database (MSigDB).[17] FDR *q* value threshold of ≤ 0.25, recommended by the developer, was used. Transcriptome analysis using data from ProteomicsDB[18] was performed with Enrichr.[19] Gene ontology annotation of rearranged genes was performed using Panther (v12.1).[20]

2.4 Validation of structural genomic alterations and small-scale mutations

Select rearrangements, interstitial deletions and SNVs were validated by Sanger sequencing on the ABI3730xl platform (Applied Biosystems). Select CNAs were validated by digital droplet PCR (ddPCR) using QX200 ddPCR system (Bio-Rad).

2.5 Fluorescent *in situ* hybridization

Break-apart fluorescent *in situ* hybridization (FISH) for *JAK2* was performed on formalin-fixed paraffin embedded (FFPE) tumor tissue as previously described.[21] Probe RP11-1151G4 (telomeric/5′) and probes RP11-140C18 and RP11-635N21 (centromeric/3′) were used to detect DNA breaks at *JAK2* (Figure S3, Table S10).

2.6 Methylation specific melting curve analysis

Methylation specific melting curve analysis (MS-MCA) was performed as previously described[22] on bisulfite-converted genomic DNA using primers directed at CpG islands within the promoter and gene body of *SH2B3* (Figure S4). Bisulfite converted CpGenome Universal Methylated DNA (Chemicon, Hampshire, United Kingdom) and unmethylated male DNA were used as positive and negative controls, respectively.

2.7 Immunohistochemistry

FFPE tissue sections were immunohistochemically stained with primary antibodies against phospho-STAT3 (Cell Signaling Technology, Cat.No. 9145) or phospho-STAT5 (Cell Signaling Technology, Cat.No. 9359) using Dako REAL detect system (Dako, Cat.No. K5005), and subsequently, counterstained in Mayer's Hematoxylin solution and coverslipped using Vectamount (Vector Laboratories, code H5000).

2.8 Cell culture, fusion gene construction and viral transduction

Ba/F3 cells (DSMZ, Cat.No. ACC-300, murine pro-B cells) were used for functional experiments given the inexistence of pcAECyTCL cell lines and unavailability of patient-derived cells. Parental Ba/F3 cells were cultured in RPMI-1640 (10% h.i. FBS, 10 ng/mL IL3) at 37° C with 5% CO_2 in a humidified atmosphere. Fusion genes were constructed and inserted into a lentiviral vector using the method described by Lu et al.[23] Lentiviral particles were produced in HEK 293T cells, quantified by p24 ELISA, and transduced into Ba/F3 cells at MOI 9 with lipofectamine. Successfully transduced Ba/F3 cells were selected with puromycin (2.5 µg/mL) for 3 days.

2.9 Ba/F3 cell viability and inhibitor assays

Cell viability of parental and transduced Ba/F3 cells was determined 7 days after IL3 withdrawal by MTT assay (Promega, Cat.No. G4000). Inhibitor assays were performed by treating IL3-independent Ba/F3 cells expressing fusion genes with ruxolitinib or AZD1480 at seven different concentrations for 72h and measuring cell viability by MTT assay.

2.10 Western Blots

The effect of JAK1/2 inhibitors Ruxolitinib and AZD1480 on JAK2 and STAT5 phosphorylation was evaluated by western blot. Cells were washed to remove traces of serum and incubated with inhibitor for 90 min. Cells were lysed in SDS lysis buffer containing protease inhibitors and separated by SDS-PAGE. Antibodies employed were anti-JAK2 (Abcam, Cat.No. ab108596), anti-phospho-JAK2 (Cell Signaling Technology, Cat.No. 3776), anti-STAT5 (Cell Signaling Technology, Cat.No. 94205), anti-phospho-STAT5 (Cell Signaling Technology, Cat.No. 9351) and anti-GAPDH (Cell Signaling Technology, Cat.No. 2118).

3. RESULTS

3.1 *JAK2 f*usions are prominent in a complex landscape of rearrangements

The analysis revealed a heterogeneous and complex landscape of genomic rearrangements (total events, 426; range, 10-65; mean/patient ± SD, 36±21) (Figure 1, Figure 2A, Figure S5). Fifty-three percent of events were interchromosomal (range/patient, 27%-80%) (Figure 2B). The majority of rearrangements (77%) disrupted either one or two annotated genes, while the rest (23%) disrupted nongenic regions (Figure 2C). Four patients, AEC6, AEC7, AEC10 and AEC12, displayed complex rearrangements (chromothripsis/chromoplexy-like) affecting chromosomes 13, 10, 1/9/12 and 4, respectively (Figure 2D, Figure S5-S6). We observed a total of 305 rearranged genes, 59 of which are implicated in neoplasms at present (Table S11). Gene ontology analysis revealed that rearranged genes

encode principally (n_{genes} = 91/305) proteins with roles in signal transduction (i.e. hydrolases, transferases, enzyme modulators, receptors) and transcriptional regulation (i.e. transcription factors, chromatin regulators) (Figure 2E, Table S12-S13). Out of seventeen recurrently rearranged genes detected in our cohort ($n_{patients}$ = 2 or 3) (Table S14), six are established cancer genes with important functions in the regulation of the cell cycle (i.e. *MYC*, *RB1*), chromatin remodeling (i.e. *BAZ1A*) and the JAK-STAT pathway (i.e. *JAK2*, *PTPRC*, *SH2B3*) (Figure 1, Figure S7).

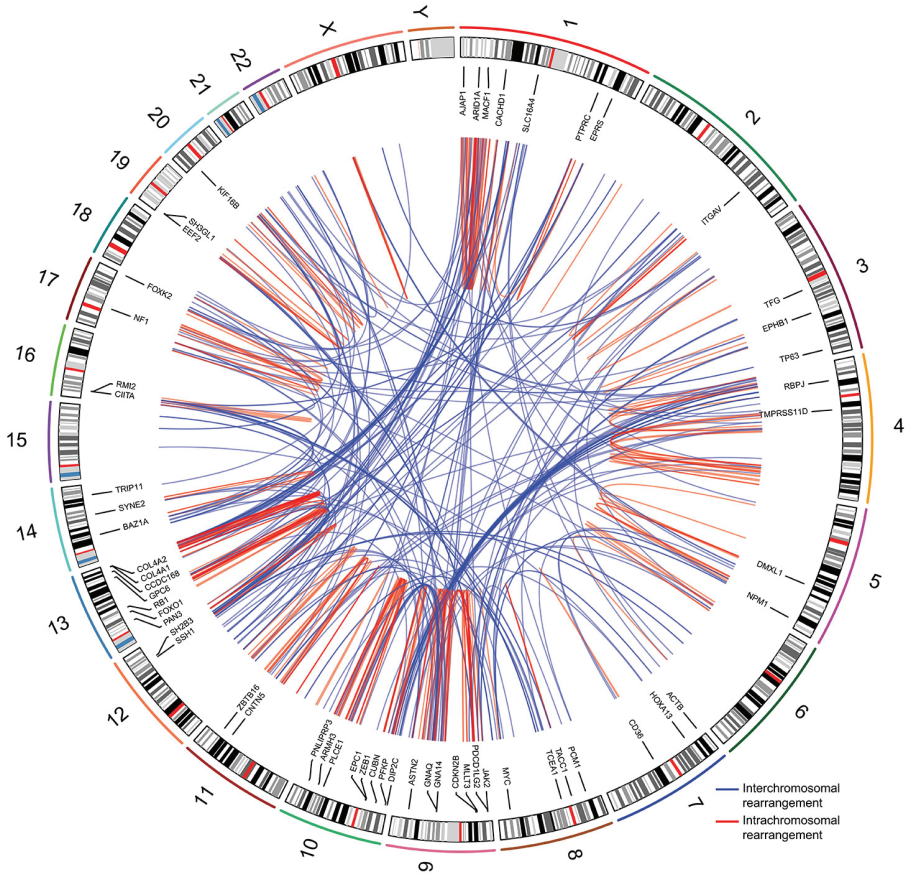


Figure 1. Landscape of genomic rearrangements in pcAECyTCL. Circos plot showing 426 genomic rearrangements detected in twelve pcAECyTCL genomes by WGS. The outer ring is formed by human chromosome ideograms arranged circularly end to end. The area at the center of the plot contains arcs representing interchromosomal (blue) and intrachromosomal (red) events. The ring between the chromosome ideograms and the arcs contains rearranged genes with established roles in cancer.

The JAK-STAT pathway, a frequent driver of hematological neoplasms, was the only cytokine-elicited signal transduction pathway impacted by rearrangements in pcAECyTCL. Fusion genes involving *JAK2* were detected in three of twelve patients (i.e. AEC4: *KHDRBS1-JAK2*; AEC9: *PCM1-JAK2*; AEC11: *TFG-JAK2*) (Figure 2F-H). These events fused the tyrosine kinase domain of JAK2 with one or more oligo/dimerization domains from the fusion partner (i.e. AEC4: Qua1 domain, AEC9: coiled-coil domains, AEC11: PB1 domain) (Figure 2F-H). The resulting chimeric proteins are predicted to self-oligo/dimerize and become activated without the need of cytokine-mediated receptor stimulation, ultimately over-activating JAK2 signaling. Of note, two of three patients carrying *JAK2* fusions carried *MYC* fusions as well (i.e. AEC4: *ACTB-MYC*, AEC9: *NPM1-MYC*) (Figure S7). Interestingly, apart from acquiring the ability to self-activate, *JAK2* fusions under the transcriptional control of their partner's promoter may also experience augmented expression in comparison to wild type *JAK2*, as evidenced in patient AEC4 (Figure 2I). In contrast, rearrangements involving *PTPRC* and *SH2B3*, each observed in two of twelve patients, disrupted these two negative regulators of the JAK-STAT pathway.

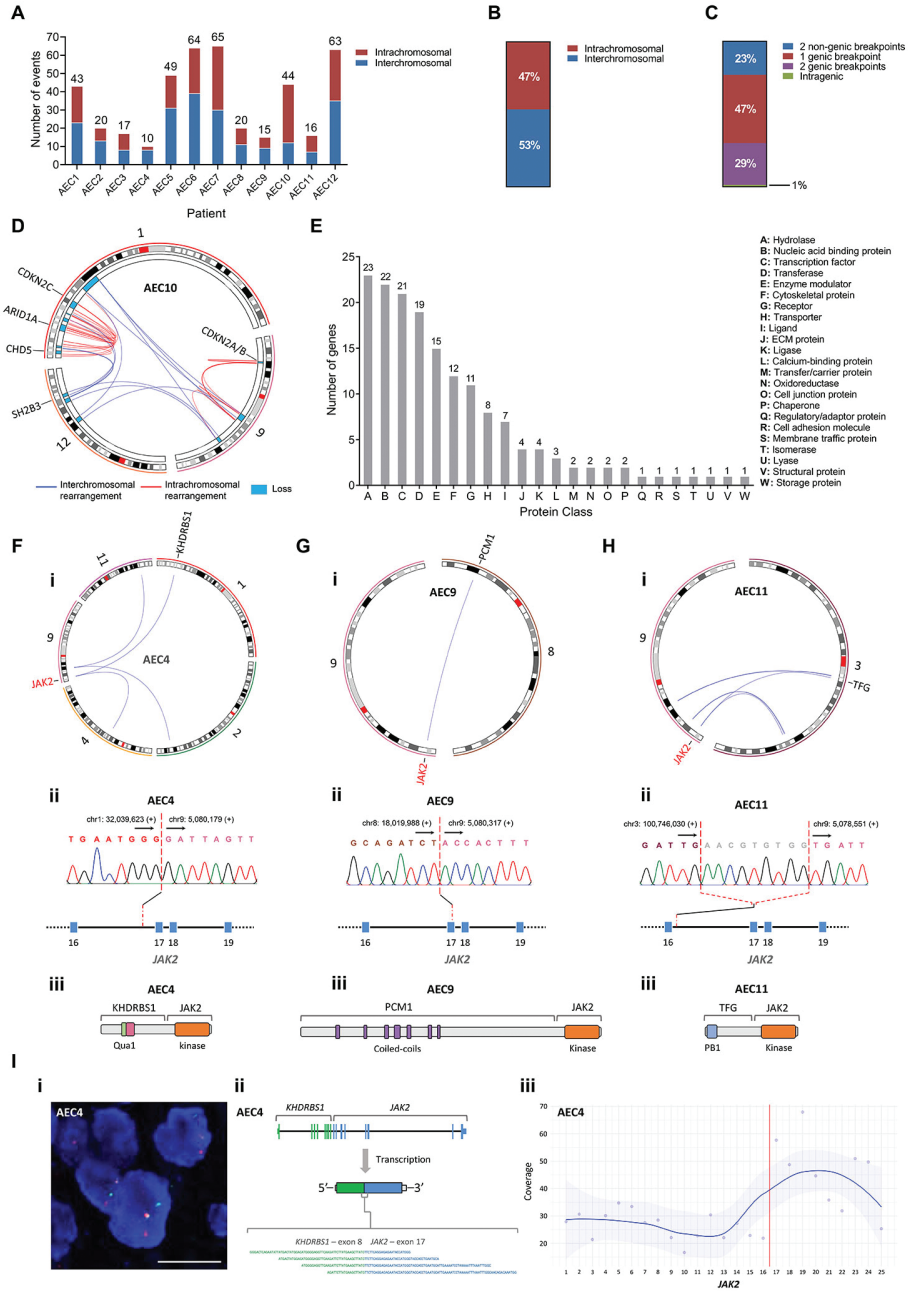


Figure 2. *JAK2* fusions are recurrent in a complex landscape of rearrangements. (A) Number of genomic rearrangements per patient. The distribution of inter- and intrachromosomal rearrangements per patient is shown too; (B) Distribution of inter- and intrach-

romosomal rearrangements (cohort); (C) Distribution of genomic rearrangements based on the type of DNA sequences (genic, nongenic) involved in the event (cohort); (D) Circos plot showing a chromoplexy-like event in patient AEC10 that mediated the loss of multiple genomic regions in chromosomes 1, 9 and 12, several of which enclosed established tumor suppressor genes; (E) Distribution of rearranged genes according to the protein class their encoded protein participates in; (F, G, H) Genomic rearrangements generated self-activating *JAK2* fusions in pcAECyTCL as evidenced in patients (F) AEC4, (G) AEC9 and (H) AEC11. (i) Circos plots showing interchromosomal rearrangements involving chromosome 9 in patients with pcAECyTCL. *JAK2* rearrangements was the common denominator between chromosome 9 events observed in these individuals. (ii) Validation of translocation breakpoints at *JAK2* by Sanger sequencing in patients with pcAECyTCL. Breakpoints occurred between exon 16 and exon 17 in all cases. (iii) Rearrangements involving *JAK2* led to the formation of fusion genes composed by the tyrosine kinase domain of *JAK2* and the oligo/dimerization domains of the fusion partners (*KHDRBS1*: Qua1 domain, *PCM1*: coiled-coil domains, *TFG*: PB1 domain), conferring the resulting chimeric protein the ability to self-activate. (F) In addition to acquiring self-activation ability, *JAK2* fusions can also experience increased expression in comparison to wild-type *JAK2*. (i) Image of break-apart FISH analysis showing a *JAK2* rearrangement in patient AEC4. Scale bar, 10 µm (ii) Active expression of fusion gene *KHDRBS1-JAK2* in patient AEC4 was detected by RNA-seq (chimeric reads shown in diagram). (iii) Plot showing mean read coverage across all exons of *JAK2* in patient AEC4. RNA expression between exon 17 and exon 25, the part of *JAK2* under the transcriptional control of *KHDRBS1*'s promoter and encoding its tyrosine kinase domain, is considerably higher compared to RNA expression between exon 1 and exon 16. The red line indicates the breakpoint site.

3.2 JAK2 signaling inhibitor *SH2B3* is focally deleted in pcAECyTCL

The most frequent broad chromosomal imbalances ($n_{patients}$ ≥4; >3 Mb) were deletions within 1p, 8p, 9q, 10p, 11q and 13q and gains within 7q, 8q, 17q and 21q (Figure 3A, Figure 4). We identified 24 recurrent focal (≤ 3 Mb) minimal common regions (MCRs) shared by CNAs between patients ($n_{patients}$ ≥ 3; deletions: 19, gains: 5) (Figure 4, Table S5), 12 of which contained cancer genes predominantly involved in the cell cycle, chromatin regulation and the JAK-STAT pathway (Figure 4).

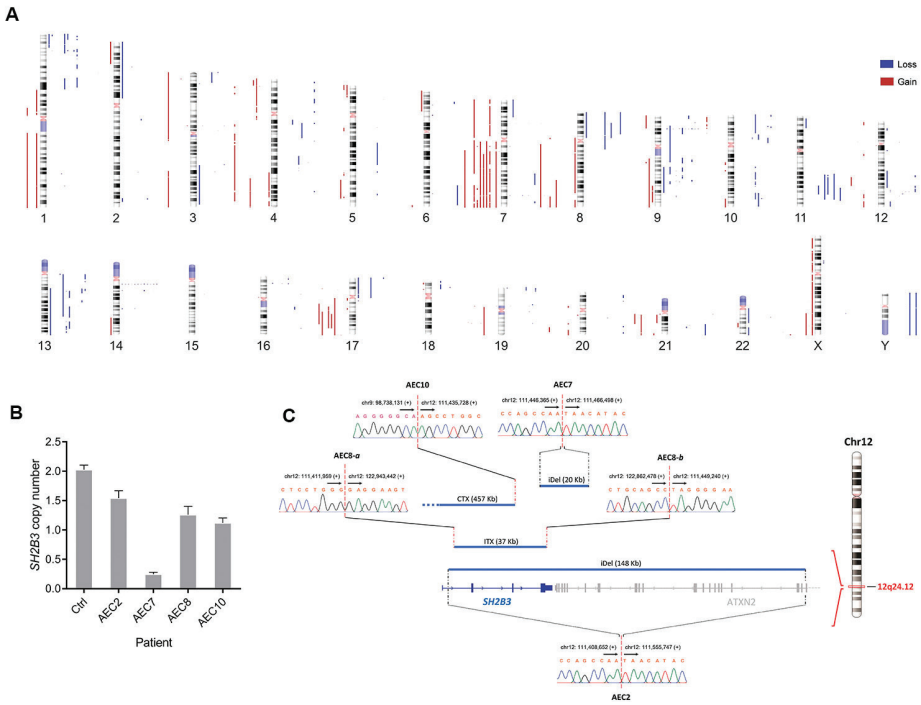


Figure 3. Landscape of CNAs in pcAECyTCL. (A) Human chromosome ideograms showing regions of gain and loss detected through WGS in twelve pcAECyTCL genomes. Blue bars to the right of the chromosomes depict regions of loss whereas red bars to the left of the chromosomes depict regions of gain. Deletions at 12q24.12, where *SH2B3* resides, were the most focal (<500 Kb) CNA events in pcAECyTCL. (B) Copy number losses involving *SH2B3* in patients with pcAECyTCL were validated by ddPCR. Ctrl, CD8+ T cells. (C) Losses at 12q24.12, which involved *SH2B3*, were mediated by interstitial deletions and unbalanced rearrangements. Breakpoints of structural alterations at 12q24.12 in all affected patients were validated by Sanger sequencing. Genomic coordinates of breakpoints according to reference genome GRCh38. Arrows indicate the direction towards which genomic coordinate numbers increase. Plus (+) and minus (-) signs specify strand polarity. iDel, interstitial deletion; ITX, intrachromosomal rearrangement; CTX, interchromosomal rearrangement.

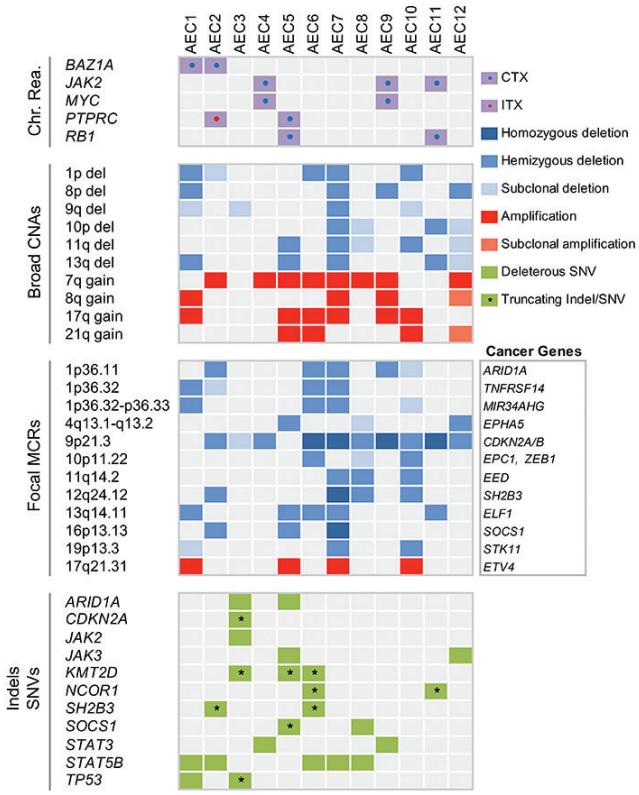


Figure 4. Distribution of recurrent chromosomal rearrangements, CNAs and deleterious indels/SNVs in pcAECyTCL. First panel: chromosomal rearrangements impacting cancer genes. Second panel: large-scale CNAs (> 3 Mb). Third panel: focal MCRs (≤ 3 Mb) shared by CNAs; *bona fide* cancer genes residing within focal MCRs are specified. Fourth panel: Indels and SNVs in cancer genes either leading to protein truncations, reported as pathogenic in literature or predicted as disease-causing (SIFT and PolyPhen-2) are shown. Only genes altered in more than one patient are indicated.

The most common focal MCR involving cancer genes was deletion at 9p21.3 (ten patients), which included cell cycle regulators *CDKN2A/B* (Figure S8). Five of twelve patients had deletions at 1p36.11 and 13q14.11, which contained chromatin remodeler *ARID1A* and candidate cancer gene *ELF1*,[24] respectively. Deletions at 1p36.32-p36.33, 1p36.22 and 12q24.12, observed in four of twelve patients, involved tumor suppressors *TNFRSF14*, *MIR34AHG* and *SH2B3*, respectively. Finally, three of twelve patients had deletions at 4q13.1-q13.2, 10p11.22, 11q14.2, 16p13.13 and 19p13.3, which contained tumor suppressors *EPHA5*, *EPC1* (alongside *ZEB1*), *EED*, *SOCS1* and *STK11*, respectively. On the other hand, gain at 17q21.31 (four patients), which enclosed *ETV4*, was the only recurrent ($n_{patients} \geq 3$) focal gain containing a cancer gene.

Remarkably, deletions at 12q24.12 were strikingly focal in all affected patients (20 Kb – 457 Kb), leading to the loss of one or more functional domains of SH2B3 (i.e. DD, PH, SH2 domains) in these individuals (Figure 3B-C, Figure S9). *SH2B3* (LNK) encodes an adaptor protein that antagonizes JAK2 signaling as part of a negative feedback loop in various hematopoietic cell types (e.g. erythroid progenitors, hematopoietic stem cells, megakaryocytes, pre-B cells, etc.) by suppressing the kinase activity of JAK2 through its SH2 domain.[25] Of note, structural alterations involving *JAK2* and *SH2B3* were mutually exclusive in our cohort, affecting altogether seven of twelve patients. In addition, we investigated the possibility of *SH2B3* silencing by promoter hypermethylation in our patients using MS-MCA and found no evidence of this inactivation mechanism (Figure S4).

3.3 Pathogenic small-scale mutations in the JAK-STAT pathway predominate in pcAECyTCL

The discovery of recurrent structural alterations affecting principally genes involved in the cell cycle, chromatin regulation and the JAK-STAT pathway (via JAK2) prompted us to search for pathogenic indels and SNVs in exonic sequences of genes with roles in the aforesaid cellular processes and additional signal transduction pathways (i.e. MAPK, NF-κB, PI-3-K/Akt and TCR pathways) (Table S7).

Besides the seven patients with structural alterations impacting the JAK2-SH2B3 signaling axis, four additional patients were found to carry *bona fide* gain-of-function SNVs either in *JAK3* (i.e. AEC5: p.R657W[26]; AEC12: p.M511I[27]) or *STAT5B* (i.e. AEC1 & AEC6: p.N642H[28]). Also, patient AEC3 bore a germline SNV in *JAK2* (p.L393V[29]) which has been reported to render JAK2 slightly hypersensitive to cytokine stimulation (EPO ligand) (Figure 4, Figure 5A-B, Table 1). Moreover, two patients with *JAK2* fusions and three patients with *SH2B3* deletions also carried SNVs affecting conserved residues in *STAT3* (i.e. AEC4: p.H19R; AEC9: p.G604A) and *STAT5B* (i.e. AEC2: p.P702S[30]; AEC7: p.Y665F[28]; AEC8: p.S434L[31]), respectively (Figure S10). Similarly, three patients carrying (putative) gain-of-function SNVs in *JAK* or *STAT* genes also had indels leading to premature stop codons either in *SH2B3* (i.e. AEC2: p.L201Sfs*78; AEC6: p.V35Afs*154) or *SOCS1* (i.e. AEC5: p.S71Rfs*14) (Figure 4, Figure 5).

Table 1. Identical variants in JAK and STAT proteins reported in other hematological malignancies.

Gene	Variant	Type	Neoplasm	Effect	Functionally Validated	Affected Patient
JAK2	p.L393V	Germline	PV[29,43]	Slightly HS	Yes[29]	AEC3
JAK3	p.M511I	Somatic	T-PLL, AML, JMML, NKTCL[27,42,43]	GoF	Yes[27]	AEC12
JAK3	p.R657W	Somatic	T-ALL[26,43]	GoF	Yes[26]	AEC5
STAT5B	p.S434L	Somatic	T-ALL[31]	Unknown	No	AEC8
STAT5B	p.N642H	Somatic	T-ALL, T-LGL, T-PLL, NKTCL, EATL[28,42,43]	GoF	Yes[28]	AEC1, AEC6
STAT5B	p.Y665F	Somatic	T-LGL, T-PLL, ALCL ALK–, NKTCL[28,42,43]	GoF	Yes[28]	AEC7
STAT5B	p.P702S	Somatic	T-PLL[30]	Possibly GoF	No	AEC2

GoF, Gain-of-function; HS, Hypersensitive. ALCL ALK–, ALK– anaplastic large cell lymphoma; AML, acute myeloid leukemia; EATL, enteropathy-associated T cell lymphoma; JMML, juvenile myelomonocytic leukemia; PV, polycythemia vera; T-ALL, T-cell acute lymphoblastic leukemia; T-LGL, T-cell large granular lymphocytic leukemia; T-PLL, T-cell prolymphocytic leukemia; NKTCL, extranodal natural killer T cell lymphoma.

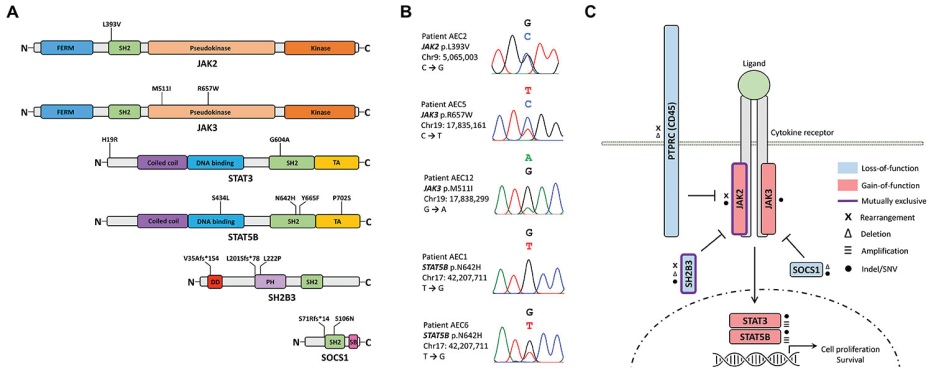


Figure 5. Small-scale mutations in genes of the JAK-STAT pathway are predominant in pcAECyTCL. (A) Diagrams showing indels and SNVs in *JAK2*, *JAK3*, *STAT3*, *STAT5B*, *SH2B3* and *SOCS1* identified in pcAECyTCL by WGS (Table 1). (B) Sanger chromatograms confirming presence of *bona fide* pathogenic SNVs in patients with pcAECyTCL. (C) Summary of genetic alterations affecting members of the JAK-STAT pathway in pcAECyTCL.

