Hematology Research Unit
Department of Clinical Chemistry and Hematology
University of Helsinki
and
Helsinki University Hospital Comprehensive Cancer Center,
Department of Hematology,
Helsinki, Finland

Doctoral School in Health Sciences,
Doctoral Program in Clinical Research

Characterization of Mature T-Cell Leukemias by Next-Generation Sequencing and Drug Sensitivity Testing

Emma Andersson
University of Helsinki
Finland

ACADEMIC DISSERTATION

To be presented, with the permission of the Medical Faculty, University of Helsinki,
for public examination in Biomedicum lecture hall 2, Haartmaninkatu 8,
on February 10th, 2017, at 12 o'clock noon.

Helsinki 2017

Supervised by:

Professor Satu Mustjoki, M.D., Ph.D.
Hematology Research Unit
Department of Clinical Chemistry and Hematology
University of Helsinki and
Helsinki University Hospital Comprehensive Cancer Center
Helsinki, Finland

Reviewed by:

Adjunct professor Eeva-Riitta Savolainen. M.D, Ph.D.
Nordlab Oulu
University of Oulu
Oulu, Finland

Adjunct professor Samuel Myllykangas, Ph.D.
Division of Genetics
Department of Biosciences
University of Helsinki
Helsinki, Finland

Official opponent:

Professor Richard Rosenquist Brandell, M.D, Ph.D.
Department of Immunology, Genetics and Pathology
Uppsala University
Uppsala, Sweden

Cover picture: Anni Moilanen
ISBN 978-951-51-2891-1 (paperback)
ISBN 978-951-51-2892-8 (PDF)
ISSN 2342-3161
http://ethesis.helsinki.fi
Hansaprint Oy
Helsinki 2017

A journey of a thousand miles starts with a single step

- Laozi

Table of contents

Abbreviations

ABL1	Abelson murine leukemia viral oncogene homolog 1
AICD	activation-induced cell death
AKT	protein kinase B
ALL	acute lymphoid leukemia
AML	acute myeloid leukemia
ANC	absolute neutrophil count
APC	antigen-presenting cell
ATLL	adult T-cell lymphoma
ATM	ataxia telangiectasia mutated gene
ATP	adenosine triphosphate
AUC	area under the curve
BCR	breakpoint cluster region protein
BM	bone marrow
BCL	B-cell lymphoma
BRAF	v-raf murine sarcoma viral oncogene homolog B1
BSA	bovine serum albumin
BWA	Burrows Wheeler aligner
BzCl	benzethonium chloride
CD	cluster of differentiation
cDNA	complementary DNA
CDK	cyclin-dependent kinase
CDR	complementarity determining regions
CLL	chronic lymphocytic leukemia
CML	chronic myeloid leukemia
CNV	copy-number variation
CTL	cytotoxic T lymphocyte
CXCR4	C-X-C chemokine receptor type 4
CyA	cyclosporine A
DC	dendritic cell
Del	deletion
DMEM	Dulbecco's Modified Eagle Medium
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotide
DSRT	drug sensitivity and resistance testing
DSS	drug sensitivity score
EDTA	ethylenediaminetetraacetic acid
EGF	epidermal growth factor

FACS fluorescence-activated cell sorting

FasL Fas ligand

FBS fetal bovine serum

FCM-A fludarabine, cyclophosphamide, and mitoxantrone followed by alemtuzumab consolidation

FGR Gardner-Rasheed feline sarcoma viral (v-fgr) oncogene homolog

GERP genomic evolutionary rate profiling

HEK-293 human embryonic kidney cell line

HSC hematopoietic stem cell

HSCT hematopoietic stem cell transplant

HTLV human T-cell lymphotrophic virus

IFN-γ interferon-gamma

IGV Integrative genomics viewer

IHC immunohistochemistry

IL interleukin

Inv inversion

IPA Ingenuity Pathway Analysis

JAK Janus kinase

LGL large granular lymphocyte

MAPK mitogen-activated protein kinase

Mb megabase

MCL-1 induced myeloid leukemia cell differentiation protein

MCM mononuclear cell media

MHC major histocompatibility complex

MNC mononuclear cell

MPD myeloproliferative disorder

mRNA messenger RNA

mTOR mammalian target of rapamycin

MTX methotrexate

NFAT nuclear factor of activated T-cells

NFκB nuclear factor kappa-B

NGS next generation sequencing

NK natural killer

NRP1 neuropilin-1

PAGE polyacrylamide gel electrophoresis

PB peripheral blood

PBS phosphate-buffered saline

PCR polymerase chain reaction

PFS progression-free survival

PI3K phosphptidylinositol 3 kinase

PIAS protein inhibitors of activated STAT
PLL prolymphocytic leukemia
PTP protein tyrosine phosphatase
PTPRT receptor-type tyrosine-protein phosphatase T
RA rheumatoid arthritis
RAD21 double-strand-break repair protein rad21 homolog
Ras rat sarcoma
RB retinoblastoma
RNA ribonucleic acid
SDS sodium dodecyl sulfate
sDSS selective drug sensitivity score
SNP single nucleotide polymorphism
SOCS suppressor of cytokine signaling
SLAMF6 SLAM family member 6
SLIT2 slit guidance ligand 2
STAT signal transducer and activator of transcription
TAS target addiction score
TCL1A T-cell leukemia 1A
TCR T-cell receptor
T-LGLL T-cell large granular lymphocytic leukemia
TNF tumor necrosis factor
T-PLL T-cell prolymphocytic leukemia
TRAIL TNF-related apoptosis inducing ligand
TYK2 tyrosine kinase 2
VAF variant allele frequency
WES whole exome sequencing
WHO World Health Organization
WT wild type

List of original publications

This thesis is based on the following original publications, which are referred to in the text by their Roman numerals:

I. **Andersson EI**, Rajala HLM, Eldfors S, Ellonen P, Olson T, Jerez A, Clemente M, Kallioniemi O, Porkka K, Heckman C, Loughran TP, Maciejewski JP, Mustjoki S. Novel Somatic mutations in large granular lymphocytic leukemia affecting the STAT-pathway and T-cell activation**.** *Blood Cancer J.* 2013; 3:e168

II. **Andersson E**, Kuusanmäki H, Bortoluzzi S, Lagström S, Parsons A, Rajala H, van Adrichem A, Eldfors S, Olson T, Clemente MJ, Laasonen A, Ellonen P, Heckman C, Loughran TP, Maciejewski JP, Mustjoki S. Activating somatic mutations outside the SH2-domain of STAT3 in LGL-Leukemia.
Leukemia 2015; 30(5):1204-1208.

III. **Andersson EI***, Tanahashi T*, Sekiguchi N, Gasparini V, Bortoluzzi Sa, Kawakami T, Sekiguchi N, Matsuda K, Mitsui T, Eldfors S, Bortoluzzi St, Coppe A, Binatti A, Lagström S, Ellonen P, Fukushima N, Nishina S, Senoo N, Sakai H, Nakazawa H, Kwong Y-L, Loughran TP, Maciejewski JP, Mustjoki S[Ŧ], Ishida F[Ŧ]. High incidence of activating STAT5B mutations in CD4-positive T-cell large granular lymphocyte leukemia. *Blood* 2016; 128(20): 2465-2468.

IV. **Andersson EI**, Pützer S, Yadav B, Dufva O, Khan S, He L, Sellner L, Schrader A, Crispatzu G, Oles M, Zhang H, Adnan S, Lagström S, Bellanger D, Mpindi JP, Eldfors S, Pemovska T, Pietarinen P, Lauhio A, Tomska K, Cuesta-Mateos C, Faber E, Koschmieder S, Brümmendorf T, Kytölä S, Ellonen P, Kallioniemi O, Wennerberg K, Ding W, Stern MH, Huber W, Anders S, Tang J, Aittokallio T, Zenz T, Herling M, Mustjoki S. Discovery of novel drug sensitivities in T-prolymphocytic leukemia (T-PLL) by high-throughput ex vivo drug testing and mutation profiling. Submitted.

*,[Ŧ] Equal contributions

Abstract

Mature T-cell malignancies comprise a heterogeneous group of diseases with widely variable clinical courses, ranging from indolent, slowly progressing to rapidly progressing disease, leading to death. While some compounds have been found to be active in these diseases, much work remains to be done to characterize the disease entities genetically and molecularly and develop effective therapeutic options. Recent studies have found recurrent genetic mutations in the JAK/STAT pathway that contribute to the pathogenesis of T-LGL leukemia and PLL leukemia. The aim of this PhD project was to reveal new somatic mutations underlying different mature T-cell malignancies and further elucidate their functional impact on the pathogenesis of the disease. Furthermore, *ex vivo* drug screening of selected patient samples was applied in an effort to find clinically relevant drugs for the individual patient.

In the first study, exome sequencing was applied to three *STAT*-mutation negative T-LGL leukemia patients to elucidate the molecular background of this subset of patients. Mutations in genes associated with either the STAT-pathway directly (such as *PTPRT*) or T-cell activation were observed in all patients. These novel mutations are potentially biologically relevant and are associated with a similar disease phenotype as that of patients with mutations in the *STAT3* gene.

In the second study, the exome sequencing of additional LGL leukemia patients revealed recurrent *STAT3* mutations outside the hotspot SH2-domain. Further targeted amplicon sequencing of the entire *STAT3* gene in a larger T-LGL cohort revealed mutations in the coiled-coil and DNA-binding domain present in 4% of patients. Similarly, in the third study, *STAT5B* mutations previously found to be rare in CD8+ LGL leukemia (2%) were discovered to be significantly over-represented in CD4+ cases (55%). For both studies, the activation and phosphorylation patterns of the new mutations were confirmed with functional assays.

In the fourth study, T-PLL leukemia patient samples were screened with a drug sensitivity and resistance testing platform to elucidate drug sensitivity patterns. The patients were divided by selective sensitivities to HDAC, PI3K/AKT/mTOR, HSP90, and JAK inhibitors. Intriguingly, all T-PLL samples were sensitive to the CDK inhibitor SNS-032, which was shown to inhibit distal TCR signaling and points to a centrally disturbed cell cycle regulation. Furthermore, different classes of p53 activators showed T-PLL specific responses, indicating a previously unexplored targeted treatment for these patients. In combination with the drug responses, a targeted sequencing of known JAK/STAT mutations was applied together with gene expression profiling to molecularly characterize this rare disease. Strikingly, drug responses did not link to the presence of *JAK/STAT* mutations, *TCL1A*-activating translocations, or *ATM* deletion status, suggesting that screening for recurrent genetic biomarkers does not readily translate into applicable deductions on effective therapeutic strategies in T-PLL. However, the *ex vivo* drug response data enabled the discovery of new drug classes with potential efficacy in T-PLL.

Introduction

Mature T-cell leukemias are a group of uncommon lymphoid neoplasms. These disorders have very variable clinical features, ranging from indolent to rapidly progressive diseases, ultimately leading to death. Cytogenetic aberrations and, more recently, recurrent genetic mutations have been found to contribute to their pathogenesis[1]. Conventional chemotherapy lacks significant effect in this group of diseases, and therapies vary from immunosuppression to treatment with antiviral agents and monoclonal antibodies.

In T-cell large granular lymphocytic leukemia (T-LGLL), the disease-driving cells are differentiated effector memory T-cells that are resistant to apoptosis by activation-induced cell-death (AICD). Many anti-apoptotic pathways are known to be activated in T-LGLL, one of the most studied being the Janus kinase (JAK)/signal transducer and activator of transcription (STAT) pathway controlling the transcription of proteins connected to cell cycle control and proliferation. The recent discovery of activating somatic mutations in the SH2-domain of *STAT3* in 40% of T-LGLL patients has further highlighted the JAK/STAT pathway as a driver of T-LGLL pathogenesis[2]. As the diagnosis of T-LGLL is often based upon the examination of clonal expansions in the T-cell population, which are also frequently seen in non-malignant conditions following antigen exposure, *STAT3* mutations provide a novel diagnostic biomarker for this disease. Even so, over half of T-LGLL cases are *STAT*-mutation negative, raising the question of what is driving their disease genetically.

T-cell prolymphocytic leukemia (T-PLL) is, in contrast to the indolent T-LGLL, an aggressive disease of mature T-cells with an urgent need for targeted therapies in light of its notorious chemo-refractory behavior. Currently, the best treatment for T-PLL is the monoclonal antibody alemtuzumab, which has prolonged the median overall survival up to 20 months, but responses are still transient and relapses inevitable[3]. Cytogenetic abnormalities are seen in all cases of T-PLL, the most commonly seen being translocations in *TCL1A* and deletions in *ATM*. The activation of the JAK/STAT pathway has also been established with activating mutations in 76% of T-PLL patients[4]. Given these recent findings, the inhibition of JAK/STAT pathway activation remains a promising treatment opportunity.

Translating the recent genomic discoveries in mature T-cell leukemias into clinically actionable therapy poses a significant challenge. Hopefully a better understanding of the pathogenesis of these disorders will lead to more precise therapy for patients afflicted with this difficult-to-treat group of diseases.

Review of the literature

1. The tale of T-cells

1.1. Lifespan of a T-cell

Peripheral T lymphocytes as a population comprise of a mixture of naïve, effector, and memory cells. T-cells develop from lymphoid progenitors in the bone marrow that also give rise to B lymphocytes. A few of these progenitors migrate through circulation to the thymus, where the T lymphocytes mature[5]. The recirculating pool of mature T-cells is formed during childhood through the continuous release of naïve T-cells from the thymus into the blood. In adults, the pool of mature T-cells is relatively self-sufficient, and the input of new T-cells from the thymus declines. Studies on T-cell turnover indicate that most peripheral T-cells can remain in a resting state for years in humans. An analysis of the phenotype of dividing vs. non-dividing cells indicates that naïve T-cells are long-lived resting cells, while most effector and memory T-cells have a faster turnover[6]. The estimated lifespan of circulating T-cells is 6 months to 1 year, while effector cells have lifespans of only a few weeks.

1.2. The T-cell receptor

The developing thymocytes pass through a series of distinct phases marked by changes in the expression of the T-cell receptor (TCR) but also different cell-surface proteins such as CD3, CD4, and CD8. These antigens can be used as markers for T-cells at different stages of development[7]. The earliest cell population in the thymus does not express any of these proteins, and therefore they are called "double negative" thymocytes (CD4-CD8-)[8]. These cells include precursors that give rise to two T-cell lineages with different TCRs; the majority population of αβ and the minority population of γδ. αβ T-cells later divide into either CD4+ helper cells or CD8+ cytotoxic cells[7].

Each T-cell harbors on its surface about 100,000 antigen receptors that all recognize the same epitope. The TCR must undergo extensive rearrangement to give rise to cells that recognize all different antigens (called the TCR repertoire). The TCR is made up of two different protein chains that are produced by a combination of the deletion and joining of variable (V), diversity (D), and joining (J) gene segments to produce a unique antigen receptor[9]. In developing thymocytes, the γ, δ, and β genes rearrange almost simultaneously. If a functional γδ receptor is formed before a functionally rearranged β-chain, the precursor commits to the γδ lineage by halting further gene arrangement[10]. The production of a functional β chain and signaling by the pre T-cell receptor commits the precursor to the αβ lineage, where both CD4 and CD8 co-receptor proteins are expressed on the cell surface (double positive thymocytes)[8]. αβ TCRs are more common and present in 90–95% of T-cells, while γδ TCRs usually present only in around 5% of mature T-cells. Mature αβ T-cells have an extensive variety of TCRs, which means that they can recognize almost any antigen and therefore

have an important role in the adaptive immune system. γδ T-cells have a more limited TCR variety, but, in contrast, they possess the ability to respond to intact antigens. This quality places them in between the innate and adaptive immune system[11].

1.3. Subtypes of T-cells

The further development of the T-cell is entirely determined by the interaction between the TCR and the self-peptides presented by the major histocompatibility complexes (MHC) on the thymic stroma. T-cells that react too strongly to the presented peptides are automatically deleted in the thymus by antigen-presenting cells (APCs) in a process of negative selection. On the other hand, CD4+ CD8+ thymocytes, whose receptors interact normally with the peptide:MHC complexes, undergo positive selection and later become mature CD4+ or CD8+ T-cells. By the end of thymic selection, mature thymocytes ready for export to the periphery will stop expressing one of the co-receptors. Simply put, the co-receptor expression depends on the class of MHC molecule the thymocyte interacted with during the selection; contact with MHC class I molecules is required for cytotoxic CD8 T lymphocyte development, while class II molecules produce CD4 helper T-cells[12].

Mature CD8 cytotoxic T-cells kill cells that display peptide fragments of intracellular pathogens (i.e., viruses) bound to MHC class I molecules at the cell surface (**Fig 1**). The naïve CD8 T-cells still need to be activated or primed by APCs before becoming cytotoxic. Mature dendritic cells often activate CD8 T-cells, because they naturally exhibit the high intrinsic co-stimulatory activity that is needed to activate cytotoxic cells. They send signals through the TCR that induces IL-2 synthesis and drives the differentiation and proliferation process. The IL-2 stimulates cytokine receptors that in turn induce the production of granule components needed for lymphocyte-mediated toxicity[13]. As opposed to the NK-cells, cytotoxic T-cells can also undergo extensive expansion after activation, giving them a major role in response to prolonged antigen stimuli or high antigen loads.

After contact with peptide/MHC class II complexes on APC, CD4+ helper T-cells initiate specific immune responses by secreting soluble proteins such as IL-2 or IFN-γ. These cytokines provide the required assistance for the activation and differentiation of other lymphoid cells. CD4+ T-cells also regulate immune responses through direct cell–cell contact and not only through secreted factors.

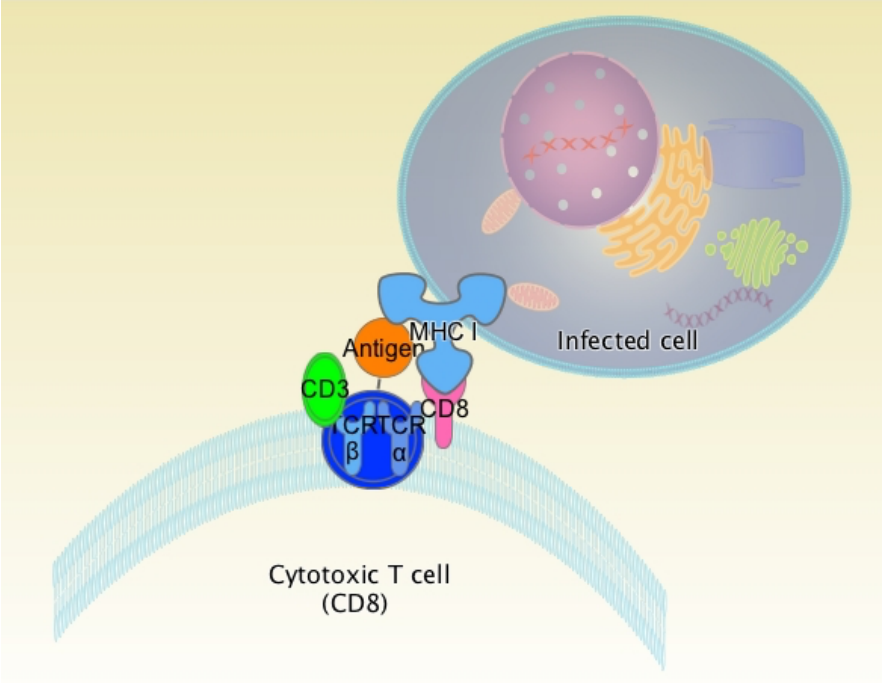


Figure 1. *Mature CD8 cytotoxic T-cells kill cells that display peptide fragments of intracellular pathogens (i.e., viruses) bound to MHC class I molecules at the cell surface. Image edited from Ingenuity Pathway Analysis software.*

1.4. T-cell elimination

When a T-cell encounters its antigen, it results in rapid proliferation of clonal effector cells over a few days. If the response continues uncontrolled it could cause harm, and therefore activated T-cells undergo activation-induced cell death (AICD)[14]. During the activation of cytotoxic T-cells, the transcription of both FAS and FAS ligands (FASL) is upregulated. If T-cells expressing FAS, a member of the tumor necrosis factor receptor family, meet other cells expressing FASL on their surface, Fas+ T-cells will undergo apoptosis. Without such a mechanism of T-cell elimination, the immune compartment would soon be overpopulated. Unfortunately, some tumors also express FASL, which provides them with an escape mechanism from immune surveillance by activated tumor-specific T-cells.

2. Mature T-cell malignancies

T-cell neoplasms occur at various stages of T-cell maturation, with, for example, T-cell acute lymphoblastic leukemia (T-ALL) developing at the early T-cell progenitor stage and peripheral T-cell lymphomas harboring neoplastic cells with markers of their more mature post-thymic counterparts (**Table 1**).

Table 1. *WHO classification of mature T and NK cell neoplasms 2016*[15]

Mature T-cell and NK-cell neoplasms
T-cell prolymphocytic leukemia
T-cell large granular lymphocytic leukemia
Chronic lymphoproliferative disorder of NK cells
Aggressive NK cell leukemia
Systemic EBV+ T-cell Lymphoma of childhood*
Hydroa vacciniforme-like lymphoproliferative disorder*
Adult T-cell leukemia/lymphoma
Extranodal NK/T-cell lymphoma, nasal type
Enteropathy-associated T-cell lymphoma
Monomorphic epitheliotropic intestinal T-cell lymphoma*
Indolent T-cell lymphoproliferative disorder of the GI tract *
Hepatosplenic T-cell lymphoma
Subcutaneous panniculitis- like T-cell lymphoma
Mycosis fungoides
Sézary syndrome
Primary cutaneous CD30 positive T-cell lymphoproliferative disorders
Lymphomatoid papulosis
Primary cutaneous anaplastic large cell lymphoma
Primary cutaneous gamma-delta T-cell lymphoma
Primary cutaneous CD8 positive aggressive epidermotropic cytotoxic T-cell lymphoma
*Primary cutaneous acral CD8+ T-cell lymphoma**
*Primary cutaneous CD4 positive small /medium T-cell lymphoproliferative disorder**
Peripheral T-cell lymphoma, NOS
Angioimmunoblastic T-cell lymphoma
*Follicular T-cell lymphoma**
*Nodal peripheral T-cell lymphoma with TFH phenotype**
Anaplastic large cell lymphoma, ALK positive
Anaplastic large cell lymphoma, ALK negative *
*Breast implant-associated anaplastic large cell lymphoma**

**Changes from the 2008 classification; provisional entities are listed in italics.*

Mature T-cell leukemias are a heterogeneous group of neoplasms with the common clinical feature of a mature neoplastic T-cell population in the blood circulation. The involvement of the bone marrow, lymph nodes, and spleen are also frequently present in these diseases. T-cell leukemias have widely different clinical behaviors. In T-cell large granular lymphocytic leukemia (T-LGLL), the clinical course is usually indolent, with the primary clinical manifestation being an associated cytopenia. On the other hand, T-cell prolymphocytic leukemia (T-PLL) is an extremely malignant process, with poor overall survival. Recent studies have added to our knowledge of the genetic background of these leukemias (**Table 2**) and have created a roadmap for potential future targeted therapies.

Table 2. *Summary of recurrent genetic alterations in mature T-cell leukemias. Data compiled from references* [2,4,16-23]

Disease	Genetic/molecular alteration	Frequency	Effect
T-cell prolymphocytic leukemia	TRA/D-TCL1 fusion	75%	Aberrant expression of TCL1 and activation of the AKT pathway
	TRA/D-MTCP1 fusion	10%	Aberrant expression of MTCP1 and activation of the AKT pathway
	JAK1/JAK3/STAT5b mutations	75%	Constitutive activation of the JAK/STAT pathway
	ATM deletions/mutations	70%	Impaired DNA damage response
T-cell large granular lymphocytic leukemia	STAT3 mutations	30-40%	Constitutive activation of the JAK/STAT pathway
	STAT5b mutations	Rare	Constitutive activation of the JAK/STAT pathway
Adult T-cell lymphoma	JAK3 mutations	10%	Constitutive activation of the JAK/STAT pathway
	CCR4 mutations	25%	PI3K/AKT activation
Sezary syndrome	TCF3 deletion	70%	Enhanced cell cycle progression
	TP53 deletion	75%	Impaired DNA damage response
	ARID1A mutations	40%	Chromatin remodelling

2.1 T-LGL leukemia

2.1.1. Clinical characteristics and diagnosis

In contrast with the other mature T-cell leukemias, T-LGLL is characterized as an indolent clonal proliferation of cytotoxic T-cells. The progression of T-LGLL is very slow, with a median survival after diagnosis >10 years. Although the disease can stay asymptomatic for a long time after diagnosis, about 60% of the patients become symptomatic during the course of the disease[24]. The most common complication associated with T-LGLL is neutropenia, and recurrent bacterial infections are therefore common initial clinical manifestations of LGL leukemia. Other complications associated with T-LGLL are anemia (50%), thrombocytopenia (20%), and splenomegaly, which is seen in two-thirds of patients. Autoimmune disorders are seen in 25–35% of T-LGLL cases, and especially common are Sjögrens syndrome and rheumatoid arthritis (RA) as well as pure red cell aplasia (in Asian populations)[25].

The diagnosis rests first and foremost on the observation of a persistent increase of LGLs in the peripheral blood. The phenotype of the cells are then analyzed to confirm the type of LGL leukemia, and the clonality is confirmed by performing TCRβ and γ gene rearrangement studies (PCR). Morphologically, the LGLs are large blood cells (15–18 μm) with a reniform or round nucleus surrounded by abundant cytoplasm containing azurophilic granules. Most patients present with an LGL count greater than 2.0×10^9/L, while the normal range is around 0.2–0.4×10^9/L[26]. A lower count may also be compatible with LGL, though establishing a diagnosis in such patients is more difficult.

LGL leukemia cells usually show a mature postthymic phenotype and typically co-express CD3+CD8+CD16+CD57+ antigens[27]. T-LGLs are also CD45RA+ and CD62L-, which is consistent with the effector memory T-cell phenotype. The T-cell receptor is typically of the αβ subtype although γδ-cases are clinically very similar to TCRαβ, with favorable survival of 85% at 3 years[28].

Although some numerical and structural chromosomal abnormalities have been reported in individual cases of T-LGLL, there are no known recurrent abnormalities to make cytogenetic examination necessary.

2.1.2. Pathogenesis of T-LGL leukemia

T-LGLL is not a typical leukemia for two reasons. Firstly, the disease rarely progresses to a more malignant phenotype, and, secondly, no excessive accumulation of malignant cells usually exists. Actually, the clonal cells resemble normal antigen-activated CD8+ CD57+ effector cells. Both normal and malignant clones express perforin and Fas-L and have the ability to suppress neutrophil development *in vitro*[29,30]. While the etiology of the disease remains to be elucidated, there are a few theories about the origin of the clonal expansion seen in the cytotoxic T-cells.

As the disease-causing cells in T-LGLL are effector-memory cytotoxic lymphocytes (CTLs), this suggests that there could be some exogenous antigens causing the leukemic expansion. These could be of viral origin, as patients with NK-LGLL are often seropositive for p21e, an envelope protein found in human T-cell lymphotrophic virus (HTLV-1). As the patients are negative to the virus itself, it has been hypothesized that a cellular or retroviral protein similar to this envelope epitope may play a significant role in LGL leukemia[31]. A few studies have analyzed the TCRβ complementarity determining region 3 (CDR3) repertoire in an effort to elucidate a possible antigen target. However, although the data suggest restricted use of the TCR repertoire in T-LGLL, no common antigens have been recognized[32].

