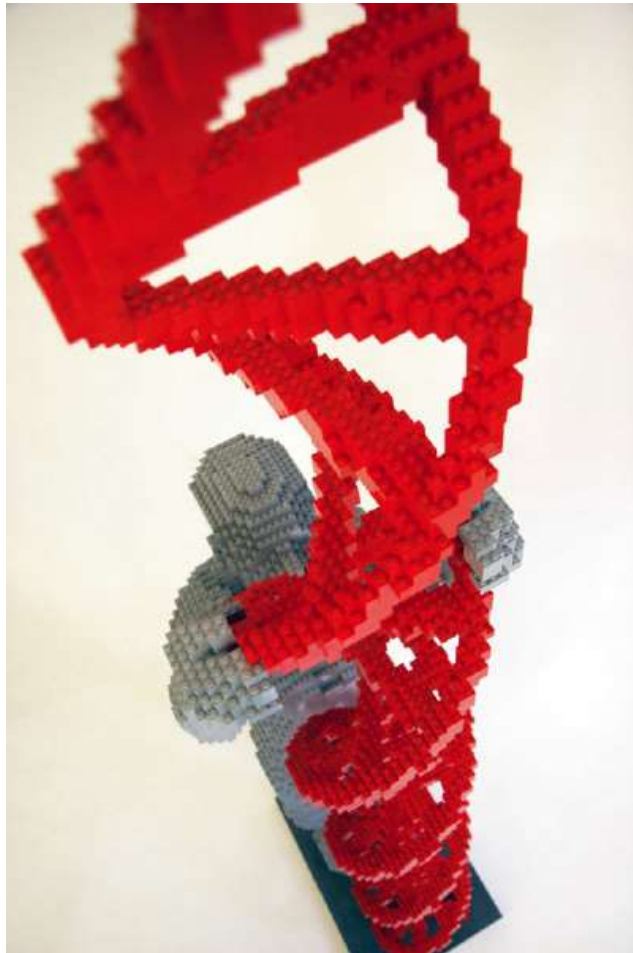


Unravelling Pathobiological Molecular Mechanisms of T-Cell Acute Lymphoblastic Leukemia



Rui D. Mendes

Unravelling Pathobiological Molecular Mechanisms of T-Cell Acute Lymphoblastic Leukemia

Rui D. Mendes

The research here described was conducted in the department of Pediatric Oncology/ Hematology at Erasmus Medical Center - Sophia Children's Hospital, Rotterdam, the Netherlands. The studies described in this thesis were financially supported by Stichting Kinderen Kankervrij (KiKa).

Publication of this thesis was financially supported by:

Cover: Nathan Sawaya. The original picture displays a two metre sculpture named 'Building Bricks of Life' made by the artist Nathan Sawaya.

"Just as each brick plays an important part in holding the sculpture together, each gene within the human genome helps us understand more about ourselves as humans and how we are all connected" - Nathan Sawaya

Book design: Rui D. Mendes

Printed by:

ISBN: XXX-XX-XXXX-XXX-X

Copyright © 2016 by Rui D. Mendes. All rights reserved. No part of this book may be reproduced, stored in a retrieval system, or transmitted in any form or by any means without prior permission of the author.

Unravelling Pathobiological Molecular Mechanisms of T-Cell Acute Lymphoblastic Leukemia

Het ontrafelen van pathobiologische moleculaire mechanismes in T-cel
acute lymphoblastische leukemie

**Thesis
to obtain the degree of doctor from**

Erasmus University Rotterdam
by command of the
rector magnificus

Prof.dr. H.A.P. Pols

and in accordance with the decision of the Doctorate Board.

The public defence shall be held on
Tuesday December 13, 2016 at 13:30 hours
by

Rui Daniel Martins Nunes Mendes
born in Lisbon, Portugal

Erasmus University Rotterdam

The logo of Erasmus University, featuring a stylized, handwritten-style script of the word "Erasmus" in a dark color.

Doctoral Committee:

Promotor: Prof.dr. R. Pieters

Leescommissie:

Prof.dr. C.M. Zwaan

Prof.dr. F.J. Staal

Prof.dr. F.G. Grosveld

Copromotor: Dr. J.P.P. Meijerink

To my beloved ones:

Mariana, Joaquim e André,

and in memory of my cousin João Martins.

Contents

| | | |
|------------------|---|-----|
| Chapter 1 | General Introduction & Aims of the Thesis | 9 |
| Chapter 2 | The relevance of PTEN-AKT in relation to NOTCH1-directed treatment strategies in T-cell acute lymphoblastic leukemia <i>Haematologica. 2016 Sep;101(9):1010-7</i> | 30 |
| Chapter 3 | PTEN microdeletions in T-cell acute lymphoblastic leukemia are caused by illegitimate RAG-mediated recombination events. <i>Blood, 2014, 124(4):567-78</i> | 49 |
| Chapter 4 | Lentiviral gene transfer into human and murine hematopoietic stem cells: size matters <i>BMC Res Notes. 2016 Jun 16;9:312</i> | 77 |
| Chapter 5 | Downregulation of CD44 functionally defines human T-cell lineage commitment <i>Manuscript submitted</i> | 88 |
| Chapter 6 | T-Cell Acute Lymphoblastic Leukemia Oncogenes hijack T-Cell Developmental Programs for Leukemogenesis: the arrest of Early T-cell programs by MEF2C, LYL1 or LMO2 <i>Manuscript in preparation</i> | 113 |
| Chapter 7 | General discussion & Future perspectives | 135 |
| Chapter 8 | | |
| | Summary/ Samenvatting | 145 |
| | Supplementary data | 151 |
| | Curriculum Vitae | 181 |
| | List of Publications | 182 |
| | Portfolio | 183 |
| | Acknowledgements | 184 |

Chapter 1

General Introduction & Aims of the Thesis

1. General introduction

1.1. Hematopoiesis and leukemia

Daily, each person produces approximately 10^{11} – 10^{12} blood cells to maintain the number of cells in the peripheral circulation.[1] The continuous production of mature blood cells is essential as these cells have a limited lifespan. This renewal process is supported by hematopoietic stem cells (HSC) that are located in the medulla of the bone (bone marrow), especially in the pelvis, femur, and sternum. Although in smaller numbers, they can also be found in umbilical cord blood (UCB) and in peripheral blood. HSCs give rise to all types of mature blood cells (Figure 1) and have the capacity of self-renewal, since upon cellular division some of the daughter cells remain as HSCs.[2] Instead, the other daughters of HSCs differentiate into myeloid and lymphoid progenitor cells that further give rise to the development of all blood cell types. These changes can often be tracked by monitoring the presence of proteins on the surface of the cell, usually determined as Cluster of Differentiation (CD) markers. HSCs express CD34 and lack markers that are lineage-specific, such as the expression of CD3 in T-cells, CD19 in B-cells, CD33 in myeloid and CD56 in NK-lineage cells. The process of maturation from HSC to effector blood cells occurs in the bone marrow and/or secondary lymphoid organs, including thymus and spleen. Following specific stimuli, such as growth factors, chemokines and contact with cells present in different microenvironments, the hematopoietic precursor cell undergoes changes in chromatin organization and gene expression that promote the differentiation of the cell towards a specific cell type, limiting the potential for differentiation into alternative lineages. [2]

However, during the process of maturation, the blood precursor cells may acquire mutations, chromosomal rearrangements and epigenetic changes that prompt the cells to uncontrollable proliferation called leukemia.[3] As leukemic cells continue to grow and divide, they eventually outcompete the normal blood cells, which results in defective protection against infections and pathogens, and hemorrhages.[4] Clinically and pathologically, leukemias can be classified in four main categories. The first division is between acute and chronic leukemias. Acute leukemias are the most common forms of leukemia in children, which are characterized by a rapid increase in the number of leukemic cells.[5] Immediate treatment is required due to the rapid progression and accumulation of malignant cells that are released into the bloodstream and spread to other organs. Other symptoms can be manifested according to the tissue occupied; infiltration of the lymph nodes causes lymphadenopathy and infiltration of the central nervous system may provoke headaches. Chronic leukemia is characterized by the excessive

accumulation of relatively more mature leukemic cells, and typically it takes months or years to progress. Chronic leukemia mostly occurs in adults. In addition, leukemias can be classified into myeloid or lymphoblastic leukemia, according to the blood lineage of the affected cell. Furthermore, lymphoblastic leukemias can be further divided into B-cell or T-cell lymphoblastic leukemia.

1.2. Acute lymphoblastic leukemia

Acute lymphoblastic leukemia (ALL) is the most common type of cancer in children, comprising approximately 25% of all childhood malignancies.[6] Almost three-fourths of childhood leukemia (ages 0-18) is ALL, while the most common form of acute leukemia in adults is acute myeloid leukemia (AML). In the majority of the cases the leukemic blasts are detected on a blood smear. Blood and bone marrow are used for further classification by immuno-phenotyping and genotyping. A lumbar puncture is performed to detect central nervous system involvement. Immunophenotypic analysis of the blasts indicates whether the leukemia is from myeloid (neutrophils, eosinophils, or basophils) or lymphoblastic origin (B- or T-lymphocytes). Genomic analysis is used to determine the underlying genetic abnormalities, which are associated with specific risks for relapse and survival.[6] Treatment protocols are based on combination chemotherapy, including different classes of drugs. Survival rates of ALL patients improved from less than 10% before 1970 to over 80%.[7]

1.3. T-cell acute lymphoblastic leukemia

T-cell acute lymphoblastic leukemia (T-ALL) results from the malignant transformation of lymphoid precursor cells residing in the thymus that are primed towards T-cell development. T-ALL represents about 15% of pediatric ALL cases and 25% of adult ALL cases,[8] and is typically more frequent in males than females. Clinically, T-ALL patients show diffuse infiltration of the bone marrow by immature T-cell blasts, high white blood cells counts, mediastinal masses with pleural effusions, and frequent infiltration of the central nervous system at diagnosis.[9] Despite the advances in our understanding of the molecular genetics of T-ALL, current treatments are mostly based on multi-agent chemotherapy and stem cell transplantation in part of the patients. Although these treatment regimens are highly effective with overall survival reaching 85% on current treatment protocols, they are often associated with severe toxicity and

long-term side effects such as cardiomyopathy, bone necrosis, chronic graft versus host disease, infertility or secondary malignancies, impairing the quality of adult life. In addition, about 15% of pediatric T-ALL patients still relapse due to therapy resistance and those cases are associated with very dismal survival perspectives.[5, 10] Therefore, current research efforts are focused on the search for targets that will eventually lead to more effective and less toxic antileukemic drugs. In order to achieve more specific or individualized therapies it is crucial to improve our understanding of the molecular events that lead to the disease, and the mechanisms involved in disease progression and relapse.

1.3.1. T-cell development in the thymus

T-cells can be distinguished between $\alpha\beta$ - or $\gamma\delta$ -cells based on the composition of the T-cell antigen receptor (TCR). While $\alpha\beta$ T-cells comprise the majority (95%) of T-cell populations in lymphoid organs, $\gamma\delta$ T-cells only represent a small fraction of T-cells that are mainly located in the gut mucosa. Normal T-cell development is a strictly regulated, multistep process in which lymphoid precursor cells differentiate into functionally diverse T-lymphocyte subsets in the thymus microenvironment. The different checkpoints are orchestrated by diverse transcriptional regulatory networks[11] and transitions between epigenetic states[12] that are triggered through the activation of membrane receptors by signalling molecules present in thymus, such as cytokines, chemokines and ligands from stromal cells. During this fine-tuned developmental process, inappropriate activation of T-ALL oncogenes and loss of tumor suppressor genes in thymic precursors may provoke uncontrolled clonal expansion and T-ALL. To better understand the disease, it is therefore essential to understand the normal T-cell development in the thymus and to distinguish eventual functions of proto-oncogenes in the transcriptional networks that regulate normal developmental progression.

Currently, our understanding of normal T-cell development in the thymus is mostly derived from mouse model studies. Based on these studies,[11] Yui and Rothenberg suggest that T-cell development can be divided in three major regulatory phases, which are divided by the T-cell commitment and the β -selection checkpoint. Those phases are named pre-commitment (phase-1), T-cell identity (phase-2) and post β -selection (phase-3), and are characterized by unique gene networks and cellular features (Figure 2). Each phase is characterized by the expression of certain transcription factors that ensure key developmental and consecutive processes, namely the recruitment of progenitor cells to the thymus, their proliferation, the commitment to T-cell lineage, the response to T-cell receptor signals, and

finally the restrain to a particular effector programme.[11] In addition, these developmental stages can be divided according to their dependency to NOTCH- and TCR- signalling programs. For instance, in mice cells in phase-1 and -2 do not express a functional TCR and depend on Notch-signalling to proliferate and differentiate, while cells in phase-3 require the expression of a functional TCR rather than NOTCH signalling (Figure 2A).[11]

Although the similarity in the core T-cell transcriptional networks that are activated in humans and mouse, there are some species-specific differences in kinetics and function. For instance, there are clear differences in NOTCH activation status between human and mice T-cell development. The highest level of NOTCH activation in humans occurs at T-cell commitment and decreases thereafter, while in mice there is an increase of NOTCH signalling until the cells reach the β -selection (Figure 2A-B). In line with this, while uncommitted early T-cell precursor cells (ETPs) in humans differentiate preferentially into $\alpha\beta$ - lineage DP thymocytes after T-cell commitment [13, 14], in mice it only occurs after β -selection stage[15].

The expression of characteristic cell surface markers during T-cell development allows to distinguish the different phases. Thymocytes can be subdivided into double-negative (DN), double-positive, or single-positive (SP) based on the expression of both T-cell markers CD4 and CD8. In mice, the DN stage can be divided into four distinct subsets based on the expression of CD44 and CD25 (Figure 2A). Although the DN stage in humans has not been clearly defined, it is characterized by consecutive acquisition of CD7, CD5, and CD1a (Figure 2B). [16, 17] In addition, the immature single-positive stage that precedes the DP stage is characterized by the expression of CD8 in mice and CD4 in humans. [18, 19]

1.3.1.1. Programs in pre-committed, early T-cell progenitor cells

In phase-1, human thymus-seeding progenitors are DN, express CD34 and CD7 and do not express CD1a. The mouse counterpart is the DN1 stage, when cells are CD44⁺CD25⁻. The T-cell specific developmental program is activated when NOTCH-1 receptors on the surface of the lymphoid precursors interact with Notch ligands in the thymic microenvironment.

During phase 1 until beginning of phase 2, Notch signalling interacts with a stem cell and/or progenitor cell gene network that is inherited from multipotent precursors, resulting in the downregulation of phase 1- restricted genes and the activation of Notch target genes. The most restricted initial subset of phase 1 regulators (*GATA2*, *MEIS1* and *HOXA9*) is predominantly expressed in the early T-cell precursor cells (ETPs) and is possibly involved in supporting

engraftment in the thymus. Compared to ETP genes, expression of genes such as *MEF2C*, *LMO2*, *SPI1/PU.1*, *HHEX*, *BCL11A*, *GFI1B*, *ERG* and *LYL1* are expressed throughout the whole pre-commitment stage (phase-1) and persist until the transition to the commitment phase (phase-2). *LMO2*, *MEIS1*, and *NMYC* are associated with stem-cell like features in early T-cells,[20, 21] similar to *ERG* and *HHEX* that are described to contribute for self-renewal of normal and malignant cells.[21, 22] *LYL1* is associated with the expression of the growth factor receptor *KIT*, and the regulatory factor *GFI1* that is essential for the generation of ETPs and for the regulation of Notch signalling.[23, 24] Therefore, albeit transient, the expression of those transcription factors is essential for normal T-cell development.

The major landmark in the developmental process, which occurs at the transition between phase-1 and -2, is known as T-cell lineage commitment. According to literature the commitment occurs in humans when cells acquire CD1 (Figure 2B), and in mice when *KIT* and/or *CD44* expression is downregulated between DN2a and DN2b stages (Figure 2A). Although expression of *CD1a* has been used to define human T-lineage commitment,[25] the exact T-cell commitment point has not been clearly defined.

Upon commitment, developing cells adopt an intrinsically irreversible T-cell fate that is caused by the “shut down” of transcription factors and gene regulatory networks, impeding cells to differentiate into alternative lineages even under permissive conditions. (reviewed in [26]) This process is initiated by existing ligands in the thymus that support differentiation of progenitor cells in the T-cell lineage. The activation of NOTCH signalling results in upregulation of many T-cell genes and the loss of progenitor-specific gene expression, and therefore cells cannot differentiate into alternative lineages, such as dendritic cell, myeloid cell, NK cell, innate lymphoid cell or mast cell.

1.3.1.2. Programs that promote T-cell commitment

The gene network exhibited in phase-2 is initiated during phase-1, as the activation of NOTCH signalling in ETPs prompts the expression of crucial regulatory transcription factors, such as *TCF1* and *GATA3*. [27, 28] Besides their role in inhibiting phase-1 progenitor specific factors, *TCF1* and *GATA3* also collaborate with regulators derived from a pre-thymic stage (e.g. *E2A*, *GFI1*, *RUNX1*, *MYB* and *Ikaros*) that are essential for T-cell development. *TCF1* is a “gatekeeper” of T-cell fate that is directly activated by NOTCH signals and drives the expression of important T-cell specific regulators including *GATA3* and *BCL11B*. [28, 29] *GATA3* has critical roles during different developmental stages, namely in early T-cell survival and growth, T-cell

lineage commitment, and in T-cell subset diversification that follows TCR expression and antigen stimulation. ([30], reviewed in [31]) Accordingly, GATA3 is responsible to exclude access to B-cell lineage explaining the loss of B-cell potential nearly at the ETP stage, [32] and it is necessary to trigger BCL11B expression during T-cell commitment as well as to prompt the cells for β -selection. The transition from phase-1 to phase-2 is especially well characterized in mice, and seems marked by various events: silencing of the majority of phase-1 genes, cells become strictly dependent on NOTCH signalling to survive, proliferation rates are reduced and BCL11B expression is activated.[33] While cells progress to the DN2b stage (equivalent to CD1⁺ stage in humans), *ETS1* and *ETS2* are immediately expressed, *LEF1* is induced by increased NOTCH signalling, E2A activity is boosted, and the expression of E-dependent and NOTCH-dependent genes (e.g. *RAG1*, *RAG2*, *PTCRA*, *TDT* and *CD3E*) is considerably higher.[34] This transition from phase-1 to phase-2 is not well described in human thymocytes, because one of the requirements for T-cell commitment is reduction of NOTCH signalling. However, recent studies indicate that GATA3 is responsible to induce human T-cell commitment by upregulation of T-cell lineage genes, restraining of Notch activity and repressing NK-cell fate. [30] Consistently, Notch target genes that require strong NOTCH signalling are downregulated during T-cell commitment (e.g. NRARP, DTX1), while others such as HES1 and MYC are expressed until the β -selection checkpoint. [30]

During phase-2, many cells enter a G1 arrest and RAG-mediated TCR gene recombinations occur until cells successfully rearrange their TCR- β gene. Beyond the key roles at phase-2 gene network, E proteins, NOTCH signalling and BCL11B seem to have a tight collaboration as well. For instance, E proteins are required to promote and sustain NOTCH-1 signalling, which can be antagonized by ID2, an E protein antagonist. [35, 36] In turn, BCL11B may support the phase-2 stage by suppressing ID2 expression.[37] Importantly, the potential of those cells to become $\alpha\beta$ instead of $\gamma\delta$ T-cells can be related to their distinctive requirement for BCL11B and NOTCH signalling, which is different in mouse and human.