Overall, nine of twelve patients had either structural or small-scale genetic alterations impacting the JAK2-SH2B3 signaling axis whereas the remaining three patients carried pathogenic indels/SNVs in other JAK-STAT pathway genes (Figure 4, Figure 5C). In addition, cancer genes involved in the cell cycle (i.e. *TP53*) and chromatin regulation (i.e. *ARID1A*, *KMT2D*, *NCOR1*) were found to be recurrently impacted either by truncating mutations (i.e. nonsense, frameshift) or SNVs predicted as deleterious (Figure 4). Finally, we observed 34 additional patient-specific small-scale mutations of unknown significance in reputable cancer genes (Table S6).

3.4 Transcriptome analysis uncovers up-regulation of JAK2 signaling in pcAECyTCL

At present it is widely accepted that malignant T cells in pcAECyTCL derive from CD8+ T cells;[1] however, to date no specific CD8+ T-cell subtype has been proposed as the cell of origin of this lymphoma. Since a hallmark of malignant T cells in pcAECyTCL is their distinctive epidermotropism,[3,4] we compared gene expression in pcAECyTCL with gene expression in normal skin-resident CD8+ T cells, which are characterized by a marked preferential tropism to the epidermal layer of the skin.[32] This analysis identified 1603 differentially expressed (DE) genes (1076 up-regulated, 527 down-regulated, FDR < 0.01) in the disease (Figure 6A, Table S8). We next performed GSEA using annotated gene sets from MSigDB to search for deregulated pathways/processes. Up-regulated canonical signaling profiles included the JAK-STAT pathway (via STAT3, and to a lesser extent, via STAT5) and the TNF-α/NF-κB pathway. In addition, pcAECyTCL was characterized by the up-regulation of the cell cycle (i.e. E2F targets, G2/M checkpoint, mitotic spindle) and high inflammatory response (Figure 6B, Table S15).

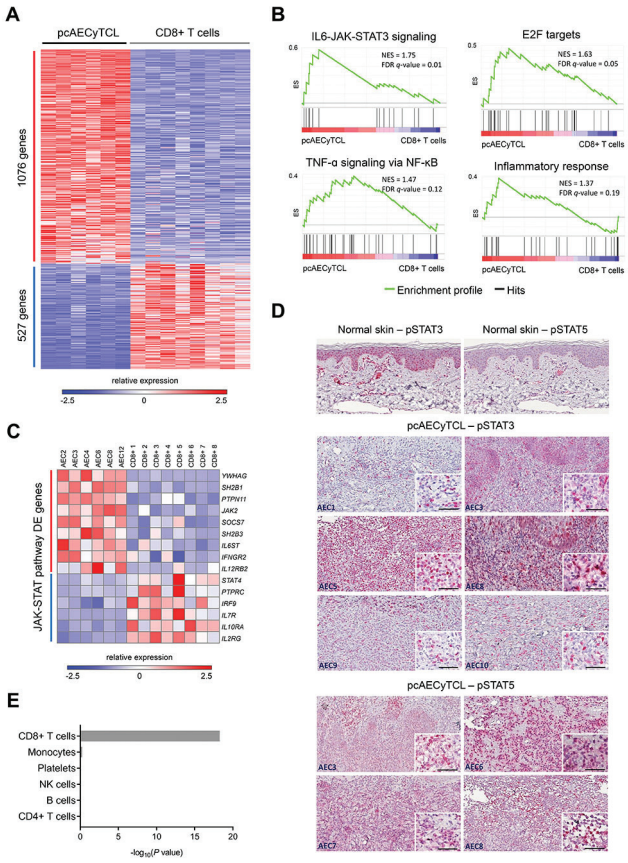


Figure 6. RNA-seq confirms up-regulation of JAK2 signaling in pcAECyTCL and supports CD8+ T cells as the cell of origin of the disease. (A) Heat map showing 1603 differentially expressed genes (1076 up-regulated, 527 down-regulated, FDR <0.01) in pcAECyTCL when compared to skin-resident CD8+ T cells. (B) Gene set enrichment analysis (GSEA) uncovered up-regulation of the JAK-STAT pathway, the cell cycle (E2F targets), the NF-κB pathway and high inflammatory response in pcAECyTCL. NES, normalized enrichment score; FDR *q*-value, false discovery rate *q*-value. (See Table S15 for a complete list of GSEA signatures) (C) Examination of differentially expressed genes involved in the JAK-STAT pathway revealed that *JAK2* itself, enhancers of JAK2 signaling (i.e. *PTPN11*, *SH2B1*) and components of cytokine receptors that signal predominantly via JAK2 are up-regulated in pcAECyTCL. In contrast, down-regulated genes were associated with signaling via JAK1, JAK3 or TYK2. (D) Activation of the JAK-STAT pathway (via STAT3 and/or STAT5) in pcAECyTCL was confirmed by IHC on tumor tissue from sequenced patients (i.e. AEC1/3/5-10). Neoplastic cells exhibit activated STAT3 and/or STAT5 in the nucleus. Normal skin (control) displays STAT3 activation in keratinocytes and endothelial cells as well as STAT5 activation in melanocytes and endothelial cells. Scale bar, 50 µm. (E) Gene set enrichment analysis showed that genes known to be expressed by CD8+ T cells were enriched among up-regulated genes in pcAECyTCL, supporting the former as the cell of origin of this lymphoma.

Further examination of DE genes involved in the JAK-STAT pathway revealed that JAK2 signaling was specifically deregulated in pcAECyTCL. Up-regulated genes included among others *JAK2* itself, components of type I and II cytokine receptors that signal predominantly via JAK2 (i.e. *IFNGR2*, *IL12RB2*) and established enhancers of JAK2 signaling (i.e. *PTPN11*, *SH2B1*). In contrast, down-regulated JAK-STAT pathway genes included *PTPRC* and genes encoding receptors exclusively associated with signal transduction via JAK1, JAK3 or TYK2 (Figure 6C).

To validate JAK-STAT pathway activation in pcAECyTCL, we investigated the presence of activated STAT proteins (pSTAT3 and pSTAT5) by immunohistochemistry (IHC) in eight sequenced patients with available tumor tissue. Robust activation of JAK-STAT signaling (via STAT3, STAT5 or both) was confirmed in all evaluated patients (Figure 6D).

3.5 Transcriptome supports CD8+ T cells as the cell of origin of pcAECyTCL

Prior characterization of the immunophenotype of malignant T cells in pcAECyTCL using IHC had strongly suggested that this lymphoma derives from CD8+ T cells.[3,4] To obtain additional insight into the cellular origin of this cancer, we compared the expression profile of pcAECyTCL with expression profiles of an assortment of human tissues and cell types deposited in ProteomicsDB using Enrichr (Table S16). Among hematopoietic cell types with available expression data, expressed genes in pcAECyTCL matched almost exclusively with expressed genes in normal CD8+ T cells, further supporting the latter as the cell of origin of pcAECyTCL (Figure 6E).

3.6 *JAK2* fusions identified in pcAECyTCL confer cytokine-independent survival ability to cells

To validate the predicted effects of the *JAK2* fusions found in pcAECyTCL (i.e. *PCM1-JAK2*, *KHDRBS1-JAK2*, *TFG-JAK2*) on cell survival, we engineered these fusion genes into murine pro-B Ba/F3 cells which die in the absence of exogenous IL3 (Figure 7A). Because self-oligo/dimerizing JAK2 fusions were predicted to activate downstream STAT proteins without the need of upstream cues elicited by cytokine stimulation, these chimeric proteins were expected to increase survival of Ba/F3 cells in the absence of IL3. Seven days after IL3 withdrawal, survival of Ba/F3 cells expressing each of the three engineered JAK2 fusions was noticeably higher ($P<0.05$, student's *t*-test) than survival of the parental Ba/F3 cells (wild-type control) and Ba/F3 cells expressing eGFP (negative control) (Figure 7B).

We next evaluated the effect of FDA-approved JAK1/2 inhibitor ruxolitinib on each of the three IL3-independent cell lines carrying *JAK2* fusions. Ruxolitinib inhibited the growth of all cell lines in a dose-dependent manner (Figure 7C-D) with IC_{50} values in the low nanomolar range (9nM – 15nM), in concordance with the reported inhibitory activity of this drug.[33] Since fusion partners *PCM1* and *TFG* have extensively been proven by others to confer chimeric kinases (including JAK2) the ability to trans-autophosphorylate via self-oligo/dimerization,[34-39] we carried on further validation with *JAK2* fusion containing novel kinase fusion partner *KHDRBS1*. For extra verification, we treated Ba/F3 cells expressing *KHDRBS1-JAK2* with inhibitor AZD1480, which has higher specificity for JAK2 than ruxolitinib,[40] and confirmed that cytokine-independent survival of these cells depends on JAK2 signaling (Figure 7D). Finally, we corroborated by western blot that growth inhibition exerted by ruxolitinib and AZD1480 was accompanied by a dose-dependent inhibition of JAK2 and STAT5 phosphorylation in Ba/F3 cells driven by *KHDRBS1-JAK2* (Figure 7D).

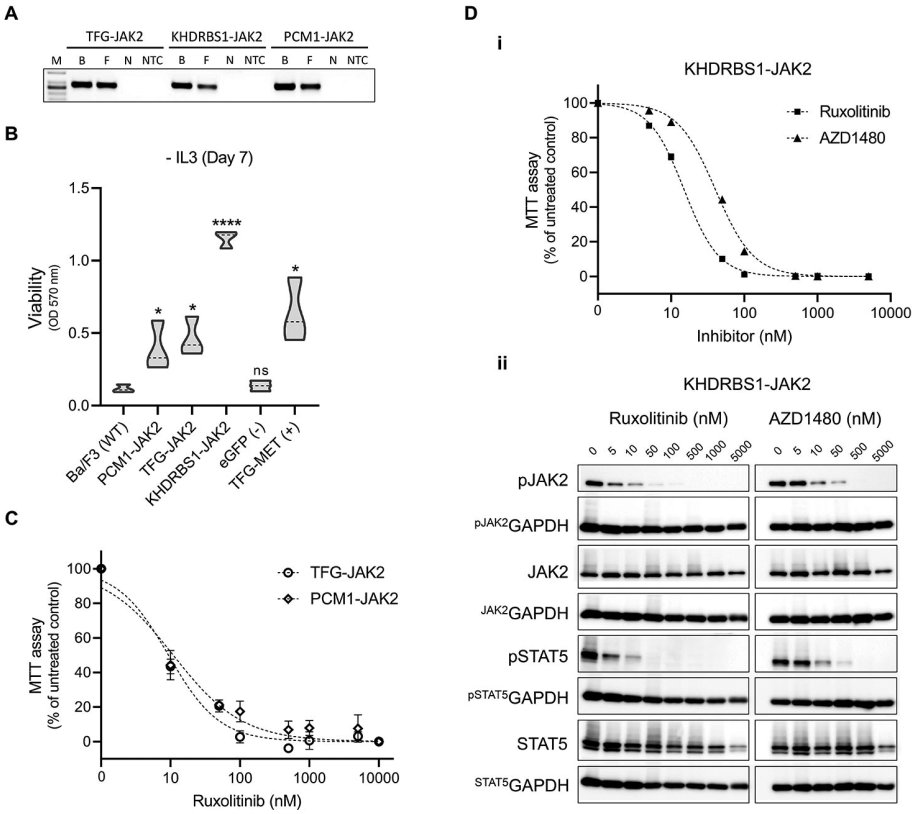


Figure 7. Oncogenicity validation of *JAK2* fusions identified in pcAECyTCL. (A) Expression of *JAK2* fusions in transduced Ba/F3 cells was verified by RT-PCR. M, molecular-weight marker. B, Fusion DNA in backbone (positive control). F, cDNA from Ba/F3 cells transduced with *JAK2* fusion gene. N, cDNA from Ba/F3 cells transduced with eGFP gene (negative control). NTC, non-template control (H_2O). (B) Violin plots showing viability of Ba/F3 cells expressing fusion genes *KHDRBS1-JAK2*, *PCM1-JAK2* or *TFG-JAK2* 7 days after IL3 withdrawal (mean OD, n = 3). Viability of all cell lines expressing *JAK2* fusions was noticeably higher ($P < 0.05$, student's *t*-test) compared to wild-type and negative control cells. Control samples: parental Ba/F3 cells (wild-type control), eGFP-expressing Ba/F3 cells (negative control), Ba/F3 cells expressing fusion gene *TFG-MET* (positive control). *, $P < 0.05$; ****, $P < 0.0001$. (C) Dose-response curves of Ba/F3 cells expressing *PCM1-JAK2* (IC_{50} = 11 nM) or *TFG-JAK2* (IC_{50} = 9 nM) when exposed to various concentrations of JAK1/2 inhibitor ruxolitinib (mean OD; error bars, SD, n = 3). (D) Validation experiments with JAK2 fusion containing novel kinase fusion partner *KHDRBS1*. (i) Dose-response curves of Ba/F3 cells expressing *KHDRBS1-JAK2* when exposed to various concentrations of JAK1/2 inhibitors ruxolitinib (IC_{50} = 15 nM) or AZD1480 (IC_{50} = 40 nM) (mean OD; error bars, SD, n = 3). (ii) Western blot analysis of *KHDRBS1-JAK2*–expressing Ba/F3 cells showing a dose-response reduction in phosphorylation of JAK2 and STAT5 with increasing concentrations of ruxolitinib or AZD1480.

4. DISCUSSION

This study describes the first high resolution genetic profiling of pcAECyTCL using next-generation sequencing. The landscape of structural genomic alterations of pcAECyTCL was characterized by considerable genomic instability and inter-patient heterogeneity. Most rearrangements (328 of 426) identified in pcAECyTCL disrupted annotated genes, and approximately one-third of all rearranged genes (91 of 305) were found to play roles in signal transduction and transcriptional regulation. In addition, 4 out of 12 patients experienced chromothripsis/chromoplexy-like events which mediated the deletion of relevant tumor suppressors (e.g. *CDKN2C*, *CHD5*, *FAS*, *PTEN*, etc.). In full agreement with previously published data,[5] we found that gains within 7q and 17q as well as losses within 1p and 13q were the most common large-scale chromosomal imbalances. Interestingly, *CDKN2A/B*, the two most frequently impaired cancer genes in pcAECyTCL, were found to be inactivated as a result of SNVs, interstitial deletions, unbalanced rearrangements and presumably even the action of long noncoding RNA *ANRIL* (*CDKN2B-AS1*), a PRC2-dependent silencer of *CDKN2B* (Figure S11).[41]

Our analysis identified a group of *bona fide* oncogenes and tumor suppressors with central roles in the cell cycle (i.e. *CDKN2A/B*, *MIR34AHG*, *MYC*, *RB1*, *TP53*), chromatin regulation (i.e. *ARID1A*, *BAZ1A*, *EED*, *EPC1*, *KMT2D*, *NCOR1*, *ZEB1*) and the JAK-STAT pathway (i.e. *JAK2*, *JAK3*, *PTPRC*, *SH2B3*, *SOCS1*, *STAT3*, *STAT5B*) whose copy number, sequence organization and/or nucleotide composition were found to be recurrently altered in our pcAECyTCL cohort. Genetic alterations involving JAK-STAT pathway genes were the most notable due to their predominance, likely proliferation-promoting effects and known causative roles in hematological cancers. A subset of SNVs affecting JAK-STAT pathway genes in pcAECyTCL have confirmed oncogenic activity in other T-cell lymphomas (Table 1).[42,43]

Remarkably, *JAK2* and *SH2B3*, which govern the activation and termination of JAK2 signaling in normal hematopoietic cells, respectively, underwent mutually exclusive alterations in nine of twelve patients from our cohort. Mutations in these two genes are associated with BCR-ABL1– myeloproliferative neoplasms (MPNs), a group of myeloid malignancies driven by overactive JAK2 signaling.[25,44] However, unlike BCR-ABL1– MPNs where *JAK2* and *SH2B3* are mainly affected by pathogenic SNVs and/or indels, these two genes experienced predominantly structural alterations in pcAECyTCL. On one hand, *JAK2* formed fusion genes encoding self-activating chimeras. On the other hand, *SH2B3* was inactivated by focal interstitial deletions and unbalanced rearrangements. The previous suggests that pcAECyTCL is mainly driven by aberrant JAK2 signaling resulting

from oncogenic changes leading to JAK2 overactivation or SH2B3 deficiency. Moreover, we demonstrated that *JAK2* fusions found in pcAECyTCL promote cytokine-independent cell survival and their oncogenic activity was shown to be successfully inhibited by ruxolitinib. Of note, *JAK2* fusions functionally analogous to the ones identified in pcAECyTCL have been previously described and confirmed as oncogenic in other hematological malignancies (e.g. B- and T-cell acute leukemias, MPNs).[34] Also, recurrent deletion of *SH2B3* has been reported in an aggressive subtype of B-cell precursor acute lymphoblastic leukemia (pre-B ALL).[45]

We found that genetic alterations involving *JAK2* and *SH2B3* coexisted with SNVs predicted or confirmed as pathogenic in *STAT3* or *STAT5B* in six out of nine affected patients. Previous functional *in vitro* studies with cell lines have suggested that mutations in STAT proteins (esp. dimerization-enhancing SNVs) observed in T-cell lymphomas operate as aberrant amplifiers of upstream signals from cytokines, overactive receptors or deregulated JAK proteins, rather than as initiators of deregulated JAK-STAT signaling themselves.[42] In this scenario, mutations in STAT3/5B would contribute to pcAECyTCL progression by making the pre-existing overactive JAK2 signaling more robust and severe. However, recent evidence derived from a murine model suggests that at least gain-of-function mutation STAT5B (p.N642H), one of the most common pathogenic SNVs in human T-cell lymphomas,[46] is sufficient by itself to promote the development of neoplasms primarily derived from mature CD8+ T cells.[47] Remarkably, malignant CD8+ T cells in these animals showed preferential migration to the skin, lung and the central nervous system, all of which are commonly affected body sites in pcAECyTCL.[47] Consistent with this evidence, patient AEC1, the only individual in our cohort who had a single JAK-STAT pathway gene mutated, carried the STAT5B (p.N642H) mutation biallelically.

Noteworthily, several pathogenetic features found in pcAECyTCL have been reported in mycosis fungoides (MF) and/or Sézary syndrome (SS) too. Genetic alterations common to pcAECyTCL, MF and SS include recurrent inactivation of *ARID1A*, *CDKN2A*, *CDKN2B*, *NCOR1*, *PTPRC*, *TP53* and *ZEB1* as well as occasional activating mutations in *JAK3*, *MYC* and *STAT3*.[21,48-53] Other genetic alterations observed in pcAECyTCL have previously been found either in MF (e.g. *SOCS1* and *STK11* inactivation) or SS (e.g. *RB1* inactivation, *STAT5B* mutations).[21,48,49] In contrast, *JAK2* fusions and *SH2B3* inactivation have not been reported in other CTCL variants to the best of our knowledge and appear to be characteristic features of pcAECyTCL.

In agreement with the recurrent genetic alterations involving the JAK2-SH2B3 signaling axis observed in pcAECyTCL, transcriptome analysis revealed up-

regulation of JAK2 signaling. *SH2B1* and *PTPN11*, which encode two proteins with the ability to enhance JAK2 signaling,[54,55] stood out among up-regulated JAK-STAT pathway genes. Adaptor protein SH2B1 has been proven to bind to JAK2 and stimulate its kinase activity.[56] Similarly, phosphatase PTPN11 (SHP-2) has been shown to positively regulate JAK2-mediated STAT5 phosphorylation.[57] In contrast, phosphatase PTPRC (CD45), whose expression has been shown to attenuate JAK2 signaling in hematopoietic and lymphoma cells,[58,59] was down-regulated in pcAECyTCL. Yet, the exact molecular interactions underlying the action of these three regulators of JAK2 signaling remain to be fully elucidated.

Transcriptome analysis also revealed up-regulation of the cell cycle, the TNF-α/NF-κB pathway and a high inflammatory response in pcAECyTCL. Notably, the co-activation (crosstalk) of JAK-STAT signaling (esp. via STAT3) and NF-κB signaling is a well-documented phenomenon in cancer, and it has been shown to promote a pro-oncogenic inflammatory microenvironment in the tumor.[60] For instance, aberrant JAK2 signaling (via STAT3) in MPNs promotes chromatin changes that induce NF-κB signaling; and the resulting combined action of these two pathways, appear to drive the characteristic chronic inflammatory state observed in these neoplasms.[61] Our data, in full concordance with the previous, suggest that co-activation of JAK2 signaling and NF-κB signaling operates in pcAECyTCL as well, and their joint action might be responsible for the inflammatory state detected in pcAECyTCL tumors.

Taken together, our findings strongly suggest that overactivation of JAK2 signaling plays a pivotal role in the pathogenesis of pcAECyTCL. Therefore, patients with this lymphoma would likely benefit from treatment with JAK2 inhibitors (e.g. FDA-approved ruxolitinib). In addition, the potential combination of JAK2 inhibitors with NF-κB inhibitors (e.g. bortezomib,[62] dimethyl fumarate[63]) represents an attractive possibility since targeting both pathways might have a synergic effect and reduce the chance of resistance acquisition.

Acknowledgements
The authors thank Tim van Groningen and Yixin Luo for providing valuable technical support. This study was funded by the Dutch Cancer Society (KWF, grant UL2013-6104) and Associazione Amici di Sabrina Fadini Onlus (A.S.F.O., grant 0800000-PR-LUMC).

Conflict of interest
The authors declare no conflicts of interest.

Author contributions

A.N.B.T., D.F., E.B., and C.P.T. designed the project. A.N.B.T. and C.P.T wrote the manuscript. A.N.B.T., D.C. and H.M. performed the bioinformatic analyses. A.N.B.T., C.O. D.F and L.V. performed the experiments. A.N.B.T. analyzed the results, interpreted the data and produced figures and tables. D.F., L.V., R.W., M.V. and E.B. provided valuable biological specimens. A.N.B.T, D.C., C.O., H.M., D.F., L.V., R.W., M.V., E.B., and C.P.T revised and approved the final manuscript.

REFERENCES

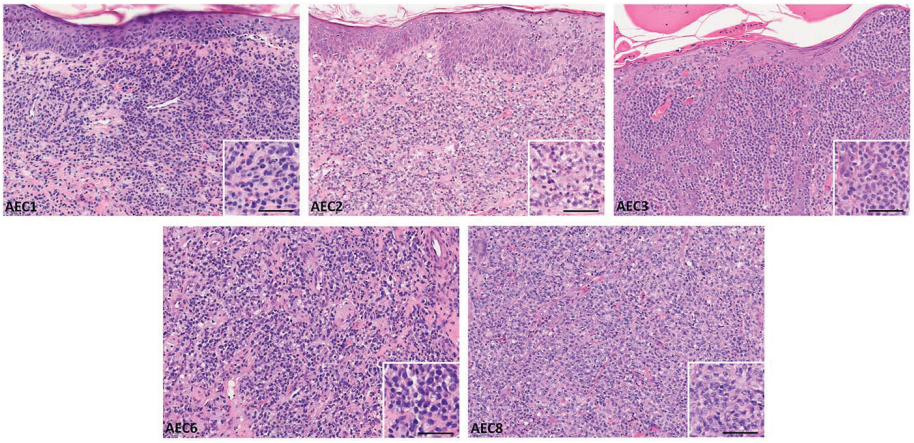
1. Berti E, Gaulard P, Willemze R, Petrella T, Jaffe E. Primary cutaneous CD8+ aggressive epidermotropic cytotoxic T-cell lymphoma. In: Swerldow SH, Campo E, Harris N, et al., eds. WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues 4th ed. Lyon: IARC; 2017.
2. Willemze R, Jaffe ES, Burg G, et al. WHO-EORTC classification for cutaneous lymphomas. *Blood*. 2005;105(10):3768-3785.
3. Berti E, Tomasini D, Vermeer MH, Meijer CJLM, Alessi E, Willemze R. Primary Cutaneous CD8-Positive Epidermotropic Cytotoxic T Cell Lymphomas. *The American Journal of Pathology*. 1999;155(2):483-492.
4. Guitart J, Martinez-Escala ME, Subtil A, et al. Primary cutaneous aggressive epidermotropic cytotoxic T-cell lymphomas: reappraisal of a provisional entity in the 2016 WHO classification of cutaneous lymphomas. *Mod Pathol*. 2017;30(5):761-772.
5. Fanoni D, Corti L, Alberti-Violetti S, et al. Array-based CGH of primary cutaneous CD8+ aggressive EPIDERMO-tropic cytotoxic T-cell lymphoma. *Genes Chromosomes Cancer*. 2018;57(12):622-629.
6. Kato K, Oh Y, Takita J, et al. Molecular genetic and cytogenetic analysis of a primary cutaneous CD8-positive aggressive epidermotropic cytotoxic T-cell lymphoma. *Int J Hematol*. 2016;103(2):196-201.
7. Tomasini C, Novelli M, Fanoni D, Berti EF. Erythema multiforme-like lesions in primary cutaneous aggressive cytotoxic epidermotropic CD8+ T-cell lymphoma: A diagnostic and therapeutic challenge. *J Cutan Pathol*. 2017;44(10):867-873.
8. Bastidas Torres AN, Cats D, Mei H, et al. Whole-genome analysis uncovers recurrent IKZF1 inactivation and aberrant cell adhesion in blastic plasmacytoid dendritic cell neoplasm. *Genes Chromosomes Cancer*. 2020;59(5):295-308.
9. Repana D, Nulsen J, Dressler L, et al. The Network of Cancer Genes (NCG): a comprehensive catalogue of known and candidate cancer genes from cancer sequencing screens. *Genome Biol*. 2019;20(1):1.
10. Boeva V, Popova T, Bleakley K, et al. Control-FREEC: a tool for assessing copy number and allelic content using next-generation sequencing data. *Bioinformatics*. 2012;28(3):423-425.
11. Mermel CH, Schumacher SE, Hill B, Meyerson ML, Beroukhim R, Getz G. GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome Biol*. 2011;12(4):R41.
12. MacDonald JR, Ziman R, Yuen RK, Feuk L, Scherer SW. The Database of Genomic Variants: a curated collection of structural variation in the human genome. *Nucleic Acids Res*. 2014;42(Database issue):D986-992.
13. Thorvaldsdóttir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief Bioinform*. 2013;14(2):178-192.
14. Park J, Yang J, Wenzel AT, et al. Genomic analysis of 220 CTCLs identifies a novel recurrent gain-of-function alteration in RLTPR (p.Q575E). *Blood*. 2017.
15. Bastidas Torres AN, Najidh S, Tensen CP, Vermeer MH. Molecular advances in cutaneous T-cell lymphoma. *Semin Cutan Med Surg*. 2018;37(1):81-86.
16. McCully ML, Ladell K, Andrews R, et al. CCR8 Expression Defines Tissue-Resident Memory T Cells in Human Skin. *J Immunol*. 2018;200(5):1639-1650.
17. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A*. 2005;102(43):15545-15550.
18. Schmidt T, Samaras P, Frejno M, et al. ProteomicsDB. *Nucleic Acids Res*. 2017;46(D1):D1271-D1281.
19. Kuleshov MV, Jones MR, Rouillard AD, et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res*. 2016;44(W1):W90-97.