Under normal physiological circumstances, activated CTLs are removed after antigen-driven expansion by apoptosis induced by different signaling pathways. One theory is that the cytotoxic T-cells in LGL leukemia fail to undergo apoptosis, which results in a persistent expansion of CTLs in the blood (**Fig 2**). This resistance results from the constitutive activation of a number of signaling pathways, such as the Fas/FasL[33] and the JAK/STAT pathways[34]. For example, LGL leukemia cells have been shown to express high levels of Fas/FasL on their surface, yet they are resistant to the homeostatic apoptosis mediated by Fas[33]. The clonal transformation of LGLs may also be explained by a constitutive overexpression of pro-survival and anti-apoptotic transcription factors. One example is STAT3, which is a transcription factor activated by phosphorylation leading to the activation of genes, affecting cell growth and apoptosis. STAT3 has been shown to be constitutively activated in T-LGLL, which may cause the inhibition of homeostatic apoptosis, such as the inhibition of the downstream Fas receptor signaling through the initiation of anti-apoptotic MCL-1[35]. As additional proof, the blocking of the STAT3-signaling in T-LGLL is shown to lead to the reversal of the Fas resistance in leukemic cells[35,36].

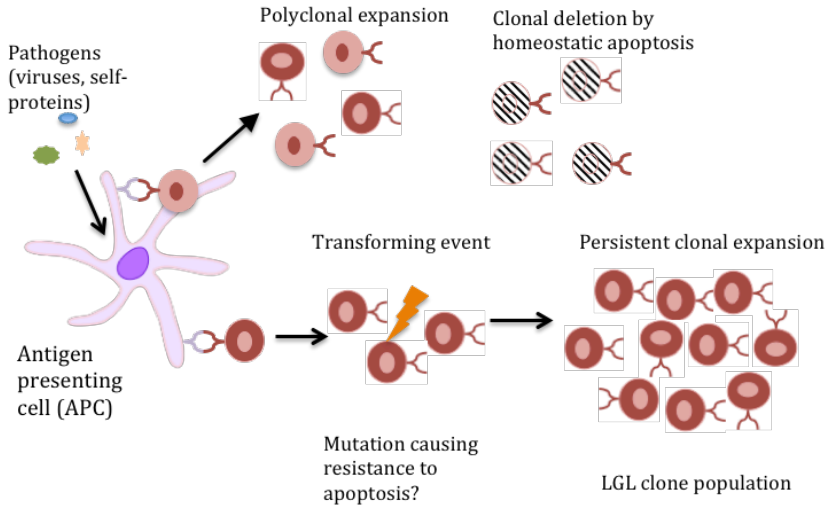


Figure 2. *Inciting events and clonal transformation in T-LGL leukemia. In normal circumstances, activated CTLs are removed by activation-induced cell death (AICD) to maintain homeostasis. In T-LGL an early transforming event may hit the expanded CTL clone, which leads to persistent clonal expansion.*

The constitutive activation of STAT3 in T-LGLL patients is in part explained by the high frequency (40%) of activating mutations clustered in the Src-like homology 2 (SH2) domain of *STAT3*[2]. In a smaller fraction of patients (2%), the JAK/STAT pathway is activated through STAT5B mutations[19], causing a more aggressive disease course than is typical in T-LGL leukemia. However, the JAK/STAT pathway is also kept activated by non-mutational mechanisms. Increased IL-6 secretion by normal mononuclear cells has been described in T-LGLL patients, leading to increased JAK/STAT signaling. This activation is also supported by the frequent epigenetic inactivation of the suppressor of cytokine signaling 3 (SOCS3), an inhibitor of the JAK/STAT pathway[37]. In some patients, multiple different *STAT3*-mutated clones are present at diagnosis[38]. These data suggest that T-LGLL is initially an oligoclonal process arising out of a background of dysregulated immune stimulation and STAT3 activation, with eventual dominance by a single clone.

2.1.3. Treatment

There is no standard treatment for patients with T-LGL leukemia. The management of asymptomatic patients is based on careful observation, as up to half of the patients may not require therapy. When there is a need for treatment, the goal is to alleviate the symptoms of the cytopenias rather than to eliminate all the malignant clones. Indications for treating patients are the development of recurrent infections, severe neutropenia (< 0.5×10^9/L), symptomatic anemia or thrombocytopenia, and symptomatic slenomegaly[25]. Other indications include autoimmune conditions (e.g., RA) associated with T-LGL leukemia requiring therapy and the presence of systemic symptoms.

To date, immunosuppressive therapy remains the main treatment, including in particular single agents methotrexate (MTX)[39], oral cyclophosphamide, and cyclosporine (CyA)[40]. In patients with autoimmune disease, a common therapeutic approach is MTX with or without steroids[41], but this is associated with frequent relapse[42]. Recent retrospective data suggest that cyclophosphamide may generate a more durable response in T-LGL leukemia patients[43]. In a limited number of patients it was found that complete molecular response (based on *STAT3* mutation detection) may be more common in cyclophosphamide-treated patients than in those treated with MTX[38]. Novel therapeutic approaches indicated by preclinical data include STAT3 inhibition[35] and different demethylating agents aimed at overcoming the loss of SOCS3 activity[37].

2.2. PLL leukemia

2.2.1. Clinical characteristics and diagnosis

T-cell prolymphocytic leukemia (T-PLL) is a rather infrequent, mostly aggressive post-thymic T-cell neoplasm. Comprising ~2% of mature lymphoid tumors, it is yet the most frequent of their primary leukemic forms. T-PLL typically follows an aggressive clinical course, but a minority of patients (~15%) may be asymptomatic at diagnosis. This indolent phase can prevail for a variable length of time and may extend to several years[44]. However, progression is unavoidable and may be very rapid when it occurs. This is a disease of older adults with a median age of onset of 61 years and occurs more frequently in males than females[45].

Clinical features include rapidly progressive lymphocytosis (>100×10^9/L), splenomegaly (66% of cases), hepatomegaly and/or lymphadenopathy (50% of cases), and some patients present with skin manifestations (~20%). The neoplastic cells in T-PLL are intermediate-sized lymphocytes, with variably prominent nucleoli. Cytoplasmic blebbing is a characteristic morphologic feature. Phenotypically T-PLL tumor cells are CD2+CD3+CD5+CD7++CD52++ and CD4/8 variable. Most cases are CD4-positive; however, a significant subset exhibits the co-expression of both CD4 and CD8, a feature that is uncommon in other mature T-cell malignancies[46].

Due to the rareness of T-PLL, clinicians will often only see a case of T-PLL once every 5–10 years, which makes the recognition of the disorder challenging. Confirmation of the diagnosis requires a systematic approach and careful integration of morphology test results with immunophenotyping and cytogenetics.

2.2.2. Recurrent genetic abnormalities

Cytogenetic and/or molecular genetic abnormalities are seen in all cases of T-PLL[47]. Almost 75% of T-PLL cases harbor chromosome 14 inversions or translocations, resulting in aberrant activation of the proto-oncogene T-cell leukemia 1A (*TCL1A*)[48]. *TCL1A* is located in the 14q32 region and has been shown to promote malignant transformation of T lymphocytes in transgenic mice carrying the *Tcl1* gene[49]. However, as most cases show complex karyotypes (>3 aberrations), *TCL1A* does not seem sufficient to drive the leukemogenesis of T-PLL alone. Furthermore, in the majority of T-PLL cases the ataxia telangiectasia mutated (*ATM)* gene, involved in the activation of the DNA damage checkpoint, is mutated or deleted. *ATM* is thought to be a tumor suppressor gene, as DNA double strand breaks result in rapid autophosphorylation and the activation of ATM, leading to checkpoint activation and the phosphorylation of substrates that regulate cell-cycle progression, DNA repair, and cell death. ATM is also known to interact with TCL1, which leads to the enhancement of nuclear factor kappa-B (NF-κB) and increased cell proliferation[50].

Recently it was also reported that mutations in genes involved in the *JAK/STAT* pathway were found in 76% of T-PLL cases (**Fig 3**)[4,51,52]. A high frequency of mutations was found specifically in the pseudokinase domain of *JAK3,* which regulates the signaling activity of the protein. This leads to recurrent activation of JAK3, which causes constitutive STAT5 hyperphosphorylation and the deregulation of the JAK/STAT pathway.

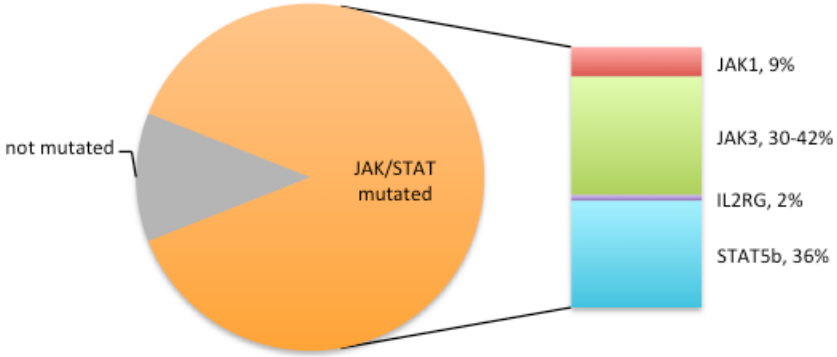


Figure 3. *Mutations of the JAK/STAT pathway in T-PLL. Activating mutations of the JAK3 (30–42% of cases), STAT5B (36%), JAK1 (9%), and IL2RG (2%) genes were identified by Bellanger et al.,* *Bergmann et al., and Kiel et al. in 2014.*

2.2.3. Treatment

The prognosis for T-PLL patients is poor due to non-responsiveness to conventional chemotherapy, resulting in a median survival of only 4 months for non-responders[45]. Currently, the best treatment for T-PLL is alemtuzumab, a monoclonal antibody that targets CD52+ mature lymphocytes for destruction but not the stem cells from which they are derived[53]. This treatment has prolonged the median survival to up to 20 months, but the response is transient and should be followed by consolidation with a hematologic stem cell transplant (HSCT) when possible[3]. This regimen extends the median survival to >4 years, but relapse and toxicity post-HSCT remain problematic.

Recently, sequential combination chemoimmunotherapy using fludarabine, mitoxantrone and cyclophosphamide followed by alemtuzumab consolidation (FCM-A) was examined in both previously treated and untreated T-PLL patients[54]. FCM-A was associated with a favorable overall response rate of 92% in this study, with median progression-free survival (PFS) of approximately 1 year. Single-agent bendamustine has some efficacy in alemtuzumab-refractory T-PLL, with an overall response rate of 53% and a PFS of 5 months[55]. Given the suboptimal long-term response of T-PLL to current therapy and the increased understanding of its molecular pathogenesis, targeted therapies are being evaluated. Given the recent finding of frequent mutational JAK/STAT pathway activation, STAT inhibition seems to be a promising treatment opportunity. In one encouraging study, the treatment of primary T-PLL cells with pimozide, a selective STAT5 inhibitor, resulted in a significant reduction of cell proliferation and viability[4].

3. JAK/STAT signaling in hematologic malignancies

The JAK/STAT pathway has a central role in the signaling of cytokines by regulating cell growth, differentiation, survival, and immune response[56]. Due to its critical roles in hematopoietic development, it is not surprising that it plays an important oncogenic role in lymphoproliferative malignancies[57]. Many cancers and blood malignancies have been associated with the constitutive activation of the STAT family of proteins, which as a rule is dependent upon JAK-mediated tyrosine phosphorylation for transcriptional activation[58]. In particular, activated JAK1/STAT3 and JAK2/STAT5 have been shown to facilitate T-cell transformation[59].

3.1 Overview of the JAK/STAT pathway

The human genome codes for four JAK (JAK1, JAK2, JAK3, and tyrosine kinase 2 [TYK2]) and seven STAT (STAT1, STAT2, STAT3, STAT4, STAT5A, STAT5B, and STAT6) proteins, which mediate signaling through more than 30 cytokine receptors[60]. A recognizable feature of cytokine receptors is their pre-association with JAK proteins, which enhances traffic and stability[61,62] and strict cytokine dependency for signaling. The receptors have high affinity for cytokines, and up to 50% of the biologic activity can be obtained at less than 10% receptor occupancy rates[63]. The binding of a cytokine to its receptor activates the pathway, which basally is completely inactive. Cytokine binding induces a conformational change that either changes the orientation of a preformed dimer[64] or causes receptor dimerization/oligomerization[65]. As a consequence, JAKs, which are pre-bound to the receptors' cytoplasmic region, become activated. This leads to cross-phosphorylation and receptor tyrosine phosphorylation, which creates binding sites for cytoplasmic proteins such as the STAT proteins[66]. Phosphorylation leads to STAT homo-dimerization and translocation to the nucleus, where they act as transcription factors (**Fig 4**). In addition to STAT activation, JAK signaling also activates other pathways such as phosphatidylinositol-3'-kinase (PI3K), AKT/mammalian target of rapamycin (mTOR), and mitogen-activated protein kinase (MAPK).

The JAK/STAT pathway regulates the formation of lymphocytes. Cytokines that bind to receptors containing the common γ chain, such as IL-2, IL-4, IL-7, IL-9, IL-15, and IL-21, are essential for the formation and function of T-, NK-, and B-cells[67]. These receptors activate JAK1 through cytokine-binding chains and JAK3 through the common γ chain[68]. Most receptor signaling via JAK1-JAK3 activate STAT3 and STAT5[69]. STAT activation can be reversed by at least three different mechanisms: (1) protein tyrosine phosphatases (PTPs) dephosphorylating the tyrosine residues on STATs[70], (2) SOCS suppressors binding to JAKs and cytokine receptors causing negative feedback[71], and (3) the protein inhibitors of activated STATs (PIAS) blocking the DNA binding of STATs in the nucleus[72]. Substantial evidence has revealed that constitutively active cytokine receptors (obtained by *in vitro* mutagenesis[73]) lead to permanent signaling. Therefore, constitutively active STAT

proteins in the nucleus somehow overcome these mechanisms of negative regulation, eventually leading to oncogenesis[74].

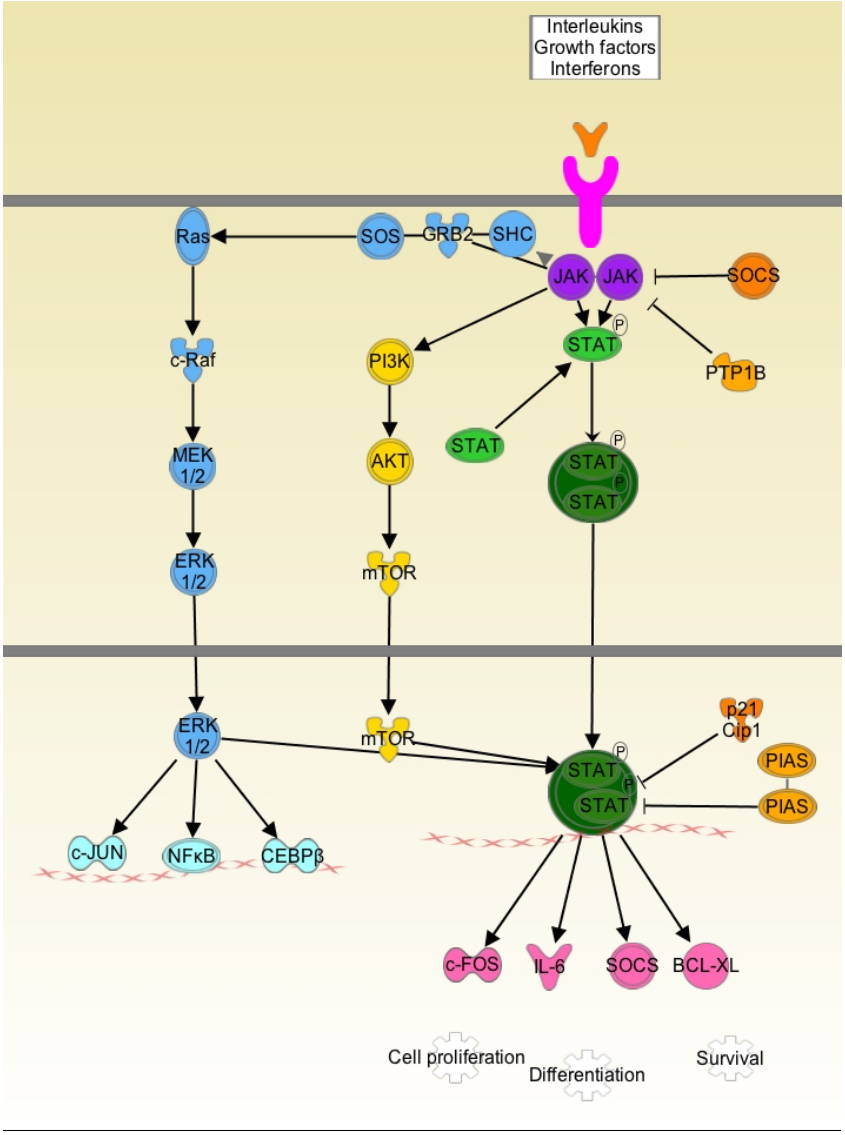


Figure 4. *Cytokine receptors exist on the cell surface in an inactive state bound to JAKs via their cytosolic domains. The binding of a specific ligand (e.g., interleukins) induces a conformational change in the preformed dimer, leading to tyrosine phosphorylation and the cross-activation of JAKs, which phosphorylate intracellular receptor tyrosine residues. In turn, the phosphorylated residues attract signaling adaptor proteins that recognize specific tyrosine phosphorylated sequences. Various adaptor proteins become substrates of JAKs, triggering signaling cascades. Cytokine receptors are linked to the STAT, Ras–MAPK, and PI3K–AKT pathways, which converge at the nucleus and regulate gene expression. Pathway edited from Ingenuity Pathway Analysis software.*

3.2. The importance of JAK/STAT signaling in hematopoiesis

The importance of the JAK/STAT pathway in the hematopoietic system has been demonstrated with various murine knockout models. For instance, *Jak1*-deficient mice die shortly after birth with immunological abnormalities phenotypically

resembling severe combined immunodeficiency, while *Jak2*-deficient mice die *in utero* as a result of fetal anemia[75,76]. The loss of *JAK3*, which is predominately expressed in NK-, T- and B-cells, results in a loss of both T- and B-cells due to the absence of IL-7 receptor signaling both in mice and humans[77]. STAT-deficient mice also exhibit a variety of immunodeficiency phenotypes. *Stat1*-deficiency leads to a failure to efficiently respond to microbial and viral infections as a result of impaired interferon signaling[78]. IL-6 and IL-21 signaling is crucial for T_H17 differentiation, which explains why mice lacking *Stat3* present with defects in T_H17 differentiation, and mice with high STAT3 activity show an accumulation of T_H17 cells[79,80]. The STAT protein family is part of many transcription factors that are involved in the determination of T-cell fates[81]. STAT4 and STAT6 are respectively considered to be T_H1- and T_H2-signature transcription factors[82,83]. The functions of STAT5 in different cell types have been extensively studied, revealing that STAT5 is a key regulator of myelopoiesis and lymphocyte development, proliferation, and differentiation. *STAT5*-deficient mice suffer from a severe combined immunodeficiency-like phenotype[84]. On the other hand, overexpression or hyper-activation of STAT5 results in an expansion of $CD8^+$ T-cells in mice[85].

3.3. STAT3 and STAT5 activation in cancer

Tyrosine kinases (JAKs and TYK2) are often activated by genetic or epigenetic alterations in cancer, causing STAT activation as a secondary event. The effects of this constitutive activation depends upon the cell type and co-activating proteins. STAT3 can promote tumor growth through downregulation of the transcription of tumor suppressor TP53, upregulation of the transcription of antiapoptotic proteins (BCL-XL and MCL-1), and prevention of cell cycle arrest[86]. Aberrant STAT3 activation has been observed in head and neck cancer, multiple myeloma, and several hematological malignancies such as AML, where it is associated with a shorter time of disease-free survival[87,88]. In autoimmune disorders, IL6-STAT3 activation has been observed in RA, psoriasis, and multiple sclerosis[89]. In hyper-IgE syndrome patients, germline mutations are seen in the DNA-binding and SH2-domain of *STAT3*[90,91], where the protein has a dominant negative effect; co-expression of wt and mutated *STAT3* leads to the impaired DNA-binding and transcriptional ability of STAT3. Also, patients with severe autoimmune diseases, including type 1 diabetes[92] and immunodysregulation polyendocrinopathy enteropathy X-linked–like syndrome (IPEX)[93], have been found to harbor activating heterozygous germline mutations in the DNA-binding domain [93,94]. In addition to the high frequency of *STAT3* mutations in T-LGLL, somatic alterations of *STAT3* have also been found in other diseases—including acute T-cell lymphoma, hepatosplenic T-cell lymphoma, γδ-T-cell-derived lymphomas, NK leukemia, myelodysplastic syndrome, NK/T-cell lymphomas, peripheral T-cell lymphoma not otherwise specified, diffuse large B-cell lymphoma, and angioimmunoblastic T-cell lymphoma—making up for more than 250 identified cases in total[95-100]. Interestingly, all these mutations are found within the SH2-domain of *STAT3*[57].

Constitutive STAT5 activation is seen in many myeloproliferative diseases (MPDs), such as chronic myeloid leukemia (CML), where the oncogenic BCR-ABL1 tyrosine kinase activates STAT5 and blocks apoptosis through the transcription of BCL-XL[101]. In other MPDs, such as polycythemia vera, myelofibrosis and essential thrombocytemia, STAT5 activation is commonly caused by an activating V617F mutation in the JH2 domain of *JAK2*[102]. Other somatic *JAK*-mutations have been seen in both adult and pediatric ALL-patients, and, more recently, activating *STAT5B* mutations were also found in T-ALL cases[103]. In comparison to the large number of patients who were identified with *STAT3* mutations, the somatic N642H mutation in *STAT5B* was reported in less than 100 cases in different hematopoietic malignancy, inducing γδ hepatosplenic T-cell lymphomas, LGLL, enteropathy-associated T-cell lymphoma, and T-PLL[4,19,95,104].

3.4. Targeting the JAK/STAT pathway

As described above, STAT proteins are constitutively activated in a majority of hematological malignancies. Although the impact of the JAK/STAT pathway on cancer development and progression has become apparent, targeting this pathway is challenging due to its pleiotropic nature. Considerable effort has been put forth to develop JAK/STAT inhibitors, and currently many JAK kinase inhibitors are being tested in clinical trials for leukemia[105]. Two of the most promising JAK inhibitors are ruxolitinib and tofacitinib. Ruxolitinib potently targets JAK1 and JAK2, while tofacitinib inhibits JAK1, JAK2, and JAK3 with nanomolar potency[106]. In early 2016, ruxolitinib was used in 27 active clinical trials for leukemia (mostly phase II), targeting MPDs, ALL, and Hodgkin's lymphoma. Tofacitinib entered clinical trials as a treatment for RA, and many phase II trials have been completed[107].

In addition to JAK-targeting drugs, the targeting of members of the STAT protein family is very promising in the search for effective cancer therapies. Recently, it was shown that a molecule already developed, pimozide, was capable of inhibiting STAT5 activation in BCR-ABL cell lines[108]. There are many strategies for targeting STAT proteins, including using peptides to bind and inhibit STAT production and activation or using small molecules to bind directly to the SH2 domain[109]. Since STAT3 and STAT5 are the most active family members of all seven STAT proteins in many cancers, many attempts are being made to produce specific STAT3 or STAT5 inhibitors. From the many inhibitors developed, only a few have entered phase I clinical trials, such as AZD9150 and OPB-31121[110,111]. In addition, a number of inhibitors targeting the SH2 domain of STAT5 are presently in development with promising potential[112].

4. Functional genomics for personalized therapy

The molecular characterization of leukemic malignancies is changing not only the understanding of these diseases but also the way the patients are diagnosed and treated. Already a few of the current diagnostic subsets set by the World Health Organization (WHO) have been defined by genetic parameters instead of the more traditional criteria such as the morphology and histology of the cells and tissues. The development towards the genetic profiling of cancer subtypes has been spurred on by the technological revolution of deep sequencing, and, although this knowledge is by itself exciting from a biological point of view, genetically driven medicine requires the translation of this information to personalized therapeutic solutions for patients. This requires not only the subdivision of diseases into subgroups with operationally important mutant genes but also knowledge of pathways that are dysregulated as a consequence and the ability to match the right cocktail of drugs with the pathways driving the pathogenesis. This leap will require the integration of genomic information with functional information about the drug sensitivity patterns of cancer cells (**Fig 5**).

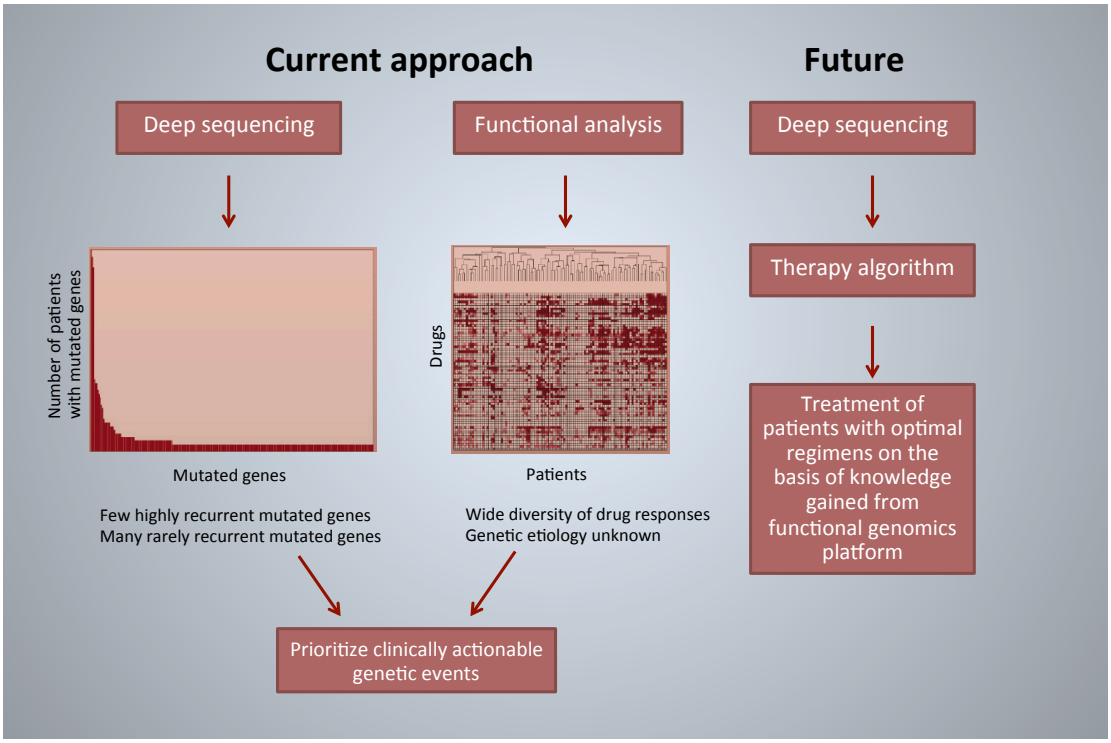


Figure 5. *Functional genomics approach: current and future. To define the functional genomic landscape of leukemia, genomic and drug sensitivity analyses of primary leukemia are being performed in parallel. Sequencing reveals a large number of mostly rare genetic events, while functional screening shows patterns of drug sensitivity with an unclear genetic background. By integrating these data streams, it is possible to identify pathogenetically important and clinically actionable genetic events. This builds towards a future in which personalized, targeted therapies are prescribed solely on the basis of the rapid genomic analysis of individual patient tumor specimens.*[113]

4.1. Next-generation sequencing

For over two decades the gold standard of sequencing was Sanger sequencing, with which the major milestone of sequencing the first human genome was accomplished[114]. The major benefit of second generation or next generation sequencing (NGS) is the ability to produce a large amount of sequencing data faster and cheaper than was ever possible with Sanger sequencing. Commonly used applications of NGS are whole genome sequencing, whole exome sequencing (WES), RNA sequencing, and targeted amplicon sequencing. All these NGS methods have several steps in common that can be grouped as template preparation, sequencing and imaging, genome alignment and assembly, and data analysis[115].