In mouse, the inhibition of NOTCH signalling by ID3 or ID2 factors promotes the development of $\gamma\delta$ T-cell lineages.[38] For that reason, the impact of *BCL11B* deletion is especially prominent in $\alpha\beta$ T-cells, in contrast to $\gamma\delta$ T-cell lineages,[39] and consequently the timing of *BCL11B* activation and expression is crucial for T-cell lineage commitment. Furthermore, BCL11B is possibly involved in direct repression of KIT, and BCL11B expression itself may be activated by TCF1, RUNX1 and perhaps GATA3, since the *BCL11B* promoter and a far-distant enhancer contain binding sites for these factors.[40] Although these positive regulators are expressed before *BCL11B*, perhaps it needs to reach a certain level of combined

expression to remove initial DNA methylation and H3K27me3 marks from both the promoter and the enhancer of *BCL11B*. Contrarily to mouse thymocytes, high NOTCH activation in human thymocytes during phase 2 promotes $\gamma\delta$ T cell development, while lower levels of NOTCH activation promote $\alpha\beta$ -lineage differentiation. In this case, specific Notch receptor-ligand interactions control human TCR- $\alpha\beta$ or - $\gamma\delta$ development, as they exhibit different Notch signal strength.[13] At the final stage of phase-2 regulatory activity, the cells are committed to T-cell lineage, $\alpha\beta$ -lineage biased and undergo further TCR- α gene rearrangements.

1.3.1.3. The β -selection program

The transition to phase-3 occurs at the β - selection checkpoint, when thymocytes that have not functionally rearranged the TCR- β gene are eliminated. The expression of a functional TCR- β chain activates the phase-3 gene network and promotes growth and differentiation beyond β - selection. At the checkpoint, TCR β proteins assemble at the cell membrane beside TCR complex components that are already expressed during phase-2, namely pre-T α , CD3 co-receptor and TCR ζ . The newly assembled pre-TCR complex stimulates accelerated cell cycle, enlargement of the cells and increased expression of the co-stimulatory molecule CD28.[41, 42]

Analysis of TCR arrangements indicates that recombination events occur in the following order during human and mouse T-cell development: TCR- δ , TCR- γ , TCR- β and TCR- α .[43] In recent studies, transplantation of CD34+ cells from patients with severe combined immunodeficiency (SCID) in xenograft mouse models revealed that rearrangement of TCR- γ starts at the DN CD7+ stage. The TCR- β rearrangements are already initiated at the DN CD7+CD5+ stage (before T-cell commitment),[44] which is an earlier stage than the previously suggested immature SP CD4⁺.[45, 46] In mice, rearrangement of TCR- γ starts at DN1 stage and TCR- β rearrangements start at DN2, but occur mainly at the DN3 stage (after T-cell commitment). [47]

Following β - selection, cells are able to signal through pre-TCR, which results in loss of NOTCH-dependency (in mice) and downregulation of NOTCH-target genes and IL7R. *Ikaros* plays a crucial function at this stage, operating as a tumor suppressor gene by suppressing Notch target genes.[48, 49]. As a response to the inhibition of those signaling pathways that promote growth, the cells activate PI3K by becoming highly responsive to chemokine signaling through CXCR4, [50] and therefore initiate a new stage of very rapid proliferation. Until the end of phase-3 and start of the DP stage, cells further increase the expression of *ETS1*, *TCF7*, *LEF1*, *ETS2* and *TCF12* (canonical *HEB*). Also, the stable expression of new transcription

factors is activated (*ROR γ t*, *NFAT3*, *POU6F1* and *IKZF3*), while others (*EGR2* and *ID3*) are only transiently activated in response to the TCR signal during the rapidly proliferating stage.

When cells reach the DP stage, the interaction of ROR γ t with MYB promotes cell survival through the induction of BCLXL and impairs conventional effector responses such as proliferation and cytokine production.[51] At this stage, the regulatory functions of TCF1 and LEF1 may be altered by specific interactions with β -catenin, which is a mediator of the canonical WNT pathway. This alteration is responsible for the expression of CD4 and CD8, and supports the replacement of the pre-TCR α -chain with a new functionally rearranged TCR α -chain, which yields a complete $\alpha\beta$ TCR. [52, 53] As a result, DP thymocytes become prepared for the positive and negative selection events. DP $\alpha\beta$ TCR⁺ cells that are positively selected on the basis of specific TCR recognition can further differentiate into effector T cell subsets, such as CD4⁺ T cells, CD8⁺ T cells, natural killer T (NKT) cells or regulatory T cells.

1.3.2. Genetic aberrations in T-ALL

T-ALL is a heterogeneous disease that arises from the oncogenic transformation of immature thymocytes. Different subtypes of T-ALL have been identified on the basis of genetic aberrations (Figure 3).[54] The genetic aberrations are characterized by chromosomal rearrangements that lead to the aberrant activation of oncogenic transcription factors, including TAL1 and LMO2 (and related family members), TLX1, TLX3, NKX2-1, HOXA, and MEF2C, and also by particular oncogenic fusion proteins that directly activate the *HOXA* or *MEF2C* genes.[54-56] We have denoted these chromosomal rearrangements as type A aberrations, because they are generally considered to be the driving oncogenic event associated with unique expression profiles.[55] Based on their gene expression signatures, T-ALLs can be classified into four major subtypes: ETP-ALL, TLX, proliferative, and TALLMO (Figure 4).[56] These subtypes are associated with specific and sequential T-cell maturational arrest.[8, 45, 54] For instance, the immature cluster is associated with a very early arrest (CD34⁺cells), similar to an early T-cell precursor (ETP) profile.[57] In addition, the TLX cluster includes $\gamma\delta$ -lineage arrested cells; the proliferative cluster is associated with a cortical stage arrest (CD1⁺cells); and the TALLMO cluster consists of cells arrested at a mature stage (CD4⁺ and/or CD8⁺cells).[54]

The maturational arrest of the cells generates a premalignant condition that facilitates the rapid acquisition of additional genetic hits, which we denoted as type B mutations.[55] These mutations are prevalent among all T-ALL subtypes and do not necessarily occur in the entire leukemic cell population. Instead, they can appear in subclones that are often either selected or

lost during disease progression or post-treatment relapse. [58, 59] Type B mutations affect different signal transduction pathways, including the NOTCH1, IL7R-JAK-STAT, RAS-MEK-ERK and PTEN-PI3K-AKT pathways, and are therefore associated with a variety of cellular processes, including survival and proliferation, cell cycle progression, and epigenetic events. The most frequent genetic hit in pediatric T-ALL is the deletion of the *CDKN2A* locus, which encompasses the *p16/INK4A* and *p14/ARF* suppressor genes, and is present in more than 70% of T-ALL cases. In addition, constitutive activation of NOTCH1 signaling is one of the most prominent oncogenic pathway in T-cell transformation, represented in more than 60% of T-ALL cases.

However, increasing evidence suggests that some of the signaling pathways are preferentially mutated in certain T-ALL subtypes. This may be related to the natural activation of a certain pathway at the stage where the pre-leukemic clone is arrested. For instance, IL7-Receptor (IL7R) signaling is required for early thymocyte progenitor (ETP) development,[60] and mutations in IL7R—or the downstream molecules JAK or RAS—are prevalent among ETP-ALL and TLX patients.[61, 62](our own observations) Moreover, PTEN inactivation events are preferentially associated with the TALLMO subtype,[63] while NOTCH1 mutations are more prevalent in the TLX cluster, but have a lower incidence in ETP-ALL and TALLMO subtypes.[64] Although strategies targeting oncogenic type A transcription factor complexes are emerging,[65] multiple selective compounds inhibiting pathways that are activated by type B hits are readily available.

1.3.3. The importance of an *in vitro* culture system to study T-ALL

Although T-ALL genetic signatures provide clues about the primary and cooperating genetic lesions in the multi-step process that drives malignant transformation, *in vivo* and *in vitro* models remain valuable tools to investigate the sequential genetic lesions that can initiate and sustain T-ALLs. In this respect, it is crucial to compare human and mouse T cell development not only at the levels of surface marker expression (e.g. CD markers), but also at the requirements of NOTCH and cytokines signalling as well as the underlying transcriptional networks during differentiation. This will assess the conservation of T-cell developmental programs between both species and is also important to understand the relevance of mouse model studies for human T-ALL. While T-cell development has been extensively studied in mice, the restricted availability of human thymus material has limited a detailed insight of human T-cell development. For that reason, research has focused on mimicking the thymus environment using *in vitro* differentiation

cultures starting with hematopoietic stem cells (HSCs) derived from human cord blood or bone marrow. Early studies were performed by seeding of HSCs in lobes of the mouse fetal thymus, cultures that are denoted as fetal thymus organ cultures (FTOCs) and allow intrathymic T-cell development *in vitro*. [66] Although this system provided important understanding of developmental processes, the capacity to evaluate specific progenitor populations has remained difficult, since this approach only provides a limited number of T cells. In addition, the procedure with FTOCs is a time-consuming and technically demanding method.

As an alternative, the OP9 stromal cells co-culture system is available to support T-cell development *in vitro*. In particular, the expression of the Delta-like 1 ligand on these bone marrow-derived stromal cells from op/op M-CSF deficient mice induces NOTCH signalling in target cells of the hematopoietic lineage. [67] NOTCH signalling results in inhibition of B-cell differentiation while T-cell differentiation is promoted. Umbilical cord blood-derived HSCs in this co-culture system are primed to develop into the T-cell lineage, and the OP9-DL1 system recapitulates *in vivo* T-cell development as measured by sequential acquisition of surface markers CD7, CD5, CD1a, and the achievement of the CD4, CD8 double-positive (DP) stage. [16, 17] These results establish an efficient and versatile *in vitro* system for the generation of relatively larger numbers of human T-lineage cells and offer a model system to investigate the factors that influence human T-cell–lineage commitment and differentiation. In addition, this system may be a valuable tool for investigating the sequential genetic lesions that can initiate and sustain T-ALL, as it is easy to manipulate many factors involved in lymphocyte differentiation, including cytokines, Notch signaling pathway and the expression of specific (onco)genes.

However, the OP9-DL1 system also has some limitations like any *in vitro* system. T-cell development can only induce efficient lymphopoiesis until the CD4+ CD8+ DP stage, and does not support positive/negative selection due to lack of proper MHC expression. T-cell development in this system is very sensitive to subtle differences in cytokines and to culture media changes. [68] Also, the effect of thymic epithelial cells (TECs) such as the presentation of different ligands and release of chemokines that create the thymus environment is ignored. Other issues as cell emigration and trafficking cannot be addressed using this system. To address these questions different immunodeficient mice models can be used, namely NOD/SCID γ c $^{-/-}$ (NSG) and Rag2 $^{-/-}$ γ c $^{-/-}$ strains. Transplantation of human CD34+ HSCs into these conditioned neonatal mice supports multi-lineage human hematopoiesis leading to the production of T cells, B cells and dendritic cells. [69]

2. Aims of the Thesis

The aims of this thesis are:

- I. to identify additional PTEN aberrations that may provide survival and proliferation advantage to (pre)-leukemic cells, and understand their clinical impact in T-ALL (Chapters 1 and 2),
- II. to optimize a lentivirus system that supports transduction and stable expression of oncogenes in human and mouse HSCs (Chapter 3),
- III. to establish an *in vitro* co-culture system (OP9-DL1) that supports T-cell development from human umbilical cord blood-derived hematopoietic stem cells (HSCs) in order to define distinct human T-cell development stages in relation to the expression of cell surface markers and underlying transcriptional programs, and in relation to mouse and human *in vivo* T-cell development (Chapter 4),
- IV. to study the impact of ETP-ALL expressed oncogenes (such as MEF2C, LYL1 or LMO2) on human T-cell development (Chapter 5).

3. Short outline of the Thesis

In chapter 2, we review the incidence of inactivating mechanisms of the PTEN tumor suppressor and we discuss relationships to other recurrent aberrations in T-ALL, such as driving oncogenic rearrangements and mutations affecting the NOTCH1, MYC and AKT pathways. We further investigate the role of PTEN aberrations in causing cellular resistance to NOTCH1-directed therapy, suggesting a new hypothesis that explains conflicting data in literature. Finally, we suggest the use of specific signaling pathway inhibitors to prevent resistance to NOTCH-directed therapy in NOTCH1-mutated T-ALL.

In chapter 3, we identify microdeletions as an additional and novel mechanism for *PTEN* inactivation in T-ALL. The breakpoints of these microdeletions are caused by illegitimate RAG-

mediated recombination events. The identification of subclonal *PTEN* microdeletions indicates that RAG activity may be ongoing in (at least part of) the leukemic cell population, and could therefore be responsible for the clonal diversity and/or selection observed in T-ALL during disease progression. The discovery of *PTEN* microdeletions adds a new level of complexity that should be addressed in the development of future antileukemic strategies for ALL.

In chapter 4, we generate a strategy for the cloning, production and transduction of a lentiviral vector with the aim of studying the leukemogenic effects of T-ALL oncogenes during T-cell development. Hence, we optimize the transduction of murine and human HSCs by optimizing both vector design and serum-free virus production as well as by developing a qRT-PCR to quantify viral batches. We suggest that proviral RNA length is an important factor that determines transfection and viral productions as well as transduction efficiencies.

In chapter 5, we validate the OP9-DL1 co-cultures as an *in vitro* system that mimics thymic $\alpha\beta$ T-cell lineage development. In particular, we perform gene expression arrays of *in vitro* differentiated early T-cell subsets at consecutive stages of development. As validation, we compare expression profiles of the *in vitro* generated T-cell subsets with different stages of human and mouse thymocyte development *in vivo*. In addition, we define a gene signature that clearly distinguishes two major T-cell differentiation stages, and that correlates with *in vivo* pre- and post-commitment T-cell profiles. We further demonstrate that loss of CD44 functionally marks human T-cell commitment, which is supported by analysis of TCR- β rearrangements as well as by multi-lineage differentiation experiments.

In chapter 6, we show that T-ALL subtypes are strongly associated with consecutive T-cell developmental programs. We demonstrate that immature and TLX subtypes are associated with a pre-commitment program, and the proliferative and TALLMO subtypes with a post-commitment program. Detailed analysis of T-ALL gene expression profiles indicates that T-ALL oncogenes interfere with T-cell developmental programs. Finally, we demonstrate that ectopic expression of typical ETP-ALL oncogenes (such as MEF2C, LYL1 or LMO2) promotes developmental arrest of early thymic progenitor (ETP) cells.

In Chapter 7, we discuss all findings presented in the thesis and put these into perspective. In Chapter 8 we summarize the content of the thesis (in English and Dutch).

References

1. Orkin, S.H. and L.I. Zon, *Hematopoiesis: an evolving paradigm for stem cell biology*. Cell, 2008. **132**(4): p. 631-44.
2. Morrison, S.J. and J. Kimble, *Asymmetric and symmetric stem-cell divisions in development and cancer*. Nature, 2006. **441**(7097): p. 1068-74.
3. Sachs, L., *The control of hematopoiesis and leukemia: from basic biology to the clinic*. Proc Natl Acad Sci U S A, 1996. **93**(10): p. 4742-9.
4. Hoffman, R., *Hematology : basic principles and practice*. Buch Buch. 2005, New York, NY [u.a.]: Elsevier, Churchill Livingstone.
5. Pui, C.H. and W.E. Evans, *Treatment of acute lymphoblastic leukemia*. N Engl J Med, 2006. **354**(2): p. 166-78.
6. Pui, C.H., M.V. Relling, and J.R. Downing, *Acute lymphoblastic leukemia*. N Engl J Med, 2004. **350**(15): p. 1535-48.
7. Pui, C.H., et al., *Biology, risk stratification, and therapy of pediatric acute leukemias: an update*. J Clin Oncol, 2011. **29**(5): p. 551-65.
8. Ferrando, A.A., et al., *Gene expression signatures define novel oncogenic pathways in T cell acute lymphoblastic leukemia*. Cancer Cell, 2002. **1**(1): p. 75-87.
9. Pui, C.H., L.L. Robison, and A.T. Look, *Acute lymphoblastic leukaemia*. Lancet, 2008. **371**(9617): p. 1030-43.
10. Pieters, R. and W.L. Carroll, *Biology and treatment of acute lymphoblastic leukemia*. Pediatr Clin North Am, 2008. **55**(1): p. 1-20, ix.
11. Yui, M.A. and E.V. Rothenberg, *Developmental gene networks: a triathlon on the course to T cell identity*. Nat Rev Immunol, 2014. **14**(8): p. 529-45.
12. Zhang, J.A., et al., *Dynamic transformations of genome-wide epigenetic marking and transcriptional control establish T cell identity*. Cell, 2012. **149**(2): p. 467-82.
13. Van de Walle, I., et al., *Specific Notch receptor-ligand interactions control human TCR-alpha/beta/gammadelta development by inducing differential Notch signal strength*. J Exp Med, 2013. **210**(4): p. 683-97.
14. Van de Walle, I., et al., *An early decrease in Notch activation is required for human TCR-alpha/beta lineage differentiation at the expense of TCR-gammadelta T cells*. Blood, 2009. **113**(13): p. 2988-98.
15. Ciofani, M. and J.C. Zuniga-Pflucker, *Notch promotes survival of pre-T cells at the beta-selection checkpoint by regulating cellular metabolism*. Nat Immunol, 2005. **6**(9): p. 881-8.
16. La Motte-Mohs, R.N., E. Herer, and J.C. Zuniga-Pflucker, *Induction of T-cell development from human cord blood hematopoietic stem cells by Delta-like 1 in vitro*. Blood, 2005. **105**(4): p. 1431-9.
17. Awong, G., et al., *Characterization in vitro and engraftment potential in vivo of human progenitor T cells generated from hematopoietic stem cells*. Blood, 2009. **114**(5): p. 972-82.
18. Rothenberg, E.V., *Transcriptional control of early T and B cell developmental choices*. Annu Rev Immunol, 2014. **32**: p. 283-321.
19. Weerkamp, F., K. Pike-Overzet, and F.J. Staal, *T-sing progenitors to commit*. Trends Immunol, 2006. **27**(3): p. 125-31.

20. Riddell, J., et al., *Reprogramming committed murine blood cells to induced hematopoietic stem cells with defined factors*. Cell, 2014. **157**(3): p. 549-64.
21. McCormack, M.P., et al., *The Lmo2 oncogene initiates leukemia in mice by inducing thymocyte self-renewal*. Science, 2010. **327**(5967): p. 879-83.
22. Thoms, J.A., et al., *ERG promotes T-acute lymphoblastic leukemia and is transcriptionally regulated in leukemic cells by a stem cell enhancer*. Blood, 2011. **117**(26): p. 7079-89.
23. Zohren, F., et al., *The transcription factor Lyl-1 regulates lymphoid specification and the maintenance of early T lineage progenitors*. Nat Immunol, 2012. **13**(8): p. 761-9.
24. Phelan, J.D., et al., *Growth factor independent-1 maintains Notch1-dependent transcriptional programming of lymphoid precursors*. PLoS Genet, 2013. **9**(9): p. e1003713.
25. Weerkamp, F., et al., *Human thymus contains multipotent progenitors with T/B lymphoid, myeloid, and erythroid lineage potential*. Blood, 2006. **107**(8): p. 3131-7.
26. Rothenberg, E.V., *T cell lineage commitment: identity and renunciation*. J Immunol, 2011. **186**(12): p. 6649-55.
27. Hosoya, T., et al., *GATA-3 is required for early T lineage progenitor development*. J Exp Med, 2009. **206**(13): p. 2987-3000.
28. Weber, B.N., et al., *A critical role for TCF-1 in T-lineage specification and differentiation*. Nature, 2011. **476**(7358): p. 63-8.
29. Germar, K., et al., *T-cell factor 1 is a gatekeeper for T-cell specification in response to Notch signaling*. Proc Natl Acad Sci U S A, 2011. **108**(50): p. 20060-5.
30. Van de Walle, I., et al., *GATA3 induces human T-cell commitment by restraining Notch activity and repressing NK-cell fate*. Nat Commun, 2016. **7**: p. 11171.
31. Hosoya, T., I. Maillard, and J.D. Engel, *From the cradle to the grave: activities of GATA-3 throughout T-cell development and differentiation*. Immunol Rev, 2010. **238**(1): p. 110-25.
32. Garcia-Ojeda, M.E., et al., *GATA-3 promotes T-cell specification by repressing B-cell potential in pro-T cells in mice*. Blood, 2013. **121**(10): p. 1749-59.
33. Yui, M.A., N. Feng, and E.V. Rothenberg, *Fine-scale staging of T cell lineage commitment in adult mouse thymus*. J Immunol, 2010. **185**(1): p. 284-93.
34. Ikawa, T., et al., *E proteins and Notch signaling cooperate to promote T cell lineage specification and commitment*. J Exp Med, 2006. **203**(5): p. 1329-42.
35. Del Real, M.M. and E.V. Rothenberg, *Architecture of a lymphomyeloid developmental switch controlled by PU.1, Notch and Gata3*. Development, 2013. **140**(6): p. 1207-19.
36. Yashiro-Ohtani, Y., et al., *Pre-TCR signaling inactivates Notch1 transcription by antagonizing E2A*. Genes Dev, 2009. **23**(14): p. 1665-76.
37. Li, L., M. Leid, and E.V. Rothenberg, *An early T cell lineage commitment checkpoint dependent on the transcription factor Bcl11b*. Science, 2010. **329**(5987): p. 89-93.
38. Lauritsen, J.P., et al., *Marked induction of the helix-loop-helix protein Id3 promotes the gammadelta T cell fate and renders their functional maturation Notch independent*. Immunity, 2009. **31**(4): p. 565-75.
39. Wakabayashi, Y., et al., *Bcl11b is required for differentiation and survival of alphabeta T lymphocytes*. Nat Immunol, 2003. **4**(6): p. 533-9.
40. Li, L., et al., *A far downstream enhancer for murine Bcl11b controls its T-cell specific expression*. Blood, 2013. **122**(6): p. 902-11.