20. Mi H, Huang X, Muruganujan A, et al. PANTHER version 11: expanded annotation data from Gene Ontology and Reactome pathways, and data analysis tool enhancements. *Nucleic Acids Res*. 2017;45(D1):D183-d189.
21. Bastidas Torres AN, Cats D, Mei H, et al. Genomic analysis reveals recurrent deletion of JAK-STAT signaling inhibitors HNRNPK and SOCS1 in mycosis fungoides. *Genes Chromosomes Cancer*. 2018;57(12):653-664.
22. van Doorn R, Slieker RC, Boonk SE, et al. Epigenomic Analysis of Sezary Syndrome Defines Patterns of Aberrant DNA Methylation and Identifies Diagnostic Markers. *J Invest Dermatol*. 2016;136(9):1876-1884.
23. Lu H, Villafane N, Dogruluk T, et al. Engineering and Functional Characterization of Fusion Genes Identifies Novel Oncogenic Drivers of Cancer. *Cancer Res*. 2017;77(13):3502-3512.
24. Paczkowska J, Soloch N, Bodnar M, et al. Expression of ELF1, a lymphoid ETS domain-containing transcription factor, is recurrently lost in classical Hodgkin lymphoma. *Br J Haematol*. 2019;185(1):79-88.
25. Maslah N, Cassinat B, Verger E, Kiladjian JJ, Velazquez L. The role of LNK/SH2B3 genetic alterations in myeloproliferative neoplasms and other hematological disorders. *Leukemia*. 2017;31(8):1661-1670.
26. Degryse S, de Bock CE, Cox L, et al. JAK3 mutants transform hematopoietic cells through JAK1 activation, causing T-cell acute lymphoblastic leukemia in a mouse model. *Blood*. 2014;124(20):3092-3100.
27. Yamashita Y, Yuan J, Suetake I, et al. Array-based genomic resequencing of human leukemia. *Oncogene*. 2010;29(25):3723-3731.
28. Rajala HL, Eldfors S, Kuusanmaki H, et al. Discovery of somatic STAT5b mutations in large granular lymphocytic leukemia. *Blood*. 2013;121(22):4541-4550.
29. Lanikova L, Babosova O, Swierczek S, et al. Coexistence of gain-of-function JAK2 germ line mutations with JAK2V617F in polycythemia vera. *Blood*. 2016;128(18):2266-2270.
30. Andersson EI, Putzer S, Yadav B, et al. Discovery of novel drug sensitivities in T-PLL by high-throughput ex vivo drug testing and mutation profiling. *Leukemia*. 2018;32(3):774-787.
31. Bandapalli OR, Schuessele S, Kunz JB, et al. The activating STAT5B N642H mutation is a common abnormality in pediatric T-cell acute lymphoblastic leukemia and confers a higher risk of relapse. *Haematologica*. 2014;99(10):e188-192.
32. Takamura S. Niches for the Long-Term Maintenance of Tissue-Resident Memory T Cells. *Front Immunol*. 2018;9:1214.
33. Quintas-Cardama A, Vaddi K, Liu P, et al. Preclinical characterization of the selective JAK1/2 inhibitor INCB018424: therapeutic implications for the treatment of myeloproliferative neoplasms. *Blood*. 2010;115(15):3109-3117.
34. Smith CA, Fan G. The saga of JAK2 mutations and translocations in hematologic disorders: pathogenesis, diagnostic and therapeutic prospects, and revised World Health Organization diagnostic criteria for myeloproliferative neoplasms. *Hum Pathol*. 2008;39(6):795-810.
35. Reiter A, Walz C, Watmore A, et al. The t(8;9)(p22;p24) is a recurrent abnormality in chronic and acute leukemia that fuses PCM1 to JAK2. *Cancer Res*. 2005;65(7):2662-2667.
36. Lierman E, Selleslag D, Smits S, Billiet J, Vandenberghe P. Ruxolitinib inhibits transforming JAK2 fusion proteins in vitro and induces complete cytogenetic remission in t(8;9)(p22;p24)/PCM1-JAK2-positive chronic eosinophilic leukemia. *Blood*. 2012;120(7):1529-1531.
37. Stransky N, Cerami E, Schalm S, Kim JL, Lengauer C. The landscape of kinase fusions in cancer. *Nat Commun*. 2014;5:4846.
38. Greco A, Fusetti L, Miranda C, et al. Role of the TFG N-terminus and coiled-coil domain in the transforming activity of the thyroid TRK-T3 oncogene. *Oncogene*. 1998;16(6):809-816.
39. Hernandez L, Bea S, Bellosillo B, et al. Diversity of genomic breakpoints in TFG-ALK translocations in anaplastic large cell lymphomas: identification of a new TFG-ALK(XL) chimeric gene with transforming activity. *Am J Pathol*. 2002;160(4):1487-1494.
40. Hedvat M, Huszar D, Herrmann A, et al. The JAK2 inhibitor AZD1480 potently blocks Stat3 signaling and oncogenesis in solid tumors. *Cancer Cell*. 2009;16(6):487-497.

41. Kotake Y, Nakagawa T, Kitagawa K, et al. Long non-coding RNA ANRIL is required for the PRC2 recruitment to and silencing of p15(INK4B) tumor suppressor gene. *Oncogene*. 2011;30(16):1956-1962.
42. Waldmann TA, Chen J. Disorders of the JAK/STAT Pathway in T Cell Lymphoma Pathogenesis: Implications for Immunotherapy. *Annu Rev Immunol*. 2017;35:533-550.
43. Hammarén HM, Virtanen AT, Raivola J, Silvennoinen O. The regulation of JAKs in cytokine signaling and its breakdown in disease. *Cytokine*. 2019;118:48-63.
44. Levine RL, Pardanani A, Tefferi A, Gilliland DG. Role of JAK2 in the pathogenesis and therapy of myeloproliferative disorders. *Nat Rev Cancer*. 2007;7(9):673-683.
45. Baughn LB, Meredith MM, Oseth L, Smolarek TA, Hirsch B. SH2B3 aberrations enriched in iAMP21 B lymphoblastic leukemia. *Cancer Genet*. 2018;226-227:30-35.
46. Pham HTT, Maurer B, Prchal-Murphy M, et al. STAT5BN642H is a driver mutation for T cell neoplasia. *J Clin Invest*. 2018;128(1):387-401.
47. de Araujo ED, Erdogan F, Neubauer HA, et al. Structural and functional consequences of the STAT5B(N642H) driver mutation. *Nat Commun*. 2019;10(1):2517.
48. Woollard WJ, Pullabhatla V, Lorenc A, et al. Candidate driver genes involved in genome maintenance and DNA repair in Sezary syndrome. *Blood*. 2016;127(26):3387-3397.
49. Wang L, Ni X, Covington KR, et al. Genomic profiling of Sezary syndrome identifies alterations of key T cell signaling and differentiation genes. *Nat Genet*. 2015;47(12):1426-1434.
50. McGirt LY, Jia P, Baerenwald DA, et al. Whole-genome sequencing reveals oncogenic mutations in mycosis fungoides. *Blood*. 2015;126(4):508-519.
51. Kiel MJ, Sahasrabuddhe AA, Rolland DC, et al. Genomic analyses reveal recurrent mutations in epigenetic modifiers and the JAK-STAT pathway in Sezary syndrome. *Nat Commun*. 2015;6:8470.
52. da Silva Almeida AC, Abate F, Khiabanian H, et al. The mutational landscape of cutaneous T cell lymphoma and Sezary syndrome. *Nat Genet*. 2015;47(12):1465-1470.
53. Choi J, Goh G, Walradt T, et al. Genomic landscape of cutaneous T cell lymphoma. *Nat Genet*. 2015;47(9):1011-1019.
54. Maures TJ, Kurzer JH, Carter-Su C. SH2B1 (SH2-B) and JAK2: a multifunctional adaptor protein and kinase made for each other. *Trends Endocrinol Metab*. 2007;18(1):38-45.
55. Liu X, Qu CK. Protein Tyrosine Phosphatase SHP-2 (PTPN11) in Hematopoiesis and Leukemogenesis. *J Signal Transduct*. 2011;2011:195239.
56. Rui L, Carter-Su C. Identification of SH2-bbeta as a potent cytoplasmic activator of the tyrosine kinase Janus kinase 2. *Proc Natl Acad Sci U S A*. 1999;96(13):7172-7177.
57. Ali S, Nouhi Z, Chughtai N, Ali S. SHP-2 regulates SOCS-1-mediated Janus kinase-2 ubiquitination/degradation downstream of the prolactin receptor. *J Biol Chem*. 2003;278(52):52021-52031.
58. Irie-Sasaki J, Sasaki T, Matsumoto W, et al. CD45 is a JAK phosphatase and negatively regulates cytokine receptor signalling. *Nature*. 2001;409(6818):349-354.
59. Wu L, Bijian K, Shen SH. CD45 recruits adapter protein DOK-1 and negatively regulates JAK-STAT signaling in hematopoietic cells. *Mol Immunol*. 2009;46(11-12):2167-2177.
60. Yu H, Pardoll D, Jove R. STATs in cancer inflammation and immunity: a leading role for STAT3. *Nat Rev Cancer*. 2009;9(11):798-809.
61. Kleppe M, Koche R, Zou L, et al. Dual Targeting of Oncogenic Activation and Inflammatory Signaling Increases Therapeutic Efficacy in Myeloproliferative Neoplasms. *Cancer Cell*. 2018;33(1):29-43.e27.
62. Zinzani PL, Musuraca G, Tani M, et al. Phase II trial of proteasome inhibitor bortezomib in patients with relapsed or refractory cutaneous T-cell lymphoma. *J Clin Oncol*. 2007;25(27):4293-4297.
63. Nicolay JP, Muller-Decker K, Schroeder A, et al. Dimethyl fumarate restores apoptosis sensitivity and inhibits tumor growth and metastasis in CTCL by targeting NF-kappaB. *Blood*. 2016;128(6):805-815.

SUPPLEMENTARY FIGURES

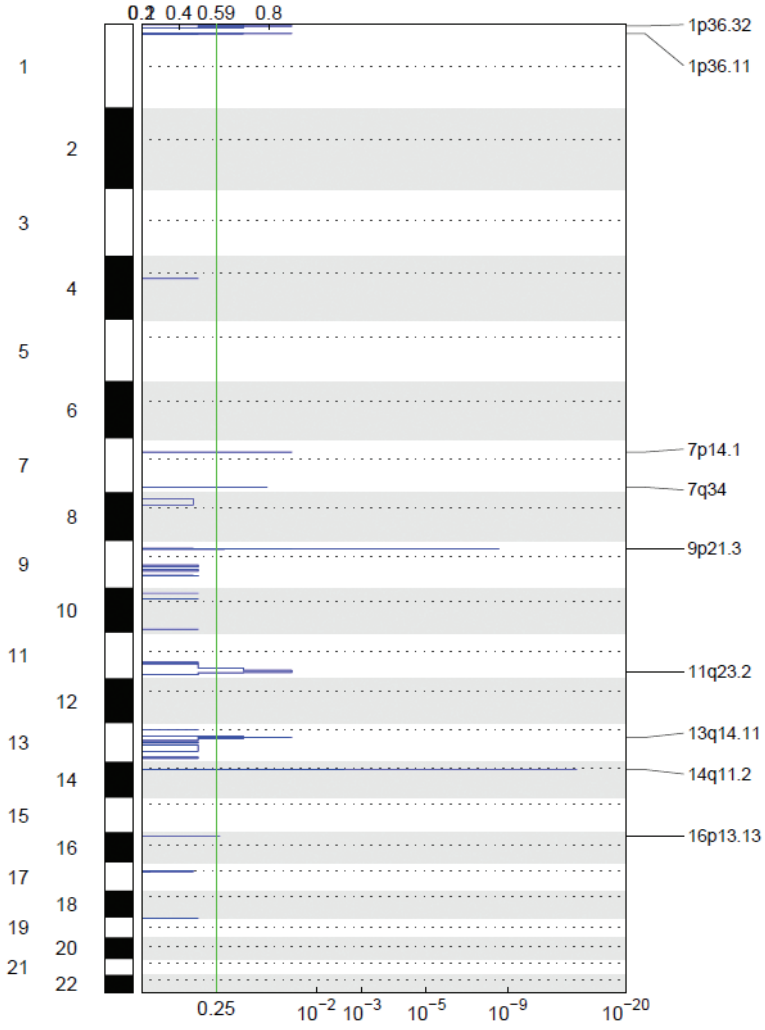
Supplementary Figure S1. Microscopy images of H&E-stained tumor tissue from representative pcAECyTCL patients included in the study.



Supplementary Figure S1. Microscopy images (magnification, 200X) of H&E-stained tumor tissue from patients with pcAECyTCL subjected to NGS. Scale bar, 50 µm.

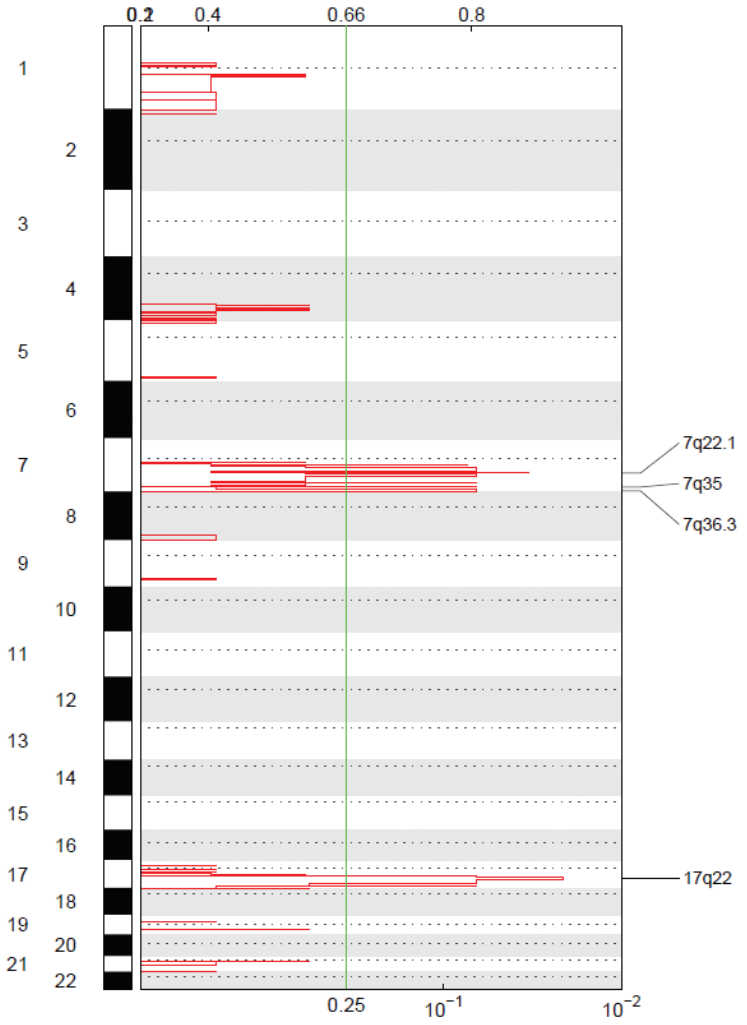
Supplementary Figure S2. Common focal CNAs in pcAECyTCL detected by GISTIC 2.0. analysis.

A)



Common focal losses in pcAECyTCL detected by GISTIC 2.0. analysis. GISTIC 2.0 detected common losses shared by pcAECyTCL patients within chromosomes 1, 7, 9, 11, 13, 14 and 16. Peaks at 1p36.11, 1p36.32, 9p21.3, 13q14.11 and 16p13.13 were all ≤ 3Mb long (as verified with IGV) and enclosed (putative) tumor suppressors *ARID1A TNFRSF14*, *CDKN2A/B*, *ELF1* and *SOCS1*, respectively. Peaks at 7p14.1, 7q34 and 14q11.2 contained loci encoding the α (TRA), β (TRB) and γ (TRG) chains of the T-cell receptor (TCR), which are rearranged during T-cell development. Loss of the exact same region within *TRA*, *TRB* and *TRG* in tumor DNA from patients suggests that pcAECyTCL is a clonal disease.

B)



Common focal gains in pcAECyTCL detected by GISTIC 2.0. analysis. GISTIC 2.0 detected common gains between pcAECyTCL patients. Peaks within chromosomes 7 and 17 were either > 3Mb long or devoid of (putative) cancer genes, and consequently, excluded.

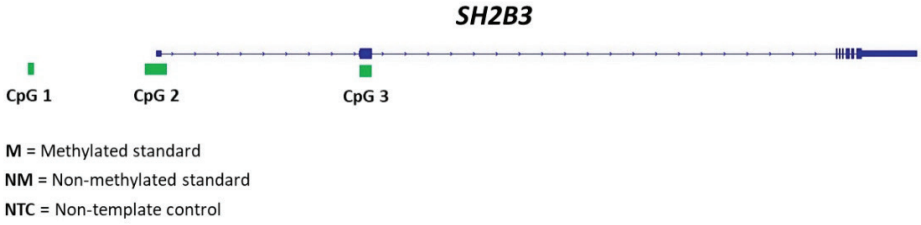
Supplementary Figure S3. Genomic location of BAC probes used in FISH experiments.



Supplementary Figure S3. (A) Diagram of break-apart BAC probes used for the validation of DNA breaks at *JAK2* (9q24.1) in pcAECyTCL. See Supplementary Table S6 for exact genomic positions of BAC probes used in this study.

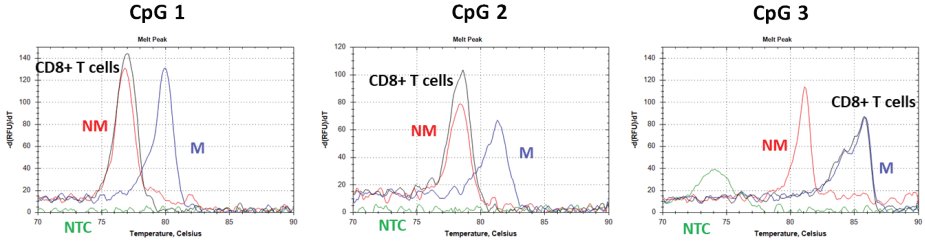
Supplementary Figure S4. Methylation-specific melting curve analysis (MS-MCA) of CpG islands at SH2B3 in pcAECyTCL.

A)



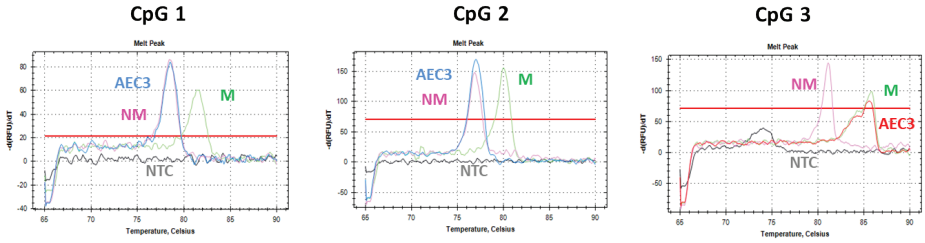
B)

CD8+ T cells (control):

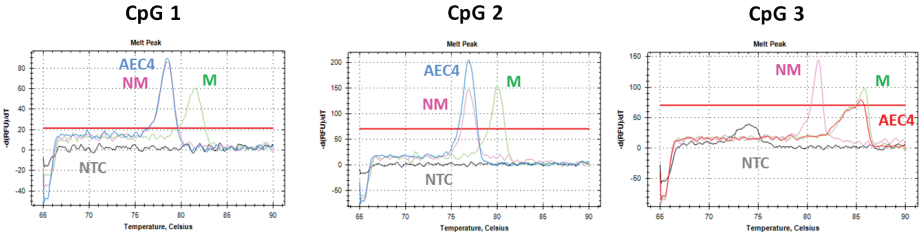


C)

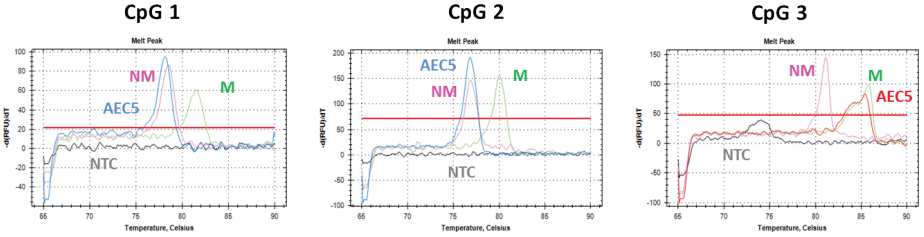
Patient AEC3:



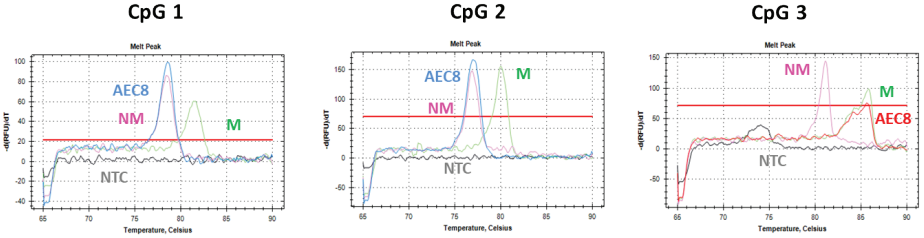
Patient AEC4:



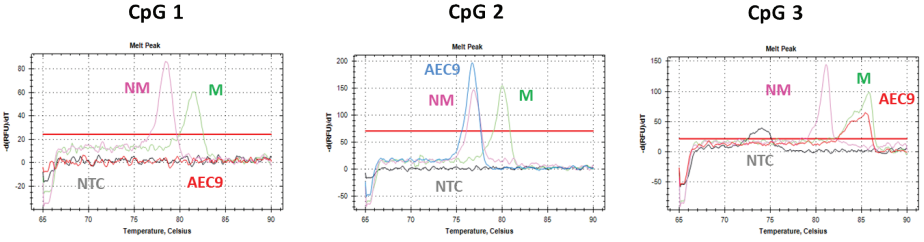
Patient AEC5:



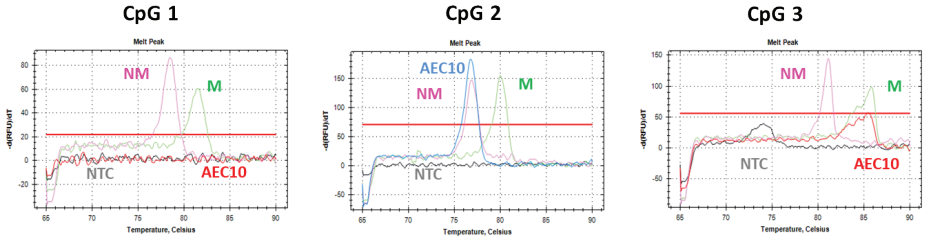
Patient AEC8:



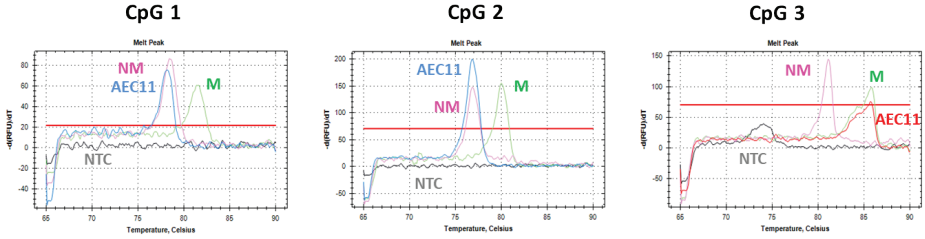
Patient AEC9:



Patient AEC10:

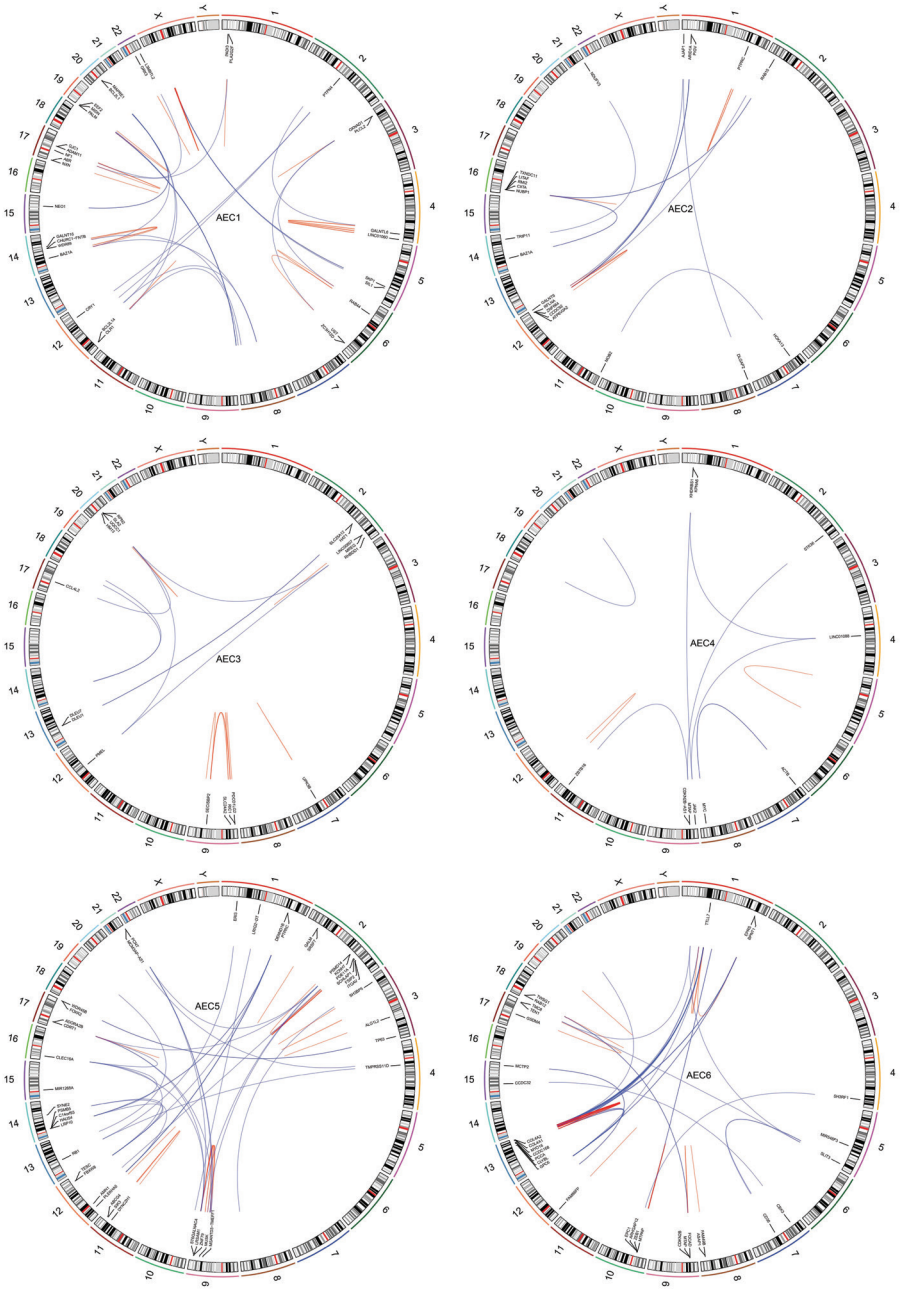


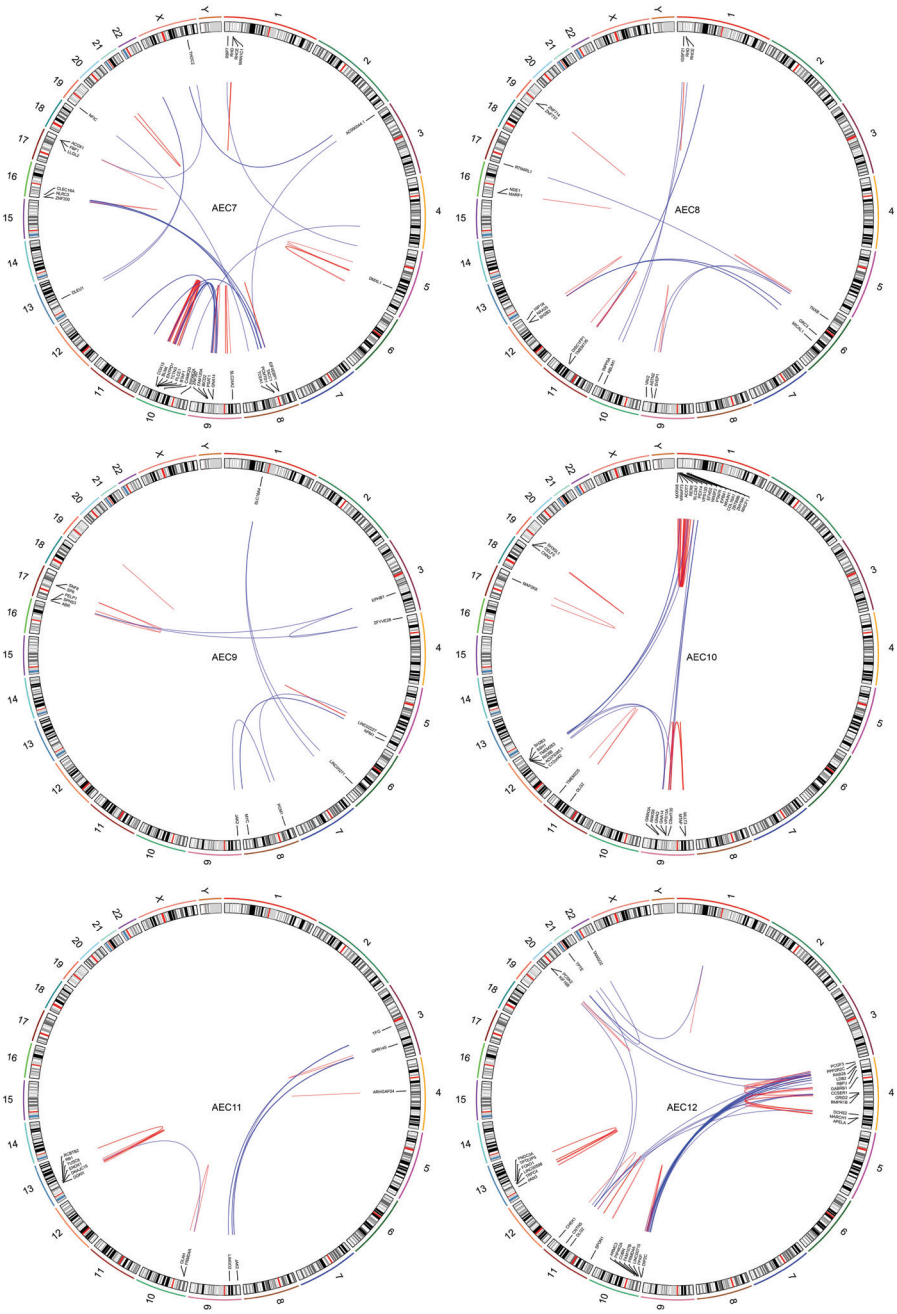
Patient AEC11:



Supplementary Figure S4. A) Diagram showing the location of the 3 CpG islands near/at *SH2B3* interrogated by MS-MCA in patients with pcAECyTCL. B) MS-MCA curve profiles of DNA from normal CD8+ T cells (control). CD8+ T cells show full unmethylation at CpG1 and CpG2, and full methylation at CpG3. C) MS-MCA curve profiles of tumor DNA from patients AEC3, AEC4, AEC5, AEC8, AEC9, AEC10 and AEC11. Just like normal CD8+ T cells, tumor DNA from patients with pcAECyTCL show full unmethylation at CpG1 and CpG2, and full methylation at CpG3, suggesting that hypermethylation at CpG3 in patients with pcAECyTCL is not associated with silencing of *SH2B3*, and therefore, it is unlikely to play a pathogenic role in the disease.