4.1.1. Exome sequencing

Over the last 10 years, exome sequencing has been applied to thousands of patients suffering from genetic diseases. Sarah B. Ng et al. were the first to report about the success of the technique in Nature (2009)[116]. They managed to rediscover a previously known mutation that causes Freeman-Sheldon syndrome by sequencing the exomes of a four-member family. Since then the technique has been responsible for a string of successful new discoveries, including mutations in different forms of cancer and leukemia. Both Yan et al. (Nature Genetics 2011)[117] and Grossmann et al. (Blood 2011)[118] reported the discovery of recurrent somatic mutations in different subtypes of acute myeloid leukemia (AML) by exome sequencing.

With exome sequencing, the amount of sequence that is covered is reduced from 3,000 Mb to about 30 Mb scattered across 180,000 exons in the genome[116]. The sequencing of the protein-coding regions of the genome is not only timesaving compared to whole-genome sequencing but it is also more likely to detect relevant mutations. It has been estimated that up to 85% of disease-causing mutations can be found in the exome regions alone[119]. Research suggests that the sequencing of exomes in small numbers of unrelated individuals with a shared monogenic disorder is as efficient as a genome-wide scan for the causative gene[116].

To perform exome sequencing, it is crucial to efficiently isolate the exons that are spread out in the genome. There are a few different methods for capturing this genomic subset, and companies such as Agilent, Roche/Nimblegen, and Febit offer commercial kits implementing these methods. The most commonly used methods are solid-phase and liquid-phase hybridization. In solid-phase hybridization, probes complimentary to the sequences of interest are fixed to a solid support (i.e., microarray or filter) and allowed to hybridize with total DNA[120]. The desired fragments are thereby attached via the probes to the solid support while the non-targeted fragments are washed away. The enriched exome DNA can then be eluted and sequenced. The principle is the same for liquid-phase hybridization, but instead of the probes being attached to a solid matrix, they are biotinylated. The probes are allowed to hybridize with their target DNA and are then bound to magnetic

streptavidin beads[121]. The commercial kits that are currently available have chosen to target the human consensus coding sequence regions[122], covering about 83% of the RefSeq coding exon bases, or 29 Mb of the genome. No exome capture method is fully efficient, which needs to be considered in the interpretation of the sequencing results.

In the past it was always problematic to sequence cancer genomes due to the fact that any given tumor sample would always be a mixture of cancerous and healthy cells. Using Sanger sequencing was inefficient, as it could not detect genomic alterations prevalent at a low frequency in clinical samples[123]. Exome sequencing, on the other hand, enables the deep coverage that is needed for studying cancer tissue. Any given sequence is read between 10–100 times, which means that the probability for the cancer cell genome to be sequenced is much higher than in Sanger sequencing. Exome sequencing can be used to detect mainly different point mutations, small insertions and deletions (indels), and copy number variations (CNVs).

4.1.2 Targeted amplicon sequencing

Targeted next-generation sequencing panels to identify genetic alterations in cancers are increasingly becoming an integral part of clinical practice. The deep sequencing of PCR amplicons (> 1000x coverage) allows for the efficient characterization of known gene targets and is especially useful for the discovery of rare somatic mutations in complex samples (such as tumors mixed with germline DNA). Illumina amplicon technology uses a pair of oligonucleotide probes designed to target and capture regions of interest, followed by NGS. The multiplexing of hundreds to thousands of amplicons reduces the sequencing costs and turnaround time compared to broader approaches. This highly multiplexed approach facilitates research by sequencing multiple genes simultaneously at approximately the same cost as a single-gene assay.

Amplicon sequencing allows researchers to sequence 16–1,536 targets at a time, spanning 150–1,500 bp per target, depending on the library preparation method used. To enrich the desired amplicons, primers with adapter tails are used, and the DNA fragments are attached to a glass slide by the adapter sequences[124]. The adapter-specific primers form template clusters through the process of bridging amplification. These clusters of identical templates are essential, as most imaging systems are unable to detect single fluorescent signals[115]. Next, a fluorescent nucleotide with a terminating group is added to the template, and lasers identify the nucleotide via fluorescence detection. Finally, the terminating group is cleaved to enable the start of the next nucleotide incorporation in a process called cyclic reversible termination.

4.1.3. NGS coverage and error profiles

For a variant to be detected with high reliability the site needs to be adequately covered by a sufficient amount of reads. The detection of variants in a low-coverage area is more difficult, and the minimum median coverage recommended for NGS experiments is 20–30-fold to ensure the proper detection of variants[125]. However, it is

important to note that minimum coverage is highly dependent on many aspects of the platform and assay, including base-call error rates, quality parameters, and analytical pipeline performance[126].

Every NGS platform has its own error profile caused by the introduction of erroneous mutations during the template amplification stage. These are indistinguishable from true variants of the source DNA and can also affect the correct alignment of the reads[127]. A better knowledge of the error patterns is essential for sequence analysis and vital for drawing valid conclusions. For amplicon sequencing with the Illumina MiSeq platform, the library preparation method and the choice of primers are the most significant sources of bias and cause distinct error patterns[128]. These errors can be filtered out by sequencing a reference DNA sample with a known sequence in parallel to give an estimate of the error profile of the actual sample sequenced and analyzed.

4.1.4. Analyzing sequence data

Sequencing results in a huge amount of data, which to be significant needs to be processed with bioinformatical tools. The nucleotide sequence of a template is called a read (usually between 25–250 bp long) and these are aligned to a known reference sequence (in this case the human genome) with different aligner tools (e.g., Burrows Wheeler aligner [BWA][129]). Next, variants are called using tools such as Varscan[130], which is an identification algorithm developed for the nomination of putative mutations, indels, and CNVs. When analyzing cancer exome data, a comparison of the tumor sample variants with the healthy germline data is essential to distinguish somatic mutations from germline variants. Putative somatic variants can be filtered with, for example, Varscan2[131] and subsequently annotated for functional consequences at the gene/protein level and the evolutionary conservation of the mutation site[132]. A major dilemma in all sequencing studies is to distinguish driver mutations—which contribute to carcinogenesis—from passenger mutations that are an incidental consequence of the many cell divisions and genomic instability of cancer cells. Passenger mutations are thought to greatly outnumber driver mutations. Different databases can be used to filter out common variants seen in the population (SISu, dbSNP, ExAC) and to annotate the variants further for clinical relevance (Clinvar, COSMIC, OMIM)[133]. Sequencing data can also be visualized with different programs such as the Integrative Genomics Viewer (IGV) for manual curation of putative mutations.

4.2. Drug sensitivity and resistance testing of cancer cells

While the molecular characterization of a disease can give us clues into which pathways are activated and driving the disease, functional data from drug sensitivity assays inform about which drugs are effective, but the etiology of the mechanisms behind the response remains unknown. Combining both genetic and functional

information could hopefully bring us closer to new forms of treatment and also illuminate the biological background of pathogenesis. Lately, evidence has been accumulating that kinases are frequently deregulated in human cancer and leukemia, with one study reporting that as much as 70% of leukemia samples show hypersensitivity to one or more kinase inhibitors[134]. Unbiased cancer drug sensitivity studies focusing on targeted agents could therefore be used to link drug responses to genetic and clinical features. Previously, this strategy has proven successful in BCR-ABL-driven leukemia, where axitinib was found to be effective in patients harboring the resistance-causing T315I mutation in ABL1[135,136]. On the other hand, the diagnosis of particularly drug resistance in individual patients would improve cancer treatment by avoiding ineffective treatments.

Various assays for fresh tumor cell culture assays have been developed, but the principal steps—the isolation of cells, incubation of cells with drugs, and assessment of cell survival—are the same[137]. All methods measure the molecular processes of cancer cells, revealing cell activity and indicating cell growth or death. Frequently used methods are thymidine incorporation into cell DNA[138] and the loss of cell ATP[139]. The comparison of different read-out methodologies has revealed the ATP-based luminescent assay to be very sensitive but also susceptible to the underestimation of drug potency and efficacy, especially for compounds interfering with DNA synthesis[140]. A major issue with drug screening assays is reproducibility; the assay set-up has a significant impact on the quantification of the drug response. This has been seen in the comparison of two major cell line screening programs (the Cancer Cell Line Encyclopedia [CCLE] and the Cancer Genome Project [CPG]), where discrepancies were found in the drug sensitivity data that were not mirrored in the gene expression profiles[141,142]. The observed differences were likely due to the lack of standardized protocols for the handling of the cells and the different concentration ranges, readouts, and scoring of the drug responses[143]. Several other factors, such as cell culture conditions and compound handling and storage as well as assay conditions, should be taken into account when comparing outcomes from different platforms.

The primary data output of drug screening studies are compound dose response curves and associated drug response metrics, such as the half-maximal inhibitory concentration (IC_{50}) and area under the curve (AUC). The IC_{50} is the most commonly used to evaluate drug sensitivity, but it has limited ability to fully capture the drug response information. In an effort to quantify the drug sensitivity data and make them more comparable between different samples, a single measurement termed drug sensitivity score (DSS) has been developed[144]. It takes into account all four curve-fitting parameters used to calculate the AUC in relation to the total area between 10% threshold and 100% inhibition (**Fig 6**). Furthermore, to detect cancer selective drug responses, DSS values detected in patient samples can be compared with DSS values from healthy samples to obtain a selective DSS (sDSS).

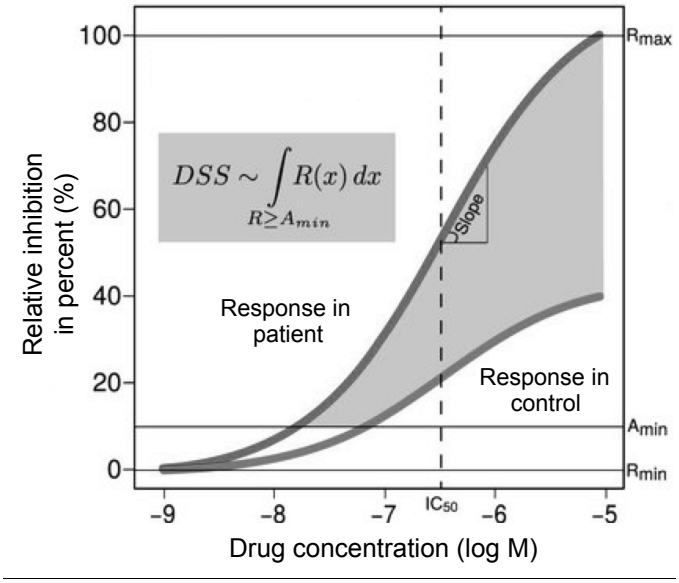


Figure 6. *Schematic illustration of the selective DSS calculation (sDSS, the grey area). Dose-response parameters estimated through a logistic function model include IC_{50} , slope of the curve at IC_{50}, and the bottom and top asymptotes of the curve (R_{min} and R_{max}).The two dose-response curves clearly show differential activity patterns, yet their relative IC_{50} is equal, showing an example in which IC_{50} is not informative enough for detecting selective responses in patient samples.*

Aims of the study

The aim of this PhD project is to reveal the somatic mutations underlying different T-cell malignancies, such as LGL leukemia and PLL leukemia, and further elucidate their functional impact on the pathogenesis of the disease. Furthermore, in an effort to discover new treatment options for these patients, the *ex vivo* drug screening of selected patient samples was applied in an effort to find clinically relevant drugs for the individual patient.

The project consisted of four parts:

1. The pathogenesis of T-LGLL in *STAT3* mutation-negative patients
2. Further characterization of *STAT3* mutations outside of the SH2-domain in T-LGLL
3. The discovery and characterization of *STAT5B* mutation spectrum in CD4+ T-LGL leukemia
4. Functional genomics with the amplicon screening of known mutations paired with drug sensitivity testing in T-PLL leukemia patients to discover novel drug targets

Patients and methods

6. Patients and ethical permissions

T-LGLL and PLL patients were diagnosed based on the WHO 2008 guidelines (studies I-IV). The samples were collected from Helsinki University Central Hospital (Finland), Oulu University Hospital (Finland), Penn State Hershey Cancer Center (USA), Cleveland Clinic (USA), Shinshu University Hospital (Japan), Queen Mary Hospital (China), Hospital Universitario de la Princesa (Spain), University Hospital Heidelberg (Germany), University Hospital Olomouc (Czech Republic), University Hospital Aachen (Germany), Mayo Clinic (Minnesota US), University Hospital Cologne (#11-319, Germany), and Saint-Louis Hospital (France). All studies were conducted in accordance of the principles of the Helsinki declaration and was approved by the Helsinki University Central Hospital Ethics Committee. Written informed consents were obtained from all patients and healthy controls in all studies.

In study I, the molecular background of STAT-mutation negative T-LGLL patients was analyzed by WES of CD8+ tumor LGLs and healthy CD4+ cells. The mutations found were subsequently screened in a larger cohort of 113 STAT-negative T-LGLL patients. DNA used for the validation analyses was extracted from whole blood or mononuclear cells (MNCs).

Study II included the three T-LGLL patients previously exome sequenced and four new CD8+ STAT-negative T-LGLL cases. The STAT3 mutations outside of the SH2-domain were analyzed through amplicon sequencing from a cohort of 99 T-LGLL patients previously confirmed to be *STAT3* hotspot mutation negative. DNA was extracted from whole blood or MNCs.

In study III, WES was performed on CD4+ tumor cells and CD8+ healthy cells from three CD4+ T-LGLL patients. The STAT5B SH2-domain and transactivation domain was screened from CD4+ (n=8), *STAT3*-mutated CD8+ (n=37) and STAT-mutation negative CD8+ (n=58) T-LGLL patients. Additionally the same regions were screened with Sanger sequencing in Japanese and Chinese LGL-leukemia cohorts consisting of CD8+ and CLPD-NK cases (n=57).

Study IV analyzed the molecular mutations and drug responses of T-PLL patients to find novel therapeutic options for this aggressive disease. From a cohort of 70 T-PLL patients, DNA was extracted from MNC samples and known mutations of the JAK/STAT pathway were screened with targeted amplicon sequencing. 39 patient samples were further screened with the DSRT system, either fresh (n=4) or viably frozen (n=35). MNC samples from four healthy subjects were also screened and used as controls. For *in vitro* experiments, the Jurkat cell line was used.

7. Mononuclear cell separation

Peripheral blood (PB) samples were collected from leukemia patients and MNCs were separated by using Ficoll-Paque^{TM} PLUS (GE Healthcare). The cells were either used immediately for CD8+ and CD4+ selection or viably frozen in liquid nitrogen. If frozen, FBS containing 10% dimethyl sulfoxide (DMSO) was added to protect the cells and the samples were stored in -70˚C for 24-48 h before transferring them into liquid nitrogen.

8. Selection/sorting of lymphocytes

In all studies, the cells were further sorted into CD4+ and CD8+ populations with either AutoMACS magnetic bead sorting (Miltenyi Biotech) or sorting by flow cytometry using BD Biosciences Multitest CD45, CD3, CD8 and CD4 antibodies. A small amount of cells from the sorted fractions were used to check the purity of the sorting with flow cytometry (FACS Aria, Beckman-Coulter Immunotech).

9. TCR-Vβ flow cytometric analysis of peripheral blood

In studies I-IV, clonal expansions in PB samples were assessed with the IO Test® Beta Mark TCR Vβ Repertoire Kit (PN IM3497, Beckman Coulter Immunotech), which is a kit for the quantitative determination of the TCR Vβ repertoire of human T lymphocytes using flow cytometry. The kit consists of 8 vials containing mixtures of conjugated TCR Vβ antibodies corresponding to 24 different specificities covering approximately 70 % of the normal human TCR Vβ repertoire. To further characterize the expansions, the samples were also stained with CD3, CD4 and CD8 antibodies (BD Biosciences).

10. DNA and RNA extraction

DNA from whole MNC, CD4+ or CD8+ cells was isolated according to the Nucleospin® Tissue Kit instructions (Macherey-Nagel). RNA was extracted with the miRNeasy kit including DNAse I digestion (Qiagen) according to the manufacturer's instructions. DNA and RNA concentrations were measured using Qubit® 2.0 Fluorometer (Invitrogen) and the quality was assessed with the Agilent 2100 Bioanalyzer (Agilent Technologies).

11. Sequencing methods

11.1. Capillary sequencing

Based on exome sequencing results, specific primers were designed for candidate mutations using the Primer Blast search (http://blast.ncbi.nlm.nih.gov/, National Center for Biotechnology Information, Bethesda, MD, USA). PCR was performed on CD8+ and CD4+ DNA. PCR products were separated by gel electrophoresis and extracted from the gel using the Nucleospin Gel and PCR Clean-up Kit (Macherey

Nagel) or ExoSAP-IT PCR Product Clean-up (Affymetrix). The purified PCR products were then sequenced with BigDye™ v.1.1 Cycle Sequencing kit and the ABI PRISM® 3730xl DNA Analyzer. Sequences were analyzed using 4Peaks (v. 1.7.2.) and Blast search.

11.2. Exome sequencing

Paired matched samples from each patient were used for exome sequencing in all studies. In studies I-II CD8+ cells were used as tumor samples while CD4+ cells represented the normal controls. In studies III-IV CD4+ cells represented the tumor fraction while CD8+ cells were used as control.

3µg of each DNA library was fragmented with a Covaris S2 instrument. Sample libraries were processed according to NEBNext DNA Sample Prep Master Mix Set 1 manual (New England BioLabs) with some exceptions; All primers and oligonucleotides were custom ordered from Sigma Aldrich. Agilent captures were processed according to the Agilent SureSelect Target Enrichment System for Illumina Paired End Sequencing Library SureSelect Human All Exon and Human All Exon Plus v2.0.1 protocol with the following exceptions: 1µl of the captured library was used in five parallel 50µl amplification reactions in postcapture PCR. 1µl of 100µM NimbleGen PE POST1 and PE POST2 primers were used in each reaction. 12 cycles were used. Sequencing of exome and RNA sequencing libraries was performed using the Illumina GAII instrument. Exome sequencing libraries were sequenced as 82 bp paired end reads.

11.3. Targeted deep amplicon sequencing

Locus-specific primers were designed for study II, III and IV to cover the desired gene areas using Primer3 with user-defined parameters (http://primer3.wi.mit.edu/). After designing the locus-specific primer sequences (**Table 3**), sequence tails corresponding to the Illumina adapter sequences, were added to the 5′ end of the forward and reverse locus-specific primers, respectively. All oligonucleotides were ordered from Sigma-Aldrich. Sample preparation was performed according to an in-house targeted PCR amplification protocol. Each amplicon was amplified in multiplexed PCR-containing locus-specific PCR primers carrying Illumina-compatible adapter sequences, Illumina TruSeq Universal Adapter primer, and Illumina TruSeq Adapter primer with a sample-specific 6-bp index sequence (used for genetic barcoding of each sample). The locus-specific primers are present in limiting quantities and are thus consumed in initial cycles of amplification forming intermediate amplification products. In subsequent amplification cycles, the adapter primers further amplify these intermediate products forming an Illumina-compatible paired-end sequencing template.

Following PCR amplification, samples were purified using a Performa V3 96-Well Short Plate (EdgeBio) and QuickStep 2 SOPE Resin (EdgeBio) and then pooled together without exact quantification. Purified sample pools were analyzed on an

Agilent 2100 Bioanalyzer (Agilent Technologies) to quantify amplification performance and yield. Sequencing of PCR amplicons was performed using the Illumina MiSeq instrument with MiSeq Control Software (version 1.2; Illumina). Samples were sequenced as 151-bp paired-end reads and two 8-bp index reads.

The sensitivity and reliability of the STAT amplicon sequencing platform was validated in a previous publication[38].

***Table 3**. Primer sequences for amplicon sequencing*

Site	Forward primer	Reverse primer	Study
STAT3 exon 2	TCCCCATCACCTGTACCCAT	TGACACCTGTGTTGGGCAAT	II
STAT3 exon 3	ACACTAACACCCGACTCTGC	TGTATGCGTCGGCTTCAGAG	
STAT3 exon 4	TCCATTCCTCCCAGACCAGG	GCTCTGAAGCCTTTGTTCCG	
STAT3 exon 5	CCGAGGCTTGTAACTTGCAT	TTCCCTTCCTCTTGTGATGG	
STAT3 exon 6	GACCAGGCTCCTTTGAGGAC	CTCTTGGGGATACTGCCTGC	
STAT3 exon 7	CCGATCTAGGCAGATGTTGG	TTCCCTCAGGTCAAGGAGTTT	
STAT3 exon 8	CTGTGGGCCTGCAGTTAAGA	GTTCCTGCTCTGGAGTTGACT	
STAT3 exon 9	AAGAGAAGATGGGCTCACGC	TCCCCTTCTCCATCTCACCT	
STAT3 exon 10	TGGAAAGAATGACCCTGGCC	CACGTGGTAGAGTGAGAGGC	
STAT3 exon 11	AATGCACCCCAAGGCTTTTG	CCTCCCACAGTGCTGAGATT	
STAT3 exon 12-14	CAAGGAAAACACCCCAGTTG	AAATAACAGGTGGTCAAAGTAGGC	
STAT3 exon 15	ATTGCCAGATGGGATGCCAA	CCACACCTGGCCTAAGAGTG	
STAT3 exon 16	GAGGAGAACTGCCAGCTCAG	CCTTTCATTCTGAGCCCCGT	
STAT3 exon 17	AGGGAGAAGGGGTGAAATGC	TGCCCCCTCCTTTTAGTTGG	
STAT3 exon 18	CCTTGCCAGCCATGTTTTCC	AACCTCTTGACCCCAAGCTG	
STAT3 exon 19	GTGCACACTCTGTCCAACCT	GCTTGAAGGCCCTGAACTCT	
STAT3 exon 20	GGAGTCAAGGCCATCTCCAC	TGGATGCCCTGTTAGCAATA	
STAT3 exon 21	CCCAAAAATTAAATGCCAGGA	GGTTCCATGATCTTTCCTTCC	
STAT3 exon 22	CTCACCCAGTGTCCCATTCC	GGCAGATGGAGCTTTCCAGA	
STAT3 exon 23	GACCAGCTCTCGGTGTGTAC	TGGAGACCAGAGTTTGATGGC	
STAT3 exon 24	GGCACTTGTCTAAGAACAACAACA	AGTTGCAGAGGGTGGACAAC	
STAT5B exons 14-15	CTGACCTCAACAAATAGTAAGTACCC	TTCCAAGTGTTACTCTGGTGTTC	III, IV
STAT5B exon16	TGTTGGGGTTTTAAGATTTCC	CAAATCAGAATGCGAACATTG	
STAT5B exon 17	CCCAGGGCTGAGACAGTTT	AGATTGCACCACCGTACTCC	
STAT5B exon 18	TGGATTCCTTTGACCCCAGC	AGATACCCCTTGGTCCCCTC	III
STAT5B exon 19	CTGGTGGCCTGTGGGGCTTG	TCTGTCTGTGGCCCCTCTGCT	
JAK3 exon 11	TCTCTGGACCCCAGACTGAG	GTCTCAACAGCAGCAGCAAC	IV
JAK3 exon 12	CAAAGTGCTGGGATGACAGG	GGCCTCAGGCATATGCTATAAT	
JAK3 exon 15	GGATCCACTTCCTTGCCCTG	GGTGTTTGAGAAGGGGAGGG	
JAK3 exon 16	CTGAGGGTGAGAGGAGCAGT	CCTGGAGAGTAAGTTCCTGGAG	
JAK3 exon 18	TGAAAGTCCCTCTGCTGGTC	CCTGCCATAATGCACAGAGA	
JAK3 exon 19	TCACGTTCCCAGCCTACCTA	TCACAGCTGGGCAAGGTAAG	
JAK1 exon 13	TCTCAACCCATTGTGTTCCA	GTTCCCATTGAGGACCCATT	
JAK1 exon 14	CCACCCACCCCTTTGAAAGA	ATTGATGTTCAGGGCCTGGG	
JAK1 exon 15	ACCAGGCACACCTTTGTTCA	AGGGGATGAAGGAGAGGACC	
JAK1 exon 20	CCACCAGCTAGCATGTCAGA	CTGTGTGTTCCGTGGCCTA	
IL2RG exon 8	ATGGCAACTGGTATTGGGGG	GAGTATGAGACGCAGGTGGG	

11.4. Real-time quantitative PCR

In study II, total RNA from the transfected HEK293 cells was extracted with AllPrep DNA/RNA/miRNA Universal Kit (Qiagen) after which the RNA was converted to cDNA (1 µg) with Superscript III reverse transcriptase (Thermo Fischer Scientific) using random primers. Expression levels of six known STAT3 target genes (*SOCS3*, *JAK2*, *MYC*, *JUNB, BCL3, CCL2*) were measured using iQ SYBR Green Supermix (Bio Rad) and the CFX96 Touch™ Real-Time PCR Detection System (Bio Rad).

Results were normalized against four housekeeping genes (NONO, PGK-1, GAPDH, LDHA), which showed uniform expression across all samples. Primer pairs are listed in **Table 4**. All reactions were run in triplicate wells and gene expression was quantified using the delta-delta Cq method.

Table 4. *Primer sequences for qPCR assay*

Target gene	Forward	Reverse
GAPDH	AGCACCCCTGGCCAAGGTCA	CCGGAGGGGCCATCCACAGT
SOCS3	CTTCGATTCGGGACCAGCCCCC	GGCGGCGGGAAACTTGCTGT
STAT3	AGCGAGGACTGAGCATCGAGCA	GCCCTTGCCAGCCATGTTTTCTTTG
MYC	CAGCTGCTTAGACGCTGGATT	GTAGAAATACGGCTGCACCGA
JAK2	GCTCAGTGGCGGCATGAT	CACTGCCATCCCAAGACATTC
JUNB	TGGTGGCCTCTCTCTACACGA	GGGTCGGCCAGGTTGAC
CCL2	GTCTCTGCCGCCCTTCTGT	TTGCATCTGGCTGAGCGAG
BCL3	CCACAGACGGTAATGTGGTG	TATTGCTGTGGTGCAGGGTA
PGK1	CCCAGCTGTATTTCCAAAATGTCGC	ACAGCAGCCTTAATCCTCTGGTTGT
LDHA	CTCTGAAGACTCTGCACCCA	GCCCAGGATGTGTAGCCTTT
NONO	GCTCTGGACAGATGCAGTGAA	CTGCTCTCGTTCCTTGTGAA

11.5. RNA sequencing

In study IV, 1-3 ug of CD4+ total-RNA was used for depletion of ribosomal RNA (Ribo-Zero, Epicentre). The ribodepleted and purified (Nucleospin cleanup XS, Machery-Nagel) RNA was reverse transcribed to ds cDNA (SuperScript™ Double-Stranded cDNA Synthesis Kit, Life Technologies). Random hexamers (New England BioLabs) were used for priming the first strand synthesis reaction and SPRI beads (Agencourt AMPure XP, Beckman Coulter) for purification of cDNA.