41. Taghon, T., et al., *Notch signaling is required for proliferation but not for differentiation at a well-defined beta-selection checkpoint during human T-cell development*. Blood, 2009. **113**(14): p. 3254-63.
42. Kreslavsky, T., et al., *beta-Selection-induced proliferation is required for alphabeta T cell differentiation*. Immunity, 2012. **37**(5): p. 840-53.
43. Dik, W.A., et al., *New insights on human T cell development by quantitative T cell receptor gene rearrangement studies and gene expression profiling*. J Exp Med, 2005. **201**(11): p. 1715-23.
44. Wiekmeijer, A.S., et al., *Identification of checkpoints in human T-cell development using severe combined immunodeficiency stem cells*. J Allergy Clin Immunol, 2016. **137**(2): p. 517-526.e3.
45. Soulier, J., et al., *HOXA genes are included in genetic and biologic networks defining human acute T-cell leukemia (T-ALL)*. Blood, 2005. **106**(1): p. 274-86.
46. Blom, B. and H. Spits, *Development of human lymphoid cells*. Annu Rev Immunol, 2006. **24**: p. 287-320.
47. Staal, F.J., et al., *Transcriptional control of t lymphocyte differentiation*. Stem Cells, 2001. **19**(3): p. 165-79.
48. Chari, S. and S. Winandy, *Ikaros regulates Notch target gene expression in developing thymocytes*. J Immunol, 2008. **181**(9): p. 6265-74.
49. Kleinmann, E., et al., *Ikaros represses the transcriptional response to Notch signaling in T-cell development*. Mol Cell Biol, 2008. **28**(24): p. 7465-75.
50. Janas, M.L., et al., *Thymic development beyond beta-selection requires phosphatidylinositol 3-kinase activation by CXCR4*. J Exp Med, 2010. **207**(1): p. 247-61.
51. Sun, Z., et al., *Requirement for RORgamma in thymocyte survival and lymphoid organ development*. Science, 2000. **288**(5475): p. 2369-73.
52. Xu, M., et al., *Sustained expression of pre-TCR induced beta-catenin in post-beta-selection thymocytes blocks T cell development*. J Immunol, 2009. **182**(2): p. 759-65.
53. Xu, M., et al., *Pre-TCR-induced beta-catenin facilitates traversal through beta-selection*. J Immunol, 2009. **182**(2): p. 751-8.
54. Meijerink, J.P., *Genetic rearrangements in relation to immunophenotype and outcome in T-cell acute lymphoblastic leukaemia*. Best Pract Res Clin Haematol, 2010. **23**(3): p. 307-18.
55. Van Vlierberghe, P., et al., *Molecular-genetic insights in paediatric T-cell acute lymphoblastic leukaemia*. Br J Haematol, 2008. **143**(2): p. 153-68.
56. Homminga, I., et al., *Integrated transcript and genome analyses reveal NKX2-1 and MEF2C as potential oncogenes in T cell acute lymphoblastic leukemia*. Cancer Cell, 2011. **19**(4): p. 484-97.
57. Zuurbier, L., et al., *Immature/early T-cell precursor acute lymphoblastic leukemia patients treated on the COALL-97 protocol are not associated with poor outcome*. Submitted for publication, 2012.
58. Notta, F., et al., *Evolution of human BCR-ABL1 lymphoblastic leukaemia-initiating cells*. Nature, 2011. **469**(7330): p. 362-7.
59. Anderson, K., et al., *Genetic variegation of clonal architecture and propagating cells in leukaemia*. Nature, 2011. **469**(7330): p. 356-61.

60. Fry, T.J. and C.L. Mackall, *The many faces of IL-7: from lymphopoiesis to peripheral T cell maintenance*. J Immunol, 2005. **174**(11): p. 6571-6.
61. Zenatti, P.P., et al., *Oncogenic IL7R gain-of-function mutations in childhood T-cell acute lymphoblastic leukemia*. Nat Genet, 2011. **43**(10): p. 932-9.
62. Zhang, J., et al., *The genetic basis of early T-cell precursor acute lymphoblastic leukaemia*. Nature, 2012. **481**(7380): p. 157-63.
63. Zuurbier, L., et al., *The significance of PTEN and AKT aberrations in pediatric T-cell acute lymphoblastic leukemia*. Haematologica, 2012.
64. Zuurbier, L., et al., *NOTCH1 and/or FBXW7 mutations predict for initial good prednisone response but not for improved outcome in pediatric T-cell acute lymphoblastic leukemia patients treated on DCOG or COALL protocols*. Leukemia, 2010. **24**(12): p. 2014-22.
65. Filippakopoulos, P., et al., *Selective inhibition of BET bromodomains*. Nature, 2010. **468**(7327): p. 1067-73.
66. DeLuca, D., et al., *Programmed differentiation of murine thymocytes during fetal thymus organ culture*. J Immunol Methods, 1995. **178**(1): p. 13-29.
67. Schmitt, T.M. and J.C. Zuniga-Pflucker, *Induction of T cell development from hematopoietic progenitor cells by delta-like-1 in vitro*. Immunity, 2002. **17**(6): p. 749-56.
68. Six, E.M., et al., *Cytokines and culture medium have a major impact on human in vitro T-cell differentiation*. Blood Cells Mol Dis, 2011. **47**(1): p. 72-8.
69. Manz, M.G., *Human-hemato-lymphoid-system mice: opportunities and challenges*. Immunity, 2007. **26**(5): p. 537-41.

Figure 1. Schematic representation of human hematopoiesis. Multipotent hematopoietic stem cell have self-renewal capacity and differentiate into all blood cell types.

Figure 2. Schematic overview of mice (A) and human (B) $\alpha\beta$ T-cell development. The sequential stages from a pre-thymic progenitor to the double-positive (DP) stage are depicted in the figures. The mice developmental stages are accompanied of associated expression of surface markers and transcription factors. The variation in colour intensity represents their level of expression. Obtained from Yui and Rothenberg, Nat Rev Immunol, 2014. 14(8): p. 529-45.

Figure 3. Schematic overview of T-ALL genetic subgroups ETP-ALL, TLX3, TLX1, HOXA and TAL/LMO in relation to their T-cell developmental stage based on EGIL or TCR classification systems. * *LMO2* is ectopically expressed in *LMO2*-rearranged cases but is not included in this figure. Adapted from Meijerink et al, Best Pract Res Clin Haematol, 2010. Sep;23(3):307-18.

Figure 4. Unsupervised hierarchical cluster analysis of 117 pediatric T-ALL samples and 7 normal bone-marrow (NBM) controls based on 435 probesets. Cytogenetic rearrangements indicated are: S, SIL-TAL1; T, TAL1; t, TAL2; O, LMO1; L, LMO2; \$, TAL2/LMO1; N, SET-NUP214; C, CALM-AF10; M, MYB; A, Inv(7)(p15q34); 1, TLX1; 3, TLX3; and n, normal bone marrow controls. The samples with the highest TAL1 or LYL1 expression, positivity for TLX1 and TLX3 expression, and expression of the immunophenotypic markers CD13 and/or CD33, CD4 or CD8 are indicated; u, no data available. Obtained from Homminga et al, Cancer Cell, 2011. 19(4): p. 484-97.

Figure 1.

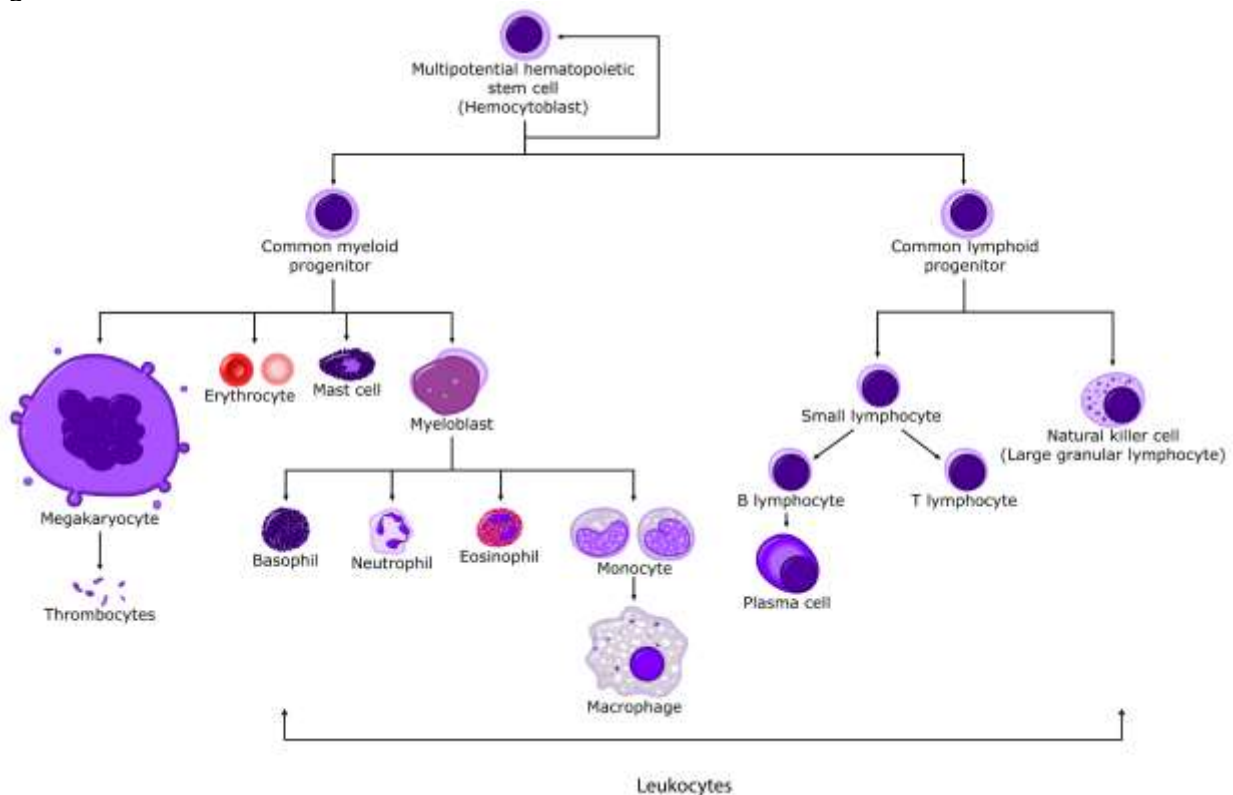


Figure 2

A

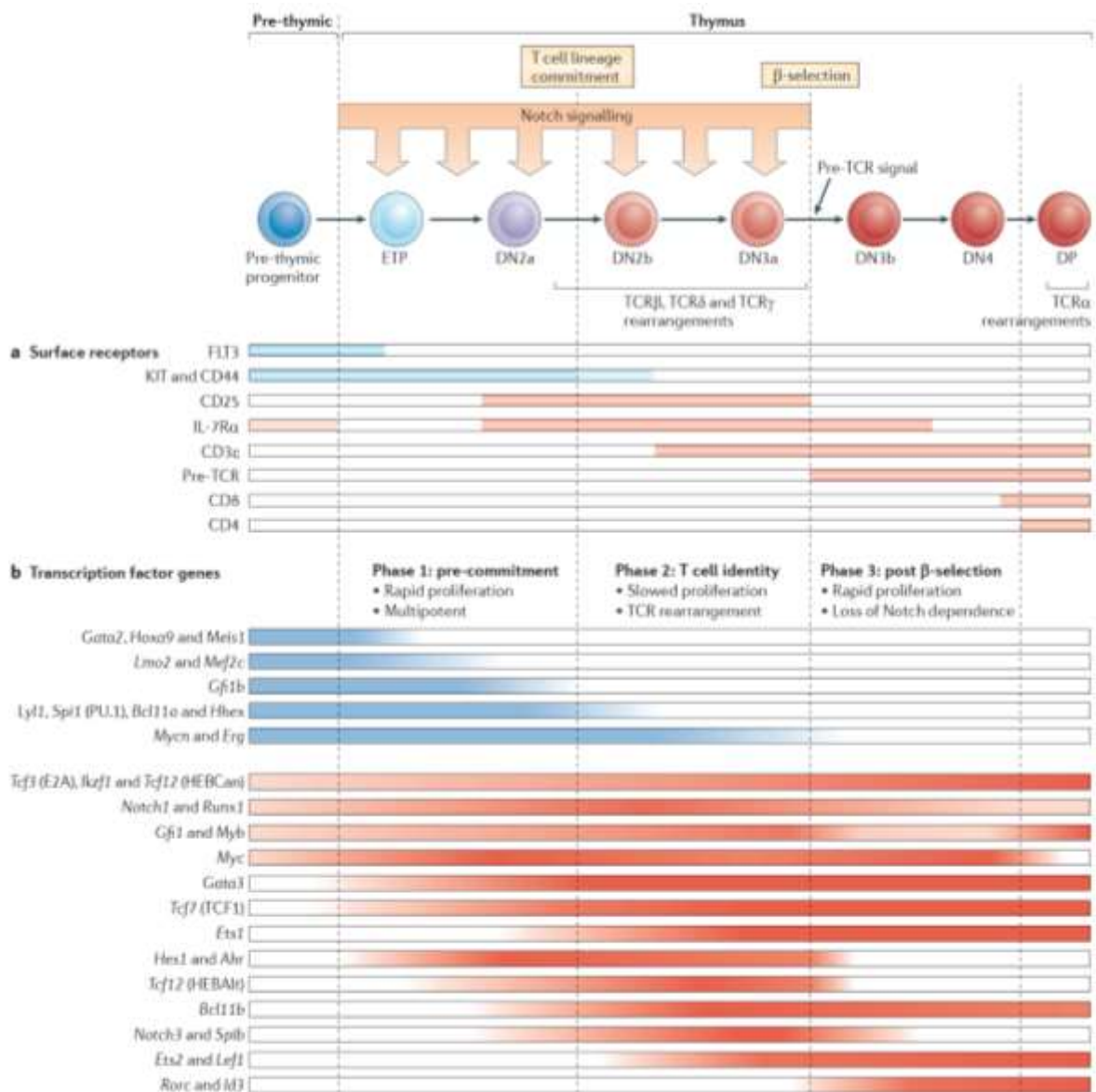


Figure 2

B

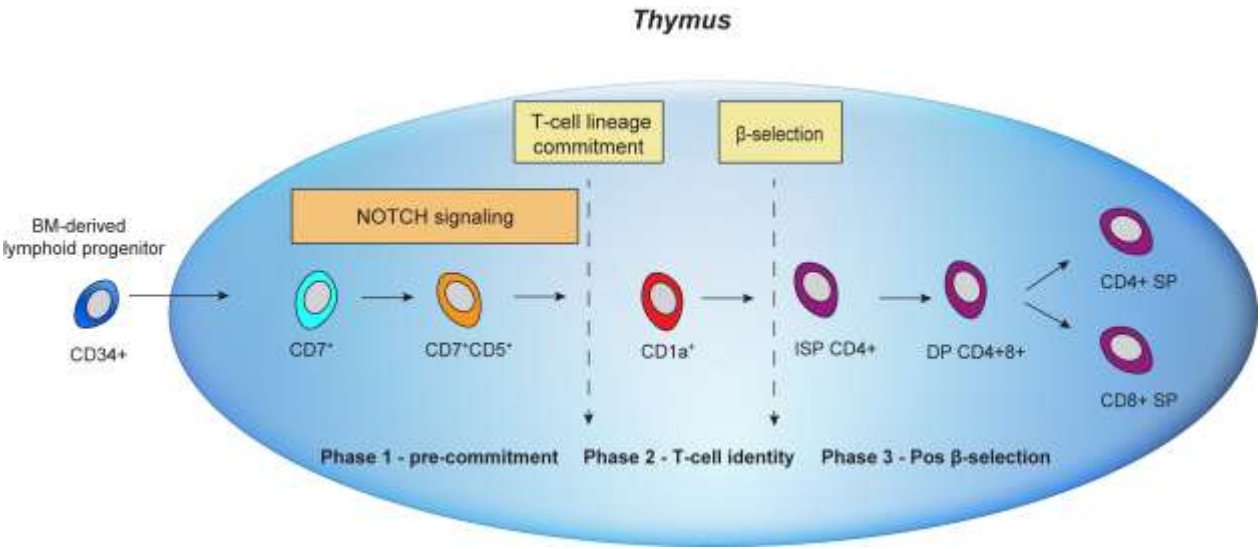


Figure 3

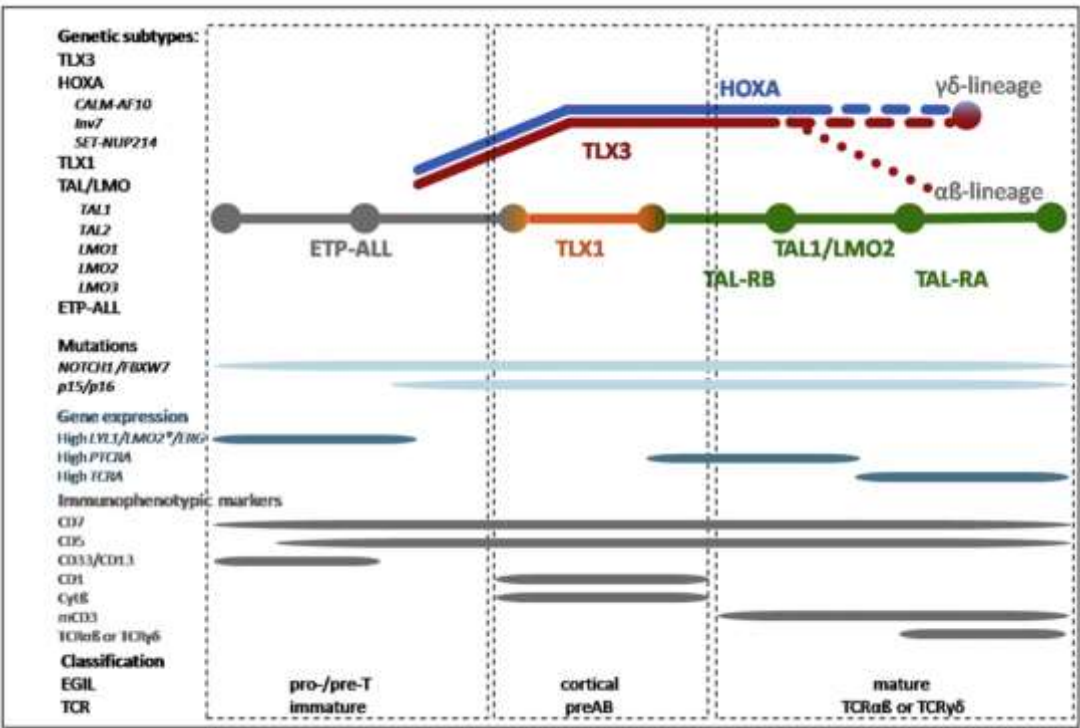
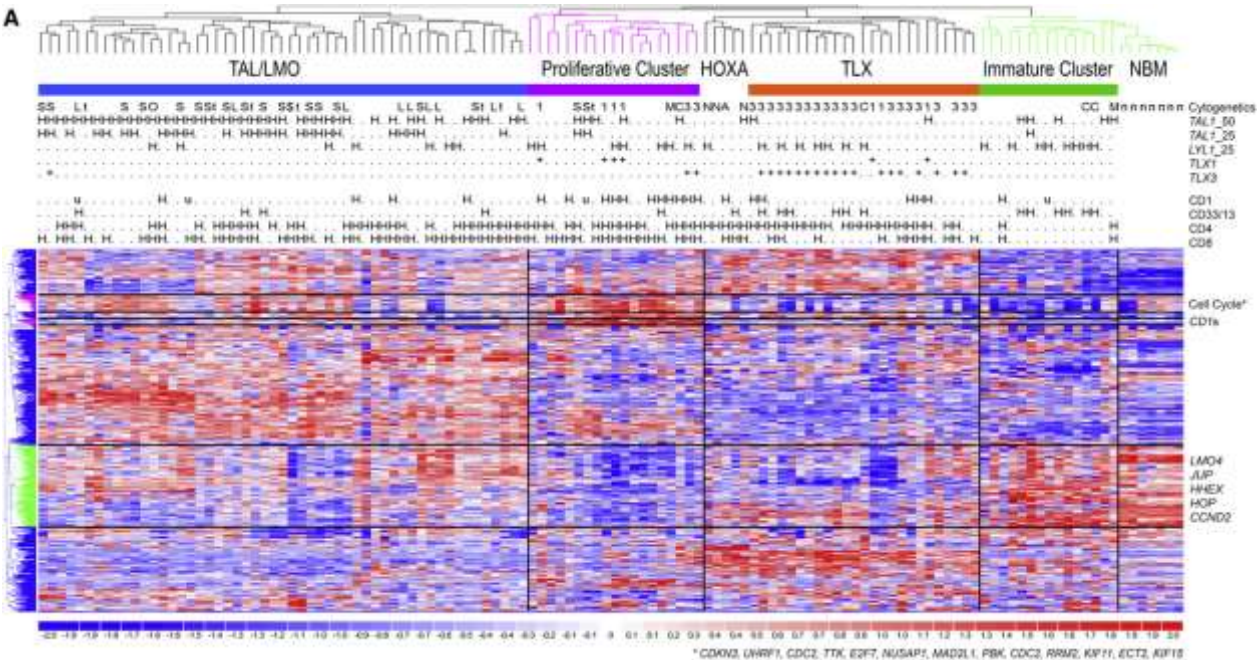