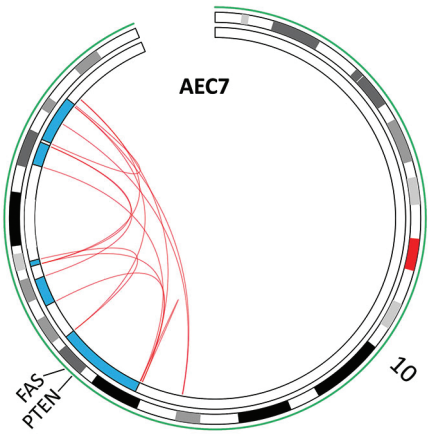
Supplementary Figure S5. Circos plots displaying genomic rearrangements per patient.





Supplementary Figure S5. Tumors from different pcAECyTCL patients contained a variety of inter- and intrachromosomal rearrangements. The number of events ranged from 10 (AEC4) to 65 (AEC7). Rearrangements involving JAK-STAT pathways genes: *JAK2* (AEC4, AEC9 and AEC11), *PTPRC* (AEC2 and AEC5) and *SH2B3* (AEC8 and AEC10) were the most prominent events in our sequenced cohort.

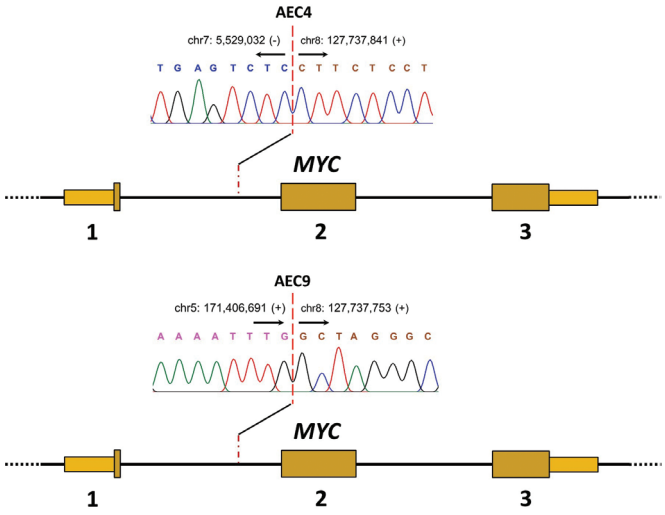
Supplementary Figure S6. Complex genomic rearrangement in patient AEC7.



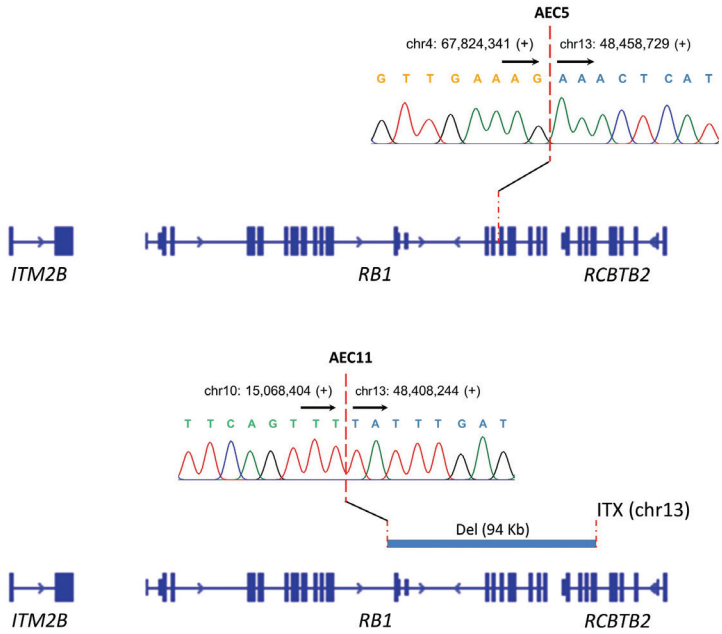
Supplementary Figure S6. Circos plot showing a chromothripsis-like event detected in chromosome 10 of patient AEC7 through WGS. This event mediated the loss of several genomic regions within chromosome 10, including the region where tumor suppressors *FAS* and *PTEN* reside.

Supplementary Figure S7. Validation of select genomic rearrangements in pcAECyTCL by Sanger sequencing.

A)

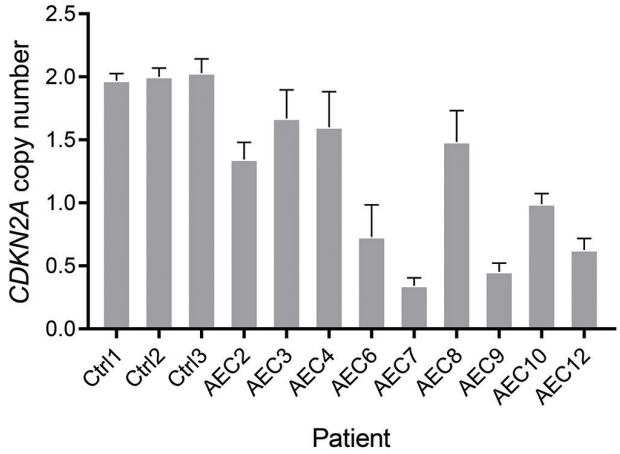


B)



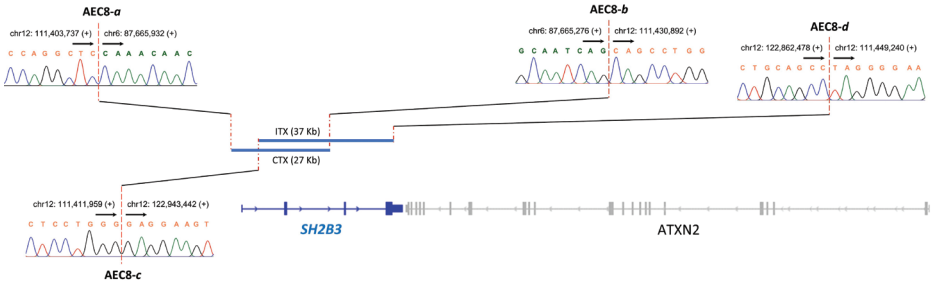
Supplementary Figure S7. Sanger chromatograms showing breakpoints of interchromosomal rearrangements at (A) *MYC* and (B) *RB1* in tumor DNA from patients with pcAECyTCL. Genomic coordinates according to reference genome GRCh38.

Supplementary Figure S8. Validation of *CDKN2A* loss in pcAECyTCL by ddPCR.



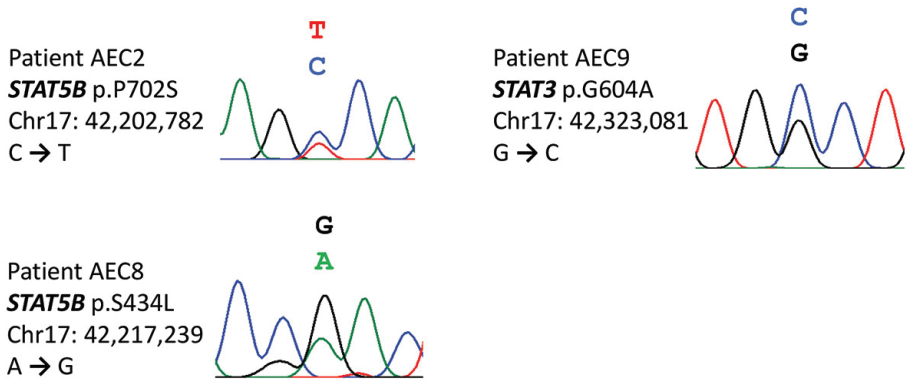
Supplementary Figure S8. Deletion of tumor suppressor *CDKN2A*, the most commonly deleted tumor suppressor in pcAECyTCL according to the WGS data, was confirmed in all patients with available material by ddPCR. Ctrl, CD8+ T cells.

Supplementary Figure S9. Validation of genomic rearrangements at *SH2B3* in patient AEC8 by Sanger sequencing.



Supplementary Figure S9. Sanger chromatograms showing breakpoints of inter- and intrachromosomal rearrangements at *SH2B3* in tumor DNA from patient AEC8. AEC8 was the only patient carrying two inactivating events at *SH2B3* in our pcAECyTCL cohort. Genomic coordinates according to reference genome GRCh38. CTX, interchromosomal rearrangement, ITX, intrachromosomal rearrangement.

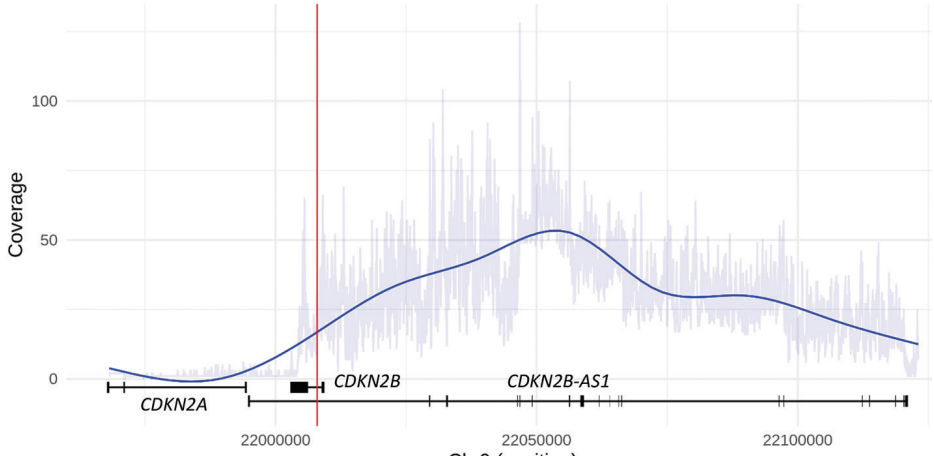
Supplementary Figure S10. Sanger sequencing validation of select SNVs.



Supplementary Figure S10. Sanger chromatograms showing SNVs in *STAT3* and *STAT5B* detected in pcAECyTCL tumors by WGS and predicted as deleterious (SIFT and PolyPhen-2). Genomic coordinates according to reference genome GRCh38.

5

Supplementary Figure S11. Interchromosomal rearrangement leads to high ANRIL (CDKN2B-AS1) expression in patient AEC6.



Supplementary Figure S11. Plot representing the mean read coverage (RNA-seq) across three cancer genes residing within 9p21.3: *CDKN2A*, *CDKN2B* and *ANRIL* (*CDKN2B-AS1*). Due to a interchromosomal rearrangement that fused *RAB12* with *ANRIL*, putting the latter under the transcriptional control of the former's promoter, patient AEC6 experienced high *ANRIL* expression when compared to the rest of patients with pcAECyTCL where ANRIL's expression was negligible (data not shown). *ANRIL* is a PRC2-dependent silencer of *CDKN2B* (p15). The red line indicates the breakpoint position within the *ANRIL* locus. Genomic coordinate numbers according to reference genome GRCh38.

6

Whole-genome analysis uncovers recurrent *IKZF1* inactivation and aberrant cell adhesion in blastic plasmacytoid dendritic cell neoplasm

Armando N. Bastidas Torres
Davy Cats
Hailiang Mei
Daniele Fanoni
Jessica Gliozzo
Laura Corti
Marco Paulli
Maarten H. Vermeer
Rein Willemze
Emilio Berti
Cornelis P. Tensen

Genes, Chromosomes and Cancer 59: 295–308 (2020)



ABSTRACT

Blastic plasmacytoid dendritic cell neoplasm (BPDCN) is a rare and highly aggressive hematological malignancy with a poorly understood pathobiology and no effective therapeutic options. Despite a few recurrent genetic defects (e.g. single nucleotide changes, indels, large chromosomal aberrations) have been identified in BPDCN, none are disease-specific, and more importantly, none explain its genesis or clinical behavior. In this study, we performed the first high resolution whole-genome analysis of BPDCN with a special focus on structural genomic alterations by using whole-genome sequencing and RNA-sequencing. Our study, the first to characterize the landscape of genomic rearrangements and copy number alterations of BPDCN at nucleotide-level resolution, revealed that *IKZF1*, a gene encoding a transcription factor required for the differentiation of plasmacytoid dendritic cell precursors, is focally inactivated through recurrent structural alterations in this neoplasm. In concordance with the genomic data, transcriptome analysis revealed that conserved IKZF1 target genes display a loss-of-IKZF1 expression pattern. Furthermore, up-regulation of cellular processes responsible for cell-cell and cell-ECM interactions, which is a hallmark of IKZF1 deficiency, was prominent in BPDCN. Our findings suggest that *IKZF1* inactivation plays a central role in the pathobiology of the disease, and consequently, therapeutic approaches directed at reestablishing the function of this gene might be beneficial for patients.

1. INTRODUCTION

Blastic plasmacytoid dendritic cell neoplasm (BPDCN) is a rare and highly aggressive hematological malignancy derived from neoplastic plasmacytoid dendritic cell (pDC) precursors. Patients with BPDCN have infiltration of malignant cells characterized by the expression of CD4, CD56, CD123 and CD303 in the skin, bone marrow, lymph nodes and peripheral blood. The median overall survival of patients with BPDCN ranges from 12 to 14 months as the disease typically relapses after initial response to chemotherapy.[1,2]

The pathogenetic basis of BPDCN has been investigated using DNA and RNA array-based methods, and more recently, next generation sequencing (NGS) approaches limited to the exome (whole-exome sequencing) or a panel of genes (targeted sequencing). Recurrent copy number alterations (CNAs) in BPDCN include broad deletions within chromosomes 7, 9, 12, 13, and 15, particularly the regions containing tumor suppressors *CDKN2A/B*, *CDKN1B* and *RB1*.[3-6] Overexpressed genes with pathogenic relevance include *BCL2* and *CCND1*.[7] Single nucleotide variants (SNVs) and indels presumed as pathogenic have been observed predominantly in genes involved in cell cycle control (e.g. *ATM*, *TP53*), chromatin regulation (e.g. *ASXL1*, *IDH2*, *NPM1*, *TET2*), signal transduction (e.g. *KRAS*, *NRAS*), splicing (e.g. *ZRSR2*) and transcriptional regulation (e.g. *IKZF1/2/3*, *ZEB2*).[8-11] Additionally, cytogenetic studies, most of them performed in the context of case reports, have revealed chromosomal rearrangements involving *ALK*, *ETV6*, *EWSR1*, *KMT2A* (*MLL*), *MYB, MYC* and *SUPT3H* in a handful of cases.[12-18]

Even though these studies have uncovered a variety of genetic abnormalities in BPDCN, none of the reported alterations has provided a clear biological rationale behind neither the genesis of the disease nor its clinical behavior. Furthermore, as yet no study has characterized the landscape of genomic rearrangements and CNAs of BPDCN using high resolution NGS.

Here, we present the first whole-genome analysis of BPDCN with a special emphasis on structural anomalies by using whole-genome sequencing (WGS) and RNA sequencing (RNA-seq). We report recurrent *IKZF1* inactivation by focal structural alterations, a loss-of-IKZF1 expression pattern and overexpression of adhesion signatures in BPDCN. Our findings on genome and transcriptome level not only support *IKZF1* inactivation as a putative driver event in the development of BPDCN, but also provide a conceptual basis for the pathobiology of the disease.

2. MATERIALS AND METHODS

2.1 Patient material

Frozen tumor biopsies from ten patients with BPDCN (Supporting Information Table 1) were subjected to WGS. Four samples of this cohort (i.e. BDN1, BDN4, BDN5 and BDN6) were additionally subjected to RNA-seq. Diagnosis was performed by an expert panel of dermatologist and pathologists according to the criteria of the WHO-EORTC classification for primary cutaneous lymphomas.[2,19,20] Frozen tumor biopsies from fifteen additional patients were used for validation purposes (extension cohort) (Supporting Information Table 1). Patient material was approved by the institutional review boards of Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Leiden University Medical Center and University of Pavia. Informed consent was obtained from patients in accordance with the declaration of Helsinki.

2.2 Nucleic acid isolation

Genomic DNA was isolated using Genomic-tip 20/G kit (Qiagen). DNA purity was evaluated with a Nanodrop One system (Nanodrop Technologies, Wilmington, CA) and DNA integrity verified by gel electrophoresis (0.7% agarose, ethidium bromide). Total RNA was isolated using RNeasy mini kit (Qiagen). RNA integrity verified with an Agilent 2100 Bioanalyzer.

2.3 Sequencing and data processing

Sequencing was performed by the Beijing Genomics Institute (BGI) and data was processed at Leiden University Medical Center (LUMC). DNA libraries were subjected to paired-end sequencing (2 x 150 bp) on the Illumina HiSeq X-Ten platform while RNA libraries derived from rRNA-depleted total RNA were subjected to paired-end sequencing (2 x 100 bp) on the Illumina HiSeq 4000 platform. Raw reads (WGS, RNA-seq) were processed using in-house pipelines and clean reads were aligned to human reference genome Hg38 (Supporting Information Table 2). WGS and RNA-seq data have been deposited in the European Genome-Phenome Archive (EGA) under study number EGAS00001003660.

2.4 Detection of genomic rearrangements and fusion transcripts

Detection of structural genomic variants (SV) was performed using an in-house pipeline that included three structural variant callers (Breakdancer-max v1.4.4, CleverSV v2.0rc3 and Delly v0.6.7) (Supporting Information Table 3). SV calls were manually verified and curated using the Integrative Genomic Viewer (IGV, v2.3.78). The expression of fusion transcripts was investigated in four BPDCN samples with available RNA-seq data using an in-house pipeline that included two fusion transcript callers (FusionCatcher v0.99.6a and Star Fusion v0.8.0)

(Supporting Information Table 4). Fusion transcript calls were contrasted with genomic SV data and visually verified on DNA level using IGV. Rearranged genes implicated in cancer were identified using the Network of Cancer Genes 6.0 (NCG 6.0)[21] and literature search.

2.5 Detection of CNAs

Control-FREEC was used to identify CNAs using a 50 Kb window. The output was subjected to a Wilcoxon rank test and a Kolmogorov-Smirnov test to generate a list of genomic regions with statistically supported CNAs, and the latter were intersected to define recurrent events. As a complement, GISTIC2.0 analysis was performed with segment files derived from Control-FREEC (Supporting Information Figure 1, Supporting Information Table 5). Recurrent CNAs derived from both analyses were visually verified and curated using IGV (Supporting Information Table 6). Only CNAs confirmed after inspection are reported.

2.6 Detection of SNVs and indels

SNVs and indels were detected using GATK (v3.5). Variants filtered out of the output included SNVs present in dbSNP, SNVs with >5% minor allele frequency in the Exome aggregation consortium (ExAC) database, and SNVs present in an in-house cohort formed by eleven healthy volunteers. We searched for deleterious SNVs in 1271 genes taking part in signal transduction pathways, the cell cycle, hematopoietic cell differentiation and cellular processes previously identified as mutated in BPDCN.[8-10] Gene lists were retrieved from the PathCards database (http://pathcards.genecards.org/). Only SNVs predicted to produce highly deleterious amino acid substitutions by both SIFT and PolyPhen-2 were further investigated on ClinVar, COSMIC, Varsome and literature.

2.7 Differential expression analysis

TMM normalization was applied to raw count data using EdgeR and differential expression (DE) analysis was performed using Limma-Voom. BPDCN samples were compared to a control group formed by resting pDCs. RNA-seq data of pDC controls were generated by Alculumbre et al. and downloaded from GEO (https://www.ncbi.nlm.nih.gov/geo/; GSE84204).[22] Genes with a FDR < 0.01 were considered differentially expressed.

2.8 Gene ontology annotation and transcriptome analysis

Gene ontology annotation of rearranged genes and pathway analysis with DE genes were performed using DAVID (v6.8) using default settings.[23] Gene set enrichment analysis (GSEA, v2.2.4) was run as a pre-ranked analysis with 1000 permutations using gene set KEEG from the Molecular Signatures Database (MSigDB).[24] Normalized enrichment scores (NES) were calculated to determine expression signatures. FDR q values were obtained. Additional functional

characterization of the transcriptome (i.e. ENCODE and CORUM databases) was performed using Enrichr.[25]

2.9 Validation of structural genomic alterations and small-scale mutations
Select genomic rearrangements, interstitial deletions, indels and SNVs were validated by Sanger sequencing. Briefly, target sequences were amplified by PCR, run on a 1% agarose gel, column-purified and sequenced on the Applied Biosystems ABI3730xl platform. Validation of select CNAs was performed using Bio-Rad QX200 droplet digital PCR (ddPCR) system. ddPCR experiments were carried out following the manufacturer's guidelines. Briefly, 20-40 ng of genomic DNA was mixed with a frequent-cutting restriction enzyme, ddPCR supermix, probes against the gene of interest (FAM-labeled) and probes against the reference gene (HEX-labeled). The reaction mix was then partitioned into 20.000 nanodroplets, and subsequently, subjected to the following cycling program: 95 °C for 10 min, 39 cycles of 94 °C for 30 seconds and 60 °C for 1 min, and 98 °C for 10 min. Copy number values were determined using Bio-Rad Quantasoft software v1.7.4. Copy number reported for *IKZF1* (at exons 1, 5 and 8) in samples from the validation cohort are the average of 3 independent measurements using different reference genes.

3. RESULTS

3.1. Landscape of genomic rearrangements

BPDCN displayed a heterogenous landscape of rearrangements. We detected a total of 86 rearrangements in ten BPDCN genomes (Figure 1). The number of events per patient ranged from 1 to 18 (mean/patient ± SD, 9 ± 5) (Figure 2A, Supporting Information Figure 2). Intrachromosomal events constituted 66% of all rearrangements (range/patient, 0%-100%) (Figure 2B). We observed events joining either two annotated genes (32% of events), a gene with a nongenic region (36%), two nongenic regions (27%), or rearranging sequences within a single gene (5%) (Figure 2C).

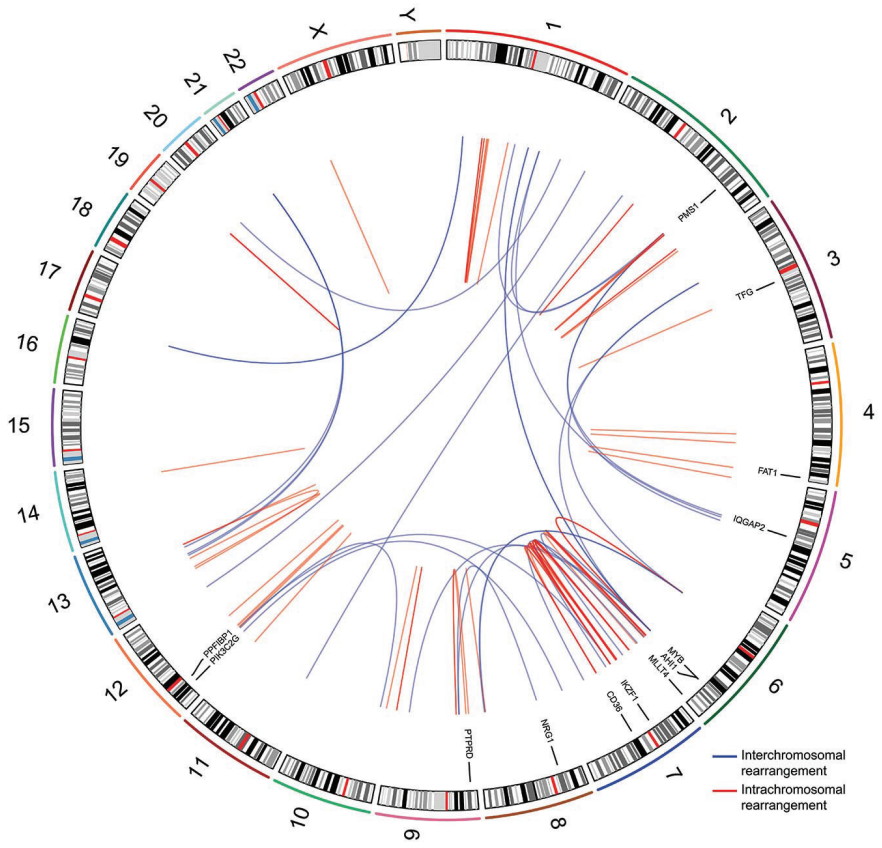


Figure 1. Landscape of genomic rearrangements in BPDCN. Circos plot showing 86 genomic rearrangements detected in 10 BPDCN genomes by WGS. The outer ring is formed by human chromosome ideograms arranged circularly end to end. The area at the center of the plot contains arcs representing interchromosomal (blue) and intrachromosomal (red) events. The ring between the chromosome ideograms and the arcs contains rearranged genes with established roles in cancer.

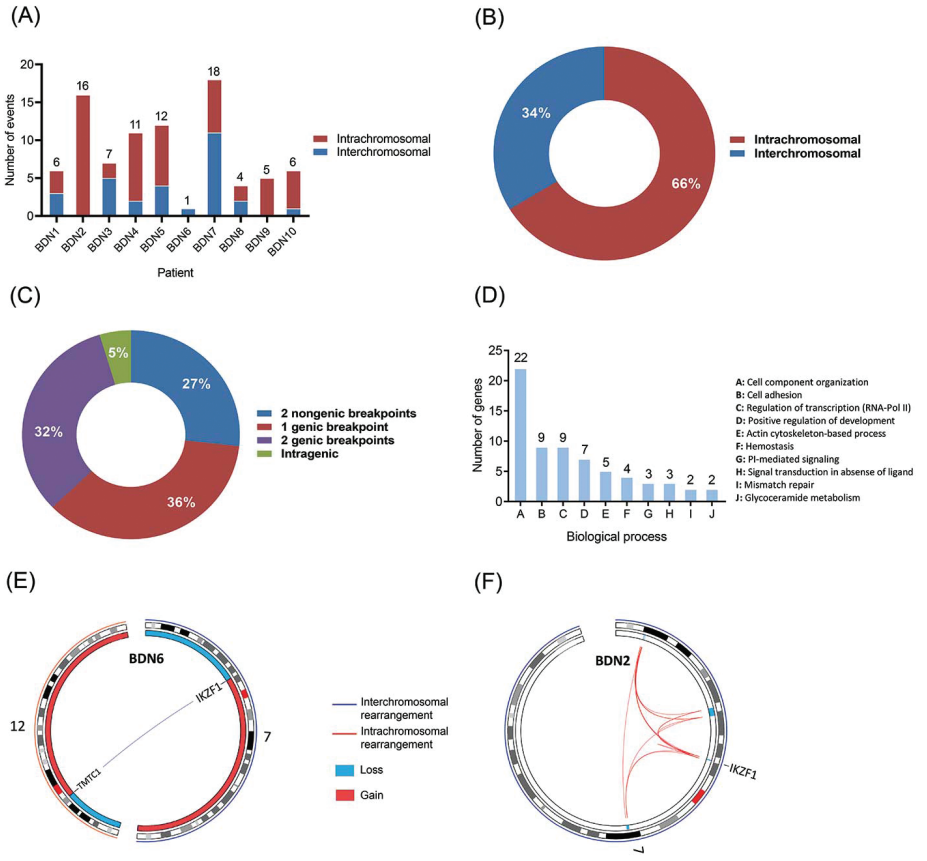


Figure 2. Distribution of genomic rearrangements and gene ontology annotation of rearranged genes in BPDCN. A, Number of genomic rearrangements per patient. The distribution of inter- and intrachromosomal rearrangements per patient is shown too; B, Distribution of inter- and intrachromosomal rearrangements (cohort); C, Distribution of genomic rearrangements based on the type of DNA sequences (genic, nongenic) involved in the event (cohort); D, Distribution of rearranged genes according to the biological process their encoded protein participates in; E, Circos plot showing interchromosomal rearrangement t(7;12)(p12;p11) in patient BDN6 which fuses *TMTC1* with *IKZF1*, causing the inactivation of the latter. F, Circos plot displaying a chromothripsis-like event in patient BDN2 which mediated the loss of several focal regions in chromosome 7, including a narrow area (450 Kb) containing *IKZF1*.