Illumina compatible Nextera™ Technology (Illumina, Inc.) was used for preparation of RNAseq Libraries. In this technology DNA fragmentation and tagging is performed by *in vitro* cut-and-paste transposition. 60ng of doublestranded cDNA was used instead of DNA. After tagmentation reaction the fragmented cDNA was purified with SPRI beads.

In order to add the Illumina specific bridge PCR compatible sites and enrich the library, limited-cycle PCR (5 cycles) was done according to instructions of Nextera system with minor modifications. For barcoded libraries, 50X Nextera Adaptor 2 was replaced with a bar coded Illumina®-compatible Adaptors from the Nextera Bar Codes kit (Illumina, Inc.) in PCR setup. SPRI beads were used for purification of the PCR-products and the library quality control was evaluated by Agilent Bioanalyzer (Agilent Technologies). Illumina HiSeq2000 platform was used for sequencing to produce more than $40x10^6$ paired end reads with 100bp read length.

12. Sequencing data analysis

12.1. Variant filtering and visualization of exome data

Raw Illumina sequence reads were trimmed of B blocks in the quality scores from the end of the read. After this, if any pair had a read shorter than 36 base pairs, the pair was removed. The quality scores were converted to Sanger phred scores using Emboss (version 6.3.1) and aligned using BWA (version 0.5.8c) against the human genome build hg19 primary assembly. After alignment, potential PCR duplicates were removed with Picard Mark Duplicates (version 1.32). A pileup file was generated with SAMtools pileup, while reads with mapping quality scores below 0 where excluded. Somatic mutations were called with the VarScan 2.2.3 somatic mutation caller using the CD8+ sample as the tumor and the CD4+ sample as the normal control. The minimum number of tumor reads required was set to 7. Putative somatic mutations were annotated for functional consequences at gene and protein sequence levels and with Genomic Evolutionary Rate Profiling (GERP) conservation scores using Annovar. Variants present in 1000 Genomes Project (May 2011 data release) or dbSNP132 data sets were excluded from further analysis. Annotations were integrated and filtering was performed using a MySQL database. Missense, nonsense, frameshift, inframe coding indels and splice site mutations with VarScan somatic p-value below 0.01 were analyzed further. Alignments at candidate mutant loci were checked visually in IGV for alignment artifacts. Mutants with alignment artifacts were excluded from further analysis. Description of cellular and biochemical function was added manually to the remaining mutations and a set of mutations in genes deemed to be biologically relevant to leukemia pathogenesis based on expert opinion were selected for validation.

12.2. Amplicon data analysis

The data was analyzed using an in-house bioinformatics pipeline developed specifically for reporting somatic variants from amplicon sequencing data. Low-quality reads were not filtered out before alignment: after alignment, quality values (i.e. Phred scores) were used to exlude error bases with low quality from further analysis. First, previously reported *STAT3/5b* mutations were verified in all STAT-mapped amplicons: all variants with variant allele count over 5 and variant allele frequency (VAF) over 0.5% were taken into consideration. Second, all variants with variant allele count over 5 and VAF over 0.5% were called from STAT-mapped amplicons. From these candidates false positives were initially filtered out based on the noise (estimated error rate) level from control sample in every run making use of a binomial distribution to compute p-value for the event that more than the frequency of alternative alleles were observed when the null hypothesis is true. However, variants with VAF over 2% were called independent of the noise. A specific frequency ratio was used to filter out false positive by dividing the ratio of variant calls/number of all the bases (at a position) by the ratio of variant allele quality sum/quality sum of all the

bases. All samples with a frequency ratio ≥0.90 were considered to be true mutations, and the variants from both scripts with a borderline frequency ratio between 0.85-0.89 were verified with IGV.

12.3. RNA sequencing analysis

RNA sequencing reads from the patient and healthy control samples were aligned using TopHat[145] and TopHat2[146] to the human genome reference hg19. Mapped reads were counted with the HTSeq Python package[147]. To analyze the differentially expressed genes between samples DESeq2 (R/Bioconductor) was used[148]. Results were visualized with Ingenuity Pathway Analysis (IPA) software.

13. Microarray expression analysis

In study I, RNA was extracted from CD8+ cells from STAT-mutation negative as well as *STAT3*-mutated LGL-patients. In study IV, RNA was extracted from PLL patients CD4+ cells. Anonymous Red Cross buffy coat CD4+, CD8+ and NK RNA were used as biological replicates and controls in the experiment. Microarray analysis was performed using the Illumina Human HT-12 v4 BeadChip expression array which targets >47000 probes in the human genome (cover content from NCBU RefSeq Release 38 and legacy UniGene content). The data was read with an iScan instrument (Illumina) and primary analysis was done with Genome Studio software v2011.1 (Illumina). The results were normalized and log2-transformed with the Chipster open source platform[149]. In order to determine the similarity of the expression profiles a distance dendrogram was constructed using Pearson correlation and the average linkage method. Differentially expressed genes were filtered from the data using the empirical Bayes test with p-value cut-off of 0.05. Microarray data are available in the ArrayExpress database (www.ebi.ac.uk/arrayexpress).

14. Drug sensitivity and resistance testing (DSRT)

14.1. Inhibitors

The oncology compound collection included 306 drugs consisting of commercially available approved, investigational and experimental anti-cancer compounds. Drug classes included in the collection were apoptotic modulators, conventional therapeutics, epigenetic modifiers, hormone therapies, HSP inhibitors, immuno-modulatory agents, kinase inhibitors, kinesin inhibitors, metabolic modifiers and rapalogs. Compounds were acquired from National Cancer Institute Drug Testing Program or purchased from Active Biochem, Axon Medchem, Cayman Chemical Company, ChemieTek, Enzo Life Sciences, LC Laboratories, Santa Cruz Biotechnology, Selleck, Sequoia Research Products, Sigma-Aldrich and Tocris Biosciences.

14.2. The platform

DSRT was performed on fresh or viably frozen MNC samples from T-PLL patients and healthy controls in study IV. The compounds were dissolved in DMSO and preprinted on 384-well plates (Corning) using Echo 550, an acoustic liquid handling device (Labcyte Inc.). Each drug was plated in five concentrations covering a 10,000-fold concentration range. The drug plates were stored in pressurized StoragePods (Roylan Developments Ltd.) before use. The compounds were dissolved in a shaker with 5 µl of Mononuclear cell media (MCM) for 10 minutes. 20 µl of single-cell suspension (equivalent to 10,000 cells) was transferred to each well using a MultiDrop Combi (Thermo Scientific) dispenser. The plates were incubated 37°C and 5% CO_2, and after 72 hours the cell viability was measured using CellTiter-Glo luminescent assay (Promega) and a Pherastar plate reader (**Fig 7**). The data were normalized to negative control wells (DMSO only) and positive control wells containing 100 µmol/L benzethonium chloride (BzCl).

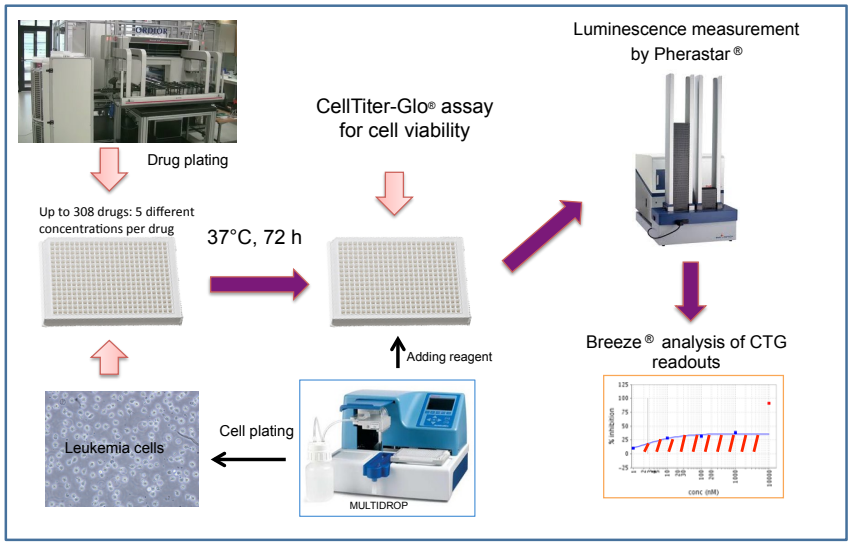


Figure 7. *Workflow of drug sensitivity and resistance testing platform and analysis using Breeze.*

14.3 Drug sensitivity analysis

Percentage of survival was calculated with Dotmatics Browser/Studies software and an in-house developed interface, Breeze, was used to analyze the data. To assess quantitative drug response a drug sensitivity score (DSS) was calculated based on the dose-response curves. DSS is an integrative and robust drug response model based on the normalized area under the curve by taking all four curve fitting parameters into account[150]. The possible range of DSS scores is between 0-50. Furthermore, selective drug sensitivity scores (sDSS), representing leukemia-specific responses, were calculated by comparing the DSS score from patient samples to the median values of healthy donors. Drugs with sDSS values above 5 were considered selective and above 10 highly selective to the cell sample tested. The DSS-based analysis platform is freely available as an R-package at: https://dss-calculation.googlecode.com/svn/trunk/

14.4. Target addiction score (TAS)

To elucidate the pathways that are active in T-PLL and to quantify the functional sensitivity of the samples to therapeutic targets, the target addiction score (TAS) was calculated for 12 samples for which drug response and microarray expression data was available. TAS quantifies the functional sensitivity of the samples to therapeutic targets[150]. The TAS algorithm integrates the DSS profiles with global compound-target interaction networks to estimate the level of addiction of each patient sample to the on- and off-targets of the targeted compounds in the drug screening library. Specifically, for a given drug target *t*, TAS was computed as an average of the DSS value over those inhibitors that target the protein *t*. The targets for TAS analysis were selected based on KIBA score of 3 and for drugs which had no targets with this criteria then the KIBA score of 4 were applied[151]. The total numbers of targets with above criteria were 614 target proteins for 243 drugs.

15. Functional assays

15.1. Mutagenesis and luciferase reporter assays

To explore the functional effects of *STAT3* and *STAT5B* mutations in study II and III, expression constructs were generated with the identified variants in either pDEST40 containing STAT3 or pCMV6-XL6 containing STAT5B. Mutagenesis PCR was performed according to instructions (Geneart® Site-Directed Mutagenesis system, Invitrogen) with mutagenesis primers ordered from Sigma (**Table 5**).

***Table 5**. STAT3 and STAT5B mutagenesis primers*

Primer target	Forward	Reverse	Study
STAT3 F174S	CCAGGATGACTTTGATTCCAACTATAAAACCCTCAA	TTGAGGGTTTTATAGTTGGAATCAAAGTCATCCTGG	II
STAT3 H410R	GCCTCTCTGCAGAATTCAAACGCTTGACCCTGAGGGAGCAGAG	CTCTGCTCCCTCAGGGTCAAGCGTTTGAATTCTGCAGAGAGGC	
STAT3 Y640F	CTGAACAACATGTCATTTGCTG	CTGCTGCTTTGTGAATGGTTCCACGG	
STAT5b Q706L	GTGAAGCCACAGATCAAGCTAGTGGTCCCTGAGTTTGTGA	TCACAAACTCAGGGACCACTAGCTTGATCTGTGGCTTCAC	III
STAT5b S715F	GAGTTTGTGAACGCATTTGCAGATGCCGGGGGC	GCCCCCGGCATCTGCAAATGCGTTCACAAACTC	
STAT5b N642H	GAAAGAATGTTTTGGCATCTGATGCCTTTTAC	AAAGGCATCAGATGCCAAAACATTCTTTTC	

For *STAT3* mutations, HEK-293 cells stably expressing a STAT3 firefly luciferase reporter (HEK293 Gloresponse SIE Luc2P Hygro cells, Promega) were grown in Dulbecco's Modified Eagle Medium (DMEM) + 10% FBS+ 200µg/ml Hygromycin B (Life Technologies). For *STAT5B* mutations, Hela cells co-transfected with the STAT5B luciferase reporter pGL4.52 (Luc2P/STAT5RE/Hygro) were used. The cells were plated in 96-well plates with 10000 cells/well the day before the luciferase readout. Transfection with the constructs was carried out 6h after plating using Fugene®HD transfection reagent (Fugene:DNA ratio 3.5:1, Promega). The next day, the HEK-293 cells were starved in DMEM+ 1% FBS for 3h, followed by stimulation with IL-6 (100ng/ml) for 3h. Luciferase activity was determined with Dual-Glo™ Luciferase Assay System (Promega) according to instructions.

In study IV, Jurkat cells containg a nuclear factor of activated T-cells (NFAT) luciferase reporter (Signosis) were cultured in RPMI1640 + 10% FBS + 2 mM L-glutamine + penicillin and streptomycin. 20,000 cells per well were incubated with

the indicated concentrations of SNS-032 on a 384-well plate for 5 hours in the presence of 0.5 µg/ml purified anti-CD3 (clone UCHT1) and 0.5 µg/ml purified anti-CD28 (clone L293) antibodies (BD Biosciences) with 4 µg/ml goat anti-mouse IgG secondary antibody (ThermoFischer Scientific). At the end of the incubation, ONE-Glo reagent (Promega) was added at a 1:1 ratio to the cell culture medium and luminescence reads were measured by a Pherastar FS. Luminescence reads were normalized to reads from wells with DMSO (negative control) and BzCl (positive control) to generate percent inhibition values.

15.2. Western blotting

To investigate the phosphorylation status of the variants, HEK-293 and Hela cells transfected with the variants were analyzed by Western blot with either a phospho STAT3 (Cell Signaling, 9131S) or a phospho STAT5 (Cell Signaling, 9314) specific antibody. Protein lysates of the different variants were separated with polyacrylamide gel electrophoresis on a 12% SDS-PAGE gel and transferred to a nitrocellulose membrane. The membrane was blocked with PBS+ 5% BSA or PBS+ 5% fat free milk containing 0.1% Tween 20. The membrane was subsequently incubated with primary antibodies, which consisted of rabbit anti-pSTAT3 1:1000 or rabbit anti-pSTAT5 antibody 1:1000 in PBS+ 5% BSA as well as mouse anti-beta actin 1:5000 (Abcam, AC-15) in PBS+ 5% milk as a loading control. Goat anti-rabbit IRDye 800 (Li-cor Odyssey 926-32211) and goat anti-mouse IRDye 680 (Li-cor Odyssey 926-68020) 1:20000 were used as secondary antibodies. The proteins were visualized with the Li-cor Odyssey Infrared Imaging system and the intensity of the pSTAT3/5 bands were normalized against STAT3/5 to compensate for the differences in transfection efficacy of the wt and variants.

15.3. Immunohistochemistry

Immunohistochemical staining with pSTAT3 and CD57 antibodies was performed with Leica BOND-MAX autostainer (Leica Microsystems) to detect the infiltration of LGL cells (CD57 staining) and phosphorylation of STAT3 (pSTAT3 staining). Paraffin sections from 5 LGL leukemia patients and 2 healthy control bone marrow (BM) biopsies were processed with Bond Polymer Refine Detection kit (Leica Microsystems) using citrate buffer for antigen retrieval. Staining was done with a STAT3 Tyr 705 antibody (9145L, Cell Signaling Technology) diluted 1:100 or a CD57 antibody (TB01, Dako) diluted 1:100.

The slides were analyzed with the Zeiss Axio Imager AX10 microscope and photographed with Nuance FX multispectral tissue imaging system (420-720nm). The pictures were managed and prepared with Nuance 3.0.0

16. Statistical methods

For the luciferase reporter assays in study II and III statistical comparisons between the variants were performed with one-way ANOVA and the Dunn's multiple comparison test. A P value <0.05 was considered statistically significant.

In study IV, correlation of the drug sensitivity profiles across all T-PLL patients was performed using the Spearman and Euclidean distance measures of the drug and sample profiles, respectively. Weak drugs (mean DSS<5) were excluded from the analysis. Hierarchical, complete linkage clustering was performed with Cluster 3.0[152] and visualized in Java Tree View (v1.1)[153].

Results

17. Characterization of STAT-mutation negative LGL patients (I)

To elucidate what other molecular factors are driving the pathogenesis of LGL leukemia, STAT-mutation negative patients were selected for exome sequencing and gene expression analysis. The chosen patients fulfilled the WHO 2008 diagnostic criteria for LGL leukemia.

17.1. Exome sequencing results of STAT-mutation negative patients

Three patients previously found not to carry *STAT3* or *STAT5* mutations were sent for exome sequencing. All patients had T-LGLL with a CD3+CD8+CD57+TCRαβ+ phenotype and harbored a CD8+ monoclonal expansion, which was detected by TCR Vβ analysis. Two of the patients harbored a major clonal expansion (73% and 89% of CD8+ cells), while one patient had a smaller expansion (28%). No clonal expansions were detected using the TCR Vβ analysis of the CD4+ lymphocytes. All three patients were diagnosed with T-LGLL at an advanced age (> 60 years). Concomitant neutropenia was seen in patients 1 and 2, and the latter also suffered from anemia and monoclonal gammopathy of unknown significance. Patient 3 had no concomitant disorders (**Table 6**).

Table 6. *Clinical characteristics of the exome sequenced STAT-mutation negative patients*

Patient	LGL leukemia type	Sex	Age at dg (y)	WBC count at dg (10E9/L)	Lymph count at dg (10E9/L)	Lymph count[1] (10E9/L)	CD8+ Vbeta[2]	Concomitant disorders	Therapy
1	CD3+CD8+ CD57+TCRαβ	F	70	7.7	6.4	3.9	Vb.7.1: 28.2%	Neutropenia, BM eosinophilia	Neutropenia: G-CSF
2	CD3+CD8+ CD57+TCRαβ	M	76	10.8	9.40	5.8	Vb.20: 73%	MGUS, neutropenia, anemia	No treatment
3	CD3+CD8+ CD57+TCRαβ	M	60	11.1	3.7	3.8	Vb.3: 89.3%	no	No treatment

Abbreviations: M, male; F, female; y, years; WBC, white blood cell; dg, diagnosis; BM, bone marrow; MGUS, monoclonal gammopathy of unknown significance; G-CSF, granulocyte-colony stimulating factor.
[1] Lymphocyte count at the time of sample preparation; [2] Proportion of Vbeta clone in CD8+ cells of the sample

The exome sequencing of CD8+ (tumor) and CD4+ (healthy) cells yielded on average 53,184,000 paired reads that mapped to the reference genome. The median coverage was not analyzed, but 55–60% of the target exons were sequenced with more than 20-fold coverage. On average, the patients had 11 non-synonymous mutation calls with high-confidence somatic p-values (< 0.01). The bioinformatics pipeline identified deleterious mutations expected to impact protein function based on conservation scores and PolyPhen predictions. Based on somatic p-values and medical relevance, variants were chosen for validation by PCR (**Table 7**).

***Table 7**. Somatic mutations identified by exome sequencing in STAT-mutation negative LGL-leukemia patients and validated by capillary sequencing*

Patient	Gene	Chr	Chr. position	Mutation type	Transcript ID	Ref. base	Var. base	Site coverage	Tumor variant freq.	Protein	Somatic p-value[1]
1	PTPRT	20	40743955	Missense	NM_133170	C	T	92	14%	V995M	0,000604
2	BCL11b	14	99723858	Missense	NM_022898.1	T	C	43	51%	H126R	$1{,}45^{-07}$
	RAD21	8	117868903	Missense	NM_006265.2	C	T	59	27%	E266K	$2{,}45^{-05}$
3	SLIT2	4	20543120	stop-gained	NM_004787.1	G	A	42	55%	W674stop	$6{,}62^{-09}$
	NRP1	10	33510758	Missense	NM_001024628.2	C	T	43	37%	V391M	$1{,}40^{-07}$

Abbreviations: Chr, chromosome; Var, variant; Ref, reference; NA, not available.
[1] Somatic p-value for somatic/LOH-events
[2] Rejected-substitution score describing the conservation of the amino acid from the program

The protein tyrosine phosphatase receptor T (*PTPRT*) was found to be mutated in patient 1 with a variant frequency of 14%. As the patient harbored a 28% clone in the leukemic CD8+ fraction, the mutation is presumed to be heterozygous. The tumor suppressor gene PTPRT is known to mediate STAT3 deactivation through the reversal of Tyr705 phosphorylation on STAT3[70]. This novel V995M mutation converts a highly conserved hydrophobic valine residue into methionine. The mutation occurs in the cytoplasmic part of PTPRT, within the tyrosine-protein phosphatase 1 domain that is actively responsible for the phosphatase activity (**Fig 8**). This *PTPRT* mutation could therefore affect the activation of STAT3 by reducing the dephosphorylation of Tyr705, leading to an increase in the expression of STAT3 target genes.

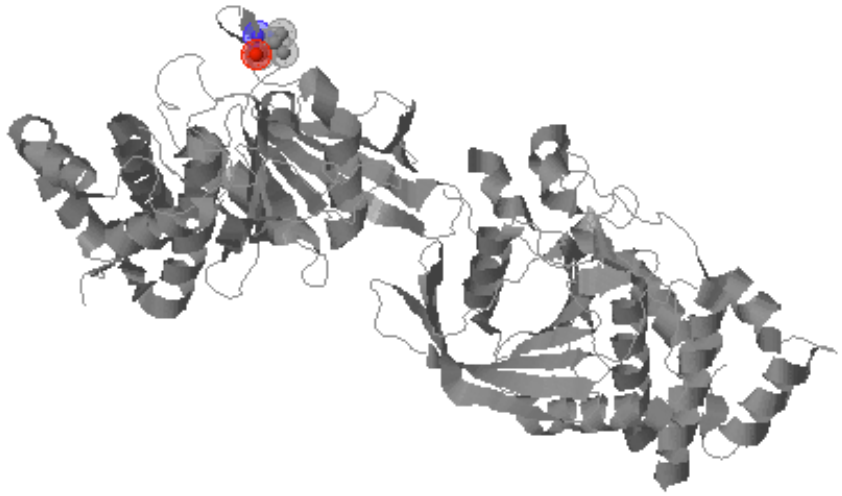


***Figure 8.** Schematic representation of the location of the V995M mutation in PTPRT from Polyphen2. The mutation is located in the catalytically active tyrosine-protein phosphatase 1 domain that is responsible for the phosphatase activity of PTPRT.*

Through exome sequencing, an H126R mutation in the *BCL11B* gene was revealed in the leukemic cells of patient 2 at a VAF of 51%. BCL11B is essential for T-cell survival, and its overexpression can lead to an increase in T-cell activation and proliferation[154]. In hematopoietic lineages, BCL11B expression is restricted mainly to T-cells and plays a key role in both the maintenance of T-cell identity and T-cell development[155]. Patient 2 also had a missense mutation in the gene *RAD21* (E266K), which mediates the repair of DNA double-strand breaks and plays a role in chromatid cohesion during mitosis. Furthermore, the deletion of *RAD21* in mouse thymocytes leads to defective chromatin architecture at the TCRα locus and limited differentiation potential[156].

The exome sequencing of the CD8+ LGL cells of patient 3 revealed a *SLIT2* mutation (W674stop) with a variant frequency of 54%. The mutation is located within a cysteine-rich domain bordering an LRR-domain, hypothesized to mediate protein-protein interactions. SLIT2 is a secreted glycoprotein that possesses anti-inflammatory properties and is, for example, able to modulate CXCR4-mediated functional effects in T-cells[157]. Another mutation, V391M in neuropilin 1 (*NRP1*), was detected with a VAF of 37% in the patient's leukemic sample. The mutation site is highly conserved and located within the discoidin domain, a major domain of many blood coagulation factors. NRP1 was originally known as a receptor for the semaphorin 3 subfamily, mediating axonal growth and neuronal guidance. NRP1 also mediates interactions between dendritic cells (DCs) and T-cells that are essential for the initiation of the primary immune response[158].

To validate the findings, the mutation sites of *PTPRT, BCL11b, SLIT2 RAD21*, and *NRP1* were screened with Sanger sequencing in a larger cohort of STAT-mutation negative LGL leukemia patients (n=113). No additional patients were found to have mutations at the targeted sites.

17.2 Gene expression analysis of STAT-mutation negative LGL-patients

Microarray gene expression analysis was performed with CD8+ RNA from three *STAT*-mutation negative patients and two patients with known STAT3 mutations. CD4+, CD8+, and NK cell fractions from healthy controls were used as biological replicates. In the distance dendrogram based on the gene expression profile (**Fig 9**), healthy CD4+, CD8+ and NK fractions clustered together, whereas LGL-leukemia patients formed a separate cluster. In the T-LGLL cluster, patients were situated independently of *STAT*-mutation status, suggesting that the expression profiles are quite similar. This was further emphasized when comparing the gene expression between T-LGLL patients with or without *STAT3* mutation: no genes were significantly over- or underexpressed.

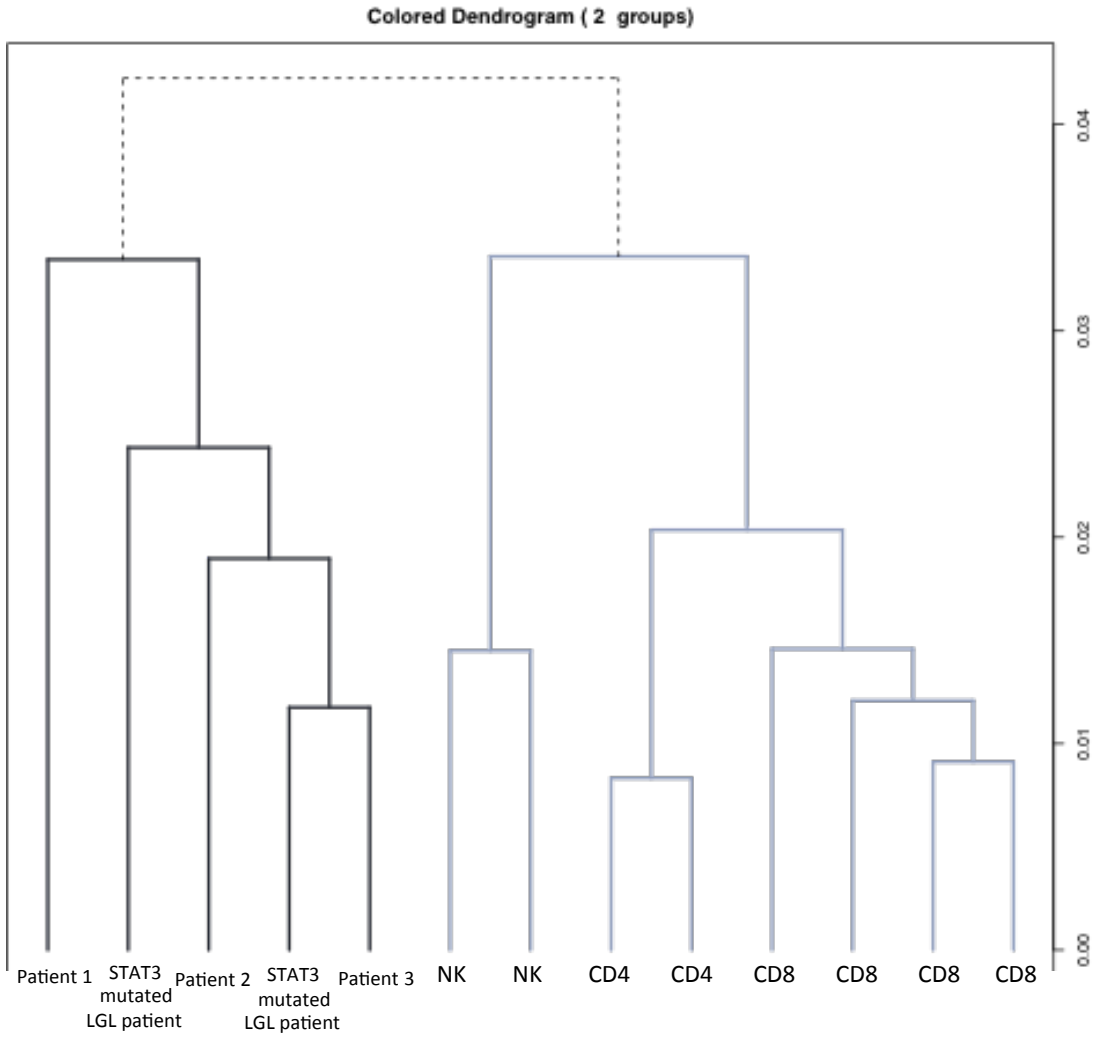


Figure 9. *Dendrogram visualizing the clustering of LGL leukemia patients and healthy controls based on their gene expression profiles.*

A comparison between all 5 T-LGLL patients and the CD8+ healthy controls revealed 39 genes to be differentially expressed (**Fig 10**). *FGR*, a member of the Src family of kinases, was 3-fold overexpressed in the T-LGLL patients when compared to healthy controls. It has been shown that FGR can phosphorylate STAT3, thereby activating the STAT3 pathway[159]. FGR is also activated by BCR-ABL in B-cells of B-ALL patients[160]. The transmembrane receptor SLAMF6, which mediates important regulatory signals between immune cells, was also 3-fold overexpressed in T-LGLL patients. SLAMF6 appears to co-stimulate particularly CD8+ and CD4-CD8- T-cells, while SLAM signaling has also been shown to be involved in the pathogenesis of autoimmune diseases, including systemic lupus erythematosus[161]. Interestingly, tumor necrosis factor (TNF) was expressed at a 2-fold lower level in T-LGLL patients when compared to healthy CD8+ cells.