Figure 4



Chapter 2

The relevance of PTEN-AKT in relation to NOTCH1-directed treatment strategies in T-cell acute lymphoblastic leukemia

Rui D. Mendes¹, Kirsten Canté-Barrett¹, Rob Pieters^{1,2} and Jules P. P. Meijerink¹

¹Department of Pediatric Oncology/Hematology, Erasmus MC Rotterdam-Sophia Children's Hospital, Rotterdam, the Netherlands, ²Princess Máxima Center of Pediatric Oncology, Utrecht, the Netherlands

Haematologica. 2016 Sep;101(9):1010-7

Abstract

The tumor suppressor phosphatase and tensin homolog (PTEN) negatively regulates phosphatidylinositol 3-kinase (PI3K)-AKT signaling and is often inactivated by mutations (including deletions) in a variety of cancer types, including T-cell acute lymphoblastic leukemia. Here, we review mutation-associated mechanisms that inactivate PTEN together with other molecular mechanisms that activate AKT and contribute to T-cell leukemogenesis. In addition, we discuss how *Pten* mutations in mouse models affect the efficacy of gamma-secretase inhibitors to block NOTCH1 signaling through activation of AKT. Based on these models and on observations in primary diagnostic samples of T-cell acute lymphoblastic leukemia patients, we speculate that PTEN-deficient cells employ an intrinsic homeostatic mechanism in which PI3K-AKT signaling is dampened over time. As a result of this reduced PI3K-AKT signaling, the level of AKT activation may be insufficient to compensate for NOTCH1 inhibition, resulting in responsiveness to gamma-secretase inhibitors. On the other hand, *de novo* acquired PTEN inactivating events in NOTCH1-dependent leukemia could result in temporal, strong activation of PI3K-AKT signaling, increased glycolysis and glutaminolysis, and consequently gamma-secretase inhibitor resistance. Due to the central role of PTEN-AKT signaling and in the resistance to NOTCH1 inhibition, AKT inhibitors may be a promising addition to current treatment protocols for T-cell acute lymphoblastic leukemia.

T-cell acute lymphoblastic leukemia

T-cell acute lymphoblastic leukemia (T-ALL) is a cancer of developing T-cells in the thymus. T-ALL is characterized by chromosomal rearrangements. These rearrangements can lead to the aberrant activation of oncogenic transcription factors by placing their genes under the control of promoters and/or enhancers of T-cell receptor genes, the *BCL11B* gene, or other genes; occasionally, these rearrangements can give rise to oncogenic fusion proteins. The activated oncogenic transcription factors include *TAL1* and *LMO2* (and related family members), *TLX1*, *TLX3*, *NKX2-1*, *HOXA*, and *MEF2C*; in addition, certain oncogenic fusion proteins can directly activate the *HOXA* or *MEF2C* genes.(1, 2) Oncogenic proteins facilitate the developmental arrest of pre-leukemic immature T-cells. We previously proposed that these chromosomal rearrangements should be classified as type A aberrations, as they are generally considered to be the driving oncogenic event associated with unique expression profiles.(2) Based upon their gene expression signatures, T-ALL can be classified into the following four major subtypes: ETP-ALL, TLX, proliferative, and TALLMO.(3-5)

Maturation arrest induces a pre-leukemic condition in which additional mutations can give rise to T-ALL.(1, 2) These secondary mutations are not necessarily clonal events and are often selected during disease progression or post-treatment relapse.(6, 7) We therefore proposed that these mutations should be classified as type B aberrations.(2) Type B mutations are prevalent among all T-ALL subtypes and affect a wide variety of cellular processes, including survival and proliferation, cell cycle progression, and epigenetic events. Type B mutations often affect signal transduction pathways, including the NOTCH1, IL7R-JAK-STAT, RAS-MEK-ERK, and PTEN-PI3K-AKT pathways. A growing body of evidence suggests that some of these signaling pathways are preferentially mutated in specific T-ALL subtypes, presumably due to the fact that developing T-cells are dependent on these pathways in specific stages. For example, mutations in IL7 receptor (IL7R) and the downstream molecules JAK or RAS are prevalent among TLX and ETP-ALL patients.(8-10) Although new therapeutic strategies that target oncogenic transcription factor complexes are emerging,(11) several compounds that selectively inhibit altered signaling pathways are currently available. Thus, inhibiting signaling proteins such as NOTCH, IL7R, RAS and/or AKT may provide a promising new therapeutic approach for T-ALL.

In this review, we describe the role of PTEN as a tumor suppressor and we discuss various PTEN inactivating mechanisms observed in different human cancers and T-ALL. Besides PTEN inactivation, we describe other mechanisms that contribute to AKT activation and leukemogenesis. Finally, we discuss PTEN-AKT signaling in relation to future NOTCH1-directed therapies and provide a rationale for the use of AKT inhibitors in addition to current treatment protocols.

The PTEN tumor suppressor

Mutations in the tumor suppressor gene *PTEN* (phosphatase and tensin homolog), which is located on chromosomal band 10q23, are highly common among a wide range of cancers.(12, 13) The *PTEN* gene contains nine exons, and the encoded protein includes an N-terminal phosphatase domain, a central C2 lipid membrane-binding domain, and a C-terminal tail domain (Figure 1). PTEN is a phosphatase that dephosphorylates PIP₃ (phosphatidylinositol (3,4,5)-triphosphate) to produce PIP₂ (phosphatidylinositol (4,5)-bisphosphate), thereby opposing the function of PI3K (phosphatidylinositol 3-kinase). PI3K converts PIP₂ into PIP₃, which in turn activates key downstream kinases, including PDK1 and AKT (Figure 2). Thus, PTEN is an important negative regulator of PI3K-AKT signaling. Because AKT plays key roles in cellular metabolism, proliferation and survival, inactivation of PTEN by genetic aberrations drives

survival and uncontrolled proliferation, ultimately leading to cancer.(14) A recent study identified an alternate translation initiation site located upstream of the coding region of canonical *PTEN* that generates a larger form of PTEN.(15) This isoform is known as PTEN α and is described to be involved in mitochondrial energy metabolism.(15)

PTEN aberrations in cancer

Heterozygous germline mutations in *PTEN* were identified initially in 60-80% of patients belonging to a group of rare syndromes, including Cowden syndrome, Bannayan-Riley-Ruvalcaba syndrome, and PTEN-related Proteus syndrome; these disorders are known collectively as PTEN hamartoma tumor syndrome (PHTS).(16) With respect to sporadic (i.e., non-hereditary) tumors, heterozygous *PTEN* mutations occur in 50-80% of prostate, glioblastoma, and endometrial cancers and 30-50% of lung, colon, and breast cancers.(17) Loss of both functional *PTEN* alleles is common among prostate and breast cancer patients, as well as melanoma and glioblastoma patients.(18) The majority of these aberrations are caused by point mutations, small insertions, or deletions, all of which can occur throughout the entire *PTEN* gene. At the transcriptional and post-transcriptional levels, PTEN inactivation can occur via promoter methylation and through the expression of *PTEN*-directed microRNAs.(19) PTEN activity is also regulated at the post-translational level: phosphorylation, ubiquitination, oxidation, and acetylation can regulate the phosphatase activity, subcellular localization, and degradation of PTEN.(17) Defects in any of these processes may explain the absence of functional PTEN in cancer patients who apparently lack genetic aberrations in *PTEN*.(20-22) Several ALL cases have been identified in which high levels of inactive PTEN are accompanied by an active PI3K-AKT pathway.(23, 24) Although the majority of prevalent pathogenic mechanisms affect the loss of one or both *PTEN* alleles, subtle changes in PTEN protein levels can have a powerful effect on cancer susceptibility and/or tumor progression, as exemplified by the *Pten* hypomorphic mouse model.(25) Therefore, the level of functional PTEN affects tumor susceptibility, and PTEN function can be compromised at the DNA, mRNA, and/or protein levels.

PTEN aberrations in T-ALL

PTEN deletions and mutations were initially identified in cell lines.(26, 27) Restoring PTEN levels in these cell lines decreases cell size and induced apoptosis by suppressing the PI3K-AKT pathway.(28) Studies by others(29-34) and our group(21, 35) revealed aberrations in the *PTEN*-

PI3K-AKT pathway in approximately 23% of primary samples obtained from pediatric T-ALL patients. With respect to T-ALL subtypes, we have shown that PTEN aberrations are strongly associated with TAL- or LMO- rearranged patients in children (21) and the same was observed in adult T-ALL cohorts.(36) The vast majority of *PTEN* aberrations include nonsense mutations in exon 7 (which truncate the C-terminal domain) and deletions that affect nearly the entire locus (Figure 1). Although truncated PTEN proteins that lack the lipid-binding C-terminal domain retain their phosphatase activity, they are highly unstable and are degraded rapidly.(37) In mice, truncated PTEN leads to decreased genomic stability and the development of multiple cancers.(38) Recently, we reported that approximately 8% of T-ALL patients have a RAG-mediated microdeletion in the phosphatase domain that disrupts the reading frame (Figure 1).(35) In addition, mutations have been identified in PI3K and AKT; specifically, 9% of pediatric T-ALL patients have a mutation in either the catalytic (*PIK3CA*) or regulatory (*PIK3R1*) subunit of *PI3K*, and 2% of patients have a mutation in *AKT* itself (Figure 2A).(21, 30) Many T-ALL patients with a heterozygous *PTEN* mutation also acquire a deletion(21) or microdeletion(35) in their remaining wild-type allele in leukemic subclones(29) that may give rise to relapse. This phenomenon was demonstrated functionally by Clappier et al. who used an elegant human T-ALL xenograft transplantation model in mice and found the selection and preferential outgrowth of PTEN-inactivated leukemic cells.(39). In line with this, heterozygous *Pten* knockout mice develop T-cell leukemia in which the remaining wild-type allele is frequently deleted.(40-42) Another leukemogenic mechanism that can inactivate PTEN at the protein level is both the increased expression of casein kinase 2 (CK2) or the production of reactive oxygen species (ROS) that stabilize inactive forms of PTEN proteins and lead to impaired phosphatase activity.(23, 43) Together, these findings indicate the existence of ongoing pathogenic pressure to inactivate both *PTEN* alleles during disease progression, and the resulting loss of PTEN activity in turn activates the PI3K-AKT pathway.

Clinical implications

We have reported that aberrations in PTEN represent a significant, independent risk factor for relapse in T-ALL patients treated using either the DCOG (Dutch Childhood Oncology Group) or COALL (German Cooperative Study Group for Childhood ALL) protocol.(21, 35) Similar results were reported for other pediatric T-ALL patient cohorts treated using other protocols.(32, 33) In the BFM (Berlin-Frankfurt-Munster) study, the presence of NOTCH1-activating mutations in addition to PTEN-inactivating mutations predicts for good outcome similar to patients harboring

NOTCH1-activating mutations only,(32) suggesting that NOTCH1-mutations can antagonize the unfavorable effect of PTEN aberrations. In the French GRAALL (Group for Research in Adult ALL) study of adult T-ALL, patients with aberrations in *RAS* and/or *PTEN* had significantly worse outcome compared to patients without such mutations.(36) This was not confirmed in the MRC UKALL2003 trial for pediatric T-ALL; *RAS* and/or *PTEN* aberrations also did not change the favorable outcome of patients with *NOTCH1/FBWX7* mutations.(44) Taken together, these findings suggest that PTEN aberrations may represent a general poor prognostic factor in T-ALL.

NOTCH1 mutations lead to activation of AKT

More than 65% of T-ALL patients have aberrant activation of the NOTCH1 pathway due to mutations either in the *NOTCH1* gene itself or in *FBWX7*, which encodes E3-ubiquitin ligase.(45, 46) Thus, the NOTCH1 pathway may be an ideal target for therapeutic intervention. Furthermore, NOTCH1-directed therapies are clinically important, as they can also boost the cellular response to steroids.(47, 48) Gamma-secretase inhibitors (GSIs), which inhibit the presenilin gamma-secretase complex, block the cleavage of NOTCH1 at its S3 site; this cleavage step is required to release the active, intracellular NOTCH1 domain (ICN1) upon ligand binding (Figure 2B). Several groups have applied GSI to cell lines derived from T-ALL patients with *NOTCH1* activating mutations; although GSI treatment initially induces cell cycle arrest, the majority of cell lines adapt and ultimately stop responding to the treatment (i.e., develop GSI resistance).(45, 49) Nevertheless, GSI treatment effectively blocks gamma-secretase activity, resulting in reduced intracellular levels of the ICN1 domain and reduced expression of NOTCH1's target genes.(29) Therefore, GSI resistance is caused by other mechanisms that circumvent NOTCH1 inhibition.(50, 51) Consistent with this notion, Palomero and co-workers found that decreased PTEN levels in cell lines are correlated with GSI resistance, and GSI-resistant lines have increased levels of activated AKT.(29) Restoring the expression of functional PTEN in these GSI-resistant lines restored a GSI sensitivity response, whereas constitutively activated AKT or using shRNA to knockdown PTEN expression provoked GSI resistance in a GSI-responsive line.(29) This seminal study identified two important NOTCH1 downstream targets that regulate *PTEN* expression: HES1 and MYC. HES1 is a robust transcriptional repressor, whereas MYC is a weak transcriptional activator. Because the negative effect of HES1 prevails over the positive effect of MYC, *PTEN* expression is suppressed (Figure 2B).(29)

However, GSI resistance by leukemic cells resulted in disappointing results upon testing the GSI inhibitor MK-0752 in clinical trial (DFCI-04-390).(52) This trial was unsuccessful due to the compound's limited efficacy in leukemic cells and severe gastrointestinal toxicity. To overcome these issues, next-generation NOTCH1 inhibitors with reduced off-target toxicity are currently in development.(53) For example, promising strategies include selectively blocking NOTCH1 using anti-NOTCH1 antibodies(54, 55) or chemically modified peptides that block the NOTCH transcriptional complex in the nucleus.(56)

PTEN is not a priori linked to GSI resistance in human T-ALL

Despite the initial report by Palomero and co-workers,(29) subsequent studies have not confirmed that loss of PTEN activity is intrinsically linked to GSI resistance.(21, 57, 58) For example, GSI sensitivity was similar between NOTCH1-driven T-cell leukemia cells obtained from wild-type mice and from PTEN knockout mice.(58) However, PTEN deficiency does accelerate the disease progression of NOTCH1-driven leukemia.(58) Using a different *Pten* knockout mouse model (*Pten*^{flox/flox}/*Lck-Cre*), Hagenbeek et al. found that PTEN-deficient thymocytes were just as sensitive to *in vitro* GSI treatment as wild-type thymocytes.(57) Moreover, several human T-ALL cell lines with mutant alleles of *PTEN*—other cell lines than those used by Palomero et al.—were actually sensitive to GSI.(21) In TALLMO diagnostic patient samples, PTEN is frequently inactivated in the absence of NOTCH-activating mutations.(21, 32, 36, 59) Thus, mutations in *PTEN* and *NOTCH1/FBXW7* mutations are frequently independent genetic events and only co-occur in a small number of primary patient samples. In those patients that harbor both *PTEN* mutations and *NOTCH1/FBXW7* mutations, the *NOTCH1* mutations are usually weakly activating mutations. Because *PTEN* and *NOTCH1* mutations are mostly independent genetic events in primary T-ALL, PTEN-deficient leukemic cells in T-ALL patients likely do not have intrinsic GSI resistance at disease presentation. Perhaps one way that PTEN-deficient T-ALL can be linked to GSI resistance is upon relapse, when NOTCH1-dependent leukemic cells may have lost PTEN activity, possibly due to clonal selection following treatment. However, there is currently no evidence to support this notion.