Genomic rearrangements disrupted a total of 54 genes across ten patients (Supporting Information Table 7). Out of 54 rearranged genes, 13 are implicated in cancer at present (Figure 1, Supporting Information Table 8). The latter group included genes implicated in hematological malignancies (i.e. *AHI1*, *CD36*, *IKZF1*, *MLLT4*, *MYB*, *TFG*) and other neoplasms (i.e. *FAT1*, *IQGAP2*, *NRG1*, *PIK3C2G*, *PMS1*, *PPFIBP1*, *PTPRD*). Gene ontology analysis showed that rearranged genes

primarily encode proteins that participate in cellular component organization (esp. cytoskeleton-related processes), cell adhesion and transcriptional regulation (Figure 2D, Supporting Information Table 9).

We observed two recurrently rearranged genes, *MYB* and *TMTC1*, both of them in two of ten patients. The type of *MYB* rearrangements found in our patients (BDN3 and BDN10) have been previously described in BPDCN and result in a gain-of-function effect due to the loss of the 3'-end of the gene which contains conserved target sites for inhibitory microRNAs (miRNAs).[16] In contrast, the observed *TMTC1* rearrangements do not seem to have analogous functional consequences (i.e. dissimilar breakpoint positions, loss vs. gain of DNA, genic vs. nongenic partner sequences). We detected additional rearrangements reported in cancer before. An oncogenic rearrangement involving *SUPT3H*, found in patient BDN1, has been previously described in BPDCN.[18] Also, fusion gene *NAV2-TCF7L1*, found in patient BDN7, is recurrent in colorectal cancer.[26]

Notably, we identified rearrangements leading to *IKZF1* inactivation. Fusion gene *TMTC1-IKZF1*, found in patient BDN6, had a truncated coding sequence (CDS) encoding a peptide constituted only by the first 77-AA of TMTC1 due to the formation of a premature stop codon; thus, preventing the synthesis of a functional IKZF1 protein (Figure 2E, Figure 5C). Another patient, BDN2, underwent a chromothripsis-like event in chromosome 7 that resulted in the deletion of several focal areas within 7p and 7q, including a narrow region (450 Kb) containing *IKZF1* (Figure 2F).

3.2. Landscape of CNAs

BPDCN exhibited a predominance of large-scale chromosomal imbalances. Broad numerical aberrations (>3 Mb) included monosomy 9 (four patients), as well as deletions within 3p (two patients), 6q (five patients), 7p (three patients), 9p (two patients), 9q (three patients), 12p (three patients), 13q (six patients), 15q (two patients) and Y (two patients). Also, broad gains within 1q (three patients), 7q (two patients) and 12q (two patients) (Figure 3, Figure 4).

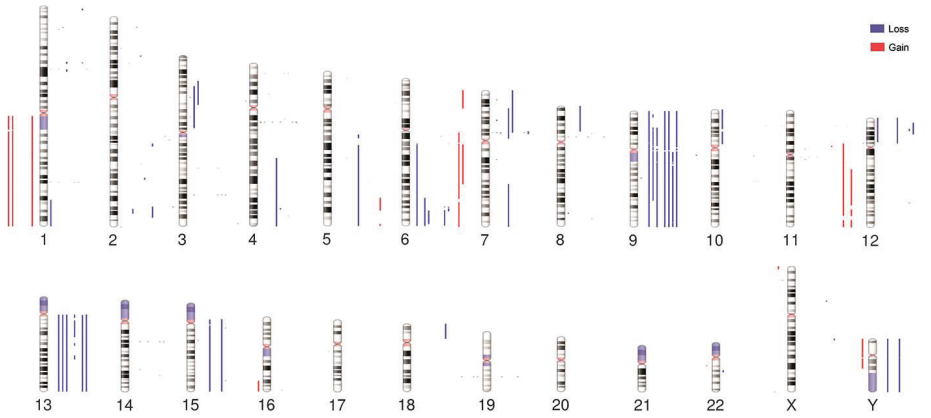


Figure 3. Landscape of CNAs in BPDCN. Human chromosome ideograms displaying regions of gain and loss detected through WGS in 10 BPDCN genomes. Blue bars to the right of the chromosomes depict regions of loss whereas red bars to the left of the chromosomes depict regions of gain.

We detected seven focal (≤3 Mb) minimal common regions (MCRs) between CNAs affecting our sequenced patients (Supporting Information Table 6). Out of seven focal MCRs, six contained either *bona fide* or candidate cancer genes. These focal MCRs were deletions at 9q21.32, 7p12.2, 13q14.2, 6q25.1, 12p13.1 and 7p14.1-p14.2 observed in eight, seven, six, five, four and three out of ten sequenced patients, respectively (Figure 4). Cancer genes residing within these regions are mainly involved in hematopoiesis and cell cycle regulation.

The most common focal MCRs were deletions at 9q21.32 and 7p12.2 which contained, respectively, tumor suppressors *HNRNPK* and *IKZF1*. *HNRNPK* encodes a nuclear ribonucleoprotein that influences multiple cellular functions (e.g. cell cycle, chromatin remodeling, RNA synthesis and processing) in hematopoietic cells.[27] *IKZF1*, on the other hand, encodes a transcription factor with critical roles in the development of various hematopoietic lineages, including pDCs.[28-31] Remarkably, deletions at 7p12.2 were considerably narrow (<500 Kb) in four of seven affected patients, ranging from loss of *IKZF1* and a few neighboring genes, to loss of the *IKZF1* locus exclusively, to loss of *IKZF1*'s exon 8 only (Figure 5A,B). Patients where deletions did not involve the whole locus, but part of it, enclosed exons encoding IKZF1's functional domains (Figure 5B). For instance, the loss of exon 8 in patient BDN1 generated a CDS encoding an IKZF1 protein that lacked its C-terminal dimerization domain, which is essential for its functionality (Figure 5D).

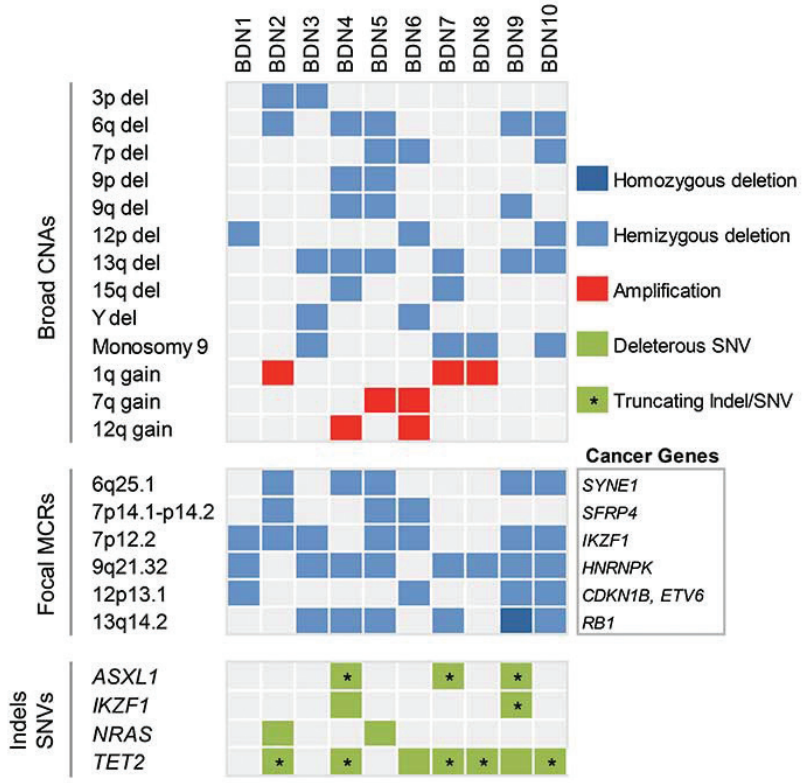


Figure 4. Distribution of recurrent CNAs and deleterious indels/SNVs in BPDCN. Upper panel: large CNAs (>3 Mb). Middle panel: focal MCRs (≤3 Mb) between CNAs; reputable tumor suppressors enclosed within focal MCRs are specified. Bottom panel: Indels and SNVs either leading to protein truncations, reported as pathogenic in literature or predicted as deleterious (SIFT and PolyPhen-2) are shown. Only genes mutated in more than one patient are indicated.

Deletions at 13q14.2 and 6q25.1 encompassed, respectively, *RB1* and *SYNE1*. *RB1* encodes an established negative regulator of the cell cycle whereas candidate tumor suppressor *SYNE1*, which is often silenced by hypermethylation in lung cancer[32], encodes a cytoskeleton-associated protein that helps maintain subcellular spatial organization.

Deletion at 12p13.1 included tumor suppressors *CDKN1B*, *DUSP16* and *ETV6* which play roles in the cell cycle, signal transduction and hematopoiesis, respectively. Finally, deletion at 7p14.1-p14.2 enclosed tumor suppressor *SFRP4*, a negative regulator of WNT signaling.

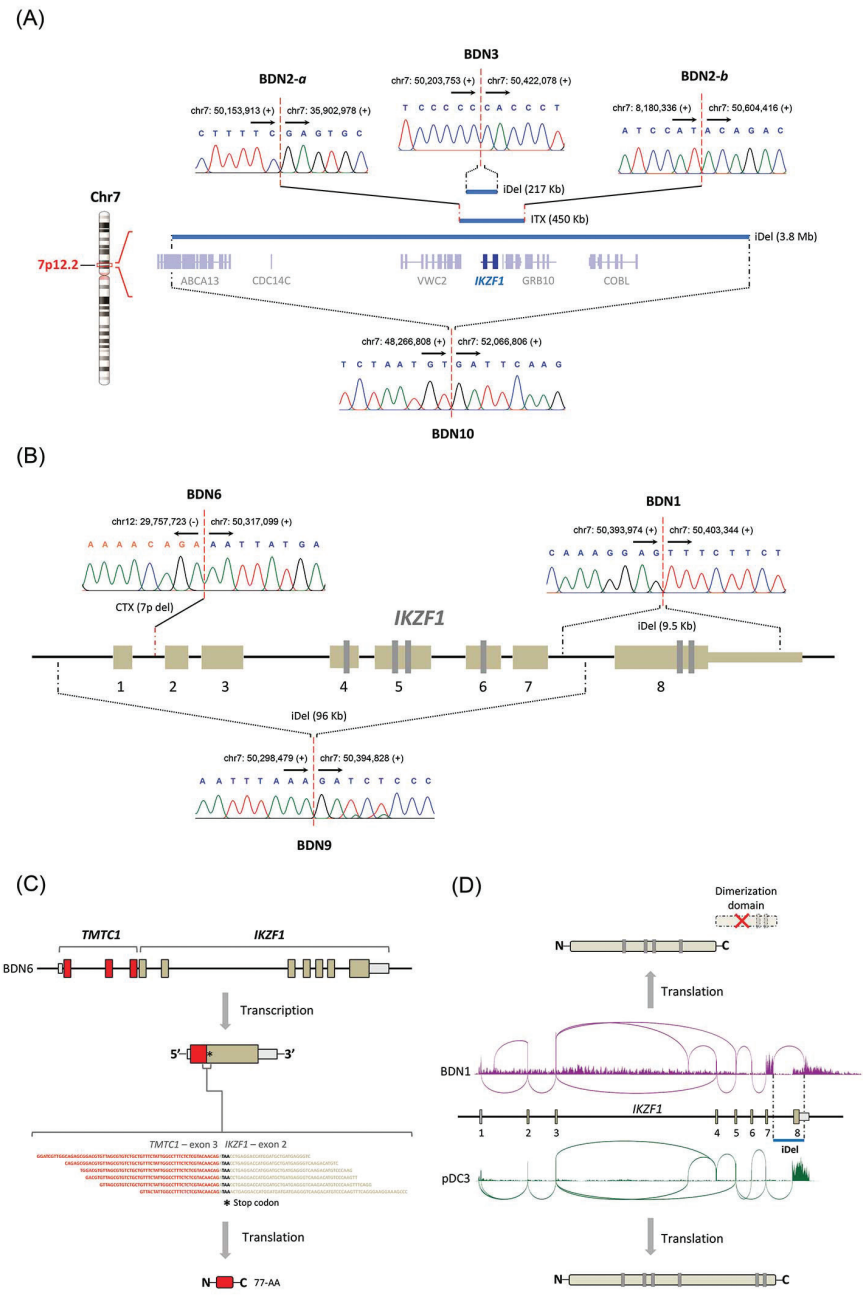


Figure 5. Interstitial deletions and rearrangements inactivate *IKZF1* in BPDCN. A, Diagram representing the hemizygous loss of *IKZF1* (217 Kb-3.8 Mb) in BPDCN tumors. Blue bars depict deleted areas at 7p12.2 in tumors from three patients with BPDCN. Monoallelic *IKZF1* deletion occurred through narrow interstitial deletions in patients BDN3 and BDN10, and through a complex intrachromosomal rearrangement in patient

BDN2 (see also Figure 2F); B, Diagram showing inactivating structural changes at the *IKZF1* locus in BPDCN tumors. These events included narrow interstitial deletions (BDN1/9) and an unbalanced interchromosomal rearrangement (BDN6). The breakpoints of all *IKZF1* -inactivating structural alterations were validated by Sanger sequencing. Genomic coordinates of breakpoints according to reference genome Hg38. Arrows indicate the direction toward which genomic coordinate numbers increase. Plus (+) and minus (−) signs specify strand polarity; C, The fusion between *TMTC1* and *IKZF1* in patient BDN6 generated a truncated coding sequence (77-AA) due to the formation of a premature stop codon (shown in RNA-seq reads), rendering IKZF1 inactive on protein level. D, Sashimi plots showing RNA expression at *IKZF1* in patient BDN1 (purple) and control sample pDC3 (green). Due to a focal interstitial deletion (blue bar), the transcript of *IKZF1* in patient BDN1 lacks exon 8, and instead includes a readthrough sequence starting from exon 7 toward the remaining part of the 3′-UTR and beyond. The resulting transcript encodes a functionally defective *IKZF1* protein that lacks its dimerization domain.

3.3. Small-scale mutations

We looked for small-scale mutations (i.e. indels, SNVs) with potential pathogenic relevance in exonic sequences of genes involved in the cell cycle, hematopoietic cell differentiation, signal transduction (i.e. JAK-STAT, MAPK, NF-κB and PI-3-K/Akt signaling pathways) and genes that have been previously found to be mutated in BPDCN (Supporting Information Table 10).

Two gain-of-function SNVs, *BRAF* (p.G464V[33]) and *NRAS* (p.E63K[34]), were observed in patient BDN5. *NRAS* (p.G12A[35]), a SNV associated with melanoma and lung cancer, was found in patient BDN2. Truncating frameshift mutations in *ASXL1* (p.E635Rfs*15, p.V624Gfs*8) were observed in BDN4/9 and BDN7, respectively. Also, seven patients carried indels and SNVs leading either to premature stop codons or amino acid substitutions predicted as deleterious in *TET2*. (Figure 4, Supporting Information Figure 3, Supporting Information Figure 4). Of note, we found mutations in *IKZF1* (i.e. BDN9: p.H421Afs*68, BDN4: p.H480Y) that impacted the dimerization domain of this gene. The former mutation generates an IKZF1 protein with a truncated dimerization domain while the latter leads to an amino acid substitution predicted as deleterious (Supporting Information Figure 3, Supporting Information Figure 4).

Finally, we found twenty-four additional patient-specific SNVs of unknown significance in cancer-associated genes, including genes previously implicated in BPDCN (i.e. *ARID1A*, *ASH1L*, *ASXL3*, *ATR*, *CBLC*, *CHD8*, *EP400*, *ETV6*, *PHC2*, *RAD52*, *SMARCD1*, *SRSF2* and *ZRSR2*) (Supporting Information Table 11).

3.4. Differentially expressed genes

Prior research using array-based gene expression profiling showed that BPDCN cells appear to originate from pDC precursors that are phenotypically close to resting pDCs.[7] Therefore, to determine differentially expressed (DE) genes

in the disease, we compared gene expression in BPDCN with gene expression in resting pDCs. The analysis uncovered 1827 DE genes (1191 up-regulated, 636 down-regulated, FDR <0.01) (Figure 6A, Supporting Information Table 12), 272 of which are currently associated with cancer (213 up-regulated, 64 down-regulated) (Supporting Information Table 13). Up-regulated genes included (proto-) oncogenes with roles in apoptosis (i.e. *BCL2*, *TERT*), cell cycle regulation (i.e. *FOXM1*), hematopoietic cell development (i.e. *AFF3*, *MYB*) and signal transduction (i.e. *BCL9*, *FGFR1*, *MET*, *PDGFRA*, *PDGFRB*, *PLCG1*, *RRAS2*) (Figure 6B). Prominent down-regulated genes included cell cycle inhibitors (i.e. *BTG2*, *IFIT3*, *TSPYL2*, *TXNIP*) and tumor suppressors implicated in hematological malignancies (i.e. *PTPRO*, *SAMD9L*) (Figure 6B).

Transcriptome analysis with Enrichr using ChIP-seq data from ENCODE revealed that genes regulated by EZH2, TCF12 and FOXM1 are greatly overrepresented in the group of up-regulated genes (Figure 6C, Supporting Information Table 14). EZH2 is part of polycomb repressive complex 2 (PRC2), a gene silencing protein complex that influences various cellular processes, including hematopoietic cell differentiation.[36] Transcription factor TCF12, on the other hand, participates in T-cell differentiation and is typically lowly expressed in mature pDCs.[37] Finally, transcription factor FOXM1, which is overexpressed in numerous cancers, takes part in multiple biological processes (e.g. apoptosis, cell cycle, DNA damage response).[38]

Since *IKZF1* deletion was the most focal recurrent CNA in our sequenced cohort, we next looked at the expression of a group of evolutionarily conserved IKZF1 target genes that are associated with the development of BCR-ABL1+ pre-B acute lymphoblastic leukemia (pre-B ALL)[39,40] (Supporting Information Table 15), an aggressive hematological cancer characterized by *IKZF1* inactivation. We detected 29 DE cancer-associated IKZF1 target genes (25 up-regulated, 4 down-regulated) in BPDCN (Figure 6D). The expression of all IKZF1 target genes matched perfectly the expected consequence of IKZF1 deficiency; that is, genes repressed by IKZF1 were up-regulated while genes activated by IKZF1 were down-regulated. Genes of interest in the group of up-regulated IKZF1 targets included among others, *CCND1*, *CDC42BPB*, and *IFITM3*. Cyclin D1 (encoded by *CCND1*) promotes G1/S transition in the cell cycle and has been previously found to be overexpressed in BPDCN.[7] *CDC42BPB* encodes a serine/threonine kinase with roles in cytoskeletal reorganization and cell migration.[41] Interestingly, overexpression of this kinase has been shown to facilitate cancer cell invasion.[42] Finally, *IFITM3* encodes an antiviral membrane protein that has been shown to trigger CD19-mediated PI-3-K/Akt signaling in pre-B ALL.[43]

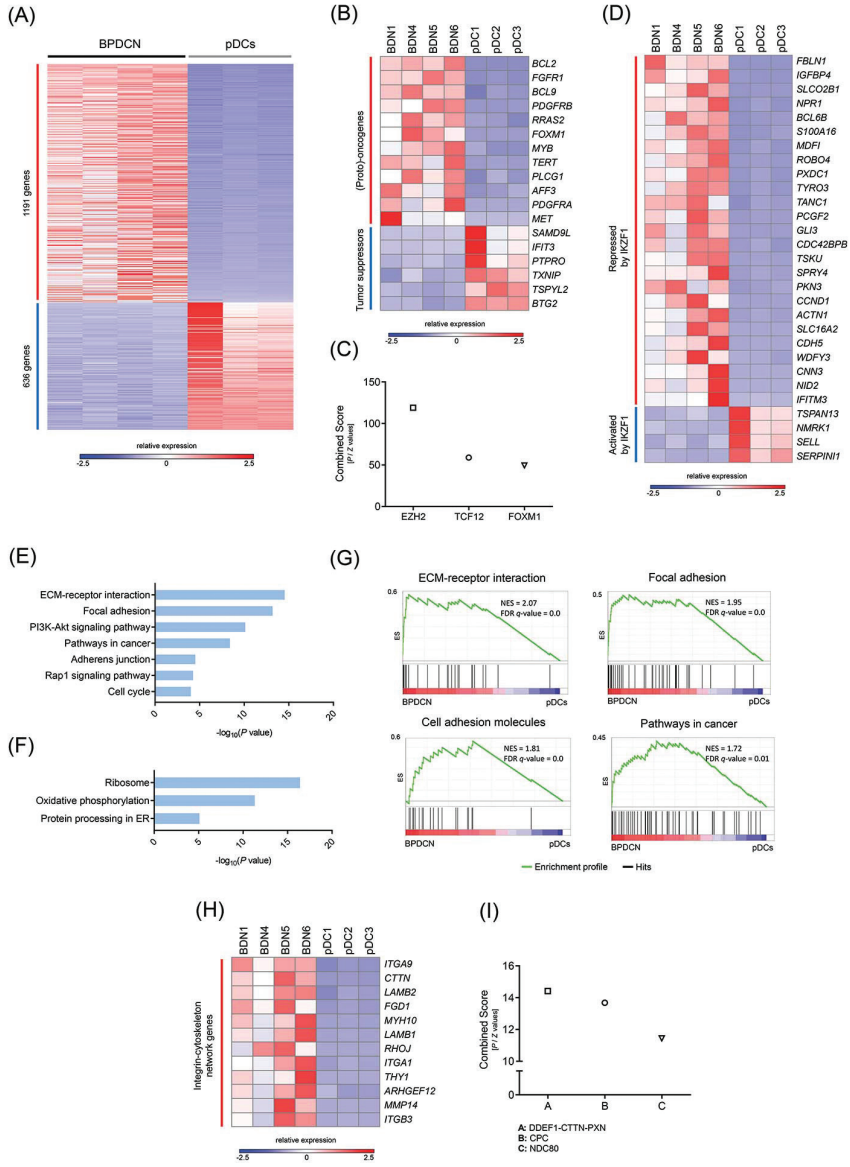


Figure 6. RNA-seq supports a loss-of-IKZF1 phenotype in BPDCN and uncovers aberrant cell adhesion in the disease. A, Heat map showing 1827 differentially expressed genes (1191 up-regulated, 636 down-regulated, FDR < 0.01) in BPDCN when compared to resting pDCs. B, (Proto)-oncogenes and tumor suppressors involved in cell cycle regulation, hematopoietic cell differentiation and signal transduction are differentially expressed in BPDCN. C, Genes regulated by EZH2, TCF12, and FOXM1 are the most overrepresented in the group of up-regulated genes. D, Conserved IKZF1 target genes follow a loss-of-IKZF1 expression pattern in BPDCN; that is, genes repressed by IKZF1 are up-regulated while genes activated by IKZF1 are down-regulated. E, F, Pathway

analysis using DAVID uncovered up-regulation of cellular processes responsible for cell-cell and cell-ECM interactions, the PI-3-K/Akt pathway, the Rap1 pathway and the cell cycle. Down-regulated profiles included ribosome, oxidative phosphorylation and protein processing in the endoplasmic reticulum (ER) (See Table S16 for a complete list of enriched terms/processes). G, Gene set enrichment analysis (GSEA) confirmed that BPDCN is characterized by the up-regulation of adhesion signatures. NES, normalized enrichment score; FDR q -value, false discovery rate q -value. (See Table S17 for a complete list of GSEA signatures). H, Genes encoding components of the integrin-cytoskeleton network, which have been shown to undergo up-regulation due to IKZF1 deficiency in BCR-ABL1+ pre-B ALL, are up-regulated in BPDCN as well. I, Protein complex DD-EF1-CTTN-PXN is predicted to exist in BPDCN cells based on transcriptome data. This complex has only been found in neoplastic cells with an invasive phenotype.

3.5. Deregulated processes/pathways

We performed pathway analysis with DAVID to identify deregulated cellular processes/pathways in BPDCN. The analysis uncovered up-regulation of several adhesion profiles: ECM-receptor interaction ($P = 2.51x10^{-15}$), focal adhesion ($P = 5.61x10^{-14}$), adherens junction ($P = 2.64x10^{-5}$) and Rap1 signaling ($P = 4.58x10^{-5}$). BPDCN also displayed up-regulation of PI-3-K/Akt signaling ($P = 6.46x10^{-11}$) and the cell cycle ($P = 8.22x10^{-5}$) (Figure 6E, Supporting Information Table 16). Down-regulated cellular processes included ribosome synthesis ($P = 3.34x10^{-17}$), oxidative phosphorylation ($P = 4.11x10^{-12}$) and protein processing in the endoplasmic reticulum (ER) ($P = 6.98x10^{-6}$) (Figure 6F, Supporting Information Table 16). To complement DAVID analysis, we performed GSEA using gene set KEEG from MSigDB. GSEA confirmed that BPDCN is characterized by the up-regulation of adhesion signatures (Figure 6G, Supporting Information Table 17). Importantly, these deregulated profiles support an IKZF1-deficient phenotype of BPDCN cells, as previous studies performed by others have established that a hallmark of IKZF1 deficiency in immune cells (i.e. pDCs, pre-B cells) is the overexpression of proteins responsible for cell-cell and cell-ECM interactions.[29,44]

Additionally, we identified a group of up-regulated genes encoding components of the integrin-cytoskeleton network that have been proven to undergo up-regulation as a result of IKZF1 deficiency in murine models of pre-B ALL (Figure 6H).[44] One of these genes, *THY1* (CD90), encodes a membrane glycoprotein involved in integrin signaling and adhesion. Notably, THY1 is one of the key surface molecules underlying the highly adhesive phenotype of IKZF1-deficient malignant cells in BCR-ABL1+ pre-B ALL.[45,46] Another gene, *CTTN*, encodes cortactin, a cytoplasmic protein that participates in the formation of podosomes in motile cells. Cortactin overexpression occurs in a number of aggressive cancers and promotes the formation of invadopodia (hyperactive podosomes) which make malignant cells highly invasive and migratory.[47] Interestingly,

transcriptome analysis with Enrichr using CORUM database predicted the existence of protein complex DDEF1-CTTN-PXN in BPDCN cells (Figure 6I, Supporting Information Table 18). This cytoskeleton-associated protein complex has been found exclusively in malignant cells with highly invasive behavior (breast cancer), and not in their non-invasive or normal counterparts.[48]

3.6. Structural alterations involving *IKZF1* are recurrent in BPDCN

Since the findings derived from our genome and transcriptome analyses strongly pointed at *IKZF1* as a central player in the pathogenesis of BPDCN, we next used ddPCR to assess the copy number of *IKZF1* in tumors from 15 additional patients. Because our genomic data showed that *IKZF1*-disrupting unbalanced rearrangements and deletions did not affect the entire locus in all patients, but deleted only the 5'-end or the 3'-end of *IKZF1* in some of them (Figure 5B, Figure 7A), we evaluated copy number at three different exons (i.e. 1, 5 and 8). Exons 1, 5 and 8 encode, respectively, the leader sequence (5'-UTR), two of the four Zn fingers forming the DNA-binding domain, and the dimerization domain of *IKZF1*.

CNAs involving *IKZF1* were detected in 6 of 15 (40%) patients from the extension cohort, 5 of which had deletions (Figure 7B). Interestingly, one patient (BDN15) displayed normal copy number at exon 1 and 8, but a gain at exon 5, suggesting a focal duplication of one or more exons within the *IKZF1* locus (Figure 7B). Of note, focal exon duplications in tumor suppressors can generate inactivating frameshifts, as documented in *BRCA1*.[49] Overall, combining the sequenced and extension cohorts, 13 of 25 patients (52%) had structural alterations involving *IKZF1*.