Figure 10. *Heatmap representing the gene expression profiles of 3 patients without STAT- mutations, 2 STAT3-mutated patients, and 4 healthy controls (CD8+). A total of 39 genes were differentially expressed when comparing all T-LGLL patients to the healthy controls ($p < 0.05$, fold change ranging from 2.7 to -2.5).*

17.3. STAT3 activation in STAT-mutation negative patients

BM-biopsy samples from healthy controls and T-LGLL patients with different mutational status were studied with immunohistochemical staining of CD57 and pSTAT3 expression (**Fig 11**). LGL cells typically express CD57 on their surface, and IHC staining showed that BM-samples from LGL leukemia patients were infiltrated with CD57-expressing LGL-cells, while no such infiltration was observed in healthy control samples. Both the STAT-mutated and STAT-negative patients presented with positive pSTAT3-staining in the infiltrated lymphocytes, indicating STAT3 activation. No pSTAT3 was observed in the healthy control BM-samples.

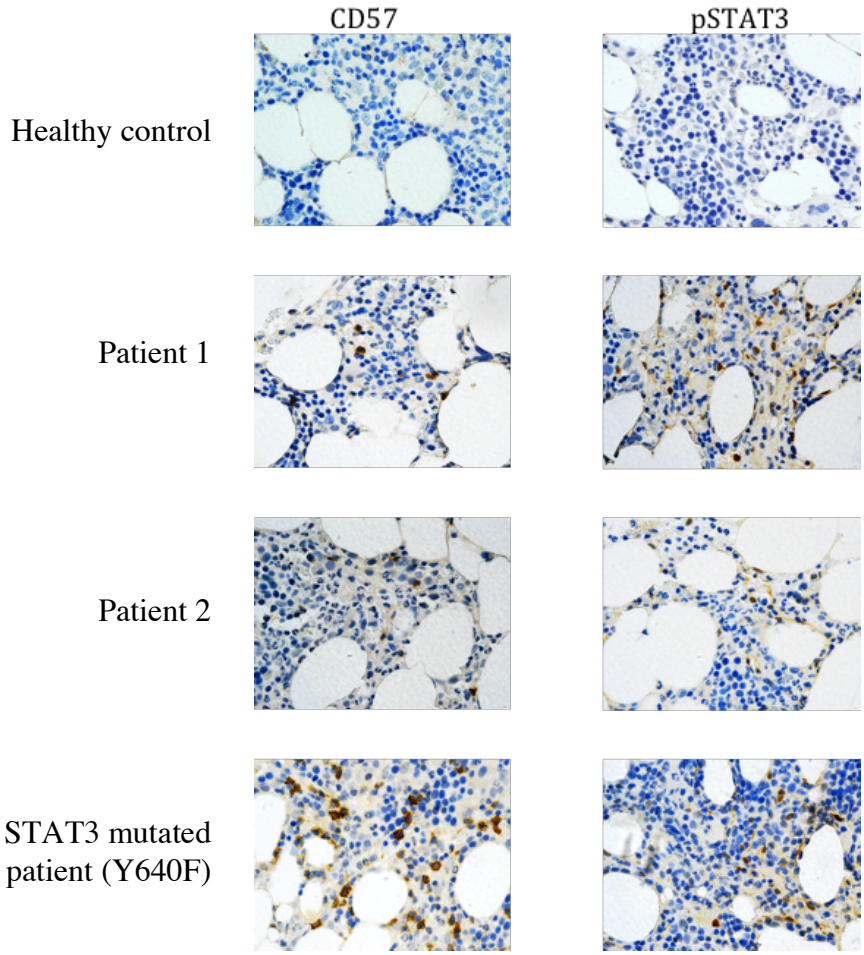


***Figure 11**. BM-biopsy samples from a healthy control and three LGL-leukemia patients stained with CD57 and pSTAT3 antibodies (patients 1 and 2 without STAT-mutations). No staining was observed in the healthy control, while the leukemic samples showed the infiltration of lymphocytes positive for CD57 and pSTAT3 (magnification, 63x).*

18. STAT3 and STAT5B mutations outside the SH2-domain in LGL-leukemia (II–III)

During the exome sequencing of additional *STAT*-mutation negative T-LGLL patients, *STAT3* and *STAT5B* mutations were found outside the hotspot SH2-domain covered by our custom STAT hotspot amplicon system. This led to the further screening of the entire *STAT3* gene and the *STAT5B* transactivation domain in a larger cohort of patients to elucidate whether mutations outside the SH2-domain are frequent in LGL-leukemia.

18.1. Somatic mutations in the coiled-coil and DNA-binding domain of STAT3 in LGL-leukemia (II)

To discover novel genetic markers, seven T-LGLL patients who were previously negative for *STAT3*- or *STAT5*-mutations in the SH2-domain were selected for exome sequencing. Patient 1 harbored a heterozygous missense mutation H410R in the DNA-binding domain. The VAF was 49% in the CD8+ fraction, which consisted of one major clone (vb.17: 95%) according to the TCR Vβ results (**Table 8**).

As the mutation spectrum outside the SH2-domain hotspot in *STAT3* has not previously been explored systematically, we developed a targeted deep amplicon sequencing platform that allows for the sensitive analysis of all the exons of the *STAT3* gene. Using this system, we sequenced MNC samples from 99 T-LGLL patients previously confirmed to be *STAT3* hotspot mutation negative. From these, three additional patients were discovered to have *STAT3* missense mutations outside of the SH2-domain (**Table 8** and **Fig 12**). Patient 2 exhibited the same H410R mutation seen in the DNA-binding domain (MNC VAF: 8.8%), and patient 3 had a S381Y mutation (MNC VAF 7%). Patient 4 had a novel F174S mutation (VAF 54% in CD8+ cells) in the coiled-coil domain of STAT3. The clinical phenotype of patients carrying *STAT3* DNA-binding and coiled-coil domain mutations did not differ from other typical LGL leukemia cases. The frequency of *STAT3* mutations was 3.8% (4 of 106 patients) in patients with no previously detected hotspot mutations in *STAT3* or *STAT5*.

Table 8. *Clinical characteristics of the patients with STAT3 mutations outside the SH2-domain*

Exome sequencing	Mutation	VAF	Sex	Age at diagnosis	Hb (g/L)	Leuk (10^9/L	Lymph (10^9/L)	Neut (10^9/L)	Trom (10^9/L)	Vbeta	Concominant disorders
Patient 1	H410R	49% CD8+	M	60	85	10.2	1.6	1.4	203	vb.17: 97%	Anemia, neutropenia, B-cell dyscracia, hypergammaglobulinemia
Amplicon sequencing											
Patient 2	H410R	9% MNC	F	75	127	4.2	3.7	0.0	260	NA	Neutropenia
Patient 3	S381Y	7% MNC	F	20	108	7	4.7	2	181	NA	Anemia
Patient 4	F174S	54% CD8+	M	68	95	8.6	7.2	1.1	211	NA	Anemia, neutropenia

Abbreviations: M, male; F, female; VAF, variant allele frequency; Hb, hemoglobin; Leuk, leukocytes; Lymph, lymphocytes; Neut, neutrophils; Trom, thrombocytes.

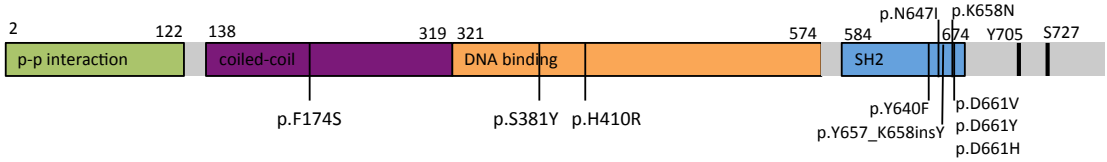


***Figure 12**. Location of novel mutations in the coiled-coil alpha domain and DNA-binding domain in addition to the previously described SH2 domain mutations in STAT3.*

18.2 Somatic STAT5B mutations in CD4+ T-cell large granular lymphocyte leukemia (III)

CD4+ T-LGL leukemia is a rare and poorly described disease entity. Through the exome sequencing of three such patients, we were able to identify somatic mutations in the transactivation domain of *STAT5B* in two CD4+ T-LGLL patients. Patient 1 had a Q706L mutation at a VAF of 45% in the CD4+CD8+ tumor fraction. Patient 2 displayed an S715F mutation (VAF 36%) in the CD4+ fraction (**Fig 13**). This mutation is located in a serine phosphorylation site in the transactivation domain.

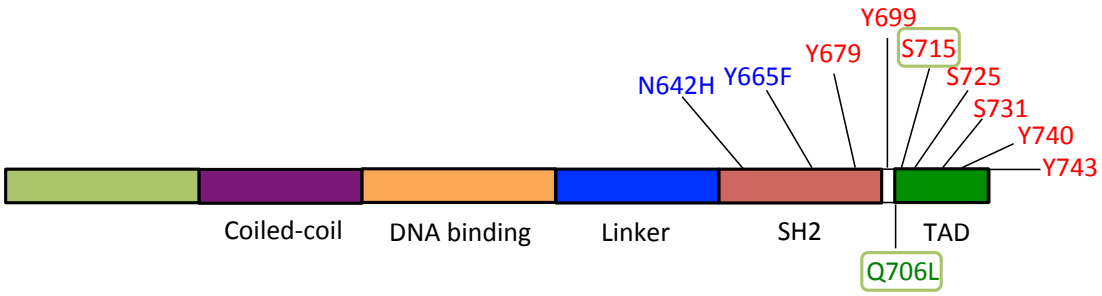


***Figure 13.** Linear representation of the STAT5B protein structure and mutations found in CD4+ LGL patients (blue and green boxes). Red represents multiple known tyrosine and serine phosphorylation sites.*

To elucidate whether *STAT5B* mutations are more prevalent in CD4+ LGLL cases, deep amplicon sequencing was used for the screening of the SH2 and transactivation domains of *STAT5B* in CD4+ (n=8), *STAT3*-mutated CD8+ (n=37), and non-mutated CD8+ (n=58) T-LGLL patients. Additionally, the same regions were screened with Sanger sequencing in Japanese and Chinese LGL-leukemia cohorts consisting of CD8+ and CLPD-NK cases (n=57). None of the patients with CD8+ T-LGLL or CLPD-NK had *STAT5B* mutations. In contrast, 4 of 8 CD4+ T-LGLL cases had *STAT5B* mutations. Of the four patients with *STAT5B* mutations, three possessed the earlier described N642H mutation and one the Y665F mutation. Altogether, the *STAT5B* mutation frequency in CD4+ T-LGLL patients in our cohort was 55% (6/11 patients). This is significantly higher than in the previous study (2%) of 211 CD8+ T- and NK-cell LGL-leukemia cases where *STAT5B* SH2 domain mutations were initially discovered[19].

Contrary to other more aggressive T-cell malignancies with *STAT5B* mutations, the disease courses of our 6 *STAT5B*-mutated patients were indolent, and

none of the patients needed treatment during the observation time (median follow-up 4 years). Two patients showed neutropenia, but anemia was not noted (**Table 9**). In our cohort, none of the 11 cases with CD4+ T-LGLL suffered from RA although it is commonly associated with CD8+ T-LGLL. All *STAT5B*-mutated CD4+ T-LGL cases possessed a TCRαβ+CD16-CD56+CD57+ T-cell phenotype in accordance with the earlier reports[162,163,164]. Interestingly, all *STAT5B*-mutated patients had mutation VAFs corresponding to large monoclonal expansions, while significant proportions of *STAT3* mutations in CD8+ T-LGL leukemia and CLPD-NK are detected in small subclones.

Table 9. *Clinical features of STAT5B-mutated CD4+ T-LGLL patients*

	Patient 1	Patient 2	Patient 3	Patient 4	Patient 5	Patient 6
STAT5b mutation	Q706L (45% VAF)	S715F (36% VAF)	N642H (25% VAF)	N642H (46% VAF)	Y665F (31% VAF)	N642H (27% VAF)
Age (years)	61	70	74	79	82	66
Sex	M	F	M	F	M	M
WBC count (10^9/L)	8.5	10.2	9.0	8.7	13.9	9.4
Neutrophil (%)	40	16	12	5	51	32
LGL (%)	52	72	71	91	39	63
Hb (g/L)	134	124	119	126	155	141
Platelets (10^9/L)	399	204	144	186	245	265
Complications	diabetes	none	none	gastrointestinal hemorrhage	none	lung cancer
Observation periods	5 years	7 years	14years	6months	3years	2years
Outcome	alive	alive	death	alive	alive	alive

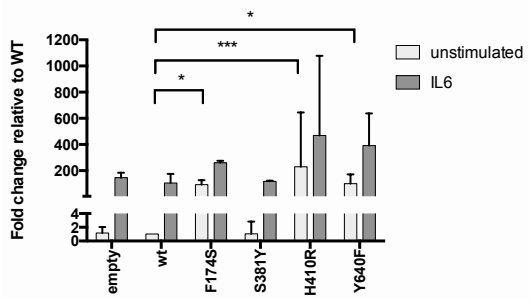
Abbreviations: VAF, variant allele frequency; M, male; F, female; WBC, white blood cell; LGL, large granular lymphocyte; Hb, hemoglobin.
Neutrophil and LGL percentage from whole white blood cell population.

18.3. Functional studies with mutated STAT3/5b (II–III)

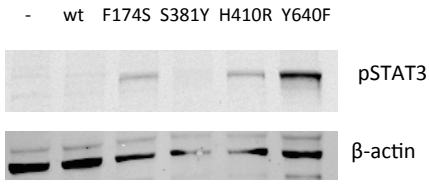
To explore the functional implications of the novel *STAT3* and *STAT5B* mutations, we generated expression constructs for WT and a variant STAT3/STAT5B protein and transfected appropriate cell lines containing reporters. Luciferase measurements of SIE-reporter HEK-293 cells transfected with these constructs revealed transcriptional activation, while the phosphorylation status of STAT3 and STAT5 was investigated with protein lysates in western blots.

In HEK-293 cells, a noticeable increase in both basal and IL6-stimulated transcriptional STAT3 activation was induced by the coiled-coil domain F174S and DNA-binding domain H410R variants compared to WT STAT3 (**Fig 14A**). The S381Y variant, on the other hand, did not show a similar increase. The phosphorylation status of the variants revealed F174S and H410R to induce a phosphorylation pattern comparable to the known activating Y640F SH2-domain mutation (**Fig 14B**). However, the S381Y variant did not seem to be highly phosphorylated, as the level of pSTAT3 was similar to that seen with the WT STAT3. Furthermore, the expression levels of six known STAT3 target genes were measured by qPCR, and *SOCS3, CCL2, JUNB,* and *BCL3* were revealed to be upregulated by the F174S, H410R, and Y640F mutations but not by the S381Y variant when compared to the WT (**Fig 14C**).

A.



B.



C.

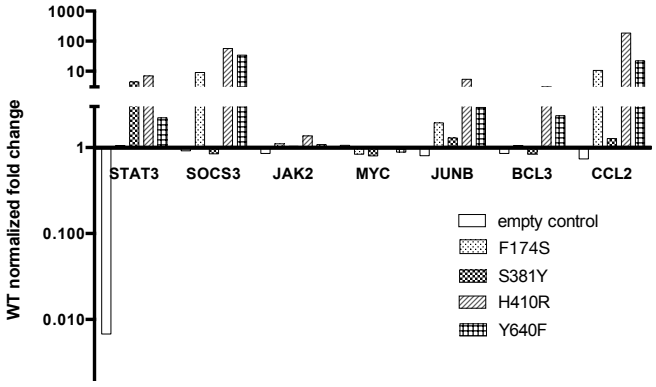
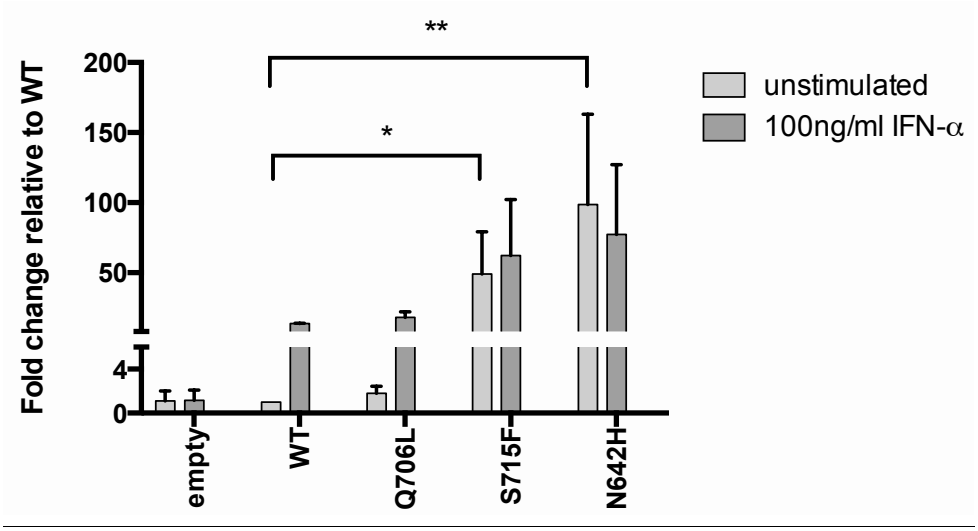


***Figure 14**. A. STAT3 reporter assay results. Each condition was tested in triplicate, and the statistical significance was calculated with a one-way ANOVA (*=p<0.05, **=p<0.001, error bars representing SD). B. Western blotting results of the phosphorylation of STAT3. β-actin was used as a loading control. C. Expression levels of six known STAT3 target genes were measured with qPCR All reactions were run in triplicate wells, and gene expression was quantified using the delta-delta Cq method.*

In HeLa cells transfected with the STAT5B constructs, the mutated STAT5B S715F construct significantly enhanced the transcription of the co-transfected STAT5 reporter (18-fold compared to WT STAT5B) similar to the N642H mutation (**Fig 15A**), while the Q706L mutation activation was equal to the WT. With the western blot analysis, the S715F and N642H mutations showed significantly increased phosphorylation when compared to the WT STAT5B (**Fig 15B**) in accordance with their activating nature.

A.



B.

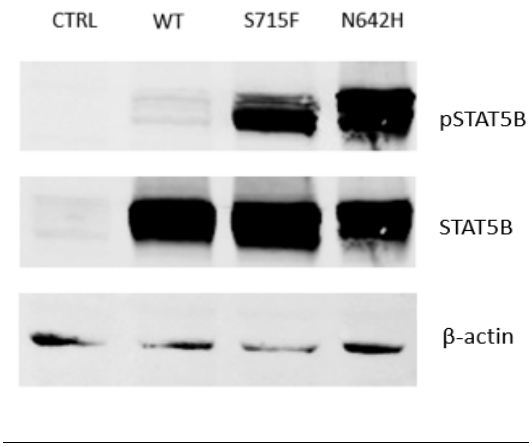


Figure 15*. A. STAT5B reporter assay results The experiment was repeated three times. Columns present mean of the fold-change activity. Error bars indicate the standard error of the mean (S.E.M.). B. Western blotting results of the phosphorylation of STAT5B. β-actin was used as a loading control.*

19. Discovery of novel drug sensitivities in T-PLL by *ex vivo* drug testing and genetic profiling (IV)

To find novel therapeutic options for the highly aggressive T-cell malignancy T-PLL, we applied *ex vivo* drug sensitivity testing to 39 patient samples. In addition, in an effort to molecularly profile the disease and find potential biomarkers for treatment, a larger cohort of T-PLL patients (n=70) underwent targeted JAK/STAT amplicon sequencing. Gene expression analysis was also applied to a smaller cohort (n=12) to find pathways addicted to certain drugs (**Fig 16**).

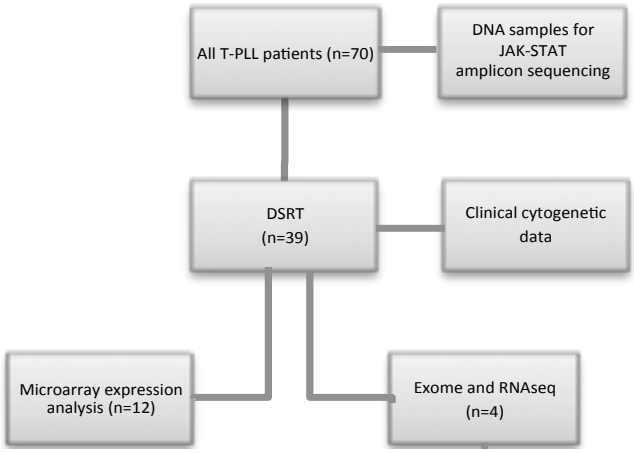


Figure 16. *Study design with relevant patient numbers.*

19.1. Targeted deep sequencing of the JAK/STAT pathway in T-PLL

Deep targeted amplicon sequencing of known recurrent genetic variants in the genes *STAT5B, JAK1,* and *JAK3* was performed with the Illumina MiSeq platform. The K315E mutation hotspot in *IL2RG* was screened with Sanger sequencing (see Table 3 for primer sequences).

The targeted amplicon sequencing revealed that 69% of the T-PLL patient samples (48/70) harbored one or several mutations in genes involved in the JAK/STAT pathway (*JAK1, JAK3, STAT5B,* or *IL2RG*; **Figs 17 and 18**). Single *JAK3* mutations were seen in 29% of T-PLL patients, and the most prevalent of these mutations was the M511I missense mutation located in the linker between the SH2 and the pseudokinase domain (37% of all patients). *STAT5B* mutations were seen in 7% of the patients, while 6% harbored single *JAK1* mutations, and 1% of the cases were *IL2RG*-mutated alone. Twenty-six percent of the patients harbored more than one mutation in the JAK-STAT pathway. The VAF of these mutations ranged from 1% to 87% in the MNC samples, with 38/48 (79%) of the mutated cases showing mutations with a frequency above 10%. Interestingly, clonal *STAT5B* mutations (9/70) did not coexist with any clonal *JAK* mutations in our cohort. In patients with multiple mutations, subclonal mutations were commonly observed.

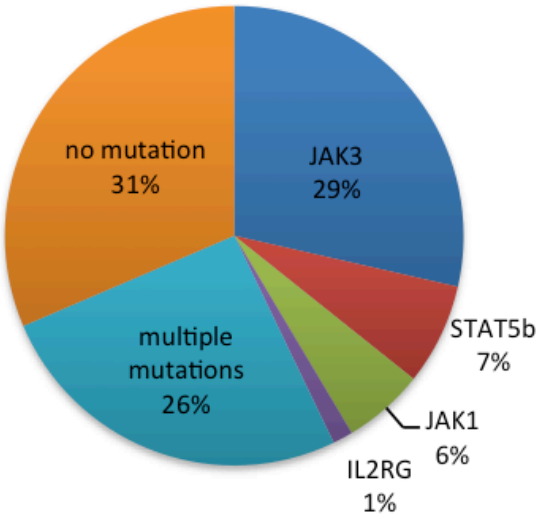


Figure 17. *Targeted amplicon sequencing was performed on the mutation hotspots in the genes JAK1, JAK3, STAT5B, and IL2RG. Based on this targeted sequencing, 69% of T-PLL patients harbor one or multiple mutations of the JAK/STAT pathway.*

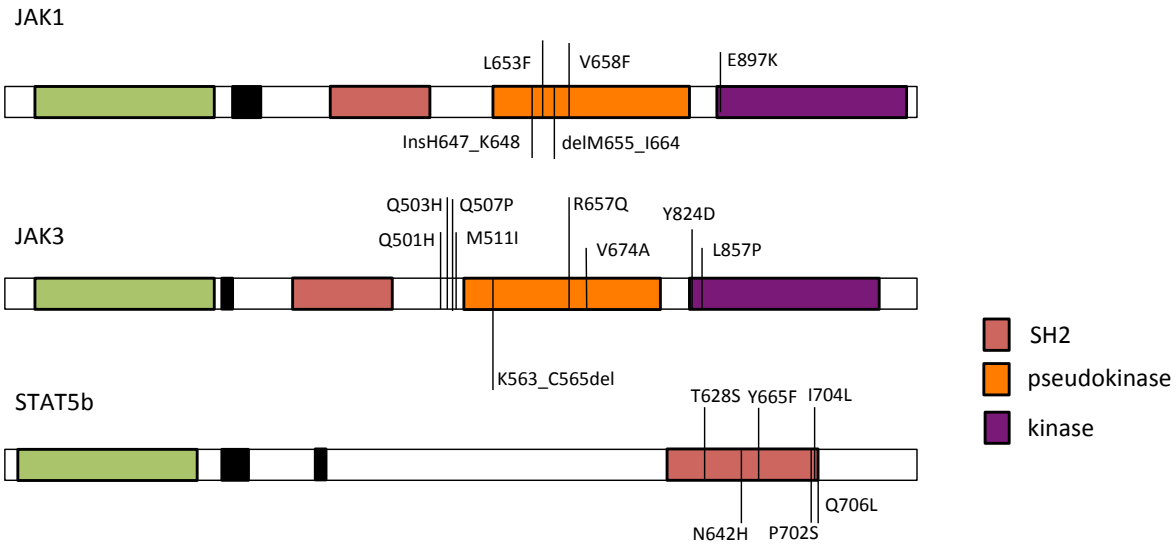
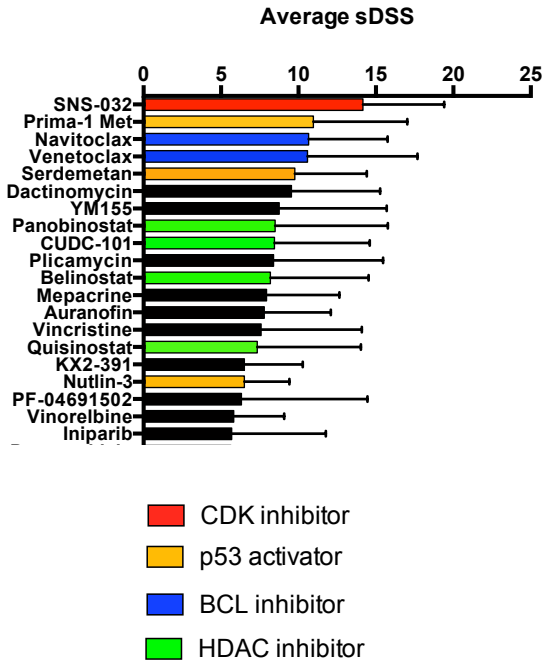


Figure 18. *Linear 2D model of JAK1, JAK3, and STAT5B with the positions of the mutations detected with deep amplicon sequencing. In IL2RG, only the K315E mutation hotspot was sequenced (not shown).*

19.2. Clustering of drug sensitivity profiles to reveal selective sensitivities

By calculating the mean of the DSS values for all T-PLL patients screened and comparing them with the median of the healthy controls, we ranked the drugs by their leukemia-specific effect on T-PLL samples. Based on DSRT analysis, the drug with the highest average sDSS score was the CDK inhibitor SNS-032 followed by the p53 re-activator Prima-1 and BCL inhibitors navitoclax and ABT-199 (**Fig 19A**). The CDK inhibitor SNS-032 is highly effective in killing cells in MNC samples from T-PLL but also in healthy PB. The p53 re-activator Prima-1 Met, on the other hand, has a more T-PLL-specific profile when compared to other leukemias and healthy control MNCs (**Fig 19B**).