The question remains, is it possible that immediately following PTEN loss, NOTCH1-dependent T-ALL becomes NOTCH1-independent and develops GSI-resistance? Recently, Adolfo Ferrando's group addressed this intriguing question by generating an elegant mouse model of NOTCH1-induced T-ALL in which the *Pten* gene is deleted only in established tumors.(60) Unlike previous *Pten* knockout models,(57, 58) deletion of *Pten* in this new model

conferred strong resistance to dibenzazepine (DBZ), a potent GSI. In this model,(60) *Pten* loss activated expression of genes involved in cell metabolism, ribosomal RNA processing, and amino acid and nucleotide biosynthesis, genes that are normally suppressed following NOTCH1-inhibiting GSI treatment. Moreover, GSI treatment increased leukemic cells' dependency on autophagy in order to recycle essential metabolites. *Pten* loss also relieved the GSI-instigated block of glycolysis and glutaminolysis, a phenotype that was copied by expressing the constitutively active myristoylated AKT. Because both the GSI DBZ and the glutaminase inhibitor BPTES (bis-2-(5-phenylacetamido-1,2,4-thiadiazol-2-yl)ethyl sulfide 3) act in a synergistic fashion in inducing anti-leukemic effects, the authors proposed glutaminolysis as a major therapeutic target for treating NOTCH-activated T-ALL.(60)

An unanswered question remains: why the loss of *Pten* in an established NOTCH1-driven tumor causes GSI resistance,(60) while NOTCH1-driven tumors that are generated in *Pten* knockout mice remain GSI-sensitive?(57, 58) The answer may lie in the ability of cells to adapt to PTEN loss by dampening PI3K-Akt signaling over time to a level that is still of advantage to leukemic cells. Unlike progressive reduction in PI3K-Akt signaling, the loss of PTEN may initially drive the rapid, high activation of Akt, resulting in cell proliferation, survival, and GSI resistance. This hypothesis predicts two consequences of GSI or other NOTCH1-inhibiting treatment. First, T-ALL patients who lack PTEN activity at disease onset may still respond to NOTCH1 inhibition. Second, NOTCH1 inhibition may trigger leukemic cells to acquire mutations such as *PTEN* deletions, which leads to activation of AKT and resistance to NOTCH1 inhibitors. Several key observations provide support for this hypothesis of reduced PI3K-AKT signaling over time in the absence of PTEN. For example, we found no difference in AKT phosphorylation between primary T-ALL patients with aberrant *PTEN* and patients without such mutations,(21) indicating that patients with *PTEN*-defective leukemia may have adapted and reduced PI3K-AKT signaling. Reduced AKT activation may explain why primary *PTEN*-defective T-ALL cells are NOTCH1-dependent and remain GSI-sensitive.(58) Although this hypothesis has not been tested formally, future NOTCH-inhibiting therapies may be more effective when combined with inhibitors of PI3K or AKT. Consistent with this, the PI3K/mTOR dual inhibitor PI-103 resulted in enhanced NOTCH-MYC activity in T-ALL cell lines.(61) Additionally, T-ALL induced by retroviral insertional mutagenesis in wild-type or *RasG12D*-mutant mice demonstrated an initial response to PI3K-inhibitor GDC-0941 treatment.(62) However, this treatment led to the survival and outgrowth of drug resistant clones with active PI3K-AKT signaling that frequently had reduced Notch1 signaling.(62) To avoid resistance when combined

with NOTCH1 inhibitors, the authors propose a sequential treatment using a NOTCH1 inhibitor at diagnosis to eliminate NOTCH1 mutant clones followed by PI3K/AKT inhibitor treatment.(62)

Other mechanisms that can activate AKT and lead to GSI resistance

Eighty-five percent of T-ALL patients have an activated AKT pathway accompanied by increased phosphorylation of AKT and its downstream targets GSK-3 β and FOXO3a.(23) Notably, this percentage is higher than the frequency of PTEN aberrations (23% of the patients) and therefore has to be explained by the activation of AKT through other mechanisms.

MYC may provide an alternative mechanism to activate AKT either directly or indirectly (e.g. MYC activates the expression of mir-17-92 and mir-19 that target *PTEN* mRNA)(63-65) (Figure 2C). Using an inducible MYC-dependent zebrafish T-ALL model, Gutierrez et al. found that established tumors regressed when MYC expression was turned off. This effect was circumvented by activating PI3K-AKT signaling, (66) showing that AKT activation is an important downstream effector of MYC that may drive GSI resistance. Moreover, the *MYC* gene is an important downstream target of NOTCH1, and T-ALL patients with activating mutations in *NOTCH1* overexpress MYC.(67, 68) NOTCH binds a distal enhancer located far downstream of the *MYC* locus.(69, 70) This NOTCH-MYC enhancer region (N-Me) is duplicated in approximately 5% of T-ALL patients, acting as a “super-enhancer”.(69) In another 6% of adult and childhood T-ALL patients, MYC is ectopically activated due to a *MYC* translocation; importantly, these patients usually do not have *NOTCH1*-activating mutations.(71) MYC may also activate NOTCH1 via a positive feedback mechanism, as MYC suppresses the expression of miRNA-30, which targets the 3' UTR of *NOTCH1* (Figure 2C).(72) Accordingly, treatment of T-ALL xenografted mice with the bromodomain protein inhibitor JQ1 results in decreased MYC levels and also reverses MYC-induced resistance to GSI. (73, 74) Also, in human T-ALL cell lines, GSI-sensitive cells can be converted to GSI-resistant by the ectopic expression of MYC.(68, 75) Under normal conditions, MYC is phosphorylated by the kinase GSK-3 β ; phosphorylated MYC is then subjected to ubiquitination by FBXW7 and proteasome-mediated degradation (Figure 2C).(76, 77) Conversely, activated AKT can stabilize MYC protein by phosphorylating—and thereby inactivating—GSK-3 β . Also, these findings may explain the observation that MYC and PTEN are reciprocally expressed in T-ALL.(78)

Apart from enhancing cellular resistance to NOTCH1 inhibitors, MYC also enhances leukemia-initiating cell (LIC) activity and poor outcome in various mouse models of T-ALL. Mutant *Fbxw7-R465C* mice develop aggressive leukemias that acquire *Notch1* mutations.(79) In

these mice, *Myc* levels are stabilized resulting in the expansion of leukemia cells that have enhanced self-renewal capacity and that express a stem cell-like expression profile.(79) The *Tal1/Lmo2* transgenic mouse model develops spontaneous T-cell tumors that also acquire *Notch1* mutations. Because *Myc* is a Notch target, Notch inhibition led to reduced LIC activity in these mice.(80) Reducing endogenous *Myc* levels led to increased survival and reduced numbers of leukemia cells with LIC potential in both models.(79, 81) Overall, these positive feedback loops between NOTCH, MYC, and AKT suggest that inhibitors of MYC or PI3K/AKT may help prevent resistance to NOTCH1-inhibiting therapies,(82) and also eliminate LIC activity in T-ALL. Co-targeting the PI3K pathway and MYC remarkably enhanced the elimination of LICs.(83)

Another AKT activation mechanism is via the gene that encodes the IL7 receptor (IL7Ra), that also represents a direct target gene of NOTCH1.(84, 85) The *IL7R* gene is mutated in nearly 10% of T-ALL patients. These mutations cause the constitutive activation of STAT5 and AKT,(8, 86, 87) and can provoke GSI resistance (Figure 2D). For instance, expressing the IL7Ra can overcome the effects of NOTCH1 inhibition on the cell cycle and survival, thereby contributing to resistance.(84) Similar results were obtained by overexpressing IGF1R, which encodes insulin-like growth factor 1 receptor and is another NOTCH1 target (Figure 2D).(88) Also in these cases, NOTCH-inhibiting therapies may be more effective when combined with AKT inhibitors. Furthermore, enhanced AKT activity may limit leukemia sensitivity to steroid treatment,(89, 90) one of the cornerstone drugs in the treatment of human T-ALL. AKT was shown to directly phosphorylate (S134) and inactivate the steroid receptor NR3C1.(89) Combined steroid treatment with the dual PI3K-mTOR inhibitor BEZ235 (91) or the MK2206 AKT inhibitor (89) sensitized AKT-activated leukemic cells to steroid treatment.

Conclusion

As a potent tumor suppressor, PTEN is considered to be the principal negative regulator of PI3K-AKT signaling. Inactivation of PTEN indirectly activates PI3K-AKT signaling, causing the uncontrolled proliferation of thymocytes, ultimately leading to T-ALL. Regardless of PTEN, AKT can be over-activated by a variety of signaling molecules, including PI3K, AKT, MYC, IL7R and IGF1R (Figure 2). Initial activation of AKT causes resistance to NOTCH1-inhibiting therapies. However, on the long-term, we suggest that AKT signaling may be dampened, thereby restoring responsiveness to NOTCH-inhibiting therapies. Overall, because AKT activation is central to a variety of leukemogenic mechanisms and crucial in the resistance to NOTCH1 inhibition, using

AKT inhibitors in current treatment protocols may be a promising strategy to treat NOTCH1-mutated T-ALL.

Acknowledgments

RDM and KC-B were financed by the Children Cancer Free Foundation (*Stichting Kinderen Kankervrij* (KiKa 2008-29 and KiKa 2013-116). We thank EnglishEditingSolutions.com for editorial assistance.

Authorship

Contribution: RDM, KC-B, RP, and JPPM wrote the manuscript.

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

References

1. Meijerink JP. Genetic rearrangements in relation to immunophenotype and outcome in T-cell acute lymphoblastic leukaemia. *Best Pract Res Clin Haematol*. 2010 Sep;23(3):307-18.
2. Van Vlierberghe P, Pieters R, Beverloo HB, Meijerink JP. Molecular-genetic insights in paediatric T-cell acute lymphoblastic leukaemia. *Br J Haematol*. 2008 Oct;143(2):153-68.
3. Homminga I, Pieters R, Langerak AW, de Rooi JJ, Stubbs A, Verstegen M, et al. Integrated transcript and genome analyses reveal NKX2-1 and MEF2C as potential oncogenes in T cell acute lymphoblastic leukemia. *Cancer Cell*. 2011 Apr 12;19(4):484-97.
4. Ferrando AA, Neuberg DS, Staunton J, Loh ML, Huard C, Raimondi SC, et al. Gene expression signatures define novel oncogenic pathways in T cell acute lymphoblastic leukemia. *Cancer Cell*. 2002 Feb;1(1):75-87.
5. Soulier J, Clappier E, Cayuela JM, Regnault A, Garcia-Peydro M, Dombret H, et al. HOXA genes are included in genetic and biologic networks defining human acute T-cell leukemia (T-ALL). *Blood*. 2005 Jul 1;106(1):274-86.
6. Notta F, Mullighan CG, Wang JC, Poepl A, Doulatov S, Phillips LA, et al. Evolution of human BCR-ABL1 lymphoblastic leukaemia-initiating cells. *Nature*. 2011 Jan 20;469(7330):362-7.
7. Anderson K, Lutz C, van Delft FW, Bateman CM, Guo Y, Colman SM, et al. Genetic variegation of clonal architecture and propagating cells in leukaemia. *Nature*. 2011 Jan 20;469(7330):356-61.
8. Zenatti PP, Ribeiro D, Li W, Zuurbier L, Silva MC, Paganin M, et al. Oncogenic IL7R gain-of-function mutations in childhood T-cell acute lymphoblastic leukemia. *Nat Genet*. 2011 Oct;43(10):932-9.
9. Zhang J, Ding L, Holmfeldt L, Wu G, Heatley SL, Payne-Turner D, et al. The genetic basis of early T-cell precursor acute lymphoblastic leukaemia. *Nature*. 2012 Jan 12;481(7380):157-63.
10. Cante-Barrett K, Spijkers-Hagelstein JA, Buijs-Gladdines JG, Uitdehaag JC, Smits WK, van der Zwet J, et al. MEK and PI3K-AKT inhibitors synergistically block activated IL7 receptor signaling in T-cell acute lymphoblastic leukemia. *Leukemia*. 2016 May 13.
11. Filippakopoulos P, Qi J, Picaud S, Shen Y, Smith WB, Fedorov O, et al. Selective inhibition of BET bromodomains. *Nature*. 2010 Dec 23;468(7327):1067-73.
12. Li J, Yen C, Liaw D, Podsypanina K, Bose S, Wang SI, et al. PTEN, a putative protein tyrosine phosphatase gene mutated in human brain, breast, and prostate cancer. *Science*. 1997 Mar 28;275(5308):1943-7.
13. Steck PA, Pershouse MA, Jasser SA, Yung WK, Lin H, Ligon AH, et al. Identification of a candidate tumour suppressor gene, MMAC1, at chromosome 10q23.3 that is mutated in multiple advanced cancers. *Nat Genet*. 1997 Apr;15(4):356-62.
14. Vogelstein B, Kinzler KW. Cancer genes and the pathways they control. *Nat Med*. 2004 Aug;10(8):789-99.
15. Liang H, He S, Yang J, Jia X, Wang P, Chen X, et al. PTENalpha, a PTEN isoform translated through alternative initiation, regulates mitochondrial function and energy metabolism. *Cell metabolism*. 2014 May 6;19(5):836-48.
16. Marsh DJ, Coulon V, Lunetta KL, Rocca-Serra P, Dahia PL, Zheng Z, et al. Mutation spectrum and genotype-phenotype analyses in Cowden disease and Bannayan-Zonana syndrome, two hamartoma syndromes with germline PTEN mutation. *Hum Mol Genet*. 1998 Mar;7(3):507-15.
17. Salmena L, Carracedo A, Pandolfi PP. Tenets of PTEN tumor suppression. *Cell*. 2008 May 2;133(3):403-14.
18. Song MS, Salmena L, Pandolfi PP. The functions and regulation of the PTEN tumour suppressor. *Nat Rev Mol Cell Biol*. 2012 May;13(5):283-96.

19. Hollander MC, Blumenthal GM, Dennis PA. PTEN loss in the continuum of common cancers, rare syndromes and mouse models. *Nat Rev Cancer*. 2011 Apr;11(4):289-301.
20. Leupin N, Cenni B, Novak U, Hugli B, Graber HU, Tobler A, et al. Disparate expression of the PTEN gene: a novel finding in B-cell chronic lymphocytic leukaemia (B-CLL). *Br J Haematol*. 2003 Apr;121(1):97-100.
21. Zuurbier L, Petricoin EF, Vuerhard MJ, Calvert V, Kooi C, Buijs-Gladdines JGCAM, et al. The significance of PTEN and AKT aberrations in pediatric T-cell acute lymphoblastic leukemia. *Haematol-Hematol J*. 2012 Sep;97(9):1405-13.
22. Mutter GL, Lin MC, Fitzgerald JT, Kum JB, Baak JP, Lees JA, et al. Altered PTEN expression as a diagnostic marker for the earliest endometrial precancers. *J Natl Cancer Inst*. 2000 Jun 7;92(11):924-30.
23. Silva A, Yunes JA, Cardoso BA, Martins LR, Jotta PY, Abecasis M, et al. PTEN posttranslational inactivation and hyperactivation of the PI3K/Akt pathway sustain primary T cell leukemia viability. *The Journal of clinical investigation*. 2008 Nov;118(11):3762-74.
24. Gomes AM, Soares MV, Ribeiro P, Caldas J, Pova V, Martins LR, et al. Adult B-cell acute lymphoblastic leukemia cells display decreased PTEN activity and constitutive hyperactivation of PI3K/Akt pathway despite high PTEN protein levels. *Haematologica*. 2014 Jun;99(6):1062-8.
25. Alimonti A, Carracedo A, Clohessy JG, Trotman LC, Nardella C, Egia A, et al. Subtle variations in Pten dose determine cancer susceptibility. *Nat Genet*. 2010 May;42(5):454-8.
26. Sakai A, Thieblemont C, Wellmann A, Jaffe ES, Raffeld M. PTEN gene alterations in lymphoid neoplasms. *Blood*. 1998 Nov 1;92(9):3410-5.
27. Shan X, Czar MJ, Bunnell SC, Liu P, Liu Y, Schwartzberg PL, et al. Deficiency of PTEN in Jurkat T cells causes constitutive localization of Itk to the plasma membrane and hyperresponsiveness to CD3 stimulation. *Mol Cell Biol*. 2000 Sep;20(18):6945-57.
28. Xu Z, Stokoe D, Kane LP, Weiss A. The inducible expression of the tumor suppressor gene PTEN promotes apoptosis and decreases cell size by inhibiting the PI3K/Akt pathway in Jurkat T cells. *Cell Growth Differ*. 2002 Jul;13(7):285-96.
29. Palomero T, Sulis ML, Cortina M, Real PJ, Barnes K, Ciofani M, et al. Mutational loss of PTEN induces resistance to NOTCH1 inhibition in T-cell leukemia. *Nat Med*. 2007 Oct;13(10):1203-10.
30. Gutierrez A, Sanda T, Grebliunaite R, Carracedo A, Salmena L, Ahn Y, et al. High frequency of PTEN, PI3K, and AKT abnormalities in T-cell acute lymphoblastic leukemia. *Blood*. 2009 Jul 16;114(3):647-50.
31. Larson Gedman A, Chen Q, Kugel Desmoulin S, Ge Y, LaFiura K, Haska CL, et al. The impact of NOTCH1, FBW7 and PTEN mutations on prognosis and downstream signaling in pediatric T-cell acute lymphoblastic leukemia: a report from the Children's Oncology Group. *Leukemia*. 2009 Aug;23(8):1417-25.
32. Bandapalli OR, Zimmermann M, Kox C, Stanulla M, Schrappe M, Ludwig WD, et al. NOTCH1 activation clinically antagonizes the unfavorable effect of PTEN inactivation in BFM-treated children with precursor T-cell acute lymphoblastic leukemia. *Haematologica*. 2013 Jun;98(6):928-36.
33. Jotta PY, Ganazza MA, Silva A, Viana MB, da Silva MJ, Zambaldi LJ, et al. Negative prognostic impact of PTEN mutation in pediatric T-cell acute lymphoblastic leukemia. *Leukemia*. 2010 Jan;24(1):239-42.
34. Remke M, Pfister S, Kox C, Toedt G, Becker N, Benner A, et al. High-resolution genomic profiling of childhood T-ALL reveals frequent copy-number alterations affecting the TGF-beta and PI3K-AKT pathways and deletions at 6q15-16.1 as a genomic marker for unfavorable early treatment response. *Blood*. 2009 Jul 30;114(5):1053-62.

35. Mendes RD, Sarmiento LM, Cante-Barrett K, Zuurbier L, Buijs-Gladdines JG, Pova V, et al. PTEN microdeletions in T-cell acute lymphoblastic leukemia are caused by illegitimate RAG-mediated recombination events. *Blood*. 2014 Jul 24;124(4):567-78.
36. Trinquand A, Tanguy-Schmidt A, Ben Abdelali R, Lambert J, Beldjord K, Lengline E, et al. Toward a NOTCH1/FBXW7/RAS/PTEN-based oncogenetic risk classification of adult T-cell acute lymphoblastic leukemia: a Group for Research in Adult Acute Lymphoblastic Leukemia study. *J Clin Oncol*. 2013 Dec 1;31(34):4333-42.
37. Georgescu MM, Kirsch KH, Akagi T, Shishido T, Hanafusa H. The tumor-suppressor activity of PTEN is regulated by its carboxyl-terminal region. *Proc Natl Acad Sci U S A*. 1999 Aug 31;96(18):10182-7.
38. Sun Z, Huang C, He J, Lamb KL, Kang X, Gu T, et al. PTEN C-terminal deletion causes genomic instability and tumor development. *Cell Rep*. 2014 Mar 13;6(5):844-54.
39. Clappier E, Gerby B, Sigaux F, Delord M, Touzri F, Hernandez L, et al. Clonal selection in xenografted human T cell acute lymphoblastic leukemia recapitulates gain of malignancy at relapse. *J Exp Med*. 2011 Apr 11;208(4):653-61.
40. Di Cristofano A, Pesce B, Cordon-Cardo C, Pandolfi PP. Pten is essential for embryonic development and tumour suppression. *Nat Genet*. 1998 Aug;19(4):348-55.
41. Suzuki A, de la Pompa JL, Stambolic V, Elia AJ, Sasaki T, del Barco Barrantes I, et al. High cancer susceptibility and embryonic lethality associated with mutation of the PTEN tumor suppressor gene in mice. *Curr Biol*. 1998 Oct 22;8(21):1169-78.
42. Di Cristofano A, Kotsi P, Peng YF, Cordon-Cardo C, Elkon KB, Pandolfi PP. Impaired Fas response and autoimmunity in Pten^{+/-} mice. *Science*. 1999 Sep 24;285(5436):2122-5.
43. Giambra V, Jenkins CR, Wang H, Lam SH, Shevchuk OO, Nemirovsky O, et al. NOTCH1 promotes T cell leukemia-initiating activity by RUNX-mediated regulation of PKC-theta and reactive oxygen species. *Nat Med*. 2012 Nov;18(11):1693-8.
44. Jenkinson S, Kirkwood AA, Goulden N, Vora A, Linch DC, Gale RE. Impact of PTEN abnormalities on outcome in pediatric patients with T-cell acute lymphoblastic leukemia treated on the MRC UKALL2003 trial. *Leukemia*. 2016 Jan;30(1):39-47.
45. Weng AP, Ferrando AA, Lee W, Morris JPt, Silverman LB, Sanchez-Irizarry C, et al. Activating mutations of NOTCH1 in human T cell acute lymphoblastic leukemia. *Science*. 2004 Oct 8;306(5694):269-71.
46. Ferrando AA. The role of NOTCH1 signaling in T-ALL. *Hematology Am Soc Hematol Educ Program*. 2009:353-61.
47. Real PJ, Tosello V, Palomero T, Castillo M, Hernando E, de Stanchina E, et al. Gamma-secretase inhibitors reverse glucocorticoid resistance in T cell acute lymphoblastic leukemia. *Nat Med*. 2009 Jan;15(1):50-8.
48. Samon JB, Castillo-Martin M, Hadler M, Ambesi-Impibato A, Paietta E, Racevskis J, et al. Preclinical analysis of the gamma-secretase inhibitor PF-03084014 in combination with glucocorticoids in T-cell acute lymphoblastic leukemia. *Mol Cancer Ther*. 2012 Jul;11(7):1565-75.
49. Lewis HD, Leveridge M, Strack PR, Haldon CD, O'Neil J, Kim H, et al. Apoptosis in T cell acute lymphoblastic leukemia cells after cell cycle arrest induced by pharmacological inhibition of notch signaling. *Chem Biol*. 2007 Feb;14(2):209-19.
50. Hales EC, Taub JW, Matherly LH. New insights into Notch1 regulation of the PI3K-AKT-mTOR1 signaling axis: targeted therapy of gamma-secretase inhibitor resistant T-cell acute lymphoblastic leukemia. *Cell Signal*. 2014 Jan;26(1):149-61.
51. Palomero T, Ferrando A. Oncogenic NOTCH1 control of MYC and PI3K: challenges and opportunities for anti-NOTCH1 therapy in T-cell acute lymphoblastic leukemias and lymphomas. *Clin Cancer Res*. 2008 Sep 1;14(17):5314-7.
52. Deangelo DJ, Stone RM, Silverman LB, Stock W, Attar EC, Fearen I, et al. A phase I clinical trial of the notch inhibitor MK-0752 in patients with T-cell acute lymphoblastic