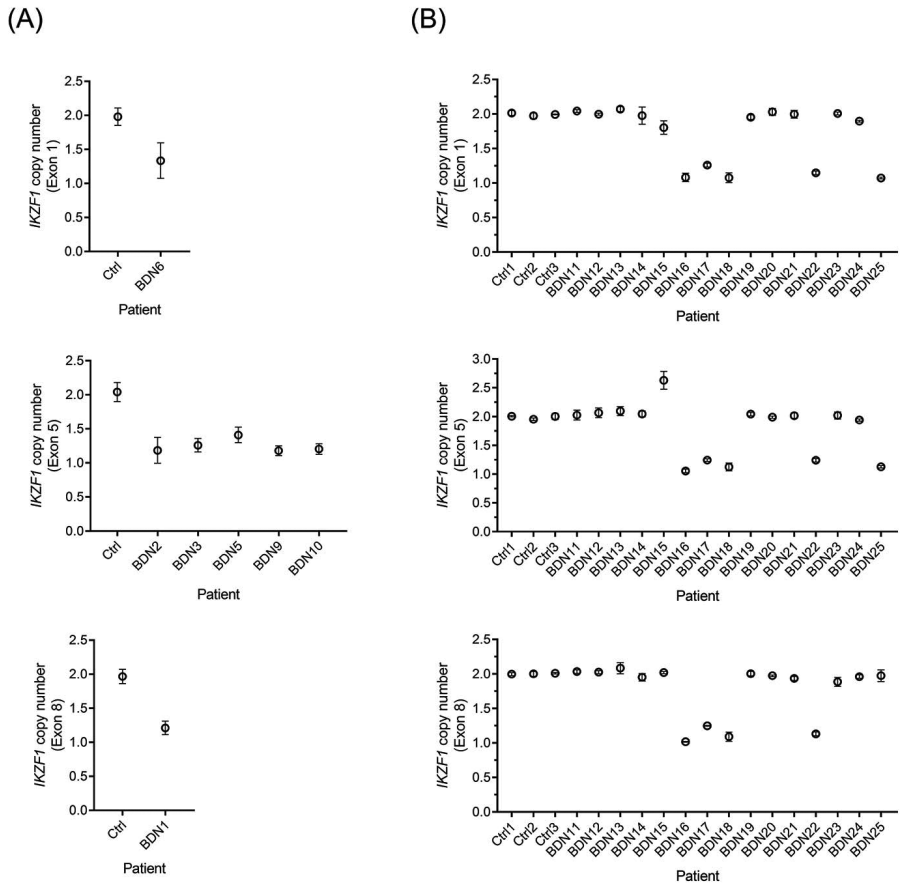


Figure 7. Evaluation of an extension cohort confirms recurrent structural alterations at IKZF1 in BPDCN. A, *IKZF1*-deleting events detected by WGS in BPDCN tumors were validated using ddPCR. Copy number was evaluated at exon 1, 5, or 8 according to the region of *IKZF1* found to be deleted by WGS in each affected patient (ie, exon 1:5′-end; exon 5: whole locus; exon 8:3′-end). B, The existence of recurrent structural alterations at *IKZF1* was confirmed in an extension cohort by ddPCR. Copy number was evaluated at three different exons in all samples to account for CNAs affecting only the 5′-end (exon 1), the middle section (exon 5) or the 3′-end (exon 8) of *IKZF1*. Exons 5 and 8 encode two of the four Zn fingers forming the DNA-binding domain and the dimerization domain of *IKZF1*, respectively, being both critical for its functionality. Ctrl, normal T cells.

4. DISCUSSION

In this study we describe the first high resolution whole-genome analysis of BPDCN using WGS. Our analysis, which focused on structural genomic alterations, identified 54 genes disrupted by rearrangements which are

predominantly involved in cytoskeleton-associated processes, adhesion and transcriptional regulation. Out of 54 rearranged genes, 13 are involved in malignancies at present, being *IKZF1* and *MYB* the only cancer genes recurrently impacted by rearrangements in our sequenced cohort. We identified 6 recurrently deleted genomic regions enclosing recognized tumor suppressors (i.e. *CDKN1B*, *ETV6*, *HNRNPK*, *IKZF1*, *RB1* and *SFRP4*) with important roles in hematopoiesis and cell cycle regulation. In addition, deleterious indels and SNVs in *ASXL1*, *IKZF1*, *NRAS* and *TET2* were recurrent in our patients.

IKZF1 inactivation, which mainly occurred through narrow interstitial deletions (9.5 Kb – 3.8 Mb) and unbalanced rearrangements, was the most prominent alteration in BPDCN due to its focality, recurrence and likely adverse effects during early development of pDCs. The latter is supported by the fact that *IKZF1* deficiency keeps pDC precursors stuck at an early differentiation stage whereas function-impairing changes in both *IKZF1* alleles lead to absence of pDCs in murine models.[29,31] Using an independent cohort, we confirmed by ddPCR that structural alterations at the *IKZF1* locus are recurrent in BPDCN. Although the incidence of CNAs involving *IKZF1* was lower in the extension cohort compared to the sequenced cohort (difference not statistically significant, $P > 0.05$, Fisher's exact test), patients with normal *IKZF1* copy number might still carry inactivating balanced rearrangements, indels or SNVs which do not entail numerical changes. In agreement with the recurrent *IKZF1* inactivation observed in BPDCN, transcriptome analysis revealed that conserved *IKZF1* target genes show a clear loss-of-*IKZF1* expression pattern.

Even though the roles of *IKZF1* in hematopoiesis and leukemogenesis are only starting to be understood, it is currently well-established that insufficiency of functional *IKZF1* proteins leads to overexpression of adhesion molecules, surface receptors and ECM proteins in immune cells, as demonstrated in murine pDC precursors carrying hypomorphic *Ikzf1* alleles (e.g. *Igll1*, *Mmp14*, *Ncam1* (*Cd56*), *Thy1*, etc.),[29] murine B cell precursors engineered to produce *Ikzf1* proteins lacking DNA-binding ability (e.g. *Itga9*, *Itgb3*, *Lamb1*, *Mmp14*, etc.)[44] and malignant cells derived from murine models of BCR-ABL1+ pre-B ALL carrying loss-of-function mutations in *Ikzf1* (e.g. *Cd28*, *Itga5*, *Thy1*, etc.).[45,46] Our data, in full concordance with this evidence, suggest that *IKZF1* inactivation lies behind the up-regulation of adhesion profiles in BPDCN.

BPDCN displayed up-regulation of the PI-3-K/Akt pathway and the Rap1 pathway. The PI-3-K/Akt pathway, often overactivated in human cancers, promotes cancer cell proliferation. Notably, IKZF1 has been shown to transcriptionally repress components of the PI-3-K/Akt pathway (e.g. *PIK3CD*, *PIK3C2B*, *PIP4K2B*, etc.),[50] and concordantly, IKZF1-deficient malignant cells from murine models of pre-B

ALL show overactivation of PI-3-K/Akt signaling.[44] Hence, IKZF1 deficiency in BPDCN might underlie or contribute to the observed overactivation of the PI-3-K/Akt pathway. In contrast, the Rap1 pathway regulates cell motility and integrin-mediated adhesion in leukocytes, lymphocytes and dendritic cells, both of which are processes hijacked by neoplastic cells to disseminate to other body sites.[51] In fact, activation of Rap1 signaling has been shown to trigger migration of T-acute lymphoblastic leukemia and B-cell lymphoma cells, *in vitro* and *in vivo*, respectively.[52,53]

We found that genes silenced by PRC2 (EZH2) were the most overrepresented in the group of up-regulated genes in BPDCN. This agrees with a previous report of a loss-of-EZH2 expression signature in tumor cells from patients with highly aggressive BPDCN.[54] Remarkably, it has been established that IKZF1, which predominantly acts as a transcriptional repressor, silences genes in erythroid and T-cell precursors by recruiting PRC2 to target sites.[55,56] Thus, up-regulation of genes silenced by PRC2 in BPDCN might result from shortage of functional IKZF1 molecules, which in turn, would hinder PRC2 recruitment to target sites.

Interestingly, analysis of DE genes in BPDCN not only confirmed up-regulation of *BCL2* and *MYB* as described in previous studies,[7,16] but also uncovered additional up-regulated (proto-) oncogenes of interest. *AFF3*, for instance, encodes a transcriptional activator expressed in early lymphoid development which appears to be leukemogenic in patients with pre-B ALL carrying *MLL-AFF3* fusions.[57] *FOXM1*, on the other hand, promotes cell proliferation and dissemination in a plethora of neoplasms by inducing cell cycle progression (esp. at G2/M) and the acquisition of an invasive phenotype.[38] *RRAS2*, which encodes a small GTPase involved in PI-3-K/Akt signaling, has been found to be overexpressed in a number of lymphoid malignancies, including cutaneous T cell lymphoma (CTCL).[58] Other prominent up-regulated oncogenes encode receptors (i.e. *FGFR1*, *MET*, *PDGFRA/B*) that signal through the PI-3-K/Akt pathway.

In addition, we identified in BPDCN a group of down-regulated genes encoding cell proliferation inhibitors. BTG2, a member of the BTG/Tob family, stops the cell cycle at G1/S by down-regulating cyclin D1, or at G2/M by preventing binding between cyclin B1 and CDK1.[59] IFIT3 and TXNIP, which increase expression of p21/p27 and p27, respectively, promote cell cycle arrest at G0/G1.[60,61] TSPYL2 triggers the inhibition of several cell cycle kinases (i.e. CDK1,2,4,6), halting the cell cycle at multiple phases (G1 to M).[62]

It has not escaped our notice the apparent parallelism between BPDCN and BCR-ABL1+ pre-B ALL in terms of deregulated genes and cellular processes. We hypothesize that these similarities are rooted in the recurrent inactivation of *IKZF1* occurring in both neoplasms. This may also explain the positive effects of ALL-like therapy for the treatment of patients with BPDCN.

In summary, our data strongly supports *IKZF1* inactivation as a key event in the development of BPDCN. Firstly, *IKZF1* was found to be inactivated by structural alterations in seven of ten patients from our sequenced cohort, and recurrent *IKZF1* inactivation was confirmed in an extension cohort. Secondly, expression of conserved IKZF1 target genes in BPDCN was fully consistent with a loss-of-IKZF1 phenotype. Thirdly, expression signatures in BPDCN matched known physiological consequences of IKZF1 deficiency (i.e. aberrant adhesion, activation of PI-3-K/Akt pathway) in immune cells. Finally, since IKZF1 is critical for pDC differentiation, its abrogation can be expected to derail normal development of pDC precursors, conceivably leading to their malignization.

Based on our findings, we suggest that counteracting the pathogenic loss-of-IKZF1 phenotype of malignant cells in BPDCN through pharmacological induction of the wild-type *IKZF1* allele (e.g. by using retinoids)[45] might have beneficial therapeutic effects on patients with BPDCN.

Acknowledgements
This study was funded by the Dutch Cancer Society (grant UL2013-6104).

Conflict of interest
The authors declare no conflicts of interest.

Author contributions
A.N.B.T., D.F., E.B. and C.P.T. designed the project. A.N.B.T. and C.P.T wrote the manuscript. A.N.B.T., D.C. and H.M. performed the bioinformatic analyses. A.N.B.T. interpreted the data, performed the experiments and produced figures and tables. D.F., J.G., L.C., M.P., M.V., R.W. and E.B. provided valuable biological specimens. A.N.B.T, D.C., H.M., D.F., J.G., L.C., M.P., M.V., R.W., E.B. and C.P.T revised and approved the final manuscript.

REFERENCES

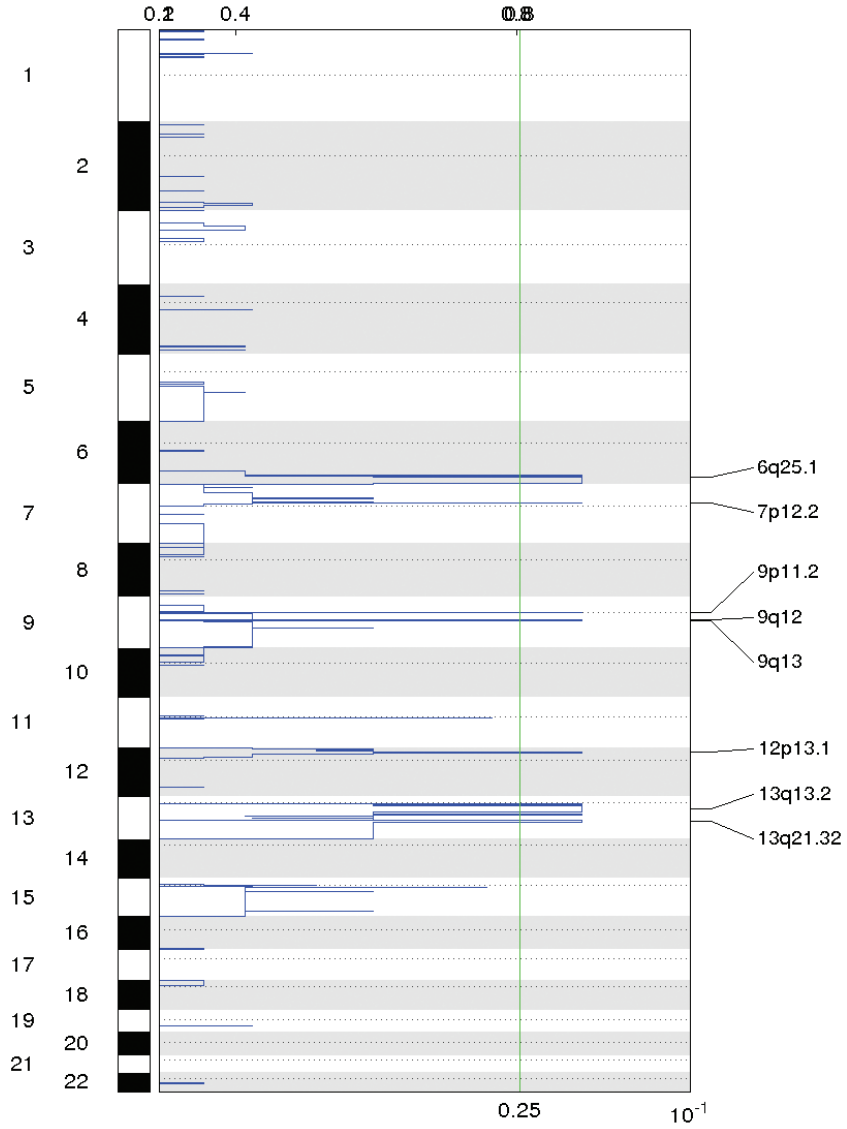
1. Pagano L, Valentini CG, Grammatico S, Pulsoni A. Blastic plasmacytoid dendritic cell neoplasm: diagnostic criteria and therapeutical approaches. *Br J Haematol*. 2016;174(2):188-202.
2. Willemze R, Jaffe ES, Burg G, et al. WHO-EORTC classification for cutaneous lymphomas. *Blood*. 2005;105(10):3768-3785.
3. Lucioni M, Novara F, Fiandrino G, et al. Twenty-one cases of blastic plasmacytoid dendritic cell neoplasm: focus on biallelic locus 9p21.3 deletion. *Blood*. 2011;118(17):4591-4594.
4. Wiesner T, Obenauf AC, Cota C, Fried I, Speicher MR, Cerroni L. Alterations of the cell-cycle inhibitors p27(KIP1) and p16(INK4a) are frequent in blastic plasmacytoid dendritic cell neoplasms. *J Invest Dermatol*. 2010;130(4):1152-1157.
5. Jardin F, Callanan M, Penther D, et al. Recurrent genomic aberrations combined with deletions of various tumour suppressor genes may deregulate the G1/S transition in CD4+CD56+ haematodermic neoplasms and contribute to the aggressiveness of the disease. *Leukemia*. 2009;23(4):698-707.
6. Dijkman R, van Doorn R, Szuhai K, Willemze R, Vermeer MH, Tensen CP. Gene-expression profiling and array-based CGH classify CD4+CD56+ hematodermic neoplasm and cutaneous myelomonocytic leukemia as distinct disease entities. *Blood*. 2007;109(4):1720-1727.
7. Sapienza MR, Fuligni F, Agostinelli C, et al. Molecular profiling of blastic plasmacytoid dendritic cell neoplasm reveals a unique pattern and suggests selective sensitivity to NF-kB pathway inhibition. *Leukemia*. 2014;28(8):1606-1616.
8. Stenzinger A, Endris V, Pfarr N, et al. Targeted ultra-deep sequencing reveals recurrent and mutually exclusive mutations of cancer genes in blastic plasmacytoid dendritic cell neoplasm. *Oncotarget*. 2014;5(15):6404-6413.
9. Menezes J, Acquadro F, Wiseman M, et al. Exome sequencing reveals novel and recurrent mutations with clinical impact in blastic plasmacytoid dendritic cell neoplasm. *Leukemia*. 2014;28(4):823-829.
10. Sapienza MR, Abate F, Melle F, et al. Blastic plasmacytoid dendritic cell neoplasm: genomics mark epigenetic dysregulation as a primary therapeutic target. *Haematologica*. 2019;104(4):729-737.
11. Alayed K, Patel KP, Konoplev S, et al. TET2 mutations, myelodysplastic features, and a distinct immunoprofile characterize blastic plasmacytoid dendritic cell neoplasm in the bone marrow. *Am J Hematol*. 2013;88(12):1055-1061.
12. Tokuda K, Eguchi-Ishimae M, Yagi C, et al. CLTC-ALK fusion as a primary event in congenital blastic plasmacytoid dendritic cell neoplasm. *Genes Chromosomes Cancer*. 2014;53(1):78-89.
13. Gao NA, Wang XX, Sun JR, Yu WZ, Guo NJ. Blastic plasmacytoid dendritic cell neoplasm with leukemic manifestation and ETV6 gene rearrangement: A case report. *Exp Ther Med*. 2015;9(4):1109-1112.
14. Cao Q, Liu F, Niu G, Xue L, Han A. Blastic plasmacytoid dendritic cell neoplasm with EWSR1 gene rearrangement. *J Clin Pathol*. 2014;67(1):90-92.
15. Toya T, Nishimoto N, Koya J, et al. The first case of blastic plasmacytoid dendritic cell neoplasm with MLL-ENL rearrangement. *Leuk Res*. 2012;36(1):117-118.
16. Suzuki K, Suzuki Y, Hama A, et al. Recurrent MYB rearrangement in blastic plasmacytoid dendritic cell neoplasm. *Leukemia*. 2017;31(7):1629-1633.
17. Boddu PC, Wang SA, Pemmaraju N, et al. 8q24/MYC rearrangement is a recurrent cytogenetic abnormality in blastic plasmacytoid dendritic cell neoplasms. *Leuk Res*. 2018;66:73-78.
18. Nakamura Y, Kayano H, Kakegawa E, et al. Identification of SUPT3H as a novel 8q24/MYC partner in blastic plasmacytoid dendritic cell neoplasm with t(6;8)(p21;q24) translocation. *Blood Cancer J*. 2015;5:e301.
19. Petrella T. FF, Pileri S.A. . Blastic plasmacytoid dendritic cell neoplasm. In: Swerdlow S.H. CE, Harris N.L., Jaffe E.S., Pileri S.A., Stein H., Thiele J., ed. *WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues*. Revised 4th edition ed. Lyon: IARC; 2017.

20. Petrella T. FF, Pileri S.A. Blastic plasmacytoid dendritic cell neoplasm. In: Elder D.E. MD, Scolyer R.A., Willemze R., ed. *WHO Classification of Skin Tumours* 4th edition ed. Lyon: IARC; 2018.
21. Repana D, Nulsen J, Dressler L, et al. The Network of Cancer Genes (NCG): a comprehensive catalogue of known and candidate cancer genes from cancer sequencing screens. *Genome Biol*. 2019;20(1):1.
22. Alculumbre SG, Saint-Andre V, Di Domizio J, et al. Diversification of human plasmacytoid predendritic cells in response to a single stimulus. *Nat Immunol*. 2018;19(1):63-75.
23. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc*. 2009;4(1):44-57.
24. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A*. 2005;102(43):15545-15550.
25. Kuleshov MV, Jones MR, Rouillard AD, et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res*. 2016;44(W1):W90-97.
26. Cancer Genome Atlas N. Comprehensive molecular characterization of human colon and rectal cancer. *Nature*. 2012;487(7407):330-337.
27. Gallardo M, Lee HJ, Zhang X, et al. hnRNP K Is a Haploinsufficient Tumor Suppressor that Regulates Proliferation and Differentiation Programs in Hematologic Malignancies. *Cancer Cell*. 2015;28(4):486-499.
28. Yoshida T, Georgopoulos K. Ikaros fingers on lymphocyte differentiation. *Int J Hematol*. 2014;100(3):220-229.
29. Allman D, Dalod M, Asselin-Paturel C, et al. Ikaros is required for plasmacytoid dendritic cell differentiation. *Blood*. 2006;108(13):4025-4034.
30. Cytlak U, Resteu A, Bogaert D, et al. Ikaros family zinc finger 1 regulates dendritic cell development and function in humans. *Nat Commun*. 2018;9(1):1239.
31. Schjerven H, McLaughlin J, Arenzana TL, et al. Selective regulation of lymphopoiesis and leukemogenesis by individual zinc fingers of Ikaros. *Nat Immunol*. 2013;14(10):1073-1083.
32. Tessema M, Willink R, Do K, et al. Promoter methylation of genes in and around the candidate lung cancer susceptibility locus 6q23-25. *Cancer Res*. 2008;68(6):1707-1714.
33. Haling JR, Sudhamsu J, Yen I, et al. Structure of the BRAF-MEK complex reveals a kinase activity independent role for BRAF in MAPK signaling. *Cancer Cell*. 2014;26(3):402-413.
34. Eberlein CA, Stetson D, Markovets AA, et al. Acquired Resistance to the Mutant-Selective EGFR Inhibitor AZD9291 Is Associated with Increased Dependence on RAS Signaling in Preclinical Models. *Cancer Res*. 2015;75(12):2489-2500.
35. Chraybi M, Abd Alsamad I, Copie-Bergman C, et al. Oncogene abnormalities in a series of primary melanomas of the sinonasal tract: NRAS mutations and cyclin D1 amplification are more frequent than KIT or BRAF mutations. *Hum Pathol*. 2013;44(9):1902-1911.
36. Di Carlo V, Mocavini I, Di Croce L. Polycomb complexes in normal and malignant hematopoiesis. *J Cell Biol*. 2019;218(1):55-69.
37. Nagasawa M, Schmidlin H, Hazekamp MG, Schotte R, Blom B. Development of human plasmacytoid dendritic cells depends on the combined action of the basic helix-loop-helix factor E2-2 and the Ets factor Spi-B. *Eur J Immunol*. 2008;38(9):2389-2400.
38. Koo CY, Muir KW, Lam EW. FOXM1: From cancer initiation to progression and treatment. *Biochim Biophys Acta*. 2012;1819(1):28-37.
39. Schjerven H, Ayongaba EF, Aghajanirefah A, et al. Genetic analysis of Ikaros target genes and tumor suppressor function in BCR-ABL1(+) pre-B ALL. *J Exp Med*. 2017;214(3):793-814.
40. Witkowski MT, Hu Y, Roberts KG, et al. Conserved IKAROS-regulated genes associated with B-progenitor acute lymphoblastic leukemia outcome. *J Exp Med*. 2017;214(3):773-791.
41. Unbekandt M, Olson MF. The actin-myosin regulatory MRCK kinases: regulation, biological functions and associations with human cancer. *J Mol Med (Berl)*. 2014;92(3):217-225.

42. Rafn B, Nielsen CF, Andersen SH, et al. ErbB2-driven breast cancer cell invasion depends on a complex signaling network activating myeloid zinc finger-1-dependent cathepsin B expression. *Mol Cell*. 2012;45(6):764-776.
43. Lee J-W, Geng H, Chen Z, et al. IFITM3 (CD225) Links the B Cell Antigen CD19 to PI3K-AKT Signaling in Human ALL Cells. *Blood*. 2015;126(23):1325-1325.
44. Joshi I, Yoshida T, Jena N, et al. Loss of Ikaros DNA-binding function confers integrin-dependent survival on pre-B cells and progression to acute lymphoblastic leukemia. *Nat Immunol*. 2014;15(3):294-304.
45. Churchman ML, Low J, Qu C, et al. Efficacy of Retinoids in IKZF1-Mutated BCR-ABL1 Acute Lymphoblastic Leukemia. *Cancer Cell*. 2015;28(3):343-356.
46. Churchman ML, Qian M, te Kronnie G, et al. Germline Genetic IKZF1 Variation and Predisposition to Childhood Acute Lymphoblastic Leukemia. *Cancer Cell*. 2018;33(5):937-948. e938.
47. MacGrath SM, Koleske AJ. Cortactin in cell migration and cancer at a glance. *J Cell Sci*. 2012;125(Pt 7):1621-1626.
48. Onodera Y, Hashimoto S, Hashimoto A, et al. Expression of AMAP1, an ArfGAP, provides novel targets to inhibit breast cancer invasive activities. *EMBO J*. 2005;24(5):963-973.
49. Puget N, Sinilnikova OM, Stoppa-Lyonnet D, et al. An Alu-mediated 6-kb duplication in the BRCA1 gene: a new founder mutation? *Am J Hum Genet*. 1999;64(1):300-302.
50. Song C, Gowda C, Pan X, et al. Targeting casein kinase II restores Ikaros tumor suppressor activity and demonstrates therapeutic efficacy in high-risk leukemia. *Blood*. 2015;126(15):1813-1822.
51. Shah S, Brock EJ, Ji K, Mattingly RR. Ras and Rap1: A tale of two GTPases. *Semin Cancer Biol*. 2019;54:29-39.
52. Infante E, Heasman SJ, Ridley AJ. Statins inhibit T-acute lymphoblastic leukemia cell adhesion and migration through Rap1b. *J Leukoc Biol*. 2011;89(4):577-586.
53. Lin KB, Tan P, Freeman SA, Lam M, McNagny KM, Gold MR. The Rap GTPases regulate the migration, invasiveness and in vivo dissemination of B-cell lymphomas. *Oncogene*. 2010;29(4):608-615.
54. Emadali A, Hoghoughi N, Duley S, et al. Haploinsufficiency for NR3C1, the gene encoding the glucocorticoid receptor, in blastic plasmacytoid dendritic cell neoplasms. *Blood*. 2016;127(24):3040-3053.
55. Ross J, Mavoungou L, Bresnick EH, Milot E. GATA-1 utilizes Ikaros and polycomb repressive complex 2 to suppress Hes1 and to promote erythropoiesis. *Mol Cell Biol*. 2012;32(18):3624-3638.
56. Oravecz A, Apostolov A, Polak K, et al. Ikaros mediates gene silencing in T cells through Polycomb repressive complex 2. *Nat Commun*. 2015;6:8823.
57. Isnard P, Core N, Naquet P, Djabali M. Altered lymphoid development in mice deficient for the mAF4 proto-oncogene. *Blood*. 2000;96(2):705-710.
58. Delgado P, Cubelos B, Calleja E, et al. Essential function for the GTPase TC21 in homeostatic antigen receptor signaling. *Nat Immunol*. 2009;10(8):880-888.
59. Boiko AD, Porteous S, Razorenova OV, Krivokrysenko VI, Williams BR, Gudkov AV. A systematic search for downstream mediators of tumor suppressor function of p53 reveals a major role of BTG2 in suppression of Ras-induced transformation. *Genes Dev*. 2006;20(2):236-252.
60. Xiao S, Li D, Zhu HQ, et al. RIG-G as a key mediator of the antiproliferative activity of interferon-related pathways through enhancing p21 and p27 proteins. *Proc Natl Acad Sci U S A*. 2006;103(44):16448-16453.
61. Jeon JH, Lee KN, Hwang CY, Kwon KS, You KH, Choi I. Tumor suppressor VDUP1 increases p27(kip1) stability by inhibiting JAB1. *Cancer Res*. 2005;65(11):4485-4489.
62. Tu Y, Wu W, Wu T, et al. Antiproliferative autoantigen CDA1 transcriptionally up-regulates p21(Waf1/Cip1) by activating p53 and MEK/ERK1/2 MAPK pathways. *J Biol Chem*. 2007;282(16):11722-11731.