A.



B.

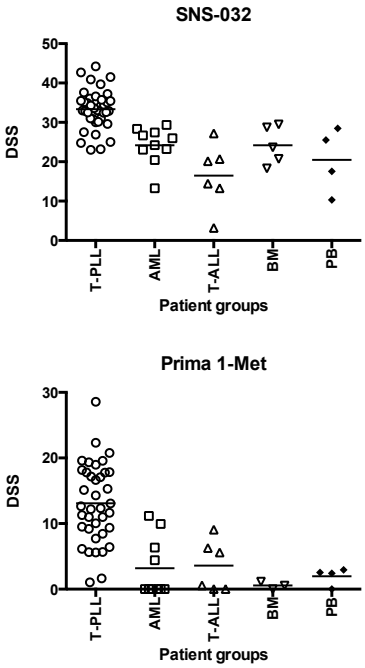


Figure 19. *A. Top 20 most effective drugs for T-PLL patients (n=39) based on the median of sDSS scores where the DSS scores have been compared with DSS from 4 healthy subjects. Drugs with median sDSS values above 5 are considered selective to the cell samples tested. B. Comparison between DSS scores of specific drugs for different leukemias and healthy controls.*

Next, we performed a drug-drug correlation of the DSS scores of all T-PLL patients to reveal effective drug groups, and PI3K/AKT/mTOR, HSP90, and HDAC inhibitors stood out as the most effective substances. The clustering of the patients' sDSS scores for these substances showed a division of samples into 3 main drug-response groups: sensitive, only sensitive to HDAC inhibitors, and resistant groups (**Fig 20**). Strikingly, the presence of *JAK/STAT* mutations, *TCL1A* translocations, or *ATM* deletion status did not predict responses to any specific substance class.

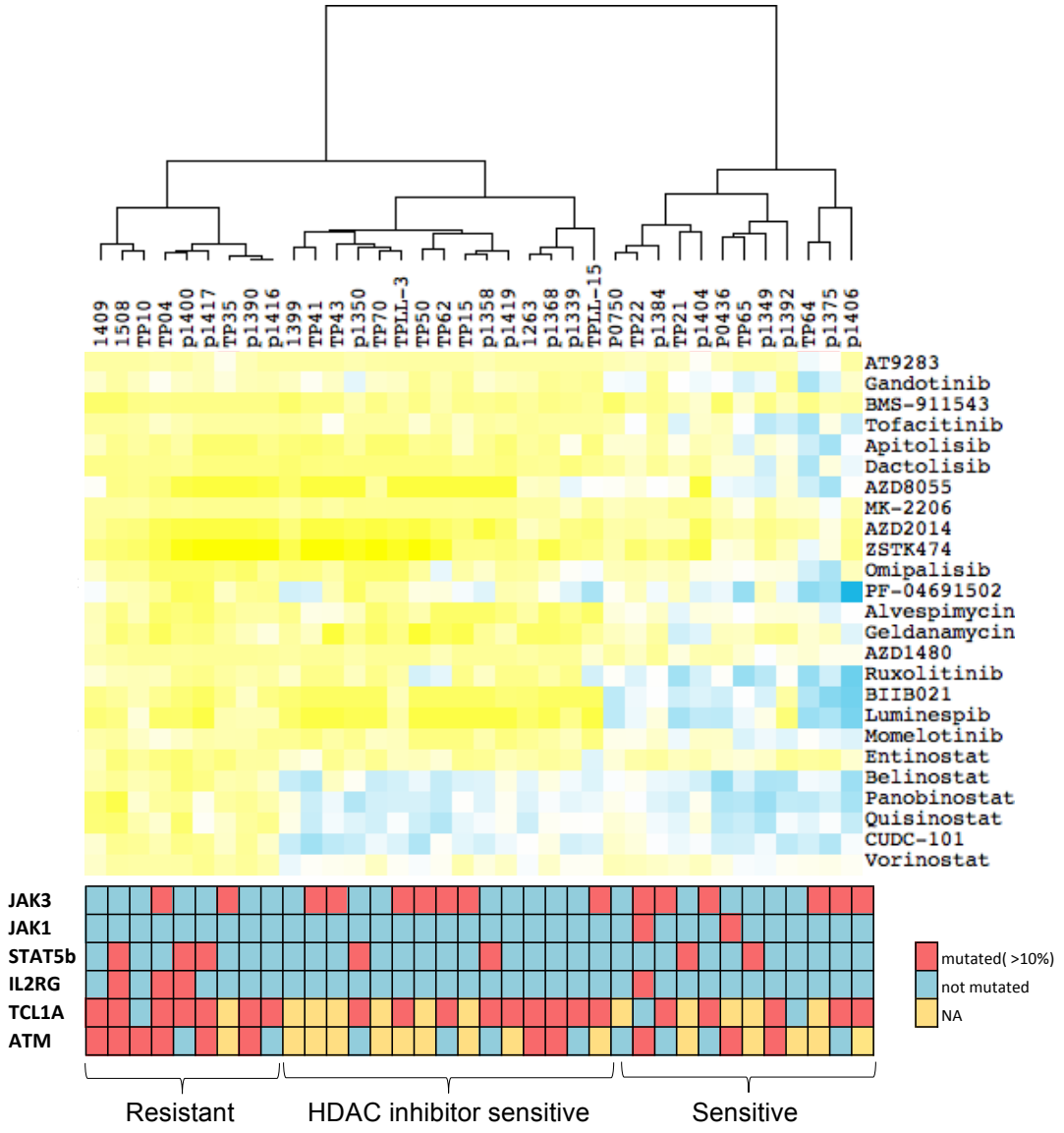


Figure 20. *Clustering of the sDSS values of the T-PLL patients was performed using unsupervised hierarchical complete-linkage clustering using Spearman and Euclidean distance measures of the drug and sample profiles. Hierarchical clustering was performed with Cluster 3.0 and visualized by Java Tree View.*

19.3. Linking drug responses to genetic profiles

The recurrent mutations in the JAK/STAT signaling elements implicate a new actionable lesion to be explored in T-PLL. A closer inspection of the JAK inhibitors ruxolitinib (JAK1/2) and tofacitinib (JAK1/3) showed that patient samples were on average more sensitive to ruxolitinib (**Fig 21**). However, drug responses to JAK inhibitors did not link to the presence of recurrent genetic aberrations in the JAK/STAT pathway. In the majority of patients with *JAK3* mutations, ruxolitinib was effective, but there were also notable exceptions. Interestingly, among the patient samples with practically no sensitivity to ruxolitinib or tofacitinib, there were two *STAT5B*-mutated patients with the highly activating N642H variant.

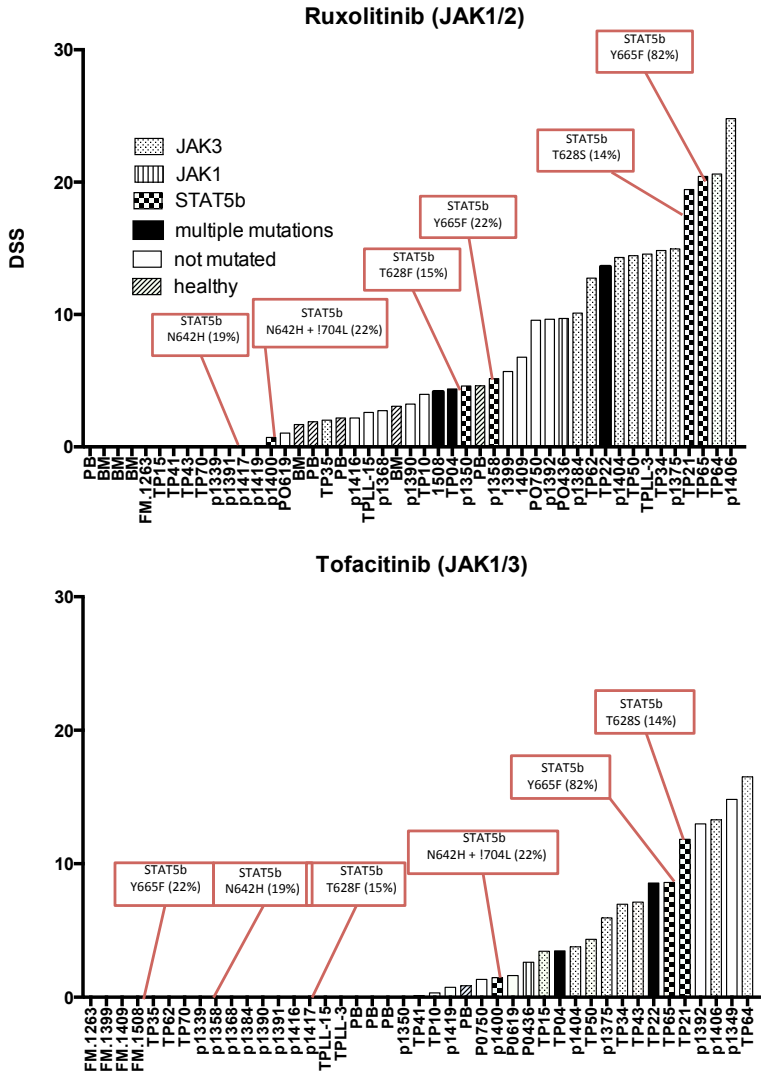


Figure 21. *Drug sensitivity of T-PLL patients to JAK inhibitors ruxolitinib (JAK1/2) and tofacitinib (JAK1/3). Treatment ranges were 0.001–10 µM for ruxolitinib and 0.0005–5 µM for tofacitinib.*

19.4. Network analysis reveals active pathways

To elucidate which pathways are active in PLL and which target genes are acted upon by certain drugs, a target addiction score (TAS) was calculated for 12 samples for which drug response and microarray expression data were available. The analysis revealed 145 addicted genes and highlighted genes related to histone deacetylation and cell cycle regulation (**Fig 22**), further supporting the use of HDAC and CDK inhibitors in T-PLL. The TAS scores and differential gene expression values were also used to analyze canonical pathways in Biocarta (**Fig 23A**). In two patients, the retinoblastoma (RB) pathway, related to checkpoint signaling in response to DNA damage, was upregulated when compared to healthy controls. In normal circumstances, the ATM protein kinase detects DNA damage and, in response to this, activates DNA repair factors and inhibits cell cycle progression. However, in T-PLL, ATM is frequently dysfunctional, leading to progression to S-phase despite defects in the genome. Compounds such as p53 activators, CDK inhibitors, and HDAC inhibitors are able to affect this pathway (**Fig 23B**).

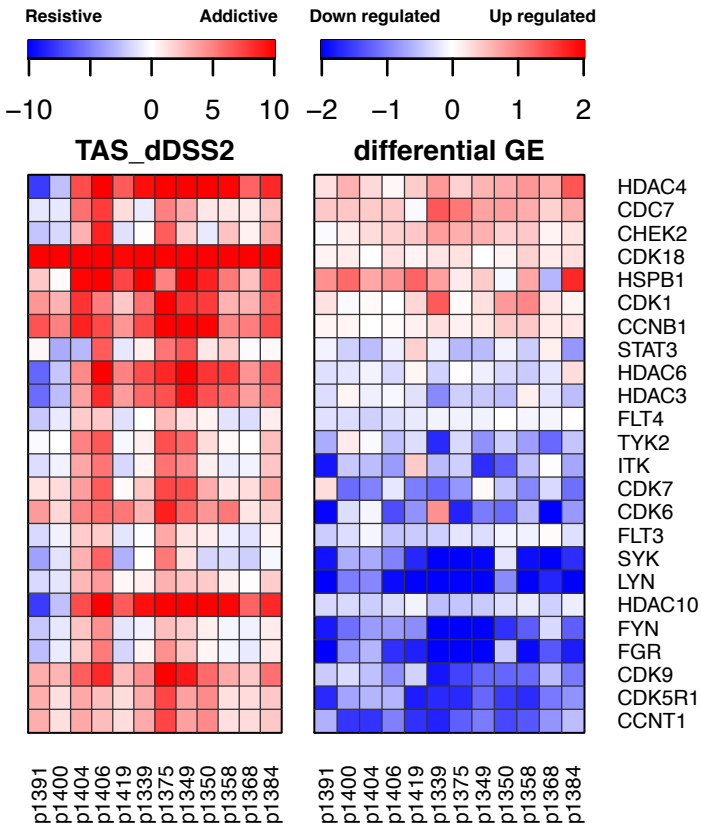
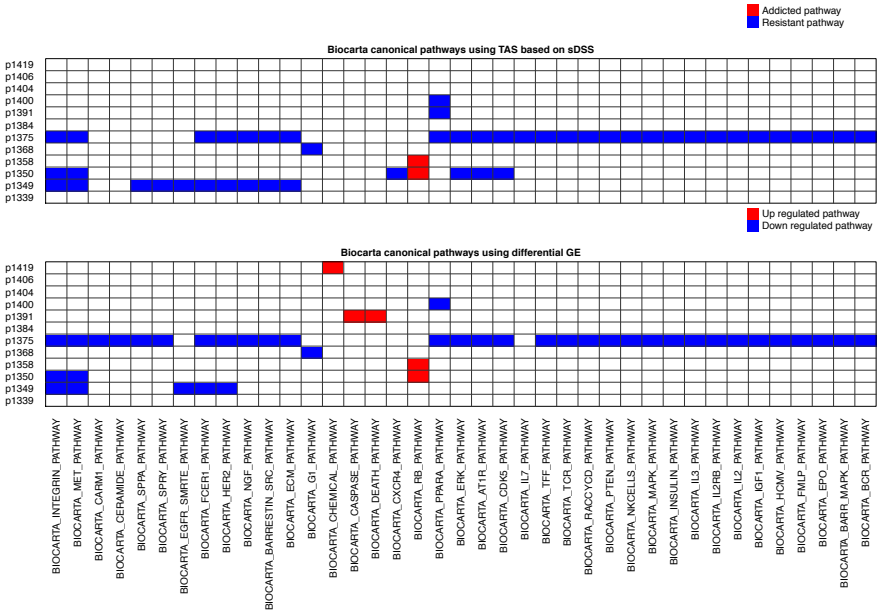


Figure 22. *TAS heatmaps based on microarray expression data and DSRT data of 12 T-PLL patients.*

A.



B.

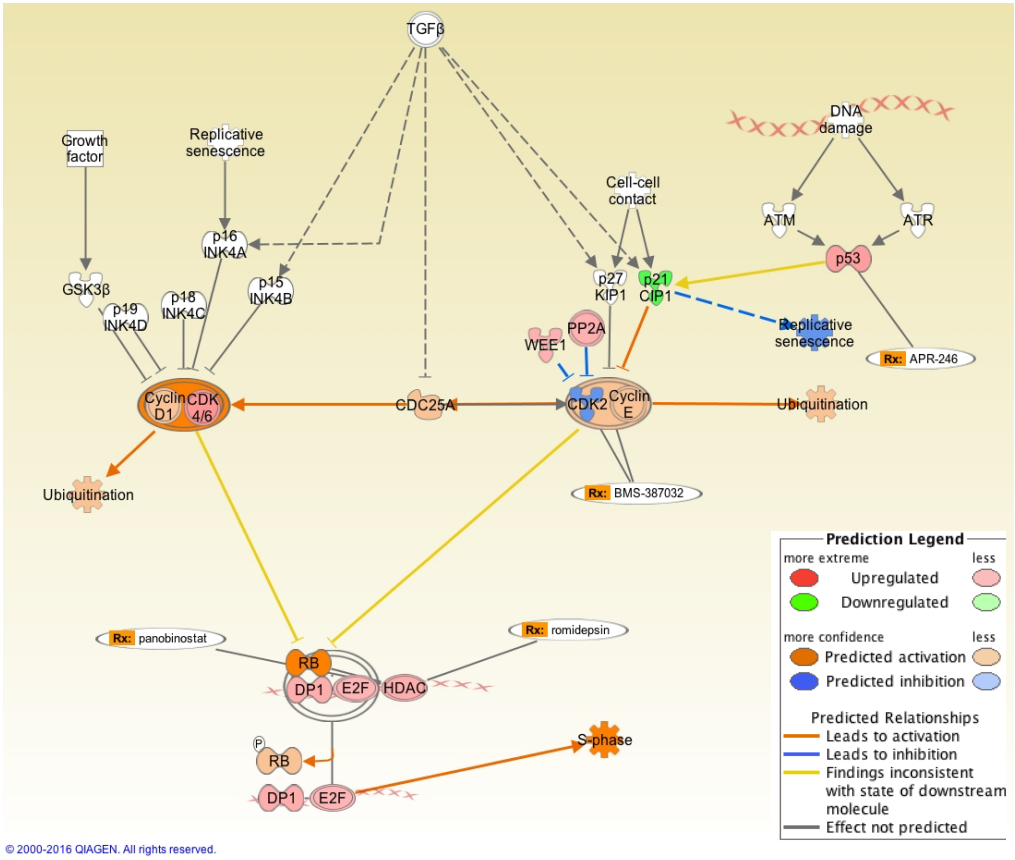
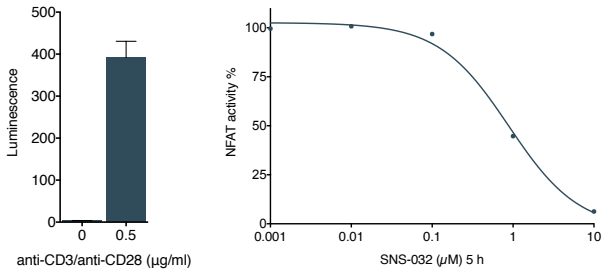


Figure 23. *A. Biocarta canonical pathways using TAS and differential gene expression.* ***B.*** *Predicted activation of cell cycle regulation pathway from IPA software overlaid with microarray expression data. Cell cycle checkpoint control at the G1 to S transition prevents the cell cycle from progressing when DNA is damaged. In response to DNA damage, ATM phosphorylates the tumor suppressor p53 that interacts with p21 to block the activity of CDK2, preventing passage from G1 to S phase and harmful replication of damaged DNA. When the tumor suppressor RB is phosphorylated by cell cycle-dependent kinases like CDK2 and CDK4, RB no longer interacts with E2F, and the cell cycle proceeds through the G1-S checkpoint. APR-246= Prima-1, BMS-387032=SNS-032.*

19.5. SNS-032 inhibits TCR activation

Across all cases, the CDK2/7/9 inhibitor SNS-032 was the most effective substance. To further elucidate the effect of SNS-032 on the most central growth-regulating receptor in T-cells, namely the TCR, Jurkat T-cells containing an NFAT-coupled luciferase reporter to read out the activation of this distal TCR signaling effector were treated with SNS-032 in the presence of TCR crosslinking anti-CD3 and anti-CD28 antibodies. There was an NFAT-inhibition down to 45% in the presence of 1µM SNS-032 (**Fig 24A**). According to our gene expression data, the TCR signaling pathway in T-PLL cells at baseline appears to be activated by the overexpression of *PI3K*, *MEKK1, Ras*, and *NFAT* in addition to the downregulation of inhibitory signals by *CTLA4* (**Fig 24B**).

A.



B.

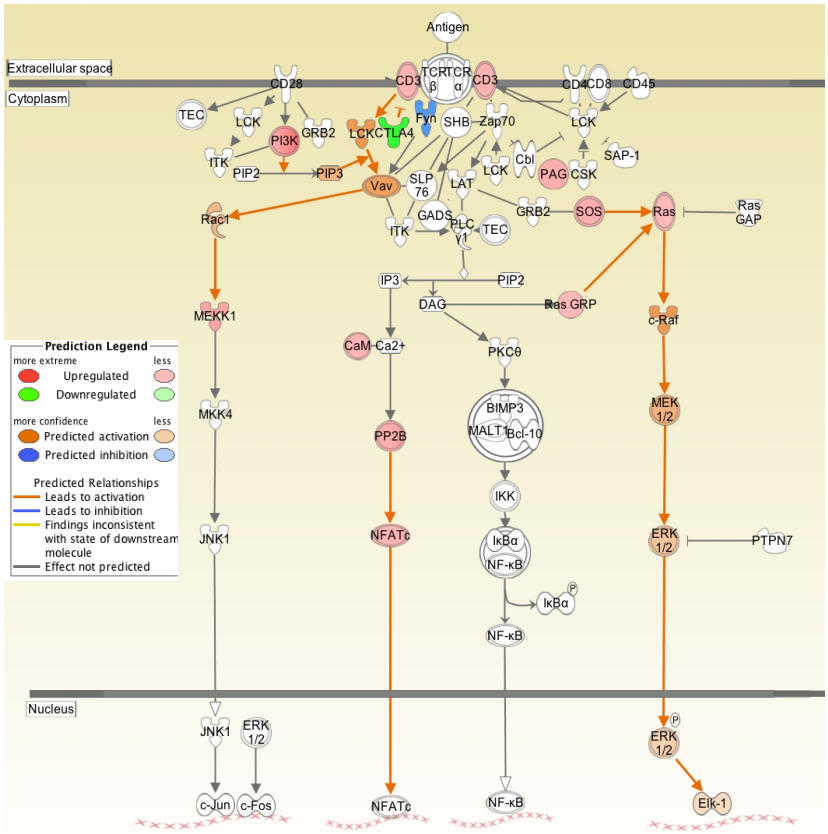


Figure 24. *A. Jurkat NFAT luciferase reporter cells were incubated with the indicated concentrations of SNS-032 on a 384-well plate for 5 hours in the presence of purified anti-CD3 and anti-CD28 antibodies. B. TCR signaling pathway from IPA software overlaid with gene expression data from RNA sequencing of 4 PLL patient samples and 2 CD4+ healthy control samples.*

Discussion

The aim of studies I–III was to characterize *STAT*-mutation negative LGL leukemia molecularly through the use of next-generation sequencing and gene expression analysis. Our previous findings revealed somatic mutations in the SH2-domain of either the *STAT3* or *STAT5B* genes in approximately 40–50% of LGL leukemia cases[2,19]. However, as the *STAT*-mutation negative patients also harbor large expansions of LGL cells, other somatic mutations may drive the development of the disease.

In study I, we showed that mutations in genes connected to the STAT-pathway and T-cell activation/proliferation represent rare genetic triggers for T-LGL leukemia. Previous studies have shown that almost all T-LGLL patients exhibit the activation of STAT3, proving that this pathway is essential for the pathogenesis of LGL leukemia[34,2,35]. In this study, we also found that STAT3 was phosphorylated in patients without actual STAT3 mutations, indicating that other underlying factors are leading to the activation of the pathway. To further confirm this finding, T-LGLL patients were also clustered independently of their mutational status in the expression analysis. Therefore, the novel V995M mutation in the *PTPRT* gene is particularly interesting, as it may directly impact STAT3 activation and is the first evidence of other mutations affecting the JAK/STAT pathway in LGL leukemia.

PTPRT is known to reverse the phosphorylation of Tyr705 on STAT3, a modification that is associated with STAT3 deactivation[70]. The *PTPRT* V995M mutation may thereby affect STAT3 activity by reducing the dephosphorylation of Tyr705, leading to an increased expression of STAT3 target genes. The mutation was found within the phosphatase domain D1 that is responsible for the proteins' phosphatase activity. *PTPRT* is also frequently disabled in other cancers, such as lung and gastric cancer, where the hypermethylation of the promoter of PTPRT increases STAT3 activation and causes sensitivity to STAT3 inhibition[165,166]. The effect of other missense mutations in the D1 domain of PTPRT have been studied[167] and the mutant proteins show a decrease in thermal stability and activation energy for phosphatase activity with respect to the WT protein. Inactivating mutations of the *PTPRT* gene may thereby have the same functional consequence as activating mutations of *STAT3* in T-LGLL patients.

To clonally expand, T-cells must have an acquired survival advantage. In addition to defects in apoptotic pathways, survival pathways such as PI3K-AKT may also be activated in LGL leukemia patients[34]. Therefore, it was interesting to observe novel missense mutations in *BCL11B* and *NRP1* in T-LGLL patients. BCL11B functions as a transcription factor required for normal T-cell development. The inactivation of *BCL11B* in mice leads to thymocyte developmental arrest and aberrant self-renewal activity[168,169]. Increased levels of BCL11B expression have been found to be associated with T-ALL[170], and the inhibition of BCL11B expression in malignant T-

cells results in apoptosis[171]. Mutations in *BCL11B* have also been reported in 9–16% of T-ALL patients[22,172,173].

NRP1, a receptor involved in axon guidance, is also expressed in human DCs and resting T-cells[158]. Interestingly, earlier studies have shown that NRP1 can play a role in the primary immune response during the formation of the immunological synapse between DCs and resting T-cells. The preincubation of DCs and resting T-cells with blocking NRP1 antibodies inhibited the DC-induced proliferation of T-cells. Thus the mutation in the *NRP1* gene may affect the proliferation capacity of T-cells.

These novel mutations affecting either the STAT3 or T-cell activation pathway were not found to be recurrent in our screening cohort (n=113). It seems that the *STAT*-mutation negative patients are a more heterogeneous patient cohort, and therefore no similar dominant mutation as in *STAT3* will be found in this group. However, in the Sanger sequencing screening assays, the primers covered only the mutation spots and nearby base pairs, and it is possible that by screening the whole genes, some additional mutations in the same genes will be found. Additionally, low frequency variants might go undetected due to the poor sensitivity of Sanger sequencing. To date, our group has performed the exome sequencing of 20 LGLL patients, and few recurrent mutations have been discovered. As mutations occurring in different genes with similar functions can drive a similar phenotype, further bioinformatical investigation of the different somatic variants should be performed to highlight pathways activated by mutations other than *STAT3/5B* in LGL leukemia.