- leukemia/lymphoma (T-ALL) and other leukemias. *J Clin Oncol* (Meeting Abstracts). 2006 June 20, 2006;24(18_suppl):6585-.
53. Tosello V, Ferrando AA. The NOTCH signaling pathway: role in the pathogenesis of T-cell acute lymphoblastic leukemia and implication for therapy. *Ther Adv Hematol*. 2013 Jun;4(3):199-210.
 54. Aste-Amezaga M, Zhang N, Lineberger JE, Arnold BA, Toner TJ, Gu M, et al. Characterization of Notch1 antibodies that inhibit signaling of both normal and mutated Notch1 receptors. *PLoS One*. 2010;5(2):e9094.
 55. Wu Y, Cain-Hom C, Choy L, Hagenbeek TJ, de Leon GP, Chen Y, et al. Therapeutic antibody targeting of individual Notch receptors. *Nature*. 2010 Apr 15;464(7291):1052-7.
 56. Moellering RE, Cornejo M, Davis TN, Del Bianco C, Aster JC, Blacklow SC, et al. Direct inhibition of the NOTCH transcription factor complex. *Nature*. 2009 Nov 12;462(7270):182-8.
 57. Hagenbeek TJ, Wu X, Choy L, Sanchez-Irizarry C, Seshagiri S, Stinson J, et al. Murine Pten(-/-) T-ALL requires non-redundant PI3K/mTOR and DLL4/Notch1 signals for maintenance and gammac/TCR signals for thymic exit. *Cancer Lett*. 2014 May 1;346(2):237-48.
 58. Madyouf H, Gao X, Armstrong F, Gusscott S, Liu Q, Gedman AL, et al. Acute T-cell leukemias remain dependent on Notch signaling despite PTEN and INK4A/ARF loss. *Blood*. 2010 Feb 11;115(6):1175-84.
 59. Zuurbier L, Homminga I, Calvert V, te Winkel ML, Buijs-Gladdines JG, Kooi C, et al. NOTCH1 and/or FBXW7 mutations predict for initial good prednisone response but not for improved outcome in pediatric T-cell acute lymphoblastic leukemia patients treated on DCOG or COALL protocols. *Leukemia*. 2010 Dec;24(12):2014-22.
 60. Herranz D, Ambesi-Impiombato A, Sudderth J, Sanchez-Martin M, Belver L, Tosello V, et al. Metabolic reprogramming induces resistance to anti-NOTCH1 therapies in T cell acute lymphoblastic leukemia. *Nat Med*. 2015 Oct;21(10):1182-9.
 61. Shepherd C, Banerjee L, Cheung CW, Mansour MR, Jenkinson S, Gale RE, et al. PI3K/mTOR inhibition upregulates NOTCH-MYC signalling leading to an impaired cytotoxic response. *Leukemia*. 2013 Mar;27(3):650-60.
 62. Dail M, Wong J, Lawrence J, O'Connor D, Nakitandwe J, Chen SC, et al. Loss of oncogenic Notch1 with resistance to a PI3K inhibitor in T-cell leukaemia. *Nature*. 2014 Sep 25;513(7519):512-6.
 63. Mu P, Han YC, Betel D, Yao E, Squatrito M, Ogradowski P, et al. Genetic dissection of the miR-17~92 cluster of microRNAs in Myc-induced B-cell lymphomas. *Genes Dev*. 2009 Dec 15;23(24):2806-11.
 64. Mavrakis KJ, Wolfe AL, Oricchio E, Palomero T, de Keersmaecker K, McJunkin K, et al. Genome-wide RNA-mediated interference screen identifies miR-19 targets in Notch-induced T-cell acute lymphoblastic leukaemia. *Nat Cell Biol*. 2010 Apr;12(4):372-9.
 65. Olive V, Bennett MJ, Walker JC, Ma C, Jiang I, Cordon-Cardo C, et al. miR-19 is a key oncogenic component of mir-17-92. *Genes Dev*. 2009 Dec 15;23(24):2839-49.
 66. Gutierrez A, Grebliunaite R, Feng H, Kozakewich E, Zhu S, Guo F, et al. Pten mediates Myc oncogene dependence in a conditional zebrafish model of T cell acute lymphoblastic leukemia. *J Exp Med*. 2011 Aug 1;208(8):1595-603.
 67. Palomero T, Lim WK, Odom DT, Sulis ML, Real PJ, Margolin A, et al. NOTCH1 directly regulates c-MYC and activates a feed-forward-loop transcriptional network promoting leukemic cell growth. *Proc Natl Acad Sci U S A*. 2006 Nov 28;103(48):18261-6.
 68. Weng AP, Millholland JM, Yashiro-Ohtani Y, Arcangeli ML, Lau A, Wai C, et al. c-Myc is an important direct target of Notch1 in T-cell acute lymphoblastic leukemia/lymphoma. *Genes Dev*. 2006 Aug 1;20(15):2096-109.
 69. Herranz D, Ambesi-Impiombato A, Palomero T, Schnell SA, Belver L, Wendorff AA, et al. A NOTCH1-driven MYC enhancer promotes T cell development, transformation and acute lymphoblastic leukemia. *Nat Med*. 2014 Oct;20(10):1130-7.

70. Yashiro-Ohtani Y, Wang H, Zang C, Arnett KL, Bailis W, Ho Y, et al. Long-range enhancer activity determines Myc sensitivity to Notch inhibitors in T cell leukemia. *Proc Natl Acad Sci U S A*. 2014 Nov 18;111(46):E4946-53.
71. La Starza R, Borga C, Barba G, Pierini V, Schwab C, Matteucci C, et al. Genetic profile of T-cell acute lymphoblastic leukemias with MYC translocations. *Blood*. 2014 Dec 4;124(24):3577-82.
72. Ortega M, Bhatnagar H, Lin AP, Wang L, Aster JC, Sill H, et al. A microRNA-mediated regulatory loop modulates NOTCH and MYC oncogenic signals in B- and T-cell malignancies. *Leukemia*. 2015 Apr;29(4):968-76.
73. Delmore JE, Issa GC, Lemieux ME, Rahl PB, Shi J, Jacobs HM, et al. BET bromodomain inhibition as a therapeutic strategy to target c-Myc. *Cell*. 2011 Sep 16;146(6):904-17.
74. Knoechel B, Roderick JE, Williamson KE, Zhu J, Lohr JG, Cotton MJ, et al. An epigenetic mechanism of resistance to targeted therapy in T cell acute lymphoblastic leukemia. *Nat Genet*. 2014 Apr;46(4):364-70.
75. Sharma VM, Calvo JA, Draheim KM, Cunningham LA, Hermance N, Beverly L, et al. Notch1 contributes to mouse T-cell leukemia by directly inducing the expression of c-myc. *Mol Cell Biol*. 2006 Nov;26(21):8022-31.
76. Sears R, Nuckolls F, Haura E, Taya Y, Tamai K, Nevins JR. Multiple Ras-dependent phosphorylation pathways regulate Myc protein stability. *Genes Dev*. 2000 Oct 1;14(19):2501-14.
77. Cross DA, Alessi DR, Cohen P, Andjelkovich M, Hemmings BA. Inhibition of glycogen synthase kinase-3 by insulin mediated by protein kinase B. *Nature*. 1995 Dec 21-28;378(6559):785-9.
78. Bonnet M, Loosveld M, Montpellier B, Navarro JM, Quilichini B, Picard C, et al. Posttranscriptional deregulation of MYC via PTEN constitutes a major alternative pathway of MYC activation in T-cell acute lymphoblastic leukemia. *Blood*. 2011 Jun 16;117(24):6650-9.
79. King B, Trimarchi T, Reavie L, Xu L, Mullenders J, Ntziachristos P, et al. The ubiquitin ligase FBXW7 modulates leukemia-initiating cell activity by regulating MYC stability. *Cell*. 2013 Jun 20;153(7):1552-66.
80. Tatarek J, Cullion K, Ashworth T, Gerstein R, Aster JC, Kelliher MA. Notch1 inhibition targets the leukemia-initiating cells in a Tal1/Lmo2 mouse model of T-ALL. *Blood*. 2011 Aug 11;118(6):1579-90.
81. Roderick JE, Tesell J, Shultz LD, Brehm MA, Greiner DL, Harris MH, et al. c-Myc inhibition prevents leukemia initiation in mice and impairs the growth of relapsed and induction failure pediatric T-ALL cells. *Blood*. 2014 Feb 13;123(7):1040-50.
82. Roti G, Stegmaier K. New Approaches to Target T-ALL. *Frontiers in oncology*. 2014;4:170.
83. Schubbert S, Cardenas A, Chen H, Garcia C, Guo W, Bradner J, et al. Targeting the MYC and PI3K pathways eliminates leukemia-initiating cells in T-cell acute lymphoblastic leukemia. *Cancer Res*. 2014 Dec 1;74(23):7048-59.
84. Gonzalez-Garcia S, Garcia-Peydro M, Martin-Gayo E, Ballestar E, Esteller M, Bornstein R, et al. CSL-MAML-dependent Notch1 signaling controls T lineage-specific IL-7R{alpha} gene expression in early human thymopoiesis and leukemia. *J Exp Med*. 2009 Apr 13;206(4):779-91.
85. Ribeiro D, Melao A, Barata JT. IL-7R-mediated signaling in T-cell acute lymphoblastic leukemia. *Adv Biol Regul*. 2013 May;53(2):211-22.
86. Barata JT, Cardoso AA, Boussiotis VA. Interleukin-7 in T-cell acute lymphoblastic leukemia: an extrinsic factor supporting leukemogenesis? *Leuk Lymphoma*. 2005 Apr;46(4):483-95.
87. Shochat C, Tal N, Bandapalli OR, Palmi C, Ganmore I, te Kronnie G, et al. Gain-of-function mutations in interleukin-7 receptor-alpha (IL7R) in childhood acute lymphoblastic leukemias. *J Exp Med*. 2011 May 9;208(5):901-8.

88. Medyouf H, Gusscott S, Wang H, Tseng JC, Wai C, Nemirovsky O, et al. High-level IGF1R expression is required for leukemia-initiating cell activity in T-ALL and is supported by Notch signaling. *J Exp Med*. 2011 Aug 29;208(9):1809-22.
89. Piovan E, Yu J, Tosello V, Herranz D, Ambesi-Impiombato A, Da Silva AC, et al. Direct reversal of glucocorticoid resistance by AKT inhibition in acute lymphoblastic leukemia. *Cancer Cell*. 2013 Dec 9;24(6):766-76.
90. Blackburn JS, Liu S, Wilder JL, Dobrinski KP, Lobbardi R, Moore FE, et al. Clonal evolution enhances leukemia-propagating cell frequency in T cell acute lymphoblastic leukemia through Akt/mTORC1 pathway activation. *Cancer Cell*. 2014 Mar 17;25(3):366-78.
91. Hall CP, Reynolds CP, Kang MH. Modulation of Glucocorticoid Resistance in Pediatric T-cell Acute Lymphoblastic Leukemia by Increasing BIM Expression with the PI3K/mTOR Inhibitor BEZ235. *Clin Cancer Res*. 2016 Feb 1;22(3):621-32.

Figure Legends

Figure 1. Schematic representation of the human *PTEN* gene located on chromosome 10q23. The *PTEN* gene contains nine exons, and the PTEN protein contains several functional domains, including a phosphatase domain (dark gray) and a C2 lipid-binding domain (light gray). The positions of nonsense insertion and deletion mutations are indicated by closed triangles, and missense mutations are indicated by open triangles. Microdeletions and deletions in the *PTEN* gene are shown below the exons. The number of patients with each mutation/deletion in our cohort of T-ALL patient samples is indicated.(21, 35)

Figure 2. Schematic overview of the upstream and downstream effectors of PTEN and associated molecular mechanisms that can activate AKT and lead to GSI resistance. (A) PTEN-PI3K-AKT mutations. (B) NOTCH mutations and AKT activation. (C) MYC signaling and AKT activation. (D) IL7R/IGF1R signaling and AKT activation. The molecules with activating and inactivating mutations are indicated in dark gray and light gray, respectively. Dashed lines/arrows represent processes that contribute to cellular GSI sensitivity, and that are frequently inactivated by inactivating mutations/rearrangements in *PTEN* or *FBXW7*. Solid lines/arrows represent GSI resistance mechanisms.

Figure 1

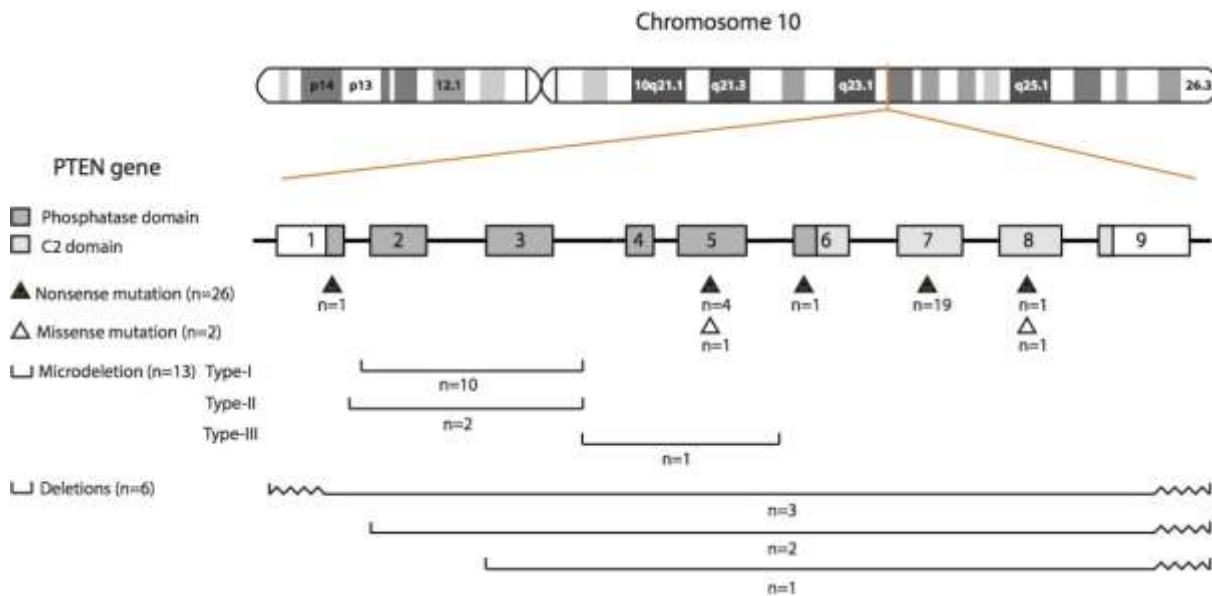
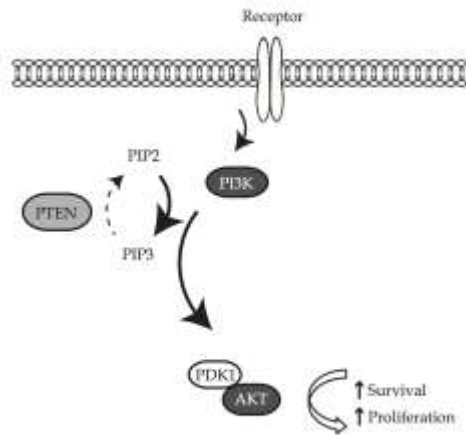
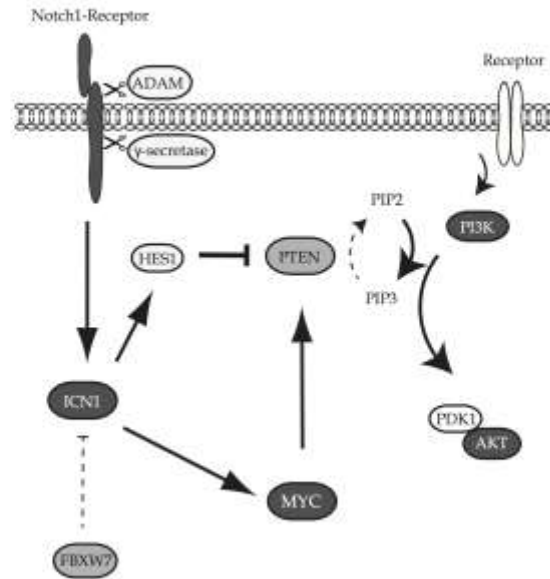


Figure 2

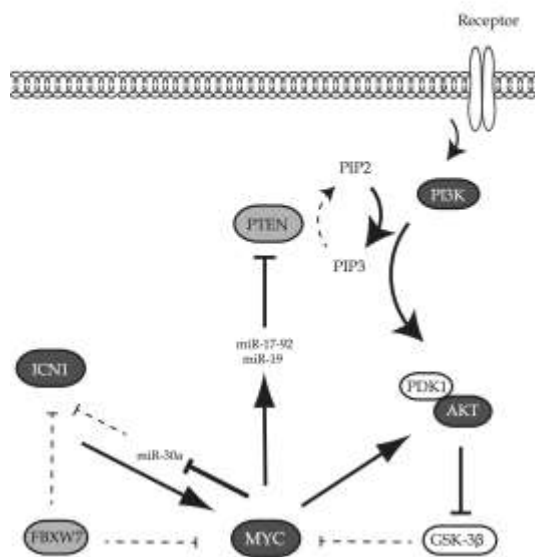
A) PTEN-PI3K-AKT mutations



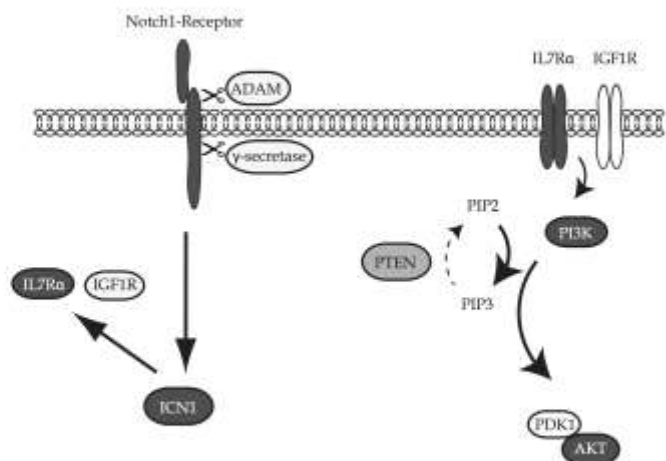
B) NOTCH1 mutations and AKT activation



C) MYC signaling and AKT activation



D) IL7R/IGF1R signaling and AKT activation



Chapter 3

PTEN microdeletions in T-cell acute lymphoblastic leukemia are caused by illegitimate RAG-mediated recombination events