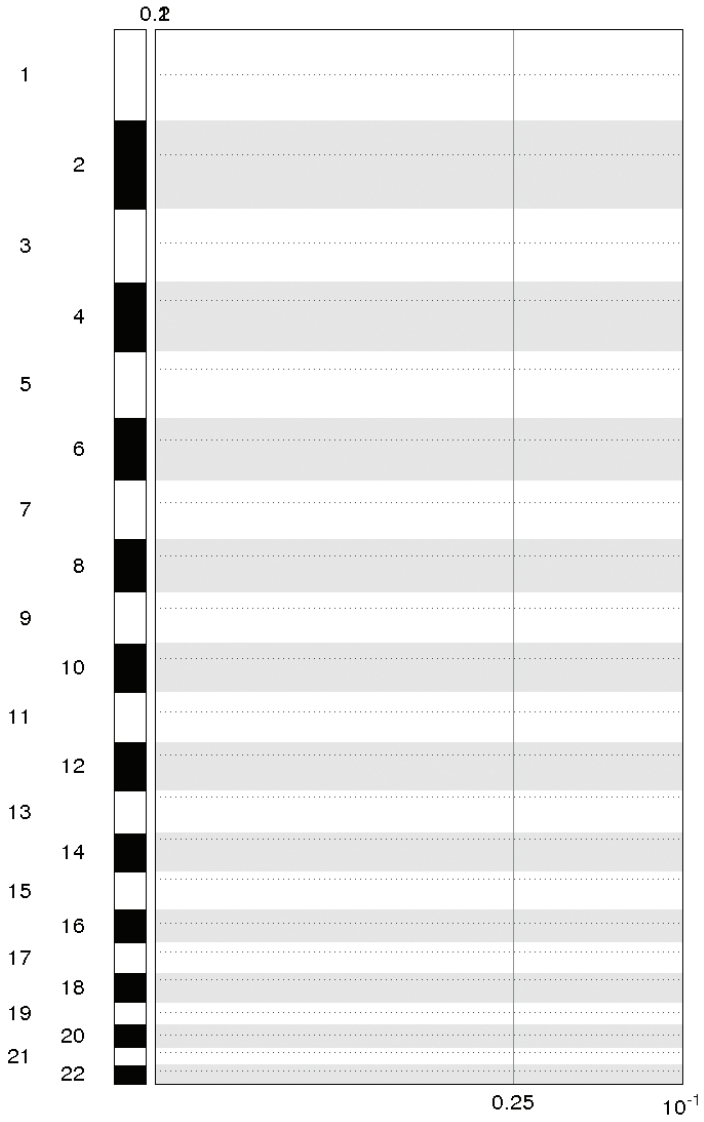
SUPPORTING INFORMATION FIGURES

Supporting Information Figure 1.
A)



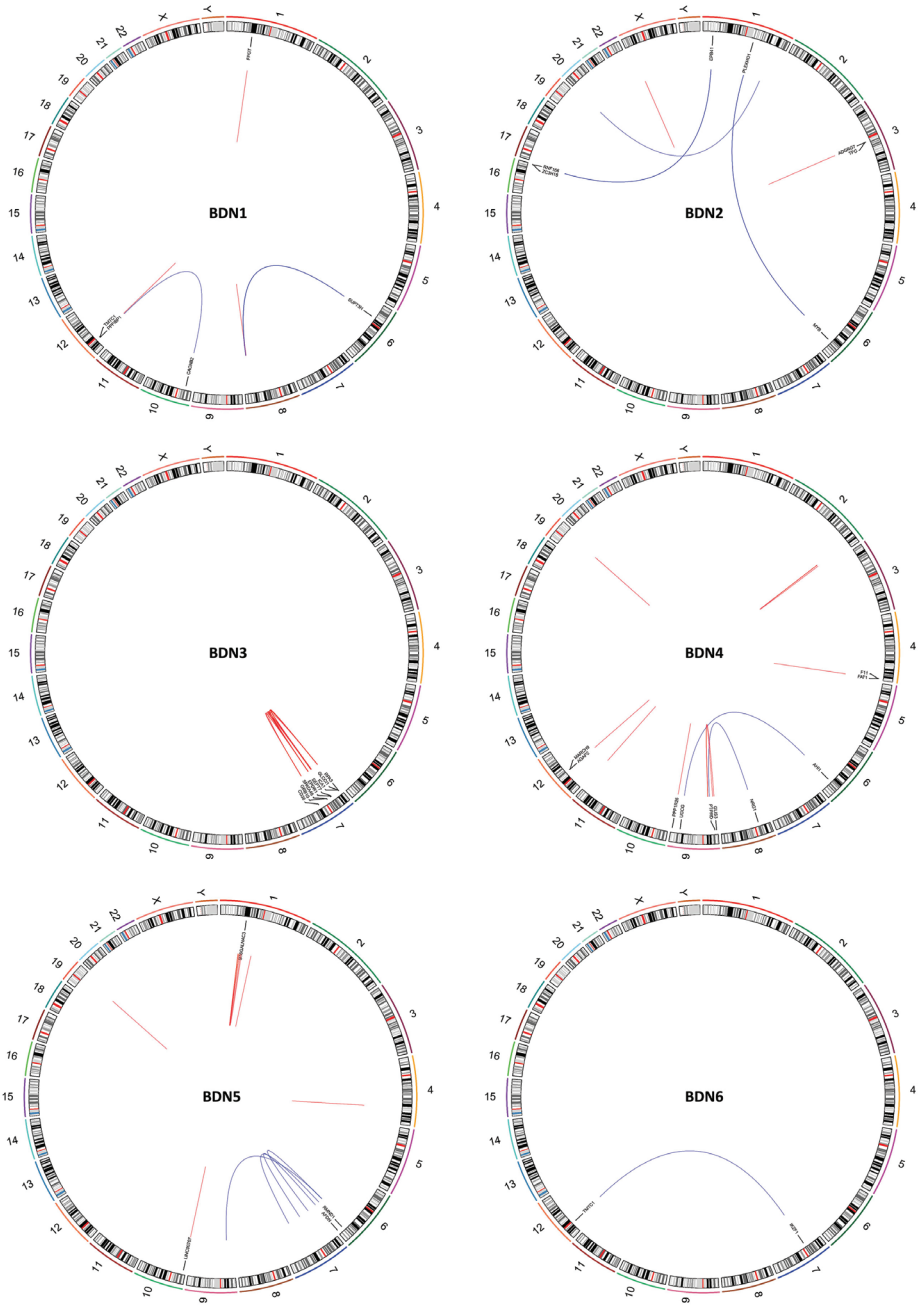
Common focal losses in BPDCN detected by GISTIC 2.0. analysis. GISTIC 2.0 detected common losses shared by BPDCN patients in chromosome 6, 7, 12 and 13. Peaks at 6q25.1, 7p12.2 and 12p13.1 were ≤ 3Mb long (as verified with IGV) and enclosed tumor suppressors *SYNE1*, *IKZF1* and *CDKN1B* (alongside *DUSP16* and *ETV6*), respectively. Peaks within chromosomes 9 and 13 were either > 3Mb long or devoid of (putative) cancer genes, and consequently, excluded.

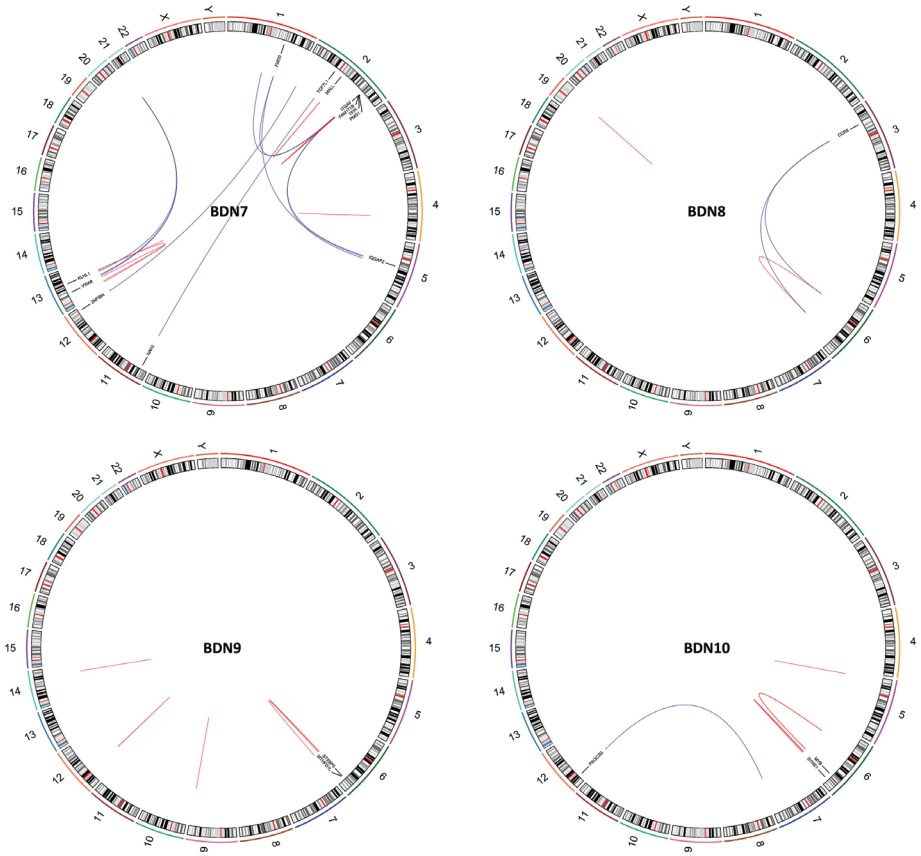
B)



Common focal gains in BPDCN detected by GISTIC 2.0. analysis. GISTIC 2.0 detected no focal common gains between BPDCN patients.

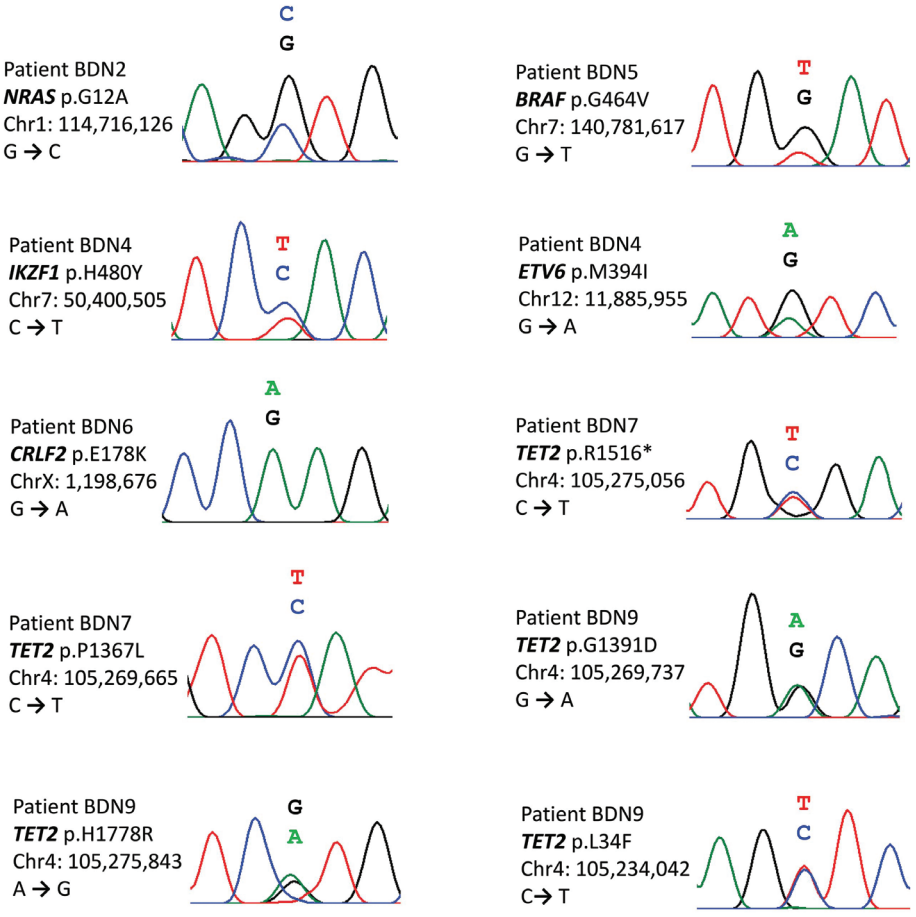
Supporting Information Figure 2.





Circos plots depicting genomic rearrangements in BPDCN per patient. Tumors from different BPDCN patients contained a variety of inter- and intrachromosomal rearrangements. The number of events ranged from 1 (BDN6) to 18 (BDN7). *IKZF1* (BDN2 and BDN6) and *MYB* (BDN3 and BDN10) were the only cancer genes recurrently disrupted by genomic rearrangements in our sequenced cohort.

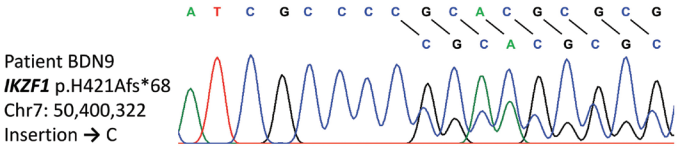
Supporting Information Figure 3.



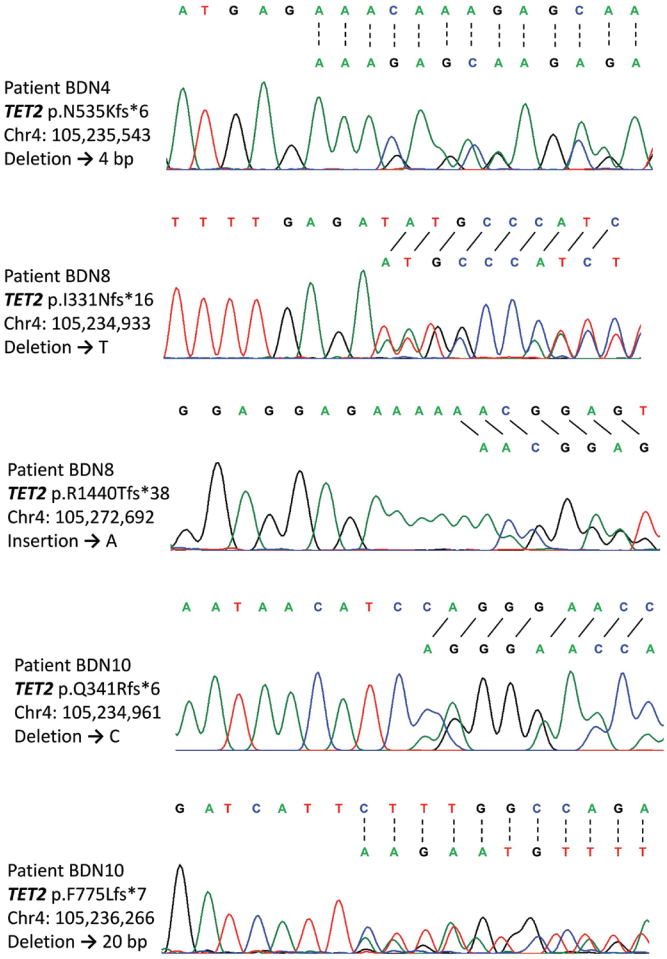
Sanger sequencing validation of select SNVs. Sanger chromatograms showing select point mutations with known or predicted deleterious effects (SIFT and PolyPhen-2) detected in BPDCN tumors by WGS. Genomic coordinates according to reference genome Hg38.

Supporting Information Figure 4.

A



B



Sanger sequencing validation of select indels. Sanger chromatograms showing indels leading to premature stop codons in (A) *IKZF1* and (B) *TET2* detected in BPDCN tumors by WGS. Genomic coordinates according to reference genome Hg38.

Note: Supporting information tables available online at https://doi.org/10.1002/gcc.22831.

7

General discussion and conclusions



The work presented in this thesis aimed at deepening the understanding of the pathogenetic basis of common and rare cutaneous lymphoma entities by using NGS. These analyses had a special focus on structural genomic alterations, since a detailed examination of this type of genetic defects in cutaneous lymphomas had been largely neglected in prior studies. This final chapter summarizes and discusses the results of the studies presented in chapters 3 4, 5 and 6 together with literature presented in chapters 1 and 2 and findings from recent studies. In the first part we will discuss the findings on each of the three CTCL entities included in this thesis (i.e. MF, pcALCL, pcAECyTCL) separately. In the second part we will examine commonalities and differences between the underlying pathogenetic basis of the analyzed CTCL variants. In the third part we will discuss the findings on precursor neoplasm BPDCN. The chapter will end with conclusions and future perspectives.

1. CUTANEOUS T-CELL LYMPHOMA ENTITIES

1.1 Mycosis Fungoides

MF, a skin lymphoma caused by malignant CD4+ T cells, is characterized by a slow progression from patches to more infiltrated plaques and eventually to tumors in the skin. While the prognosis of patients with early-stage disease is favorable, the prognosis of patients with tumor-stage disease worsens considerably. Thus, the development of more effective therapies for patients with T-MF based on molecular insights is imperative. Three main factors have hindered a better understanding of the pathogenetic basis of this malignancy. Firstly, as discussed in chapters 1 and 2, MF is largely underrepresented in molecular studies, despite being the most common CTCL variant. Secondly, traditionally MF and SS have been regarded as merely different manifestations of the same disease, and consequently, lumped together as a single entity in genetic studies. Thirdly, prior NGS analyses on MF have focused primarily on the assessment of small-scale mutations, overlooking the study of structural genomic alterations, which seem to play a leading role in the lymphomagenesis of other CTCL variants, as discussed in chapter 2.

To address the limitations of previous research on T-MF, we performed a high-resolution integrated (DNA/RNA) genome-wide study focused on the characterization of genomic rearrangements and CNAs in a cohort exclusively formed by patients with T-MF. The analysis revealed considerable genome instability, with numerous rearrangements mediating the deletion of tumor suppressors (e.g. *ARID1A*, *CDKN2A/B*, *PTPRC*, *STK11*, etc.) involved in pathways that are commonly deregulated in MF patients. The main highlight of our study

was the discovery of recurrent deletion of *HNRNPK* and *SOCS1* in MF tumors, two established tumor suppressors which are novel to MF genetics.

Functional studies carried out by others have shown that *HNRNPK* and *SOCS1* exert negative regulation on STAT3 signaling. Consistent with this evidence, our study showed that T-MF displayed up-regulation of STAT3 signaling at transcriptome level. This finding matched prior reports of robust STAT3 activation in MF cells observed *in vitro* and *in vivo* through immunocytochemistry (ICC) and IHC, respectively.[1,2] After the publication of our study, Pérez et al. reported an association between STAT3 signaling and disease progression in MF, with tumor-stage cases showing much stronger STAT3 activation than early-stage cases.[3] Altogether, this evidence suggests that STAT3 signaling plays a central role in T-MF and may determine the progression of the disease from a milder to a more aggressive lymphoma.

Different lines of evidence suggest that the etiology and/or maintenance of aberrant STAT3 signaling in MF is multifactorial. Genetic alterations, epigenetic deregulation and even tumor inflammation triggered by *Staphylococcus aureus* colonization at advanced stages of the disease can seemingly promote overactivation of STAT3 signaling in MF cells.[4] Our study showed for the first time that genes with inhibitory effects on STAT3 signaling are recurrently deleted in T-MF at substantial frequencies (~35%), constituting potential candidate events behind the initial deregulation of STAT3 signaling in this lymphoma. In contrast, prior WES/TS studies had only found solitary or few MF cases carrying gain-of-function SNVs in *JAK3* and *STAT3*, which activate STAT3 signaling.[5-7] The infrequency of activating SNVs in *JAK* and *STAT* genes in MF was confirmed in our study.

In addition, miR-155, which is the most consistently overexpressed oncomir in T-MF, has been proven to inhibit *SOCS1* in other human cancers, leading to overactivation of STAT3 signaling. In line with this evidence, our data showed that MF patients with functional *SOCS1* alleles expressed higher levels of *MIR155HG* (miR-155 precursor) than patients with *SOCS1* deletions. A final layer of complexity comes from the interaction between the microbiota and the tumor microenvironment. *S. aureus*, which often colonizes tumors in MF patients, has been shown to promote a pro-oncogenic tumor microenvironment by triggering the physical interaction between malignant T cells and non-malignant T cells, which in turn, induces the production of growth factors and the activation of STAT3 signaling.[8,9]

Taken together, our findings in combination with evidence derived from other studies strongly suggest that overactivation of STAT3 signaling plays a leading

role in MF progression towards a more aggressive form of the disease. Given that STAT3 signaling promotes proliferation of malignant T cells and the production of cytokines that mediate pro-oncogenic cell-cell communication within the microenvironment, targeting this pathway is expected to block survival of the neoplastic cells themselves and the maintenance of the inflammatory microenvironment as well. Besides STAT3 itself, additional potential drug targets proposed in our study are oncomir miR-155 and MECOM. Ultimately, future translational and clinical studies will be necessary to determine the applicability, effectivity and potential benefit to patients of therapies directed at inhibiting STAT3 signaling in T-MF.

1.2 Primary cutaneous anaplastic large cell lymphoma

pcALCL, part of the spectrum of primary cutaneous CD30+ LPDs, is one the least studied CTCL variants. The limited understanding of the molecular basis of pcALCL is the result of the restricted number of studies (cytogenetic, array-based, NGS, etc.) focused exclusively on this lymphoma. In chapter 1, we pointed out that cytogenetic and molecular studies have shown that pcALCL and sALCL share several pathogenetic features, despite having dissimilar clinical course and prognosis. These commonalities include occasional rearrangements involving *DUSP22/IRF4*, *ALK*, *TP63* and *TYK2* and overexpression of *TNFRSF8* (CD30), IRF4 and oncomir miR-155. Therefore, a deeper molecular scrutiny of pcALCL was required to uncover distinctive genetic alterations of this lymphoma and establish molecular differences with sALCL.

In chapter 4 we described the most comprehensive molecular profiling of pcALCL thus far. Our study provided for the first time a genome-wide overview of genomic rearrangements, CNAs, small-scale mutations and deregulated gene expression underlying pcALCL. We confirmed the predominance of broad losses and gains within 6q and 7q, respectively, as well as the up-regulation of several genes with pathobiological relevance (e.g. *CCR8*, *CCR10*, *EZH2*, *IRF4*, *MIR155HG*, *TNFRSF8*) reported in previous studies. Importantly, our data showed that although pcALCL is genetically heterogeneous, genetic defects cluster into four main cellular processes: the cell cycle, T-cell physiology regulation, transcription, and signaling via the PI-3-K/Akt, MAPK and G protein pathways. Recurrent alterations impacting cancer genes *GPS2*, *PDPK1*, *PIK3R1* and *TNK1* had special prominence in the disease. Gene expression patterns in pcALCL were consistent with overexpression of *PDGFRB* (also *DDR2*), suggesting that PDGFR blockade with currently available FDA-approved inhibitors (i.e. imatinib) could work as an alternative therapy for the disease. Also, drugs with the ability to block effectors of the PI-3-K/Akt and MAPK pathways could be tested for the treatment of pcALCL. In addition, this study showed that pcALCL was devoid of genetic alterations in JAK-STAT pathway genes, which on the contrary, characterize

sALCL. Overall, our work revealed key molecular features of pcALCL and supported the notion that pcALCL and sALCL are related but distinct entities.

1.3 Primary cutaneous CD8+ aggressive epidermotropic cytotoxic T-cell lymphoma

pcAECyTCL is a rare and highly aggressive variant of CTCL derived from CD8+ T cells. Patients with this lymphoma have a median overall survival of 12 months since thus far no effective therapeutic regimes have been defined. In chapter 1 we discussed the lack of molecular studies on this lymphoma as a result of its rarity, with the CNA profiling of 20 patients using a-CGH performed by Fanoni et al. being the only cohort study.[10] The urgent need of therapeutic options for patients with pcAECyTCL demanded molecular insights that could guide the selection or development of appropriate treatment.

To look for genetic alterations and deregulated expression underlying this lymphoma, in chapter 5 we analyzed the genome and transcriptome of pcAECyTCL by using WGS and RNA-seq, respectively. Our analysis revealed multiple rearrangements, CNAs and small-scale mutations impacting genes involved in the cell cycle, chromatin regulation and the JAK-STAT pathway. Notably, overactivation of JAK2 signaling due to mutually exclusive genetic alterations involving *JAK2* and *SH2B3*, were found to underlie predominantly pcAECyTCL. *JAK2* formed fusion genes encoding self-activating chimeras whereas *SH2B3* was inactivated by focal interstitial deletions and unbalanced rearrangements. Furthermore, functional studies confirmed that JAK2 fusions identified in pcAECyTCL induce cytokine-independent cell survival and their activity was shown to be successfully blocked by FDA-approved JAK1/2 inhibitor ruxolitinib.

Overall, this study achieved three important goals. Firstly, it uncovered for the first time multiple genetic defects underlying pcAECyTCL, with events leading to overactivation of JAK2 signaling being the most prominent. Secondly, it identified genetic alterations that characterize pcAECyTCL (i.e. *JAK2* fusions, *SH2B3* inactivation) which can be instrumental to distinguish this malignancy from other CTCL variants. Thirdly, it identified JAK2 inhibition with ruxolitinib as a potential therapy for the disease.

1.4 Commonalities and differences in the pathogenetic basis of CTCL variants.

A comparison of the data derived from the studies presented in chapters 3, 4 and 5 reveals molecular similarities and differences between MF, pcALCL and pcAECyTCL. Several genes involved in the cell cycle, chromatin regulation and the JAK-STAT pathway were commonly altered in T-MF and pcAECyTCL. In

contrast, pcALCL shared very few molecular alterations with the previous two CTCL variants. Pathogenic features commonly found in T-MF and pcAECyTCL included recurrent deletion of tumor suppressors *ARID1A*, *CDKN2A/B*, *PTPRC*, *SOCS1* and *STK11* and occasional rearrangements involving *CNTN5*, *MLLT3* and *TP63*. A major difference between MF and pcAECyTCL was the predominance of genetic alterations in different *JAK* and *STAT* genes. MF displayed mutations exclusively in *JAK3* and *STAT3* while pcAECyTCL carried genetic alterations mainly in *JAK2* and *STAT5B*. Molecular alterations exclusive to pcALCL involved genes with roles in T-cell physiology (i.e. *PRDM1*, *TNFRSF8*) and signaling via the PI-3-K (e.g. *PDPK1*, *PIK3R1*), MAPK (e.g. *GPS2*, *TNK1*) and G protein pathways (e.g. *RGS7*, *RGS12*). Molecular alterations shared between pcALCL and T-MF were down-regulation of *GPS2*, up-regulation of *MIR155HG* and TP63 rearrangements. Molecular alterations shared between pcALCL and pcAECyTCL were deletion of *TNFRSF14* and *TP63* rearrangements. The absence of deregulation of JAK-STAT signaling in pcALCL (an indolent CTCL variant) and the progressive activation of this pathway in MF as the disease worsens in severity, suggest that deregulation of JAK-STAT signaling is associated with more aggressive CTCL phenotypes.

2. PRECURSOR NEOPLASMS

2.1 Blastic plasmacytoid dendritic cell neoplasm

BPDCN is a rare and aggressive hematological malignancy derived from pDC precursors. In chapter 1 we discussed prior molecular studies on BPDCN which had established two main pathogenetic features of this neoplasm: 1) the predominance of small-scale mutations (indels, SNVs) in epigenetic regulators (e.g. *ASLX1*, *TET2*) and 2) the inactivation of regulators of the cell cycle at G1/S (i.e. *CDKN1B*, *CDKN2A/B*, *RB1*). Given that mutations in epigenetic regulators (esp. *TET2*) are common in malignancies of myeloid origin and the dysfunction of cell cycle regulators occurs in a plethora of blood and solid cancers, additional studies have looked for disease-specific molecular features in BPDCN. The studies carried out by Ceribelli et al., Sakamoto et al. and Suzuki et al. gave *TCF4*, *MYC* and *MYB* notoriety as potential drivers of BPDCN, respectively, as these genes were found to be highly expressed and/or dysfunctional in a considerable proportion of cases.