In studies II and III, additional *STAT*-mutations were revealed in the process of the exome sequencing of T-LGLL patients. First, a subgroup of CD8+ T-LGLL patients (4%) harboring mutations in the DNA-binding and coiled-coil domain of *STAT3* were discovered. STAT3 activation is traditionally thought to be mediated through the dimerization interface containing the SH2-domain[174]. While the constitutive activation of STAT3 in LGL leukemia patients with SH2 hotspot mutations could be explained by an increased stabilization of STAT3 homodimers through enhanced hydrophobic attraction, the mode of activation through other domains is not as straightforward. Arginine residues within the DNA binding domain of STAT3 have, however, been found to promote phosphorylation and the intracellular shuttling of STAT3[175]. An arginine–glutamine exchange at the STAT3 moieties R414 and R417 reduces the cytokine-dependent tyrosine phosphorylation of STAT3. The H410R variant seen in two LGL leukemia patients results in one extra arginine residue in this area, and it is plausible that this increase in hydrophilicity within the DNA binding domain mediates the activation of STAT3. The H410R mutation was also seen to be recurrent—one of the hallmarks of activating mutations.

Previously, germline mutations in the DNA binding domain of *STAT3* have been found in different immune disorders. In hyper IgE syndrome [176], dominant-negative mutations in the DNA binding domain leads to the diminished DNA binding ability of

the STAT3 dimer. There has also been some interest in using specific STAT3 inhibitors in the treatment of other hematological malignancies, and recently a small molecule compound specifically targeting the STAT3 DNA-binding domain was found to inhibit cancer cell proliferation, migration, and invasion[177].

One CD8+ T-LGLL patient was found to harbor an F174S mutation in the coiled-coil domain of *STAT3*. This domain consists of four antiparallel helices and has previously been shown to be essential in receptor binding mediated by the SH2-domain as well as subsequent activation. The α1 region, containing the F174S mutation, seems to be crucial for STAT3 tyrosine phosphorylation stimulated by the epidermal growth factor (EGF) and IL6 as the deletion of the α1 abolished tyrosine phosphorylation of STAT3 in EGF-induced COS-1 cells[178]. One possible mechanism for the involvement of the coiled-coil domain in STAT3 phosphorylation is through the reinforcement of the protein-protein interaction between STAT3 and the activated IL6 receptor. Mutations in the coiled-coil domain have previously been seen in patients with inflammatory hepatocellular adenomas, where the E166Q mutant was shown to be constitutively phosphorylated on Tyr705, hypersensitive to IL6, and able to translocate to the nucleus [179]. Persistent STAT3 activation has also been found in activated B-cell-like diffuse large B-cell tumors, where the M206K *STAT3* mutation leads to increased cell proliferation[180].

In study III, exome sequencing identified novel somatic missense mutations in the transactivation domain of *STAT5B* in two CD4+ T-LGLL patients. Altogether, the *STAT5B* mutation frequency in the CD4+ T-LGL leukemia patients in our cohort was 55% (6/11 patients). This is significantly higher than in the previous study of 211 CD8+ T- and NK-cell LGL-leukemia cases, where *STAT5B* SH2 domain mutations were initially discovered in 2% of the cases[19]. The reported *STAT5B*-mutated patients had a much more aggressive disease course than that of typical T-LGLL. This finding, together with the high prevalence of *STAT5B* mutations in aggressive T-cell neoplasms, such as T lymphoblastic leukemia[103,104], T-PLL[4], and hepatosplenic T-cell lymphoma[95], may suggest that *STAT5B* mutations predict a more aggressive clinical behavior than does the mutational activation of *STAT3.* However, in our cohort of *STAT5B*-mutated CD4+ T-LGLL patients, the disease course was quite indolent, and none of the patients with *STAT5B* mutations needed treatment during the observation time. The analysis of STAT5 target genes with ChIPseq has shown that STAT5B binds to molecules such as DOCK8, SNX9, FOXP3, and IL2RA, making it a key factor in T-cell development[181]. Together, these results suggest that STAT5B plays a central role in the development of T-cell neoplasms and that *STAT5B* mutations can be considered a novel diagnostic marker for CD4+ T-LGLL.

The diagnosis of T-LGL leukemia is often difficult, as symptoms and immunophenotypical findings can resemble other reactive conditions, making the discrimination between malignant lymphoproliferation and reactive processes challenging. While T-cell clonality is considered an important finding supporting diagnosis, its significance remains controversial, as a proportion of healthy elderly individuals also present with non-neoplastic clonal T-cell expansions[182]. Therefore, consistent detection of *STAT3* and *STAT5B* mutations in T-LGL leukemia represents a more definitive new diagnostic biomarker. Subsequently, the importance of screening the entire *STAT3* and *STAT5B* genes in the diagnostic workup of T-LGLL is further emphasized.

There also seems to be a clear division between the occurrence of *STAT3* mutations in T-LGLL CD8+ cases versus the occurrence of *STAT5B* mutations in CD4+ cases. What could be the biological reason for this division? For STAT3, it is known that the genes of the activated IL6-STAT3 axis are overexpressed in CD8+CD57+ LGL cells due to chronic immune activation in T-LGLL[37]. Because the transcription of a gene has been found to correlate with the rate of mutagenesis[183], this could explain the increased number of somatic *STAT3* mutations. As for *STAT5B*, the upstream regulator IL-15 has been found to stimulate the proliferation and survival of CD4+ memory T-cells[184]. In a network model for T-LGLL survival signaling, IL-15 and PDGF were able to sustain a leukemic population[185], and the overexpression of IL-15 has been observed in T-LGLL, in autoimmune disorders, and during viral infections. Cytomegalovirus-derived stimulation has been associated with CD4+ T-LGL cases[163], which may explain the higher *STAT5B* variant rate seen in CD4+ cases.

Another leukemic disease where CD4+ cells are frequently *STAT5B*-mutated is T-PLL. Although the mutation sites are overlapping, the resulting disease phenotypes are completely different, suggesting that in the more aggressive T-PLL, other aberrations such as chromosomal translocations are also driving the disease. In study IV, we systematically explored the diversity of drug responses in a large collection of T-PLL patient samples *ex vivo* using a drug screening platform and explored the associations of drug sensitivities with genetic aberrations. Such a strategy has proven successful before in BCR-ABL-driven leukemias, where axitinib was found to be effective in patients harboring the resistance-causing T315I mutation in ABL1[135].

The deep amplicon sequencing of JAK/STAT pathway genes in our patient cohort revealed recurrent mutations in 69% of cases. Single-gene mutations affected *JAK3* (29%), *STAT5B* (7%), *JAK1* (6%), and *IL2RG* (1%), while coexisting multiple mutations were found in 26% of the cases, either within more than one gene or within the same gene. Compared to the previous genetic profiling of JAK/STAT pathway genes in T-PLL patients[4,51,52,186], these results were highly concordant. For example, Kiel et al. (2014) reported a 76% JAK/STAT mutation frequency in T-PLL (38/50). Overall, the recurrent involvement of JAK/STAT signaling elements in non-

synonymous mutations and indications of an activated JAK/STAT axis implicate a new actionable lesion to be explored in T-PLL. However, our sensitivity screen revealed that the *ex vivo* sensitivity of patient samples to JAK-inhibitors, such as ruxolitinib, momelotinib, tofacitinib and gandotinib, was not directly driven by these mutations in the JAK/STAT pathway. However, it was interesting to note that, for ruxolitinib, a JAK1/2 inhibitor, patients with the highly activating N642H *STAT5B* mutation were not as sensitive to the drug as patients with other *STAT5B* mutations. This could be due to the stability of the N642H mutant homodimer compared to the wild-type STAT5B (25-fold higher binding affinity), whereby the over-activation cannot be suppressed completely with JAK-inhibitors alone[187].

Generally, the clustering of drug sensitivity values in T-PLL revealed three patient subgroups driven by their resistance/sensitivity to HDAC-, HSP90-, and PI3K/mTOR/Akt-inhibitors. Surprisingly, despite the prevalence of the signature event of the activation of TCL1A (an established AKT coactivator) in the majority of T-PLL cases, only a subset responded to PI3K/AKT/mTOR inhibitors, such as AZD-8055, MK-2206, apitolisib, and dactolisib. This might be due to the rather initiating role of TCL1A or its predominant effect in the context of TCR activation. On the other hand, the HDAC-inhibitors vorinostat, panobinostat, CUDC-101, quisinostat, and belinostat showed efficacy in the majority of PLL samples *ex vivo*. Recently, it was shown that treating relapsed or refractory PLL-patients with a combination therapy consisting of cladribine and an HDAC inhibitor re-sensitizes patients to alemtuzumab and results in better overall survival[188]. Further highlighting the importance of epigenetic regulation in PLL, mutations in genes encoding epigenetic regulators *EZH2*, *TET2*, and *BCOR* have been found in a number of T-PLL patients[189].

There was an outstandingly uniform sensitivity of virtually all cases to the CDK2/7/9-inhibitor SNS-032, which had been tested in phase I for advanced chronic lymphocytic leukemia (CLL) and multiple myeloma (MM)[190,191]. These data are intriguing, since they indicate cell cycle dysregulation as a thus-far underappreciated central feature of T-PLL cells and implicate a novel vulnerability. As a potentially underlying genomic lesion, a deletion at 12p13 is found in approximately half of T-PLL patients and has been associated with haplo-insufficiency of the *CDKN1B* gene encoding the cyclin-dependent kinase inhibitory protein p27[190]. In our expression analysis, we also found that CDKN1A, a protein that usually inhibits the activity of CDK2 and CDK4 complexes, was almost 3-fold downregulated in T-PLL patients. Also, CDK4, CDK18, and CDK20 were upregulated 2-fold. It was therefore not surprising that the combined analysis of the microarray expression data with *ex vivo* drug response data from 12 patients revealed the cell cycle regulating the RB pathway to be upregulated and addicted in some of the patients. Furthermore, SNS-032 also inhibited the activity of NFAT, a central transcription factor and effector molecule of TCR signaling and T-cell activation. It has previously been demonstrated that T-PLL cells usually express functional TCRs and that TCL1A enhances TCR signaling

mediating to a hyper-responsive and proliferative phenotype in the context of TCR engagement[192].

Another group of compounds with a striking effect in T-PLL samples *ex vivo* were the p53 re-activator Prima-1 Met and MDM2 inhibitors serdemetan and nutlin-3. While *p53* mutations are frequently seen in many other cancers, in T-PLL these are not commonly found[193]. However, the overexpression and accumulation of wild type p53 is common in T-PLL, which could also be seen in our expression data. The efficacy of the p53 activators *ex vivo* in our T-PLL cohort indicates a new actionable pathway where these compounds act as an on-switch for the accumulated p53 and induce T-PLL tumor cell specific apoptosis.

In T-PLL, pathways such as JAK/STAT, PI3K/Akt/mTOR, and cell cycle signaling can all be dysregulated by different mutations and translocations. Therefore, it is not enough to simply abolish the effect of one of these pathways to hinder the progression of the disease. Combining different compounds allows for synergy or sensitivity that could not be achieved by the single compound alone. In the clinical setting, unless the compound acts on multiple pathways itself, it is essential to find the perfect combination rather than relying on a single compound.

Overall, the CDK-inhibitor SNS-032 and p53 activators stood out as drug classes with high efficacy across nearly all T-PLL samples and should be further tested in this disease. The acknowledgement of the fact that successful *ex vivo* results are not a guarantee for a therapy with the desired effect in the patient is important. However, compounds that do not function well on *ex vivo* cells are unlikely to perform well *in vivo.* As each round of ineffective treatment narrows the window of opportunity to help the patient, it can be just as vital to avoid unnecessary therapy regimens. The concept of individualized system medicine could improve patient outcomes and reduce health care costs by eliminating these rounds of expensive therapy that have little effect on cancer but can severely diminish quality of life.

Conclusions and future perspectives

This thesis provides additional molecular information about the pathogenesis of T-LGL leukemia. The discovery of additional *STAT3* mutations outside the mutation hotspot in the SH2-domain of *STAT3* further highlights the importance of screening the entire gene for *STAT3* mutations in LGL leukemia patients, as it is a viable diagnostic marker. In the rarer CD4+ cases, the high frequency of *STAT5B* mutations presents with a novel diagnostic criteria alongside other established clinical and hematological parameters. Additionally, in STAT-mutation negative patients, the STAT-activation seems to be mediated through heterogenous genetic events only visible through exome or whole genome sequencing. As sequencing costs decline even further, applying exome sequencing to every patient becomes plausible, which would further increase our knowledge about the genetic landscapes of many diseases. In T-LGLL, further studies are warranted to establish additional genetic events or pathways that lead to T-LGL leukemia pathogenesis. A future task will be to develop new molecules or therapeutic strategies targeting the frequently activated JAK/STAT pathway.

In T-PLL, *ex vivo* drug testing of primary patient cells has the potential to provide novel personalized drug candidates such as SNS-032 and Prima-1 Met. These observations hint at a centrally disturbed cell cycle regulation as well as a central role of TCR signaling in T-PLL. Despite high incidences of activating JAK/STAT mutations and indications on pathway activations, inhibitors of the JAK/STAT pathway gave variable responses. Overall, *ex vivo* drug response patterns in T-PLL did not closely correlate with known genetic aberrations, suggesting that screening for recurrent genetic biomarkers cannot easily be turned into effective therapeutic strategies in this problematic disease. This further emphasizes the importance of performing *ex vivo* drug screenings in combination with genomic analyses, as genomic data alone rarely lead to the right cancer drug for a patient. With drug screening, the phenotypic effect is already known, making the genetic analysis more focused on determining through which pathways a given drug is acting. As our understanding increases, we can look forward to more streamlined and effective drug discovery pipelines that use this information as a foundation.

Acknowledgements

The work for this doctoral thesis was carried out at the Hematology Research Unit Helsinki (HRUH) at the Department of Clinical Chemistry and Hematology, University of Helsinki and Helsinki University Hospital Comprehensive Cancer Center from 2011 to 2016. Financial support was provided by the Finnish Cultural Foundation, Swedish Cultural Foundation, Instrumentarium Science Foundation, Biomedicum Helsinki Foundation, Signe and Ane Gyllenberg Foundation, Blood Disease Foundation, Maud Kuistila Foundation, Medicinska Understödsföreningen Liv och Hälsa, and the Finnish Cancer Societies. I would like to thank the Doctoral School in Health Sciences for all the courses and support it has offered.

My deepest thanks go to my supervisor Professor Satu Mustjoki for all your support, encouragement and guidance through the years. I have the deepest respect for your knowledge of everything related to leukemia and your great collaborative skills, without which these projects would not have been possible. I would also like to thank Professor Kimmo Porkka for his input and advice over the years. I am grateful to the official reviewers of this thesis, Adjunct Professors Eeva-Riitta Savolainen and Samuel Myllykangas, for their comments and constructive criticism that greatly improved the thesis.

During my time at HRU I got to work with so many amazing and kind people. First I would like to thank Hanna Rajala for your support and guidance from the very first LGL project. It has been a pleasure to work with you and learn from you. I also had the privilege to share this PhD adventure with Mette Ilander, Paavo Pietarinen, and Anna Kreutzman. Thank you Mette for brightening up my days with your twisted humor and for all the fun times during conference trips! More importantly, I value our friendship and will sorely miss your company in the office. Paavo, you have been a great support especially during the last months of making this thesis. Anna, thank you for all the scientific and not-so-scientific discussions through the years. It has been a joy to follow in your footsteps.

I would also like to sincerely thank:

- Minna Pajuportti and Saara Vaalas for always being so helpful with the patients and everything else I could possibly ask for.
- Hanna Lähteenmäki and Tiina Kasanen for all your kindness and help in the lab.
- Heikki Kuusanmäki for always having time to help with the functional experiments and figuring out what went wrong.
- Mika Kontro for your medical insight.
- Tiina Kelkka and Giljun Park for interesting conversations and wise advice.
- Olli Dufva, Shady Adnan and Paula Savola for all your help in the lab and the laughs outside of it.

- my Italian summer students Sabrina Bortoluzzi and Vanessa Gasparini for working with me on the LGL projects and always being enthusiastic and ready to learn new things. Mille grazie!
- Samuli Eldfors, Bhagwan Yadav, Suleiman Khan and Liye He for your expertise in data analysis and for having time to explain what it all means.
- Sonja Lagström, Maija Lepistö, Aino Palva, Pirkko Mattila and Pekka Ellonen for all your expertise and guidance with the many sequencing projects.
- Laura Turunen and Jani Saarela from the HTB unit for your drug testing expertise.
- Krister Wennerberg, Tero Aittokallio, Jing Tang, Caroline Heckman and Olli Kallioniemi from FIMM for their collaboration and insightful advice on all the projects.
- all our collaborators from the US, Japan and all over Europe for providing invaluable patient samples.

I also want to thank my greatest supporters: my family. Mamma och Pappa, jag känner mej oerhört lycklig för allt ert stöd och kärlek i vad än jag hittar på. Varmt tack till Fredde, Erica, Johanna, Ulrika, Famo och Fafa för allt stöd och alla roliga tillställningar längs med åren. Jannica, innerligt tack för att du är en så god vän och alltid har tid att lyssna.

And thank you to my "Irish" girls Anni, Ella, and Sanna for all the laughs and believing in me through the years. Special thanks to Anni for designing the cover of the thesis, you are the best!

Finally I would like to thank my new family, Juha and Pörrö, for keeping me sane and happy :)

Helsinki, January 2017

Emma Andersson

References

1. Bailey NG, Elenitoba-Johnson KS. Mature T-cell leukemias: Molecular and Clinical Aspects. *Curr Hematol Malig Rep*. 2015;10(4):421-428.
2. Koskela HL, Eldfors S, Ellonen P, et al. Somatic STAT3 mutations in large granular lymphocytic leukemia. *N Engl J Med*. 2012;366(20):1905-1913.
3. Krishnan B, Else M, Tjonnfjord GE, et al. Stem cell transplantation after alemtuzumab in T-cell prolymphocytic leukaemia results in longer survival than after alemtuzumab alone: a multicentre retrospective study. *Br J Haematol*. 2010;149(6):907-910.
4. Kiel MJ, Velusamy T, Rolland D, et al. Integrated genomic sequencing reveals mutational landscape of T-cell prolymphocytic leukemia. *Blood*. 2014;124(9):1460-1472.
5. Zlotoff DA, Bhandoola A. Hematopoietic progenitor migration to the adult thymus. *Ann N Y Acad Sci*. 2011;1217:122-138.
6. Tough DF, Sprent J. Life span of naive and memory T cells. *Stem Cells*. 1995;13(3):242-249.
7. Kisielow P. Development and selection of T cells: how many subsets? How many rules? *Arch Immunol Ther Exp (Warsz)*. 2003;51(6):407-414.
8. Lauritsen JP, Haks MC, Lefebvre JM, Kappes DJ, Wiest DL. Recent insights into the signals that control alphabeta/gammadelta-lineage fate. *Immunol Rev*. 2006;209:176-190.
9. Schwartz RS. Shattuck lecture: Diversity of the immune repertoire and immunoregulation. *N Engl J Med*. 2003;348(11):1017-1026.
10. Livak F, Petrie HT, Crispe IN, Schatz DG. In-frame TCR delta gene rearrangements play a critical role in the alpha beta/gamma delta T cell lineage decision. *Immunity*. 1995;2(6):617-627.
11. Thedrez A, Sabourin C, Gertner J, et al. Self/non-self discrimination by human gammadelta T cells: simple solutions for a complex issue? *Immunol Rev*. 2007;215:123-135.
12. Fey MF. Normal and malignant hematopoiesis. *Ann Oncol*. 2007;18 Suppl 1:i9-i13.
13. Russell JH, Ley TJ. Lymphocyte-mediated cytotoxicity. *Annu Rev Immunol*. 2002;20:323-370.
14. Zhang J, Xu X, Liu Y. Activation-induced cell death in T cells and autoimmunity. *Cell Mol Immunol*. 2004;1(3):186-192.
15. Swerdlow SH, Campo E, Pileri SA, et al. The 2016 revision of the World Health Organization classification of lymphoid neoplasms. *Blood*. 2016;127(20):2375-2390.
16. Matutes E, Brito-Babapulle V, Swansbury J, et al. Clinical and laboratory features of 78 cases of T-prolymphocytic leukemia. *Blood*. 1991;78(12):3269-3274.
17. Maljaei SH, Brito-Babapulle V, Hiorns LR, Catovsky D. Abnormalities of chromosomes 8, 11, 14, and X in T-prolymphocytic leukemia studied by fluorescence in situ hybridization. *Cancer Genet Cytogenet*. 1998;103(2):110-116.
18. Nowak D, Le Toriellec E, Stern MH, et al. Molecular allelokaryotyping of T-cell prolymphocytic leukemia cells with high density single nucleotide polymorphism arrays identifies novel common genomic lesions and acquired uniparental disomy. *Haematologica*. 2009;94(4):518-527.

19. Rajala HL, Eldfors S, Kuusanmaki H, et al. Discovery of somatic STAT5b mutations in large granular lymphocytic leukemia. *Blood*. 2013;121(22):4541-4550.
20. Elliott NE, Cleveland SM, Grann V, Janik J, Waldmann TA, Dave UP. FERM domain mutations induce gain of function in JAK3 in adult T-cell leukemia/lymphoma. *Blood*. 2011;118(14):3911-3921.
21. Nakagawa M, Schmitz R, Xiao W, et al. Gain-of-function CCR4 mutations in adult T cell leukemia/lymphoma. *J Exp Med*. 2014;211(13):2497-2505.
22. 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.
23. Steininger A, Mobs M, Ullmann R, et al. Genomic loss of the putative tumor suppressor gene E2A in human lymphoma. *J Exp Med*. 2011;208(8):1585-1593.
24. Pandolfi F, Loughran TP, Jr., Starkebaum G, et al. Clinical course and prognosis of the lymphoproliferative disease of granular lymphocytes. A multicenter study. *Cancer*. 1990;65(2):341-348.
25. Dearden C. Large granular lymphocytic leukaemia pathogenesis and management. *Br J Haematol*. 2011;152(3):273-283.
26. Sokol L, Loughran TP, Jr. Large granular lymphocyte leukemia. *Curr Hematol Malig Rep*. 2007;2(4):278-282.
27. Semenzato G, Zambello R, Starkebaum G, Oshimi K, Loughran TP, Jr. The lymphoproliferative disease of granular lymphocytes: updated criteria for diagnosis. *Blood*. 1997;89(1):256-260.
28. Bourgault-Rouxel AS, Loughran TP, Jr., Zambello R, et al. Clinical spectrum of gammadelta+ T cell LGL leukemia: analysis of 20 cases. *Leuk Res*. 2008;32(1):45-48.
29. Oshimi K, Shinkai Y, Okumura K, Oshimi Y, Mizoguchi H. Perforin gene expression in granular lymphocyte proliferative disorders. *Blood*. 1990;75(3):704-708.
30. Mollet L, Fautrel B, Leblond V, et al. Leukemic CD3+ LGL share functional properties with their CD8+ CD57+ cell counterpart expanded after BMT. *Leukemia*. 1999;13(2):230-240.
31. Loughran TP, Jr., Hadlock KG, Perzova R, et al. Epitope mapping of HTLV envelope seroreactivity in LGL leukaemia. *Br J Haematol*. 1998;101(2):318-324.
32. Wlodarski MW, O'Keefe C, Howe EC, et al. Pathologic clonal cytotoxic T-cell responses: nonrandom nature of the T-cell-receptor restriction in large granular lymphocyte leukemia. *Blood*. 2005;106(8):2769-2780.
33. Lamy T, Liu JH, Landowski TH, Dalton WS, Loughran TP, Jr. Dysregulation of CD95/CD95 ligand-apoptotic pathway in CD3(+) large granular lymphocyte leukemia. *Blood*. 1998;92(12):4771-4777.
34. Schade AE, Wlodarski MW, Maciejewski JP. Pathophysiology defined by altered signal transduction pathways: the role of JAK-STAT and PI3K signaling in leukemic large granular lymphocytes. *Cell Cycle*. 2006;5(22):2571-2574.
35. Epling-Burnette PK, Liu JH, Catlett-Falcone R, et al. Inhibition of STAT3 signaling leads to apoptosis of leukemic large granular lymphocytes and decreased Mcl-1 expression. *J Clin Invest*. 2001;107(3):351-362.
36. Wlodarski MW, Schade AE, Maciejewski JP. T-large granular lymphocyte leukemia: current molecular concepts. *Hematology*. 2006;11(4):245-256.
37. Teramo A, Gattazzo C, Passeri F, et al. Intrinsic and extrinsic mechanisms contribute to maintain the JAK/STAT pathway aberrantly activated in T-type large granular lymphocyte leukemia. *Blood*. 2013;121(19):3843-3854, S3841.

38. Rajala HL, Olson T, Clemente MJ, et al. The analysis of clonal diversity and therapy responses using STAT3 mutations as a molecular marker in large granular lymphocytic leukemia. *Haematologica*. 2015;100(1):91-99.
39. Hamidou MA, Sadr FB, Lamy T, Raffi F, Grolleau JY, Barrier JH. Low-dose methotrexate for the treatment of patients with large granular lymphocyte leukemia associated with rheumatoid arthritis. *Am J Med*. 2000;108(9):730-732.
40. Gabor EP, Mishalani S, Lee S. Rapid response to cyclosporine therapy and sustained remission in large granular lymphocyte leukemia. *Blood*. 1996;87(3):1199-1200.
41. Lamy T, Loughran TP, Jr. How I treat LGL leukemia. *Blood*. 2011;117(10):2764-2774.
42. Bareau B, Rey J, Hamidou M, et al. Analysis of a French cohort of patients with large granular lymphocyte leukemia: a report on 229 cases. *Haematologica*. 2010;95(9):1534-1541.
43. Moignet A, Hasanali Z, Zambello R, et al. Cyclophosphamide as a first-line therapy in LGL leukemia. *Leukemia*. 2014;28(5):1134-1136.
44. Garand R, Goasguen J, Brizard A, et al. Indolent course as a relatively frequent presentation in T-prolymphocytic leukaemia. Groupe Francais d'Hematologie Cellulaire. *Br J Haematol*. 1998;103(2):488-494.
45. Dearden C. How I treat prolymphocytic leukemia. *Blood*. 2012;120(3):538-551.
46. Chen X, Cherian S. Immunophenotypic characterization of T-cell prolymphocytic leukemia. *Am J Clin Pathol*. 2013;140(5):727-735.
47. Stengel A, Kern W, Zenger M, et al. Genetic characterization of T-PLL reveals two major biologic subgroups and JAK3 mutations as prognostic marker. *Genes Chromosomes Cancer*. 2015.
48. Yokohama A, Saitoh A, Nakahashi H, et al. TCL1A gene involvement in T-cell prolymphocytic leukemia in Japanese patients. *Int J Hematol*. 2012;95(1):77-85.
49. Virgilio L, Lazzeri C, Bichi R, et al. Deregulated expression of TCL1 causes T cell leukemia in mice. *Proc Natl Acad Sci U S A*. 1998;95(7):3885-3889.
50. Gaudio E, Spizzo R, Paduano F, et al. Tcl1 interacts with Atm and enhances NF-kappaB activation in hematologic malignancies. *Blood*. 2012;119(1):180-187.
51. Bellanger D, Jacquemin V, Chopin M, et al. Recurrent JAK1 and JAK3 somatic mutations in T-cell prolymphocytic leukemia. *Leukemia*. 2014;28(2):417-419.
52. Bergmann AK, Schneppenheim S, Seifert M, et al. Recurrent mutation of JAK3 in T-cell prolymphocytic leukemia. *Genes Chromosomes Cancer*. 2014;53(4):309-316.
53. Gilleece MH, Dexter TM. Effect of Campath-1H antibody on human hematopoietic progenitors in vitro. *Blood*. 1993;82(3):807-812.
54. Hopfinger G, Busch R, Pflug N, et al. Sequential chemoimmunotherapy of fludarabine, mitoxantrone, and cyclophosphamide induction followed by alemtuzumab consolidation is effective in T-cell prolymphocytic leukemia. *Cancer*. 2013;119(12):2258-2267.
55. Herbaux C, Genet P, Bouabdallah K, et al. Bendamustine is effective in T-cell prolymphocytic leukaemia. *Br J Haematol*. 2015;168(6):916-919.
56. Bromberg JF, Wrzeszczynska MH, Devgan G, et al. Stat3 as an oncogene. *Cell*. 1999;98(3):295-303.
57. Pencik J, Pham HT, Schmoellerl J, et al. JAK-STAT signaling in cancer: From cytokines to non-coding genome. *Cytokine*. 2016.