Rui D. Mendes^{1,9}, Leonor M. Sarmento^{2,9}, Kirsten Canté-Barrett¹, Linda Zuurbier¹, Jessica G.C.A.M. Buijs-Gladdines¹, Vanda Póvoa², Willem K. Smits¹, Miguel Abecasis³, J. Andres Yunes⁴, Edwin Sonneveld⁵, Martin A. Horstmann^{6,7}, Rob Pieters^{1,8}, João T. Barata^{2,10} and Jules P.P. Meijerink^{1,10}

¹Department of Pediatric Oncology/Hematology, Erasmus MC Rotterdam-Sophia Children's Hospital, Rotterdam, the Netherlands; ²Instituto de Medicina Molecular, Faculdade de Medicina da Universidade de Lisboa, Lisboa, Portugal; ³Cardiologia Pediátrica Medico Cirúrgica, Hospital Sta. Cruz, Lisboa, Portugal; ⁴Centro Infantil Boldrini, Campinas, SP, Brazil; ⁵Dutch Childhood Oncology Group (DCOG), the Hague, the Netherlands; ⁶German Cooperative Study Group for Childhood Acute Lymphoblastic Leukemia (COALL), Hamburg, Germany; ⁷Research Institute Children's Cancer Center Hamburg, Clinic of Pediatric Hematology and Oncology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany; ⁸Princess Maxima Center for Pediatric Oncology, Utrecht, Netherlands

⁹These authors are co-first authors; ¹⁰These authors contributed equally to this work

Blood, 2014, 124(4):567-78

Abstract

PTEN inactivating mutations and/or deletions are an independent risk factor for relapse just like NOTCH-activating mutations or male gender for T-cell acute lymphoblastic leukemia (T-ALL) patients treated on DCOG or COALL protocols. Some monoallelic mutated or *PTEN* wild-type patients lack PTEN protein, implying that additional *PTEN* inactivation mechanisms exist. We show that *PTEN* is inactivated by small deletions affecting only a few exons in 8% of pediatric T-ALL patients. These microdeletions were clonal in 3% and sub-clonal in 5% of patients. Conserved deletion breakpoints are flanked by cryptic RAG-recombination signal sequences (cRSS) and frequently have non-template derived nucleotides inserted in between breakpoints, implying that microdeletions are the result of illegitimate RAG recombination activity. Identified cRSSs drive RAG-dependent recombination in a reporter system as efficiently as *bona fide* RSSs that flank gene segments of the T-cell receptor locus. Remarkably, equivalent microdeletions were also detected in thymocytes of healthy individuals. Similar to other *PTEN* aberrations, microdeletions strongly associate with the TALLMO-subtype characterized by *TAL1* or *LMO2* rearrangements. Primary and secondary xenotransplantation of *TAL1*-rearranged leukemia allowed development of leukemic subclones with newly acquired *PTEN* microdeletions. Ongoing RAG activity may therefore actively contribute to the acquisition of pre-leukemic hits, clonal diversification and disease progression.

Introduction

T-cell acute lymphoblastic leukemia (T-ALL) represents 10-15% of pediatric acute leukemias. Despite major therapeutic improvements due to treatment intensification and refined risk-adapted stratification during the past decade, ~30% of T-ALL cases relapse with very poor prognosis.¹

T-cell transformation is characterized by aberrant expression of oncogenic transcription factors combined with inactivation of tumor suppressor genes (e.g. *PTEN*, *CDKN2A*) and/or activation of the NOTCH1 pathway.² The ectopic expression of oncogenes is typically caused by chromosomal rearrangements—the so-called type A hits—that place oncogenes under the control of T-cell specific promoters or enhancer elements.^{3,4} The analysis of translocation breakpoints revealed frequent involvement of illegitimate V(D)J recombination in these translocations by binding of recombination activating gene (RAG)-1/2 proteins to sequences that resemble authentic recombination signal sequences (RSS).⁵ These recurrent chromosomal rearrangements activate several oncogenes, such as *TAL1*, *LMO2*, *TLX3*, *TLX1* or *NKX2-1/NKX2-2*, which are believed to represent the clonal disease drivers.^{2,6}

Besides near mutually exclusive type A mutations, recurrent genetic aberrations that affect cell viability and/or proliferation—the so-called type B hits—are found in nearly all T-ALL genetic subgroups. Type B mutations include NOTCH1-activating mutations affecting *NOTCH1* and *FBXW7* that are found in over 60% of pediatric T-ALL patients⁷⁻¹¹ (reviewed in Ferrando¹²), as well as less frequent events such as *IL7R* mutations in around 10% of T-ALL cases.^{13,14} In addition, mutations in the phosphatase and tensin homolog (*PTEN*) tumor suppressor gene have been associated with poor prognosis,¹⁵⁻¹⁸ resulting in overactive PI3K–AKT signaling that drives enhanced cell proliferation and cell metabolism, and impairs apoptosis.^{16,19,20} *PTEN* is considered to be a haploinsufficient tumor suppressor gene, because *PTEN* dose determines cancer susceptibility.²¹⁻²³ The majority of *PTEN* aberrations in T-ALL are deletions affecting the entire *PTEN* locus or mutations that truncate the membrane binding C2-domain.^{15,18}

In our previous studies, we detected *PTEN* aberrations in 13-20% of T-ALL patients^{18,24} and revealed that those mutations are especially associated with *TAL* or *LMO* rearrangements and nearly absent in *TLX3*-rearranged T-ALL.¹⁸ In general, *PTEN* mutated T-ALL appears to be devoid of NOTCH1-activating mutations.¹⁸ Interestingly, we did not observe differential *AKT* activation when comparing *PTEN* mutant/deleted with wild-type patient samples, indicating that other mechanisms may influence the PI3K-AKT pathway. In this respect, non-deletional posttranslational inactivation of *PTEN*,²⁴ rare mutations in *PIK3CA* (encoding PI3K) and *AKT*

themselves¹⁶, or PI3K-AKT pathway activation downstream of activated NOTCH1 have been described.¹⁵ However, none of these mechanisms explain the absence of PTEN protein in some T-ALL patient samples that have retained at least one *PTEN* wild-type allele.¹⁸

In this study, we have used multiplex ligation-dependent probe amplification (MLPA) to investigate copy-number variations among *PTEN* exons to detect potential additional *PTEN* deletions. We identified *PTEN* microdeletions in T-ALL patient samples and we provide evidence that these are driven by illegitimate RAG-mediated recombination events.

Materials and Methods

Patients

A total of 146 primary pediatric T-ALL patient samples were enrolled in the Dutch Childhood Oncology Group (DCOG) protocols ($n=72$)²⁵⁻²⁷ or the German Co-Operative Study Group for Childhood Acute Lymphoblastic Leukemia study (COALL-97) ($n=74$) were included in this study.²⁸ Normal thymocytes were isolated from thymic tissue obtained from children undergoing cardiac surgery.^{24,30} Informed consents were in accordance with the Institutional Review Boards of the ErasmusMC (Rotterdam, the Netherlands), the Ethics Committee of the City of Hamburg, Germany, the Hospital Sta. Cruz, Centro Hospitalar de Lisboa Ocidental (Lisboa, Portugal) and the Declaration of Helsinki.

Computational detection of putative RAG recombination signal sequences

The human *PTEN* gene (ENSG00000171862) was screened for the presence of cryptic RAG recombination signal sequences (cRSS) using the PERL software algorithms developed by Cowell et al.²⁹

Generation of GFPi-*PTEN* cRSS reporter constructs and recombination assay

To measure efficiencies of predicted cRSSs in mediating recombination of GFPi-mRFP reporter constructs, PCR amplified *PTEN* cRSS1-4 or defined RSS control sequences were cloned into this reporter construct and recombination assays were carried out as described.³⁰

Statistics

Statistics was performed using IBM SPSS Statistics 21 software (IBM, Armonk, NY, USA). The Pearson's Chi-square was used for nominal distributed data, the Fisher's exact test was alternatively used in case the number of patients in individual groups was lower than five. The Mann-Whitney-U test was used for continuously distributed data. Differences in relapse-free survival were tested using the log-rank test. Proportional risk for relapse was done by univariate and multivariate Cox regression analyses. The recombination efficiencies of cRSSs were compared using a one-way ANOVA with the Bonferroni's multiple comparison post-test. Data were considered significant when $p \leq 0.05$ (two-sided).

See Supplementary Materials and Methods for further experimental details.

Results

***PTEN* microdeletions in T-ALL patients**

In our previous study, we identified various T-ALL primary patient samples that lack *PTEN* protein expression and seemed *PTEN* wild-type or that contained an inactivating mutation or *PTEN* deletion in only one allele (summarized in Table 1).¹⁸ To identify additional *PTEN* inactivating mechanisms, we performed MLPA analysis to screen for potential microdeletions affecting single or a few *PTEN* exons that had been missed by array-CGH and FISH analyses. We analyzed 146 T-ALL patient samples for copy-number alterations in any of all 9 coding exons. Heterozygous microdeletions were detected in 3 T-ALL patients (Figure 1A), encompassing exons 2 and 3 in 2 patients (#21, #11) and exons 4 and 5 in another patient (#20). Accordingly, 2 out of the 3 patients (#11, #20) demonstrated defective *PTEN*-splicing with previously unknown underlying genetic aberrancies.¹⁸ A fourth patient (#12) was identified with a homozygous deletion of exons 2 and 3 that was confirmed by high-resolution array-CGH analysis (Figure 1A,B).

In order to clone the breakpoint regions of these microdeletions, a PCR-based strategy was designed for introns 1 and 3 (patients #21, #11 and #12) and intron 3 and 5 (patient #20) (Figure 2A,B), and resulting positive reactions were cloned and sequenced. These analyses predicted microdeletions of ~65Kb that encompassed exons 2-3 and of ~11Kb that encompassed exons 4-5 (Figure 2B). The homozygously deleted patient (#12) revealed different breakpoints that point to independent deletion events for each allele, with insertion of random bases in between breakpoints for one allele (Figure 2B). Breakpoints for the exon 2-3

microdeletion in patient #21 were identical to breakpoints of one allele of patient #12 and also lacked insertion of random bases. T-ALL patient #11 had a similar exon 2-3 deletion that shared the identical breakpoint in intron 3 but had an alternate breakpoint in intron 1 (Figure 2A-C). Breakpoints for the fourth T-ALL patient #20 with a microdeletion that affected exons 4 and 5 were located in introns 3 and 5 (Figure 2B). All three types of microdeletions result in out-of-frame *PTEN* transcripts (Figure 2C).

As MLPA does not allow for sensitive detection of subclones with microdeletions, we performed PCR analysis to screen the T-ALL cohort for similar microdeletions. Seven additional patients were identified with deletions affecting exons 2-3 that were similar to the breakpoints as observed in patients #21 and #12 (Figure 2B). Based on the conservation of breakpoints, this microdeletion was denoted as a type I microdeletion. One additional patient was identified with breakpoints similar to patient #11, therefore this deletion was denoted as a type II microdeletion. The deletion affecting exons 4 and 5 as identified in patient #20 was accordingly denoted as a type III microdeletion. These deletions had not been detected before by array-CGH, FISH or MLPA analyses. One patient sample (#20) had already been identified by MLPA as having a clonal microdeletion affecting exons 4 and 5, but also contained a subclonal type I microdeletion in exons 2 and 3 as detected by PCR (Figure 2B and Table 1). In another case (#19), array-CGH had revealed a heterozygous *PTEN* deletion of exons 3-9, but now PCR also revealed a subclonal type I exon 2-3 microdeletion (Table 1). Sequencing of the breakpoints in these additional T-ALL cases revealed that five out of the seven type I deletions and the type II deletion involved the insertion of unique, random nucleotide sequences, thereby excluding false positives due to PCR contamination. Notably, the PCR product for these patient samples as visualized by gel electrophoresis was much weaker than those for the 4 patient samples with clonal microdeletions (data not shown). This strongly indicates that these deletions must be present on the subclonal level and therefore only detectable by specific PCRs. Overall, we have identified *PTEN* microdeletions in 11 out of 146 T-ALL patients (8%), comprising a total of 13 deletional events. Only 4 patients presented these mutations at the clonal level.

Microdeletion breakpoints are flanked by cryptic RAG recombination signal sequences

The conservation of breakpoints among patient samples and the inclusion of non-template derived nucleotides by terminal deoxynucleotidyltransferase (TdT)³¹ in most breakpoint regions pointed to a RAG-mediated deletion mechanism. We then searched for the presence of cryptic RAG recombination signal sequences (cRSS) that could function as putative RAG-mediated recombination signals, such as the RSS involved in T- or B-cell receptor gene segment

rearrangements.³² Analysis of sequences directly flanking the breakpoints immediately revealed typical CAC canonic trinucleotides (Figure 2B and Table S5), which is a hallmark of the heptamer sequences of RSSs. The search for nearby nonamer sequences with A-nucleotide enrichment revealed a putative 12-spacer cRSS in intron 3 with a 5' to 3' orientation (cRSS1). A 23-spacer RSS was identified that directly flanks the breakpoint in intron 5 (cRSS2), and two others were identified flanking both breakpoints in intron 1 (cRSS3 and cRSS4; Figure 2A,B and Table S5). All 23-spacer cRSSs (cRSS2, cRSS3 and cRSS4) are present in a 3' to 5' orientation with respect to the *PTEN* reading frame orientation, and are therefore correctly positioned to allow illegitimate RAG-mediated recombinations with cRSS1 (Figure 2A,C). In this scenario, RAG1/2 molecules bind a pair of 12- and 23-RSSs resulting in two DNA double-strand breaks adjacent to each heptamer. Most microdeletion breakpoints are the consequence of heptamer-to-heptamer sequence fusions resembling signal joints of excision circles that are generated during normal T- or B-cell receptor gene segment rearrangements (Figure 2D, top): Types I and II microdeletions result from cleaved DNA sequences 3-prime of cRSS3 or cRSS4, respectively, that are fused to sequences 5-prime of cRSS1. This retains both cRSSs in the genomic sequences that flank the deletion breakpoints as depicted in Figure 2D. The type III deletion resembles a typical coding joint that results from cleaved DNA sequences 5-prime of cRSS1 that are fused to sequences 3-prime of cRSS2 resulting in the loss of cRSSs from the genomic sequence (Figure 2D, bottom). T- or B-cell receptor coding joints give rise to fused gene segments with potential exonuclease processing of both ends and incorporation of random nucleotides whereby directly flanking RSSs and intervening DNA sequences are lost as excision circles. For the type III deletion of patient #20 (Figure 2B), this led to the fusion of sequences 5-prime of cRSS1 to sequences of 3-prime of cRSS2 with loss of 14 nucleotides and incorporation of 17 GC-rich N-nucleotides.

Prediction of cRSSs

To further characterize these cRSSs and estimate their recombination potential, we calculated RSS Information Content (RIC) scores.²⁹ Cryptic RSSs with RIC scores close to RIC score threshold levels that discriminate bona-fide functional RSSs flanking gene segments of antigen receptor loci from cRSSs (i.e. -38.81 for 12-spacer RSSs and -58.45 for 23-RSSs) were further investigated.²⁹ This search in the *PTEN* locus predicted a 12-spacer cRSS1 with a strong RIC score of -34.23 as well as 23-spacer cRSS2 (-55.59) and cRSS3 (-59.78) with RIC scores that were close to the threshold levels separating RSSs from cRSSs. A 23-spacer cRSS4 was predicted with a RIC score of -75.59 that is barely above the mean background RIC score value

for 39-nucleotide non-RSS DNA sequences (-77.76). Thus, the obtained RIC scores for cRSS1, cRSS2 and cRSS3 strongly support *PTEN* microdeletions as RAG-mediated recombination events with similar recombination potential to that of bona-fide RSSs flanking immunoglobulin V(D)J gene segments.

Cryptic RSS1-4 support RAG-mediated recombination

We then tested whether the predicted cRSS1-4 could functionally mediate RAG recombinations. We used the GFPi-mRFP RAG reporter construct³⁰ (Figure 3A), in which the 12-spacer cRSS1 was cloned in combination with a consensus 23-spacer RSS. Also, the 23-spacer RSSs (cRSS2, cRSS3 and cRSS4) were cloned in combination with a consensus 12-spacer RSS. Recombination efficiency of each variant GFPi-cRSS-mRFP construct was measured by flow cytometry as the frequency of GFP-positive (recombination-positive) HEK293T cells within the population of RFP-positive (transfected) cells (Figure 3A). Indeed, all four *PTEN* cRSSs were able to mediate RAG-dependent recombination of the GFPi substrate (Figure 3B). Recombination efficiencies were $7.5 \pm 0.19\%$ for cRSS1 (Figure 2B, left panel), $2.2 \pm 0.15\%$ for cRSS2, $4.1 \pm 0.21\%$ for cRSS3 and $2.1 \pm 0.16\%$ for cRSS4 (right panel). For comparison, the putative 12-spacer cRSS SCL(12)³⁰ or the 23-spacer cRSS SCL(23)³³ from the human *SCL* gene yielded $1.3 \pm 0.09\%$ and $1.1 \pm 0.13\%$ of GFP-positive cells, respectively. Both of these *SCL* cRSSs were used as references for the lower limit of detection in the GFPi-mRFP RAG reporter assay, as these do not give rise to distinct GFP-positive cell populations in the reporter assay. In contrast, the 12-spacer RSS that flanks the J β 2-2 gene segment of the mouse *TCR β* locus yielded $8.0 \pm 0.31\%$ of GFP-positive cells. Also, the 12-spacer cRSS that is involved in recurrent *LMO2* translocations in T-ALL⁵ yielded $11.4 \pm 0.30\%$ of GFP-positive cells. These reporters highlight the capability of the recombination assay to measure low-efficiency RSS and cRSS activities. Despite the low frequencies of recombination, cRSS2-4 reporters give rise to distinct populations of GFP-positive cells (Figure 3B), in contrast to SCL(12) and SCL(23) cRSSs. Moreover, the efficiencies of recombination of *PTEN* cRSS1-4 differed significantly from the those of SCL(12) or SCL(23) cRSSs (Figure 3C). These results strongly support the involvement of predicted cRSS1-4 in *PTEN* microdeletions in illegitimate RAG-mediated recombination events. Additionally, the recombination potential of these cRSSs are in line with the observed frequencies of type I microdeletions (cRSS3-cRSS1) versus type II (cRSS4-cRSS1) and type III (cRSS1-cRSS2) microdeletions in T-ALL patients (Figure 2B).

***PTEN* microdeletions in xenografted human T-ALL cells**

Subclonal microdeletions in *PTEN*, even in patients that already had undergone clonal inactivating events affecting one allele, strongly imply that acquisition of microdeletions is an ongoing phenomenon in T-ALL leading to clonal diversity.³⁴ To test this, we performed primary and secondary xenotransplantation experiments into NSG mice (Figure 4A) using *TAL1*-rearranged T-ALL blasts from patient (#24) at diagnosis that had a subclonal microdeletion (Figure 4B). Several months post-transplantation, mice developed overt leukemia. Primary (X1) and secondary (X2) xenotransplanted material was then analyzed for the presence of *PTEN* microdeletions in bone marrow, thymus, spleen and liver biopsies. Using MLPA analysis, no *PTEN* microdeletions were detected (data not shown), indicating that the subclonal *PTEN* microdeletion in the diagnostic patient material had not been clonally selected following xenotransplantation. Three distinct, subclonal *PTEN* microdeletions were detected by PCR in thymocyte and liver biopsies: One (X1-24 thymus-1) was identical to the microdeletion as originally identified in this patient (Figure 4B), whereas 2 novel microdeletions were detected suggesting that these had occurred upon serial retransplantation.