In chapter 6 we presented the first whole-genome analysis of BPDCN using NGS. We characterized for the first time the landscape of genomic rearrangements and CNAs in this neoplasm by employing high-resolution WGS. We confirmed that mutations in *ASXL1* and *TET2*, as well as deletion of *CDKN1B* and *RB1* are recurrent events in BPDCN. However, rearrangements affecting *MYB* and *MYC*

(*SUPT3H*) reported at high frequencies in prior studies, were found in solitary or few patients in our cohort. Notably, our analysis revealed that *IKZF1*, a gene with critical roles in the early development of pDC precursors (the cell of origin of BPDCN), is focally inactivated in a significant fraction of BPDCN cases (~50%) as a result of structural genomic defects. In agreement with this finding, we observed that cancer-associated IKZF1 target genes displayed a loss-of-IKZF1 expression pattern in this neoplasm. Moreover, deregulated cellular processes in BPDCN (e.g. cell adhesion signatures, PI-3-K/Akt signaling) matched known physiological consequences of IKZF1 deficiency in immune cells.

A large body of evidence has shown that *IKZF1* inactivation underlies the development of BCR-ABL1+ pre-B acute lymphoblastic leukemia (pre-B ALL). *IKZF1* inactivation appears to induce three oncogenic cellular processes in BCR-ABL1+ pre-B ALL: 1) maintenance of an immature cellular phenotype, 2) gain of integrin-dependent survival ability, and 3) activation of pro-survival PI3-K/Akt signaling.[11,12] These oncogenic cellular processes seemingly operate in BPDCN as well. Prior research by others showed that IKZF1 deficiency prevents pDC precursors from undergoing differentiation[13] and we detected up-regulation of cell adhesion signatures and the PI3-K/Akt pathway in BPDCN by RNA-seq. Given the importance of *IKZF1* in the biology of early pDC precursors, *IKZF1* inactivation may play a pivotal role in establishing the 'founding oncogenic program' of BPDCN. Ultimately, functional studies with cells and animal models will be necessary to confirm the hypothesized role of *IKZF1* inactivation in the genesis of the disease. From a translational viewpoint, the use of drugs (e.g. retinoids) with the ability to induce expression from the remaining functional allele in BPDCN patients with hemizygous *IKZF1* inactivation could reverse or attenuate the oncogenic effects of IKZF1 deficiency.

Since *TCF4*, *MYC* and *MYB* have been proposed as key drivers of BPDCN in prior studies, establishing their exact contribution to the development of BPDCN is relevant and merits future research. Yet, the current knowledge of the functions of these genes provides hints on their potential roles in the disease. Ceribelli et al. regarded transcription factor TCF4 (E2-2) as a "master regulator of the BPDCN oncogenic program", as TCF4 was shown to be essential for the survival of BPDCN cells.[14] However, TCF4 had already been proven to be vital for the survival of normal pDCs, as it induces the expression of genes required for the maintenance of a normal pDC phenotype (e.g. *BCL11A*, *CLEC4C*, *IRF8*, *IL3RA*, *SPIB*, *TLR9*).[15,16] Therefore, the dependence of BPDCN cells on TCF4 for survival may alternatively mean that, despite undergoing malignization, these neoplastic cells largely preserve a regular pDC expression program.

Sakamoto et al. reported that 39% of BPDCN cases carried *MYC* rearrangements which correlated with MYC overexpression and more aggressive clinical course.[17] Even though *MYC* is often overexpressed in numerous human cancers, it remains unclear whether *MYC* deregulation is primarily linked to cancer genesis or progression. The high incidence of *MYC* rearrangements in neoplasms such as Burkitt's lymphoma suggests a role in the initiation of certain cancers.[18] However, *MYC* was recently shown to be a universal amplifier of gene expression as it binds thousands of E-box motifs near or at pre-activated genes to boost their expression temporarily.[19] This mechanism of action explains the correlation between MYC expression and the activation of a wide diversity of cellular processes/pathways in different cell types, with very little overlap between them.[19,20] In cancer, MYC overexpression is thought to make the pre-existing oncogenic program more robust and severe, explaining the correlation with more aggressive clinical course.[20] The impact of MYC deregulation in cancer development and its exact role in BPDCN warrants future studies.

Finally, Suzuki et al. informed that 64% of evaluated cases (100% children, 44% adults) had *MYB* rearrangements. These authors also detected a MYB expression signature in BPDCN.[21] In agreement with the previous, we confirmed *MYB* overexpression in BPDCN by RNA-seq. MYB is a transcriptional regulator with essential roles in hematopoiesis. Several studies have shown that MYB overexpression suppresses normal differentiation of hematopoietic precursors (incl. lymphoid, myeloid and erythroid lineages) and promotes leukemic transformation.[22] The ability of MYB to promote the expansion of progenitor cells and block differentiation suggests a role in the onset of hematological cancers derived from precursors such as BPDCN. The role of MYB deregulation in the pathobiology of BPDCN deserves mechanistic research with cells and animal models.

In sum, BPDCN seems to arise primarily from defects in genes involved in the cell cycle, hematopoiesis, and the epigenetic regulation of the genome. Our work has positioned *IKZF1* as a putative key driver in the development of BPDCN, just like previously proposed for *TCF4*, *MYB* and *MYC*.

3. CONCLUSION AND FUTURE PERSPECTIVES

The NGS studies included in this thesis have expanded our understanding of the pathogenetic basis of T-MF, pcALCL, pcAECyTCL and BPDCN. Importantly, our work showed that structural genomic alterations play important roles in the pathobiology of these malignancies. For instance, recurrent oncogenic alterations affecting important cancer genes in T-MF (e.g. *SOCS1*), pcALCL (*GPS2*, *TNK1*), pcAECyTCL (e.g. *JAK2*, *SH2B3*) and BPDCN (e.g. *IKZF1*, *MYB*) resulted primarily from rearrangements or CNAs rather than indels or SNVs. These studies also identified recurrently affected genes and cellular processes/pathways in each of these malignancies, providing molecular insights that could be instrumental for the development of novel therapies in the future. In addition, a comparison of the genetic alterations underlying each of the three CTCL variants analyzed in this thesis revealed relevant commonalities and differences between them and clarified the molecular mechanisms governing each of these neoplasms.

The body of work presented in chapters 3, 4, 5 and 6 represents a starting point for future mechanistic and translational studies. We propose the following follow-up research. Firstly, the validation of our findings in larger patient cohorts. Secondly, the undertaking of mechanistic studies to unravel the physiological consequences of the recurrent genetic alterations highlighted in this thesis. Thirdly, the assessment of potential therapies for T-MF, pcALCL, pcAECyTCL and BPDCN based on the molecular insights described in our studies. Moreover, research on additional complex aspects of the pathobiology of these neoplasms will be decisive to fully elucidate the oncogenic mechanisms operative in these malignancies. The study of the epigenome, the non-protein coding transcriptome, the impact of mutations in nongenic/regulatory regions and the tumor microenvironment are some of the aspects that should be addressed in the future. Ultimately, these research efforts are expected to translate into major improvements in treatment efficacy and patient survival.

REFERENCES

1. Sommer VH, Clemmensen OJ, Nielsen O, et al. In vivo activation of STAT3 in cutaneous T-cell lymphoma. Evidence for an antiapoptotic function of STAT3. *Leukemia*. 2004;18(7):1288-1295.
2. Nielsen M, Kaltoft K, Nordahl M, et al. Constitutive activation of a slowly migrating isoform of Stat3 in mycosis fungoides: tyrphostin AG490 inhibits Stat3 activation and growth of mycosis fungoides tumor cell lines. *Proc Natl Acad Sci U S A*. 1997;94(13):6764-6769.
3. Pérez C, Mondéjar R, García-Díaz N, et al. Advanced-stage mycosis fungoides: role of the signal transducer and activator of transcription 3, nuclear factor-κB and nuclear factor of activated T cells pathways. *Br J Dermatol*. 2020;182(1):147-155.
4. Ødum N. Deregulated signalling and inflammation in cutaneous T-cell lymphoma. *Br J Dermatol*. 2020;182(1):16-17.
5. Ungewickell A, Bhaduri A, Rios E, et al. Genomic analysis of mycosis fungoides and Sezary syndrome identifies recurrent alterations in TNFR2. *Nat Genet*. 2015;47(9):1056-1060.
6. McGirt LY, Jia P, Baerenwald DA, et al. Whole-genome sequencing reveals oncogenic mutations in mycosis fungoides. *Blood*. 2015;126(4):508-519.
7. da Silva Almeida AC, Abate F, Khiabanian H, et al. The mutational landscape of cutaneous T cell lymphoma and Sezary syndrome. *Nat Genet*. 2015;47(12):1465-1470.
8. Willerslev-Olsen A, Krejsgaard T, Lindahl LM, et al. Bacterial toxins fuel disease progression in cutaneous T-cell lymphoma. *Toxins (Basel)*. 2013;5(8):1402-1421.
9. Lindahl LM, Willerslev-Olsen A, Gjerdrum LMR, et al. Antibiotics inhibit tumor and disease activity in cutaneous T-cell lymphoma. *Blood*. 2019;134(13):1072-1083.
10. Fanoni D, Corti L, Alberti-Violetti S, et al. Array-based CGH of primary cutaneous CD8+ aggressive EPIDERMO-tropic cytotoxic T-cell lymphoma. *Genes Chromosomes Cancer*. 2018;57(12):622-629.
11. Joshi I, Yoshida T, Jena N, et al. Loss of Ikaros DNA-binding function confers integrin-dependent survival on pre-B cells and progression to acute lymphoblastic leukemia. *Nat Immunol*. 2014;15(3):294-304.
12. Churchman ML, Mullighan CG. Ikaros: Exploiting and targeting the hematopoietic stem cell niche in B-progenitor acute lymphoblastic leukemia. *Exp Hematol*. 2017;46:1-8.
13. Allman D, Dalod M, Asselin-Paturel C, et al. Ikaros is required for plasmacytoid dendritic cell differentiation. *Blood*. 2006;108(13):4025-4034.
14. Ceribelli M, Hou ZE, Kelly PN, et al. A Druggable TCF4- and BRD4-Dependent Transcriptional Network Sustains Malignancy in Blastic Plasmacytoid Dendritic Cell Neoplasm. *Cancer Cell*. 2016;30(5):764-778.
15. Nagasawa M, Schmidlin H, Hazekamp MG, Schotte R, Blom B. Development of human plasmacytoid dendritic cells depends on the combined action of the basic helix-loop-helix factor E2-2 and the Ets factor Spi-B. *Eur J Immunol*. 2008;38(9):2389-2400.
16. Cisse B, Caton ML, Lehner M, et al. Transcription factor E2-2 is an essential and specific regulator of plasmacytoid dendritic cell development. *Cell*. 2008;135(1):37-48.
17. Sakamoto K, Katayama R, Asaka R, et al. Recurrent 8q24 rearrangement in blastic plasmacytoid dendritic cell neoplasm: association with immunoblastoid cytomorphology, MYC expression, and drug response. *Leukemia*. 2018;32(12):2590-2603.
18. Schmitz R, Ceribelli M, Pittaluga S, Wright G, Staudt LM. Oncogenic mechanisms in Burkitt lymphoma. *Cold Spring Harb Perspect Med*. 2014;4(2).
19. Nie Z, Hu G, Wei G, et al. c-Myc is a universal amplifier of expressed genes in lymphocytes and embryonic stem cells. *Cell*. 2012;151(1):68-79.
20. Lin CY, Lovén J, Rahl PB, et al. Transcriptional amplification in tumor cells with elevated c-Myc. *Cell*. 2012;151(1):56-67.
21. Suzuki K, Suzuki Y, Hama A, et al. Recurrent MYB rearrangement in blastic plasmacytoid dendritic cell neoplasm. *Leukemia*. 2017;31(7):1629-1633.
22. Ramsay RG, Gonda TJ. MYB function in normal and cancer cells. *Nat Rev Cancer*. 2008;8(7):523-534.

Appendix



NEDERLANDSE SAMENVATTING

Cutane lymfomen zijn hematologische maligniteiten die in de huid aanwezig zijn zonder aanwijzingen voor extracutane ziekte op het moment van diagnose. Deze groep van neoplasieën kan worden onderverdeeld in drie categorieën: cutane T-cellymfomen (CTCL), cutane B-cellymfomen en precursor neoplasieën. In dit proefschrift worden drie CTCL-varianten, tumor-stadium mycosis fungoides (T-MF), primair cutaan anaplastisch grootcellig lymfoom (pcALCL) en primair cutaan CD8+ agressief epidermotroop cytotoxisch T-cellymfoom (pcAECyTCL) en één precursor neoplasie, blastisch plasmacytoïde dendritische celneoplasmie (BPDCN) geanalyseerd met behulp van hoge resolutie genoom-brede sequentiebepaling (WGS) en volledige transcriptoom sequentiebepaling (RNA-seq). Deze analyses waren met name gericht op identificering van structurele genomische veranderingen (d.w.z. veranderingen in kopie aantal (CNA's) en chromosomale herrangschikkingen), aangezien een gedetailleerd onderzoek van dit type genetische defecten in huidlymfomen in eerdere moleculaire studies nauwelijks aan de orde kwamen. De bevindingen die in dit proefschrift worden beschreven, hebben het begrip van de pathogenetische basis van deze entiteiten aanzienlijk vergroot.

In hoofdstuk 1 geven we een overzicht van de biologie van cutane lymfomen, de techniek van next-generation sequencing (NGS), principes van oncogenese en eerdere pathogenetische studies aan T-MF, pcALCL, pcAECyTCL en BPDCN. In hoofdstuk 2 bespreken we in detail de belangrijkste pathogenetische kenmerken die ontdekt zijn door NGS-studies aan de tot dusver best bestudeerde CTCL-varianten (mycosis fungoides en Sézary syndrome). Dit hoofdstuk beschrijft ook de therapeutische implicaties van deze moleculaire bevindingen.

In hoofdstuk 3 hebben we de eerste geïntegreerde (DNA/RNA) genoom-brede analyse van T-MF uitgevoerd. Om structurele genomische defecten met pathogeen potentieel te identificeren, werd een groep van 9 MF-patiënten geanalyseerd met WGS en RNA-seq. In deze studie werd het landschap van herrangschikkingen (incl. Fusietranscripten) en CNA's bij patiënten met T-MF voor het eerst gekarakteriseerd met NGS. Een vergelijkende analyse van genexpressie in T-MF en normale CD4+ T-cellen werd uitgevoerd om ontregelde cellulaire processen/routes in de ziekte te identificeren. De analyse bracht een aanzienlijke genoominstabiliteit aan het licht, met talrijke DNA herrangschikkingen die leiden tot verlies van DNA dat codeert voor voor tumoronderdrukkers (bijv. *ARID1A*, *CDKN2A/B*, *PTPRC*, *STK11*, enz.). Deze eiwitten zijnbetrokken zijn bij signaalpaden die vaak gedereguleerd zijn in tumorcellen van MF-patiënten. De belangrijkste vinding in deze studie was de ontdekking van een terugkerende deletie van genen coderend voor *HNRNPK* en

SOCS1 in MF-tumoren (~35% als de initiële gesequentieerde en extensiecohorten gecombineerd worden), twee bekende tumoronderdrukkers die nieuw zijn voor MF-genetica. Transcriptoomanalyse liet zien dat de celcyclus, JAK-STAT, PI-3-K en ontwikkelingsroutes aberrant geactiveerd zijn.

Gezien het beperkte moleculaire inzicht dat beschikbaar is voor pcALCL, hebben we in hoofdstuk 4 de meest uitgebreide genetische studie in pcALCL tot nu toe uitgevoerd. Met behulp van een multi-platform benadering (WGS, WES, RNA-seq) onderzochten we genetische veranderingen en ontregelde genexpressie in tumoren van 12 patiënten. Deze studie bracht meerdere pathogene mutaties, CNA's en herschikkingen aan het licht die genen betreffen die betrokken zijn bij de celcyclus, regulatie van de T-celfysiologie, transcriptie en signalering via de PI-3-K-, MAPK- en G-eiwitroutes. Terugkerende gebeurtenissen die kankergeassocieerde genen aantasten, waren onder meer deletie van *PRDM1* en *TNFRSF14*, winst van *EZH2* en *TNFRSF8*, kleinschalige mutaties in *LRP1B*, *PDPK1* en *PIK3R1* en herschikkingen van *GPS2*, *LINC-PINT* en *TNK1*. PCALCL vertoonde ook een mutatiesignatuur die voornamelijk toegeschreven kan worden aan UV-straling en veroudering met geringe bijdragen van afwijkingen in dubbelstrengs breukherstel en AID-hypermutatie. Ontregelde klassieke signaalroutes werden gevonden door genexpressie in pcALCL te vergelijken met genexpressie in normale CD4+ T-cellen. pcALCL toonde verhoogde activatie van signaaltransductieroutes geassocieerd met de PI-3-K-, MAPK- en G-eiwitroutes (ERK, fosfolipase-C, AKT, enz.), hetgeen consistent is met de genomische gegevens. Genexpressiepatronen in pcALCL kwamen overeen met overexpressie van receptortyrosinekinasen, bijv. PDGFRB en DDR2, waarvan met onafhankelijke technieken werd bevestigd dat ze up-gereguleerd zijn in dit lymfoom.

In hoofdstuk 5 beschrijven we de eerste NGS-studie die ooit is uitgevoerd op pcAECyTCL, een van de zeldzaamste en meest agressieve varianten van CTCL. Door tumormateriaal van twaalf patiënten met WGS te analyseren, identificeerden we voor de eerste keer voor een groep bonafide oncogenen en tumoronderdrukkers met een centrale rol in de celcyclus (d.w.z. *CDKN2A/B*, *MIR34AHG*, *MYC*, *RB1*, *TP53*), chromatine-regulatie (d.w.z. *ARID1A*, *BAZ1A*, *EED*, *EPC1*, *KMT2D*, *NCOR1*, *ZEB1*) en de JAK-STAT-route (d.w.z. *JAK2*, *JAK3*, *PTPRC*, *SH2B3*, *SOCS1*, *STAT3*, *STAT5B*) dat het aantal kopieën, de sequentie-organisatie en/of de nucleotide volgorde herhaaldelijk bleken te zijn veranderd in pcAECyTCL. Opmerkelijk is dat JAK2 en SH2B3, die respectievelijk de activering en beëindiging van JAK2-signalering in normale hematopoëtische cellen regelen, wederzijds uitsluitende veranderingen ondergingen bij negen van de twaalf patiënten uit ons cohort. Genetische veranderingen met betrekking tot *JAK2* en *SH2B3* werden gevalideerd door FISH-, ddPCR- en/of Sanger-sequencing.

Tumormateriaal van zes patiënten werd geanalyseerd met RNA-seq en de genexpressie in pcAECyTCL werd vergeleken met genexpressie in normale CD8 + T-cellen om gedereguleerde expressiesignaturen in dit lymfoom te bepalen. Opgereguleerde cellulaire processen/routes in pcAECyTCL betreffen de celcyclus (d.w.z. E2F-doelwitten, het G2/M-controlepunt, mitotische spoel), de JAK-STAT-route (via STAT3 en in mindere mate via STAT5), de TNF-α/NF-KB-route en en een sterke ontstekingsreactie. Ten slotte bevestigden functionele studies dat JAK2-fusies geïdentificeerd in pcAECyTCL cytokine-onafhankelijke celoverleving induceren en dat deze activiteit met succes kan worden geblokkeerd door de FDA-goedgekeurde JAK1/2-remmer ruxolitinib.

BPDCN is al het onderwerp geweest van meerdere moleculaire studies, maar het landschap van genomische herrangschikkingen en CNA's is nooit onderzocht met NGS. In hoofdstuk 6 presenteren we de eerste genoombrede analyse van BPDCN gericht op structurele genomische defecten. Tumor DNA van tien patiënten werd geanalyseerd met WGS. Onze analyse identificeerde herrangschikkingen en verstoringen van 54 genen die voornamelijk betrokken zijn bij cytoskelet-geassocieerde processen, adhesie en transcriptionele regulatie. Deze groep omvatte genen die betrokken zijn bij hematologische maligniteiten (d.w.z. *AHI1*, *CD36*, *IKZF1*, *MLLT4*, *MYB*, *TFG*) en andere neoplasiën (d.w.z. FAT1, IQGAP2, *NRG1*, *PIK3C2G*, *PMS1*, *PPFIBP1*, *PTPRD*). We identificeerden verlies van DNA in 6 terugkerende genomische regio's coderend voor tumoronderdrukkers (d.w.z. *CDKN1B*, *ETV6*, *HNRNPK*, *IKZF1*, *RB1* en *SFRP4*) met een belangrijke rol in hematopoëse en celcyclusregulatie. Bovendien kwamen pathogene indels en SNV's in *ASXL1*, *IKZF1*, *NRAS* en *TET2* herhaaldelijk terug is ons tumormateriaal, wat in overeenstemming is met eerdere onderzoeken. Met name werd gevonden dat IKZF1, een gen met een cruciale rol in de vroege ontwikkeling van pDC-precursoren, focaal werd geïnactiveerd in een aanzienlijk deel van de BPDCN-gevallen (gesequentieerde en extensiecohorten gecombineerd, ~50%) als gevolg van structurele genomische defecten. Tumormateriaal van vier patiënten werd bestudeerd met RNA-seq en een vergelijkende analyse van genexpressie in BPDCN-tumoren en rustende pDC's werd uitgevoerd om ontregelde cellulaire processen/paden in dit neoplasme te detecteren. Expressiesignaturen in BPDCN kwamen overeen met bekende fysiologische gevolgen van IKZF1-deficiëntie (d.w.z. afwijkende adhesie, activering van PI-3-K/Akt-route) in immuun cellen. Bovendien was de expressie van geconserveerde IKZF1-doelwitgenen in BPDCN volledig consistent met een verlies van IKZF1-fenotype.

Hoofdstuk 7 vat de belangrijkste bevindingen van dit proefschrift samen en bespreekt deze. Het hoofdstuk eindigt met conclusies en toekomstperspectieven.

LIST OF ABBREVIATIONS

a-CGH	Array comparative genomic hybridization
BID	Benign Inflammatory dermatoses
BPDCN	Blastic plasmacytoid dendritic cell neoplasm
CBCL	Cutaneous B-cell lymphoma
CD30+ LPDs	Primary cutaneous CD30+ lymphoproliferative disorders
ChiP-seq	Chromatin immunoprecipitation sequencing
CNA	Copy number alteration
CTCL	Cutaneous T-cell lymphoma
ddPCR	Droplet digital polymerase chain reaction
FISH	Fluorescence *in situ* hybridization
IHC	Immunohistochemistry
ICC	Immunocytochemistry
LyP	Lymphomatoid papulosis
MF	Mycosis fungoides
PAT-seq	Poly(A)-test sequencing
pcAECyTCL	Primary cutaneous CD8+ aggressive epidermotropic cytotoxic T-cell lymphoma
pcALCL	Primary cutaneous anaplastic large cell lymphoma
pDC	Plasmacytoid dendritic cells
PN	Precursor neoplasm
NGS	Next-generation sequencing
RNA-seq	RNA sequencing
RT-qPCR	Reverse transcription quantitative polymerase chain reaction
sALCL	Systemic anaplastic large cell lymphoma
SKY	Spectral karyotyping
SNP	Single nucleotide polymorphism
SNV	Single nucleotide variant
SS	Sézary syndrome
T-MF	Tumor-stage mycosis fungoides
TS	Targeted sequencing
WES	Whole-exome sequencing
WGS	Whole-genome sequencing
WHO-EORTC	World Health Organization – European Organization for Research and Treatment of Cancer

CURRICULUM VITAE

Armando N. Bastidas Torres was born on December 6 1985 in Quito, Ecuador. He obtained a B.Sc. degree in Biotechnology with Summa Cum Laude honors from Universidad San Francisco de Quito. During his B.Sc. studies he performed internships at the Biotechnology laboratory of Universidad San Francisco de Quito and the Biochemistry laboratory of Hospital de los Valles. Afterwards, he worked for two and half years as research and teaching assistant at the Biotechnology laboratory of Universidad San Francisco de Quito, period during which he supervised undergraduate research projects and assisted teaching the courses molecular biology and genetic engineering. In 2012 he was awarded an excellence scholarship by the Ecuadorian government to pursue a M.Sc. degree in Life Science and Technology at Leiden University in the Netherlands, which he completed with Cum Laude honors in 2014. During his M.Sc. studies he conducted an internship at the UV carcinogenesis group of the Department of Dermatology at Leiden University Medical Center where he worked on the characterization of the transcriptome of skin Lgr6+ (cancer) stem cells under the supervision of dr. F. de Gruijl and dr. G. van de Glind. In 2014 he started his PhD at the cutaneous lymphoma group of the Department of Dermatology at Leiden University Medical Center under the supervision of dr. C.P. Tensen and Prof. dr. M.H. Vermeer. The results of this work are described in this thesis.

PUBLICATIONS

Bastidas Torres, A. N., Cats, D., Mei, H., Fanoni, D., Gliozzo, J., Corti, L., Paulli, M., Vermeer, M. H., Willemze, R., Berti, E., & Tensen, C. P. (2020). Whole-genome analysis uncovers recurrent IKZF1 inactivation and aberrant cell adhesion in blastic plasmacytoid dendritic cell neoplasm. Genes, chromosomes & cancer, 59(5), 295–308.

Bastidas Torres, A. N., Cats, D., Mei, H., Szuhai, K., Willemze, R., Vermeer, M. H., & Tensen, C. P. (2018). Genomic analysis reveals recurrent deletion of JAK-STAT signaling inhibitors HNRNPK and SOCS1 in mycosis fungoides. Genes, chromosomes & cancer, 57(12), 653–664.

Bastidas Torres, A. N., Najidh, S., Tensen, C. P., & Vermeer, M. H. (2018). Molecular advances in cutaneous T-cell lymphoma. Seminars in cutaneous medicine and surgery, 37(1), 81–86.

van de Glind, G. C.*, **Bastidas Torres, A. N.***, Zoutman, W. H., van der Zeeuw, S., Kiełbasa, S. M., de Gruijl, F. R., & Tensen, C. P. (2018). RNA-seq analysis of Lgr6+ stem cells and identification of an Lgr6 isoform. Experimental dermatology, 27(10), 1172–1175. *Shared 1st authorship.

Iżykowska, K., Przybylski, G. K., Gand, C., Braun, F. C., Grabarczyk, P., Kuss, A. W., Olek-Hrab, K., **Bastidas Torres, A. N.**, Vermeer, M. H., Zoutman, W. H., Tensen, C. P., & Schmidt, C. A. (2017). Genetic rearrangements result in altered gene expression and novel fusion transcripts in Sézary syndrome. Oncotarget, 8(24), 39627–39639.

Schrader, A. M., Chung, Y. Y., Jansen, P. M., Szuhai, K., **Bastidas Torres, A. N.**, Tensen, C. P., & Willemze, R. (2016). No TP63 rearrangements in a selected group of primary cutaneous CD30+ lymphoproliferative disorders with aggressive clinical course. Blood, 128(1), 141–143.

ACKNOWLEDGEMENTS

First and foremost, I wish to express my deepest gratitude to my promotor prof. dr. Maarten H. Vermeer and my co-promotor dr. Cornelis (Kees) P. Tensen. Maarten, I am truly grateful for the opportunities that you and the department have given me throughout my PhD journey. Kees, I will forever be indebted to you for believing in me, your guidance, support, generosity and unparalleled kindness. To both of you, I wish and hope that our paths cross again in the future. Also, I would like to pay my special regards to the Dutch Cancer Society for the financial support that made possible the research contained in this thesis.

I would like to extend my deep gratitude to Leon Mei and Davy Cats from the Sequencing Analysis Support Core of LUMC, who were actively involved in all studies contained in this thesis, and without whom the successful completion of these research projects would have not been possible. Also, I would like to express my sincere gratitude to Coby Out for the invaluable help provided in the studies presented in chapters 4 and 5. Coby, your knowledge, technical expertise and selfless willingness to contribute to both projects were key to their success, and I want you to know that it was a pleasure working with you. My gratitude also goes out to dr. Karoly Szuhai and Danielle de Jong from the department of Cell and Chemical Biology of LUMC for the technical support provided in the study presented in chapter 3.

To my colleagues at the departments of Dermatology and Ophthalmology: Wim, Safa, Yixin, Rutger, Koen, Suzanne, Remco, Rein, Eirini, Catarina, Rajiv, Arnout, Frank, Abdoel, Marion, Tim, Nelleke, Rogier, Pieter and Mieke: thank you for the moments and talks shared inside and outside LUMC throughout the years. In my memories, The Netherlands will always be linked to your voices and faces as well as the experiences that we shared.

I could not end this section without acknowledging the support and great love of my family. To my beloved parents, Rosa and Clímaco, your moral support and love have infused me with the necessary strength to overcome the challenges that life has thrown at me during all these years. To my sisters, Alejandra and Paula, thank you for always keeping me in your minds and your encouragement. Finally, Xuechan, thanks for being my biggest supporter, my emotional shelter and my best friend.