58. Vainchenker W, Constantinescu SN. JAK/STAT signaling in hematological malignancies. *Oncogene*. 2013;32(21):2601-2613.
59. Migone TS, Lin JX, Cereseto A, et al. Constitutively activated Jak-STAT pathway in T cells transformed with HTLV-I. *Science*. 1995;269(5220):79-81.
60. Watowich SS, Wu H, Socolovsky M, Klingmuller U, Constantinescu SN, Lodish HF. Cytokine receptor signal transduction and the control of hematopoietic cell development. *Annu Rev Cell Dev Biol*. 1996;12:91-128.
61. Huang LJ, Constantinescu SN, Lodish HF. The N-terminal domain of Janus kinase 2 is required for Golgi processing and cell surface expression of erythropoietin receptor. *Mol Cell*. 2001;8(6):1327-1338.
62. Royer Y, Staerk J, Costuleanu M, Courtoy PJ, Constantinescu SN. Janus kinases affect thrombopoietin receptor cell surface localization and stability. *J Biol Chem*. 2005;280(29):27251-27261.
63. Syed RS, Reid SW, Li C, et al. Efficiency of signalling through cytokine receptors depends critically on receptor orientation. *Nature*. 1998;395(6701):511-516.
64. Constantinescu SN, Huang LJ, Nam H, Lodish HF. The erythropoietin receptor cytosolic juxtamembrane domain contains an essential, precisely oriented, hydrophobic motif. *Mol Cell*. 2001;7(2):377-385.
65. Thomas C, Moraga I, Levin D, et al. Structural linkage between ligand discrimination and receptor activation by type I interferons. *Cell*. 2011;146(4):621-632.
66. Ihle JN. The Janus protein tyrosine kinase family and its role in cytokine signaling. *Adv Immunol*. 1995;60:1-35.
67. Liao W, Lin JX, Leonard WJ. IL-2 family cytokines: new insights into the complex roles of IL-2 as a broad regulator of T helper cell differentiation. *Curr Opin Immunol*. 2011;23(5):598-604.
68. Ghoreschi K, Laurence A, O'Shea JJ. Janus kinases in immune cell signaling. *Immunol Rev*. 2009;228(1):273-287.
69. Demoulin JB, Van Roost E, Stevens M, Groner B, Renauld JC. Distinct roles for STAT1, STAT3, and STAT5 in differentiation gene induction and apoptosis inhibition by interleukin-9. *J Biol Chem*. 1999;274(36):25855-25861.
70. Zhang X, Guo A, Yu J, et al. Identification of STAT3 as a substrate of receptor protein tyrosine phosphatase T. *Proc Natl Acad Sci U S A*. 2007;104(10):4060-4064.
71. Alexander WS, Hilton DJ. The role of suppressors of cytokine signaling (SOCS) proteins in regulation of the immune response. *Annu Rev Immunol*. 2004;22:503-529.
72. Shuai K, Liu B. Regulation of gene-activation pathways by PIAS proteins in the immune system. *Nat Rev Immunol*. 2005;5(8):593-605.
73. Yoshimura A, Longmore G, Lodish HF. Point mutation in the exoplasmic domain of the erythropoietin receptor resulting in hormone-independent activation and tumorigenicity. *Nature*. 1990;348(6302):647-649.
74. Longmore G, Watowich S, Pharr P, Neumann D, Lodish H. Activation of the erythropoietin receptor and leukemia induction in mice. *Leukemia*. 1993;7 Suppl 2:S113-116.
75. Rodig SJ, Meraz MA, White JM, et al. Disruption of the Jak1 gene demonstrates obligatory and nonredundant roles of the Jaks in cytokine-induced biologic responses. *Cell*. 1998;93(3):373-383.

76. Witthuhn BA, Quelle FW, Silvennoinen O, et al. JAK2 associates with the erythropoietin receptor and is tyrosine phosphorylated and activated following stimulation with erythropoietin. *Cell*. 1993;74(2):227-236.
77. Buckley RH. Molecular defects in human severe combined immunodeficiency and approaches to immune reconstitution. *Annu Rev Immunol*. 2004;22:625-655.
78. Durbin JE, Hackenmiller R, Simon MC, Levy DE. Targeted disruption of the mouse Stat1 gene results in compromised innate immunity to viral disease. *Cell*. 1996;84(3):443-450.
79. Mathur AN, Chang HC, Zisoulis DG, et al. Stat3 and Stat4 direct development of IL-17-secreting Th cells. *J Immunol*. 2007;178(8):4901-4907.
80. Chen Z, Laurence A, Kanno Y, et al. Selective regulatory function of Socs3 in the formation of IL-17-secreting T cells. *Proc Natl Acad Sci U S A*. 2006;103(21):8137-8142.
81. Zhu J, Yamane H, Paul WE. Differentiation of effector CD4 T cell populations (*). *Annu Rev Immunol*. 2010;28:445-489.
82. Kaplan MH, Sun YL, Hoey T, Grusby MJ. Impaired IL-12 responses and enhanced development of Th2 cells in Stat4-deficient mice. *Nature*. 1996;382(6587):174-177.
83. Zhu J, Guo L, Watson CJ, Hu-Li J, Paul WE. Stat6 is necessary and sufficient for IL-4's role in Th2 differentiation and cell expansion. *J Immunol*. 2001;166(12):7276-7281.
84. Yao Z, Cui Y, Watford WT, et al. Stat5a/b are essential for normal lymphoid development and differentiation. *Proc Natl Acad Sci U S A*. 2006;103(4):1000-1005.
85. Kelly J, Spolski R, Imada K, Bollenbacher J, Lee S, Leonard WJ. A role for Stat5 in CD8+ T cell homeostasis. *J Immunol*. 2003;170(1):210-217.
86. Yu H, Jove R. The STATs of cancer--new molecular targets come of age. *Nat Rev Cancer*. 2004;4(2):97-105.
87. Hodge DR, Hurt EM, Farrar WL. The role of IL-6 and STAT3 in inflammation and cancer. *Eur J Cancer*. 2005;41(16):2502-2512.
88. Benekli M, Xia Z, Donohue KA, et al. Constitutive activity of signal transducer and activator of transcription 3 protein in acute myeloid leukemia blasts is associated with short disease-free survival. *Blood*. 2002;99(1):252-257.
89. Harris TJ, Grosso JF, Yen HR, et al. Cutting edge: An in vivo requirement for STAT3 signaling in TH17 development and TH17-dependent autoimmunity. *J Immunol*. 2007;179(7):4313-4317.
90. Holland SM, DeLeo FR, Elloumi HZ, et al. STAT3 mutations in the hyper-IgE syndrome. *N Engl J Med*. 2007;357(16):1608-1619.
91. Minegishi Y, Saito M, Tsuchiya S, et al. Dominant-negative mutations in the DNA-binding domain of STAT3 cause hyper-IgE syndrome. *Nature*. 2007;448(7157):1058-1062.
92. Flanagan SE, Haapaniemi E, Russell MA, et al. Activating germline mutations in STAT3 cause early-onset multi-organ autoimmune disease. *Nat Genet*. 2014;46(8):812-814.
93. Haapaniemi EM, Kaustio M, Rajala HL, et al. Autoimmunity, hypogammaglobulinemia, lymphoproliferation and mycobacterial disease in patients with dominant activating mutations in STAT3. *Blood*. 2014.
94. Milner JD, Vogel TP, Forbes L, et al. Early-onset lymphoproliferation and autoimmunity caused by germline STAT3 gain-of-function mutations. *Blood*. 2015;125(4):591-599.

95. Nicolae A, Xi L, Pittaluga S, et al. Frequent STAT5B mutations in gammadelta hepatosplenic T-cell lymphomas. *Leukemia*. 2014;28(11):2244-2248.
96. Bamford S, Dawson E, Forbes S, et al. The COSMIC (Catalogue of Somatic Mutations in Cancer) database and website. *Br J Cancer*. 2004;91(2):355-358.
97. Couronne L, Scourzic L, Pilati C, et al. STAT3 mutations identified in human hematologic neoplasms induce myeloid malignancies in a mouse bone marrow transplantation model. *Haematologica*. 2013;98(11):1748-1752.
98. Jerez A, Clemente MJ, Makishima H, et al. STAT3 mutations indicate the presence of subclinical T-cell clones in a subset of aplastic anemia and myelodysplastic syndrome patients. *Blood*. 2013;122(14):2453-2459.
99. Morin RD, Mendez-Lago M, Mungall AJ, et al. Frequent mutation of histone-modifying genes in non-Hodgkin lymphoma. *Nature*. 2011;476(7360):298-303.
100. Odejide O, Weigert O, Lane AA, et al. A targeted mutational landscape of angioimmunoblastic T-cell lymphoma. *Blood*. 2014;123(9):1293-1296.
101. Gesbert F, Griffin JD. Bcr/Abl activates transcription of the Bcl-X gene through STAT5. *Blood*. 2000;96(6):2269-2276.
102. Levine RL, Wadleigh M, Cools J, et al. Activating mutation in the tyrosine kinase JAK2 in polycythemia vera, essential thrombocythemia, and myeloid metaplasia with myelofibrosis. *Cancer Cell*. 2005;7(4):387-397.
103. Kontro M, Kuusanmaki H, Eldfors S, et al. Novel activating STAT5B mutations as putative drivers of T-cell acute lymphoblastic leukemia. *Leukemia*. 2014;28(8):1738-1742.
104. 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.
105. Buchert M, Burns CJ, Ernst M. Targeting JAK kinase in solid tumors: emerging opportunities and challenges. *Oncogene*. 2016;35(8):939-951.
106. Flanagan ME, Blumenkopf TA, Brissette WH, et al. Discovery of CP-690,550: a potent and selective Janus kinase (JAK) inhibitor for the treatment of autoimmune diseases and organ transplant rejection. *J Med Chem*. 2010;53(24):8468-8484.
107. Curtis JR, Lee EB, Kaplan IV, et al. Tofacitinib, an oral Janus kinase inhibitor: analysis of malignancies across the rheumatoid arthritis clinical development programme. *Ann Rheum Dis*. 2016;75(5):831-841.
108. Nelson EA, Walker SR, Weisberg E, et al. The STAT5 inhibitor pimozide decreases survival of chronic myelogenous leukemia cells resistant to kinase inhibitors. *Blood*. 2011;117(12):3421-3429.
109. Furqan M, Akinleye A, Mukhi N, Mittal V, Chen Y, Liu D. STAT inhibitors for cancer therapy. *J Hematol Oncol*. 2013;6:90.
110. Hexner EO, Serdikoff C, Jan M, et al. Lestaurtinib (CEP701) is a JAK2 inhibitor that suppresses JAK2/STAT5 signaling and the proliferation of primary erythroid cells from patients with myeloproliferative disorders. *Blood*. 2008;111(12):5663-5671.
111. Hong D, Kurzrock R, Kim Y, et al. AZD9150, a next-generation antisense oligonucleotide inhibitor of STAT3 with early evidence of clinical activity in lymphoma and lung cancer. *Sci Transl Med*. 2015;7(314):314ra185.
112. Page BD, Khoury H, Laister RC, et al. Small molecule STAT5-SH2 domain inhibitors exhibit potent antileukemia activity. *J Med Chem*. 2012;55(3):1047-1055.
113. Tyner JW. Functional genomics for personalized cancer therapy. *Sci Transl Med*. 2014;6(243):243fs226.

114. International Human Genome Sequencing C. Finishing the euchromatic sequence of the human genome. *Nature*. 2004;431(7011):931-945.
115. Metzker ML. Sequencing technologies - the next generation. *Nat Rev Genet*. 2010;11(1):31-46.
116. Ng SB, Turner EH, Robertson PD, et al. Targeted capture and massively parallel sequencing of 12 human exomes. *Nature*. 2009;461(7261):272-276.
117. Yan XJ, Xu J, Gu ZH, et al. Exome sequencing identifies somatic mutations of DNA methyltransferase gene DNMT3A in acute monocytic leukemia. *Nat Genet*. 2011;43(4):309-315.
118. Grossmann V, Tiacci E, Holmes AB, et al. Whole-exome sequencing identifies somatic mutations of BCOR in acute myeloid leukemia with normal karyotype. *Blood*. 2011;118(23):6153-6163.
119. Choi M, Scholl UI, Ji W, et al. Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. *Proc Natl Acad Sci U S A*. 2009;106(45):19096-19101.
120. Albert TJ, Molla MN, Muzny DM, et al. Direct selection of human genomic loci by microarray hybridization. *Nat Methods*. 2007;4(11):903-905.
121. Teer JK, Mullikin JC. Exome sequencing: the sweet spot before whole genomes. *Hum Mol Genet*. 2010;19(R2):R145-151.
122. Pruitt KD, Harrow J, Harte RA, et al. The consensus coding sequence (CCDS) project: Identifying a common protein-coding gene set for the human and mouse genomes. *Genome Res*. 2009;19(7):1316-1323.
123. Meyerson M, Gabriel S, Getz G. Advances in understanding cancer genomes through second-generation sequencing. *Nat Rev Genet*. 2010;11(10):685-696.
124. Fedurco M, Romieu A, Williams S, Lawrence I, Turcatti G. BTA, a novel reagent for DNA attachment on glass and efficient generation of solid-phase amplified DNA colonies. *Nucleic Acids Res*. 2006;34(3):e22.
125. Nielsen R, Paul JS, Albrechtsen A, Song YS. Genotype and SNP calling from next-generation sequencing data. *Nat Rev Genet*. 2011;12(6):443-451.
126. Rehm HL, Bale SJ, Bayrak-Toydemir P, et al. ACMG clinical laboratory standards for next-generation sequencing. *Genet Med*. 2013;15(9):733-747.
127. Taub MA, Corrada Bravo H, Irizarry RA. Overcoming bias and systematic errors in next generation sequencing data. *Genome Med*. 2010;2(12):87.
128. Schirmer M, Ijaz UZ, D'Amore R, Hall N, Sloan WT, Quince C. Insight into biases and sequencing errors for amplicon sequencing with the Illumina MiSeq platform. *Nucleic Acids Res*. 2015;43(6):e37.
129. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009;25(14):1754-1760.
130. Koboldt DC, Chen K, Wylie T, et al. VarScan: variant detection in massively parallel sequencing of individual and pooled samples. *Bioinformatics*. 2009;25(17):2283-2285.
131. Koboldt DC, Zhang Q, Larson DE, et al. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res*. 2012;22(3):568-576.
132. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res*. 2010;38(16):e164.
133. Forbes SA, Bhamra G, Bamford S, et al. The Catalogue of Somatic Mutations in Cancer (COSMIC). *Curr Protoc Hum Genet*. 2008;Chapter 10:Unit 10 11.

134. Tyner JW, Yang WF, Bankhead A, 3rd, et al. Kinase pathway dependence in primary human leukemias determined by rapid inhibitor screening. *Cancer Res*. 2013;73(1):285-296.
135. Pemovska T, Johnson E, Kontro M, et al. Axitinib effectively inhibits BCR-ABL1(T315I) with a distinct binding conformation. *Nature*. 2015;519(7541):102-105.
136. Pemovska T, Kontro M, Yadav B, et al. Individualized systems medicine strategy to tailor treatments for patients with chemorefractory acute myeloid leukemia. *Cancer Discov*. 2013;3(12):1416-1429.
137. Lippert TH, Ruoff HJ, Volm M. Current status of methods to assess cancer drug resistance. *Int J Med Sci*. 2011;8(3):245-253.
138. Kern DH, Weisenthal LM. Highly specific prediction of antineoplastic drug resistance with an in vitro assay using suprapharmacologic drug exposures. *J Natl Cancer Inst*. 1990;82(7):582-588.
139. Andreotti PE, Cree IA, Kurbacher CM, et al. Chemosensitivity testing of human tumors using a microplate adenosine triphosphate luminescence assay: clinical correlation for cisplatin resistance of ovarian carcinoma. *Cancer Res*. 1995;55(22):5276-5282.
140. Chan GK, Kleinheinz TL, Peterson D, Moffat JG. A simple high-content cell cycle assay reveals frequent discrepancies between cell number and ATP and MTS proliferation assays. *PLoS One*. 2013;8(5):e63583.
141. Haibe-Kains B, El-Hachem N, Birkbak NJ, et al. Inconsistency in large pharmacogenomic studies. *Nature*. 2013;504(7480):389-393.
142. Jang IS, Neto EC, Guinney J, Friend SH, Margolin AA. Systematic assessment of analytical methods for drug sensitivity prediction from cancer cell line data. *Pac Symp Biocomput*. 2014:63-74.
143. Hatzis C, Bedard PL, Birkbak NJ, et al. Enhancing reproducibility in cancer drug screening: how do we move forward? *Cancer Res*. 2014;74(15):4016-4023.
144. Yadav B, Gopalacharyulu P, Pemovska T, et al. From drug response profiling to target addiction scoring in cancer cell models. *Dis Model Mech*. 2015;8(10):1255-1264.
145. Trapnell C, Pachter L, Salzberg SL. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics*. 2009;25(9):1105-1111.
146. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol*. 2013;14(4):R36.
147. Anders S, Pyl PT, Huber W. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics*. 2015;31(2):166-169.
148. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15(12):550.
149. Kallio MA, Tuimala JT, Hupponen T, et al. Chipster: user-friendly analysis software for microarray and other high-throughput data. *BMC Genomics*. 2011;12:507.
150. Yadav B, Pemovska T, Szwajda A, et al. Quantitative scoring of differential drug sensitivity for individually optimized anticancer therapies. *Sci Rep*. 2014;4:5193.
151. Tang J, Szwajda A, Shakyawar S, et al. Making sense of large-scale kinase inhibitor bioactivity data sets: a comparative and integrative analysis. *J Chem Inf Model*. 2014;54(3):735-743.
152. de Hoon MJ, Imoto S, Nolan J, Miyano S. Open source clustering software. *Bioinformatics*. 2004;20(9):1453-1454.

153. Saldanha AJ. Java Treeview--extensible visualization of microarray data. *Bioinformatics*. 2004;20(17):3246-3248.
154. Huang X, Du X, Li Y. The role of BCL11B in hematological malignancy. *Exp Hematol Oncol*. 2012;1(1):22.
155. Liu P, Li P, Burke S. Critical roles of Bcl11b in T-cell development and maintenance of T-cell identity. *Immunol Rev*. 2010;238(1):138-149.
156. Seitan VC, Hao B, Tachibana-Konwalski K, et al. A role for cohesin in T-cell-receptor rearrangement and thymocyte differentiation. *Nature*. 2011;476(7361):467-471.
157. Wu JY, Feng L, Park HT, et al. The neuronal repellent Slit inhibits leukocyte chemotaxis induced by chemotactic factors. *Nature*. 2001;410(6831):948-952.
158. Tordjman R, Lepelletier Y, Lemarchandel V, et al. A neuronal receptor, neuropilin-1, is essential for the initiation of the primary immune response. *Nat Immunol*. 2002;3(5):477-482.
159. Schreiner SJ, Schiavone AP, Smithgall TE. Activation of STAT3 by the Src family kinase Hck requires a functional SH3 domain. *J Biol Chem*. 2002;277(47):45680-45687.
160. Hu Y, Liu Y, Pelletier S, et al. Requirement of Src kinases Lyn, Hck and Fgr for BCR-ABL1-induced B-lymphoblastic leukemia but not chronic myeloid leukemia. *Nat Genet*. 2004;36(5):453-461.
161. Chatterjee M, Kis-Toth K, Thai TH, Terhorst C, Tsokos GC. SLAMF6-driven co-stimulation of human peripheral T cells is defective in SLE T cells. *Autoimmunity*. 2011;44(3):211-218.
162. Lima M, Almeida J, Dos Anjos Teixeira M, et al. TCRalphabeta+/CD4+ large granular lymphocytosis: a new clonal T-cell lymphoproliferative disorder. *Am J Pathol*. 2003;163(2):763-771.
163. Rodriguez-Caballero A, Garcia-Montero AC, Barcena P, et al. Expanded cells in monoclonal TCR-alphabeta+/CD4+/NKa+/CD8-/+dim T-LGL lymphocytosis recognize hCMV antigens. *Blood*. 2008;112(12):4609-4616.
164. Olteanu H, Karandikar NJ, Eshoa C, Kroft SH. Laboratory findings in CD4(+) large granular lymphocytoses. *Int J Lab Hematol*. 2010;32(1 Pt 1):e9-16.
165. Wang Z, Shen D, Parsons DW, et al. Mutational analysis of the tyrosine phosphatome in colorectal cancers. *Science*. 2004;304(5674):1164-1166.
166. Peyser ND, Freilino M, Wang L, et al. Frequent promoter hypermethylation of PTPRT increases STAT3 activation and sensitivity to STAT3 inhibition in head and neck cancer. *Oncogene*. 2016;35(9):1163-1169.
167. Pasquo A, Consalvi V, Knapp S, et al. Structural stability of human protein tyrosine phosphatase rho catalytic domain: effect of point mutations. *PLoS One*. 2012;7(2):e32555.
168. Li L, Leid M, Rothenberg EV. An early T cell lineage commitment checkpoint dependent on the transcription factor Bcl11b. *Science*. 2010;329(5987):89-93.
169. Ikawa T, Hirose S, Masuda K, et al. An essential developmental checkpoint for production of the T cell lineage. *Science*. 2010;329(5987):93-96.
170. Oshiro A, Tagawa H, Ohshima K, et al. Identification of subtype-specific genomic alterations in aggressive adult T-cell leukemia/lymphoma. *Blood*. 2006;107(11):4500-4507.
171. Grabarczyk P, Przybylski GK, Depke M, et al. Inhibition of BCL11B expression leads to apoptosis of malignant but not normal mature T cells. *Oncogene*. 2007;26(26):3797-3810.

172. Gutierrez A, Kentsis A, Sanda T, et al. The BCL11B tumor suppressor is mutated across the major molecular subtypes of T-cell acute lymphoblastic leukemia. *Blood*. 2011;118(15):4169-4173.
173. De Keersmaecker K, Real PJ, Gatta GD, et al. The TLX1 oncogene drives aneuploidy in T cell transformation. *Nat Med*. 2010;16(11):1321-1327.
174. Becker S, Groner B, Muller CW. Three-dimensional structure of the Stat3beta homodimer bound to DNA. *Nature*. 1998;394(6689):145-151.
175. Ginter T, Fahrer J, Krohnert U, et al. Arginine residues within the DNA binding domain of STAT3 promote intracellular shuttling and phosphorylation of STAT3. *Cell Signal*. 2014;26(8):1698-1706.
176. He J, Shi J, Xu X, et al. STAT3 mutations correlated with hyper-IgE syndrome lead to blockage of IL-6/STAT3 signalling pathway. *J Biosci*. 2012;37(2):243-257.
177. Huang W, Dong Z, Wang F, Peng H, Liu JY, Zhang JT. A small molecule compound targeting STAT3 DNA-binding domain inhibits cancer cell proliferation, migration, and invasion. *ACS Chem Biol*. 2014;9(5):1188-1196.
178. Zhang T, Kee WH, Seow KT, Fung W, Cao X. The coiled-coil domain of Stat3 is essential for its SH2 domain-mediated receptor binding and subsequent activation induced by epidermal growth factor and interleukin-6. *Mol Cell Biol*. 2000;20(19):7132-7139.
179. Pilati C, Amessou M, Bihl MP, et al. Somatic mutations activating STAT3 in human inflammatory hepatocellular adenomas. *J Exp Med*. 2011;208(7):1359-1366.
180. Hu G, Witzig TE, Gupta M. A novel missense (M206K) STAT3 mutation in diffuse large B cell lymphoma deregulates STAT3 signaling. *PLoS One*. 2013;8(7):e67851.
181. Kanai T, Seki S, Jenks JA, et al. Identification of STAT5A and STAT5B target genes in human T cells. *PLoS One*. 2014;9(1):e86790.
182. Posnett DN, Sinha R, Kabak S, Russo C. Clonal populations of T cells in normal elderly humans: the T cell equivalent to "benign monoclonal gammapathy". *J Exp Med*. 1994;179(2):609-618.
183. Park C, Qian W, Zhang J. Genomic evidence for elevated mutation rates in highly expressed genes. *EMBO Rep*. 2012;13(12):1123-1129.
184. Wei L, Laurence A, O'Shea JJ. New insights into the roles of Stat5a/b and Stat3 in T cell development and differentiation. *Semin Cell Dev Biol*. 2008;19(4):394-400.
185. Zhang R, Shah MV, Yang J, et al. Network model of survival signaling in large granular lymphocyte leukemia. *Proc Natl Acad Sci U S A*. 2008;105(42):16308-16313.
186. Stengel A, Kern W, Zenger M, et al. Genetic characterization of T-PLL reveals two major biologic subgroups and JAK3 mutations as prognostic marker. *Genes Chromosomes Cancer*. 2016;55(1):82-94.
187. Kucuk C, Jiang B, Hu X, et al. Activating mutations of STAT5B and STAT3 in lymphomas derived from gammadelta-T or NK cells. *Nat Commun*. 2015;6:6025.
188. Hasanali ZS, Saroya BS, Stuart A, et al. Epigenetic therapy overcomes treatment resistance in T cell prolymphocytic leukemia. *Sci Transl Med*. 2015;7(293):293ra102.
189. Lopez C, Bergmann AK, Paul U, et al. Genes encoding members of the JAK-STAT pathway or epigenetic regulators are recurrently mutated in T-cell prolymphocytic leukaemia. *Br J Haematol*. 2016;173(2):265-273.

190. Le Toriellec E, Despouy G, Pierron G, et al. Haploinsufficiency of CDKN1B contributes to leukemogenesis in T-cell prolymphocytic leukemia. *Blood*. 2008;111(4):2321-2328.
191. Tong WG, Chen R, Plunkett W, et al. Phase I and pharmacologic study of SNS-032, a potent and selective Cdk2, 7, and 9 inhibitor, in patients with advanced chronic lymphocytic leukemia and multiple myeloma. *J Clin Oncol*. 2010;28(18):3015-3022.
192. Herling M, Patel KA, Teitell MA, et al. High TCL1 expression and intact T-cell receptor signaling define a hyperproliferative subset of T-cell prolymphocytic leukemia. *Blood*. 2008;111(1):328-337.
193. Brito-Babapulle V, Hamoudi R, Matutes E, et al. p53 allele deletion and protein accumulation occurs in the absence of p53 gene mutation in T-prolymphocytic leukaemia and Sezary syndrome. *Br J Haematol*. 2000;110(1):180-187.