***PTEN* deletions are associated with *TALLMO* T-ALL patients**

PTEN aberrations have been associated with a low incidence of *NOTCH1*-activating mutations, but with a high incidence of rearrangements in *TAL1*- and/or *LMO2*-related oncogenes.¹⁸ We now extend these findings, totaling 26 out of 146 T-ALL patients (18%) that have *PTEN* aberrations including point, missense or nonsense mutations, entire locus deletions and/or microdeletions at the clonal or subclonal level as summarized in Table 1. Twelve patients had clonally inactivated *PTEN* on both alleles and 10 patients on one allele. Evidence for subclonal *PTEN* aberrations was found in 11 patients, seven of whom also had clonally inactivated *PTEN* at least in one allele. The other four patients had either a subclonal missense mutation (patient #23) or subclonal microdeletions (3 patients) only. Still, for 8 T-ALL patients for whom protein data were available, absence of *PTEN* protein could not be solely explained by the genetic aberrations found, suggesting that additional *PTEN* inactivating mechanisms await identification. Overall, our previously observed association with *TAL* or *LMO*-rearranged leukemia¹⁸ became considerably more significant ($p=0.003$; Table S6). Also, the significance levels for absence of these mutations in *TLX3*-rearranged T-ALL ($p=0.002$; Table S6) and reduced overlap with *NOTCH1*-activating mutations were further strengthened ($p=0.001$, Table S6).

***PTEN* aberrations and outcome**

Our results do not support observations by others³⁴ that *PTEN*-inactivated T-ALL subclones become selected during disease progression giving rise to relapse. Therefore, we regarded T-ALL patients with subclonal *PTEN* aberrations as wild type patients in outcome analyses. *PTEN/AKT* aberrant T-ALL patients, including patients lacking PTEN protein expression, were not significantly associated with poor outcome in both treatment cohorts (5yrs RFS for *PTEN/AKT* mutant patients is 64±15% versus 70±6% for wild type patients on DCOG protocols and 57±15% versus 76±6% for patients on COALL protocols). This is due to the fact that *PTEN/AKT* mutations and NOTCH-activating mutations predominantly behave as mutually exclusive mutations. In addition, NOTCH-activating mutations have a strong trend towards poor outcome (5yrs RFS for NOTCH-activated patients is 62±8% versus 82±8% for wild type patients on DCOG protocols and 68±8% versus 80±9% for patients on COALL protocols, ($p=0.06$ (stratified for protocol); and supplementary Table S7).³⁵ However, if NOTCH-activated and *PTEN/AKT* mutated T-ALL patients are being compared to wild type patients, wild type patients demonstrate significantly fewer relapses (stratified $p=0.04$; Figure 5), albeit having more frequent events including toxic deaths and secondary malignancies.¹⁸ Using the Cox regression proportional hazard method, NOTCH-activating and *PTEN/AKT* mutations were investigated along with male gender and the presence of *TLX3* rearrangements that both negatively relate with poor outcome (Supplementary Table S7 and Table 2). NOTCH1-activating mutations and *PTEN/AKT* mutations did not significantly predict for increased risk for relapse in univariate analyses, but both were identified as strong, independent risk factors along with male gender in multivariate analysis (Table 2).

***PTEN* microdeletions in healthy human thymocytes**

The presence of recombination-prone cRSSs in *PTEN* intron sequences led us to speculate that microdeletions may occur in healthy thymocytes. Screening DNAs that were isolated from thymocytes of non-leukemic children for the presence of *PTEN* microdeletions by PCR revealed evidence for subclonal type I deletions in 3 out of 11 (27%) thymocyte biopsies. Two of those microdeletions had unique random nucleotide sequences inserted in between the breakpoints (Figure 4C), ruling out false PCR positivity due to contamination. Thus, RAG-mediated *PTEN* microdeletions are not exclusive to T-ALL, but also occur during normal T-cell development.

Discussion

PTEN has been identified as a haploinsufficient tumor suppressor gene,²¹⁻²³ for which gene mutations and/or deletions have been associated with poor outcome in T-ALL in various^{17,18,36,37} but not all studies.^{16,38} For T-ALL patients treated on DCOG ALL-7/8/9 or COALL97/03 treatment protocols, we demonstrate that clonal PTEN inactivating aberrations or loss of PTEN protein is an independent factor that predicts for relapse just like NOTCH-activating mutations and male gender. About half of the T-ALL patients that retained one wild type allele do not express PTEN protein. This indicates that additional genetic, epigenetic or post-translational inactivating events are expected. We have identified recurrent inactivating *PTEN* microdeletions in T-ALL patients due to illegitimate RAG-mediated recombination events, mediated through cryptic RSSs (cRSSs). Taking into account all point, missense or nonsense mutations as well as deletions including microdeletions, 18% of T-ALL patients in our patient cohort harbor PTEN inactivating aberrations.

Increasing evidence suggests that cRSSs can participate in oncogenic mechanisms, including chromosomal translocations in lymphoma, B-ALL^{39,40} and T-ALL^{4,41,42}, such as *SIL-TAL1* gene fusions and *HPRT* deletions. Different chromosomal translocation mechanisms have been described as a consequence of erroneous rearrangements between cRSS that flank oncogenes with RSS-sequences of T-cell receptor gene segments. Alternatively, broken DNA strands near oncogenes become mistakenly fused through a non-cRSS mechanism to T-cell receptor gene during V(D)J-assembly (reviewed in Radich and Sala⁴³). Other illegitimate, cRSS-driven coding-joint recombination events may cause intrachromosomal deletions such as described for *IKZF1*,^{44,45} *ERG1*,⁴⁶ *BTG1*,⁴⁷ and *CDKN2A/B*⁴⁸⁻⁵⁰ in humans and *Notch1*^{51,52} and *Bcl11b*⁵³ in mice.

PTEN microdeletions occur as a consequence of aberrant RAG-mediated recombination events, and breakpoints are flanked by cRSSs containing heptamer and nonamer sequences separated by 12 or 23 nucleotide spacers. These cRSSs facilitate recombinations in *in vitro* recombination assays⁵⁴ with efficiencies that match the frequencies of different types of microdeletions in T-ALL patients. Approximately, one third of all signal joint-related type I and II microdeletions have perfect heptamer-to-heptamer fusions that lack incorporation of random nucleotides just like IgH or TCR excision circle signal-joints. Two third of microdeletion signal junctions represent atypical joints that incorporated GC-rich N-nucleotides, phenomena which is also observed in V(D)J-associated signal junctions in mouse lymphocytes.⁵⁵ The T-cell receptor-mediated translocation in T-ALL line SUP-T1 is a comparable atypical signal junction.⁵⁶ Some of these microdeletion atypical signal joints (patient #12, #26 and #11) had undergone exonuclease

processing of signal ends (Figure 2B). Non-canonic heptamer sequence variations may destabilize the RAG complex, allowing alternative joining mechanisms of coding-ends and signal-ends.⁵⁷ This may include open-and-shut joint recombinations,^{58,59} resulting from a single RAG cleavage adjacent to cRSS heptamers, exonuclease processing and insertion of random nucleotides before re-ligation of the DNA ends. This may explain one rare T-ALL case²⁴ with a mutation that replaces 13 nucleotides including the start codon by 15 random nucleotides. This mutation is flanked by a 12-spacer cRSS with a strong RIC score of -45.48 that allows RAG-mediated recombinations equal to the efficiency ($2.5 \pm 0.13\%$) of the mouse IgH locus VH/87 RSS (L.M.S., unpublished data and Davila et al⁶⁰). However, no other T-ALL patient in our current series had an equivalent mutation at this position, indicating that these events are rare.

The identification of subclonal *PTEN* microdeletions—as well as entire *PTEN* locus deletions¹⁸—indicates that RAG activity may be ongoing in (at least part of) the leukemic cell population. This may explain clonal diversity and selection that results in disease progression and relapse. This latter is also supported by *in-vitro* recombination assays using T-ALL cell lines, and demonstrate that about one percent of leukemic cells or less will undergo recombinations of the reporter construct within a one week time-frame.³⁰ Since intraclonal heterogeneity at diagnosis and clonal evolution at relapse are known to occur in ALL,^{46,61-64} we checked whether *PTEN* microdeletions in minor leukemic clones at presentation of disease become clonally selected following xenograft transplantation just like lentiviral *PTEN*-silenced T-ALL blasts.³⁴ However, we did not observe preferential selection of leukemic cells with *PTEN* microdeletions to near clonal levels following xenotransplantation. This could be explained by preferential outgrowth of leukemic subclones having other mutations that were advantageous for engraftment in mice over subclones having *PTEN* microdeletions. Furthermore, additional and new illegitimate RAG-mediated *PTEN* microdeletions possibly as consequence of ongoing RAG activity were detected that were not found in primary leukemic cells. We cannot formally rule out that subclonal selection of a leukemic subpopulation with a novel *PTEN* microdeletion occurred from a PCR-undetectable subclone that was already present at diagnosis. In addition, one patient had a subclonal missense mutation, indicating that there is an ongoing pressure on TALLMO-dysregulated leukemic cells to inactivate remaining wild-type *PTEN* alleles. RAG activity also results in *PTEN* microdeletions in developing thymocytes of healthy individuals. These rearrangements may facilitate a pre-malignant condition from which leukemia can develop. Likewise, Marlunescu et al⁵ described two mechanisms of illegitimate V(D)J chromosomal rearrangement that were found in healthy children, i.e. the D δ 2/LMO2

recombination in the t(11;14)(p13;q11) and the TAL2/TCR β translocation t(7;9)(q34;q32),⁶⁵ known as driving oncogenic lesions in T-ALL.

Overall, our discovery of *PTEN* microdeletions has reinforced the fact that *PTEN* aberrations are especially abundant in *TAL*- or *LMO*-rearranged leukemia but not in *TLX3*-rearranged patients,¹⁸ as also observed in adult T-ALL patient series³⁷. *PTEN* abnormalities seem to be associated with a reduced incidence of *NOTCH1*-activating mutations. The TALLMO subtype represents an immunophenotypically mature subtype of arrested leukemic cells in T-ALL, in which ongoing RAG-activity creates an opportunistic and extended time window for cRSS-mediated illegitimate recombination events. These may provoke disease progression and relapse in leukemia patients, adding a new level of complexity that should be addressed in the development of future antileukemic strategies for ALL. Taking into account all currently known *PTEN* inactivation mechanisms —*PTEN* mutations, entire locus deletions and *PTEN* microdeletions—, some seemingly wild type T-ALL patients still lack PTEN protein expression indicating that other *PTEN* inactivation mechanisms await identification.

Acknowledgements

The authors would like to thank Gustavo G.L. Costa and Izabella A.P. Neshich for help with computational detection of cRSSs. RM, KCB and LZ were financed by the Stichting Kinderen Kankervrij (Grant no. KiKa 2007-12, KiKa 2008-29 and KiKa 2013-116). LMS received a postdoctoral fellowship from Fundação para a Ciência e a Tecnologia (FTC). This work was supported by grants PTDC/SAU-ONC/113202/2009 from FCT (to JTB), and FAPESP 08/10034-1.

Authorship

Contribution: R.D.M., L.M.S., J.T.B. and J.P.P.M. designed the study and wrote the manuscript; R.D.M., W.K.S., L.Z. and J.G.C.A.M.B.-G. performed the MLPA analyses, breakpoint mapping and PCR-based screening of *PTEN* microdeletions; J.A.Y. performed the computational detection of putative RAG recombination signal sequences; L.M.S. and V.P. were responsible for generation of GFPi-*PTEN* cRSS reporter constructs and recombination assays; K.C.-B. and W.K.S. performed the xenotransplantation experiments; J.P.P.M. performed the statistical analyses of the data; M.A., E.S., M.H. collected and provided patient samples and their characteristics; R.D.M., L.M.S., K.C.-B., R.P., J.T.B., J.P.P.M. analyzed and interpreted data; and all authors read, revised, and approved the paper.

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

References

1. Pui CH, Evans WE. Treatment of acute lymphoblastic leukemia. *N Engl J Med*. Jan 12 2006;354(2):166-178.
2. Meijerink JP. Genetic rearrangements in relation to immunophenotype and outcome in T-cell acute lymphoblastic leukaemia. *Best Pract Res Clin Haematol*. Sep 2010;23(3):307-318.
3. Dik WA, Nadel B, Przybylski GK, et al. Different chromosomal breakpoints impact the level of LMO2 expression in T-ALL. *Blood*. Jul 1 2007;110(1):388-392.
4. Le Noir S, Ben Abdelali R, Lelorch M, et al. Extensive molecular mapping of TCRalpha/delta- and TCRbeta-involved chromosomal translocations reveals distinct mechanisms of oncogene activation in T-ALL. *Blood*. Oct 18 2012;120(16):3298-3309.
5. Marculescu R, Le T, Simon P, Jaeger U, Nadel B. V(D)J-mediated translocations in lymphoid neoplasms: a functional assessment of genomic instability by cryptic sites. *J Exp Med*. Jan 7 2002;195(1):85-98.
6. Homminga I, Pieters R, Langerak AW, et al. Integrated transcript and genome analyses reveal NKX2-1 and MEF2C as potential oncogenes in T cell acute lymphoblastic leukemia. *Cancer Cell*. Apr 12 2011;19(4):484-497.
7. Weng AP, Ferrando AA, Lee W, et al. Activating mutations of NOTCH1 in human T cell acute lymphoblastic leukemia. *Science*. Oct 8 2004;306(5694):269-271.
8. Sulis ML, Williams O, Palomero T, et al. NOTCH1 extracellular juxtamembrane expansion mutations in T-ALL. *Blood*. Aug 1 2008;112(3):733-740.
9. Malyukova A, Dohda T, von der Lehr N, et al. The tumor suppressor gene hCDC4 is frequently mutated in human T-cell acute lymphoblastic leukemia with functional consequences for Notch signaling. *Cancer Res*. Jun 15 2007;67(12):5611-5616.
10. O'Neil J, Grim J, Strack P, et al. FBW7 mutations in leukemic cells mediate NOTCH pathway activation and resistance to gamma-secretase inhibitors. *J Exp Med*. Aug 6 2007;204(8):1813-1824.
11. Thompson BJ, Buonamici S, Sulis ML, et al. The SCFFBW7 ubiquitin ligase complex as a tumor suppressor in T cell leukemia. *J Exp Med*. Aug 6 2007;204(8):1825-1835.
12. Ferrando AA. The role of NOTCH1 signaling in T-ALL. *Hematology Am Soc Hematol Educ Program*. 2009:353-361.
13. Shochat C, Tal N, Bandapalli OR, et al. Gain-of-function mutations in interleukin-7 receptor-alpha (IL7R) in childhood acute lymphoblastic leukemias. *J Exp Med*. May 9 2011;208(5):901-908.
14. Zenatti PP, Ribeiro D, Li W, et al. Oncogenic IL7R gain-of-function mutations in childhood T-cell acute lymphoblastic leukemia. *Nat Genet*. Oct 2011;43(10):932-939.
15. Palomero T, Sulis ML, Cortina M, et al. Mutational loss of PTEN induces resistance to NOTCH1 inhibition in T-cell leukemia. *Nat Med*. Oct 2007;13(10):1203-1210.
16. Gutierrez A, Sanda T, Grebliunaite R, et al. High frequency of PTEN, PI3K, and AKT abnormalities in T-cell acute lymphoblastic leukemia. *Blood*. Jul 16 2009;114(3):647-650.
17. Jotta PY, Ganazza MA, Silva A, et al. Negative prognostic impact of PTEN mutation in pediatric T-cell acute lymphoblastic leukemia. *Leukemia*. Jan 2010;24(1):239-242.
18. Zuurbier L, Petricoin EF, Vuerhard MJ, et al. The significance of PTEN and AKT aberrations in pediatric T-cell acute lymphoblastic leukemia. *Haematologica*. Apr 4 2012.
19. Palomero T, Dominguez M, Ferrando AA. The role of the PTEN/AKT Pathway in NOTCH1-induced leukemia. *Cell Cycle*. Apr 15 2008;7(8):965-970.
20. Maser RS, Choudhury B, Campbell PJ, et al. Chromosomally unstable mouse tumours have genomic alterations similar to diverse human cancers. *Nature*. Jun 21 2007;447(7147):966-971.

21. Alimonti A, Carracedo A, Clohessy JG, et al. Subtle variations in Pten dose determine cancer susceptibility. *Nat Genet.* May 2010;42(5):454-458.
22. Correia NC, Girio A, Antunes I, Martins LR, Barata JT. The multiple layers of non-genetic regulation of PTEN tumour suppressor activity. *Eur J Cancer.* Jan 2014;50(1):216-225.
23. Berger AH, Knudson AG, Pandolfi PP. A continuum model for tumour suppression. *Nature.* Aug 11 2011;476(7359):163-169.
24. Silva A, Yunes JA, Cardoso BA, et al. PTEN posttranslational inactivation and hyperactivation of the PI3K/Akt pathway sustain primary T cell leukemia viability. *The Journal of clinical investigation.* Nov 2008;118(11):3762-3774.
25. Kamps WA, Bokkerink JP, Hahlen K, et al. Intensive treatment of children with acute lymphoblastic leukemia according to ALL-BFM-86 without cranial radiotherapy: results of Dutch Childhood Leukemia Study Group Protocol ALL-7 (1988-1991). *Blood.* Aug 15 1999;94(4):1226-1236.
26. Kamps WA, Bokkerink JP, Hakvoort-Cammel FG, et al. BFM-oriented treatment for children with acute lymphoblastic leukemia without cranial irradiation and treatment reduction for standard risk patients: results of DCLSG protocol ALL-8 (1991-1996). *Leukemia.* Jun 2002;16(6):1099-1111.
27. Veerman AJ, Kamps WA, van den Berg H, et al. Dexamethasone-based therapy for childhood acute lymphoblastic leukaemia: results of the prospective Dutch Childhood Oncology Group (DCOG) protocol ALL-9 (1997-2004). *Lancet Oncol.* Oct 2009;10(10):957-966.
28. Escherich G, Horstmann MA, Zimmermann M, Janka-Schaub GE. Cooperative study group for childhood acute lymphoblastic leukaemia (COALL): long-term results of trials 82,85,89,92 and 97. *Leukemia.* Feb 2010;24(2):298-308.
29. Cowell LG, Davila M, Kepler TB, Kelsoe G. Identification and utilization of arbitrary correlations in models of recombination signal sequences. *Genome Biol.* 2002;3(12):RESEARCH0072.
30. Trancoso I, Bonnet M, Gardner R, et al. A Novel Quantitative Fluorescent Reporter Assay for RAG Targets and RAG Activity. *Front Immunol.* 2013;4:110.
31. Gellert M. V(D)J RECOMBINATION: RAG PROTEINS, REPAIR FACTORS, AND REGULATION*. *Annual Review of Biochemistry.* 2002;71(1):101-132.
32. Lewis SM, Agard E, Suh S, Czyzyk L. Cryptic signals and the fidelity of V(D)J joining. *Mol Cell Biol.* Jun 1997;17(6):3125-3136.
33. Raghavan SC, Kirsch IR, Lieber MR. Analysis of the V(D)J Recombination Efficiency at Lymphoid Chromosomal Translocation Breakpoints. *Journal of Biological Chemistry.* August 3, 2001 2001;276(31):29126-29133.
34. Clappier E, Gerby B, Sigaux F, et al. Clonal selection in xenografted human T cell acute lymphoblastic leukemia recapitulates gain of malignancy at relapse. *J Exp Med.* Apr 11 2011;208(4):653-661.
35. Zuurbier L, Homminga I, Calvert V, et al. NOTCH1 and/or FBXW7 mutations predict for initial good prednisone response but not for improved outcome in pediatric T-cell acute lymphoblastic leukemia patients treated on DCOG or COALL protocols. *Leukemia.* Dec 2010;24(12):2014-2022.
36. Bandapalli OR, Zimmermann M, Kox C, et al. NOTCH1 activation clinically antagonizes the unfavorable effect of PTEN inactivation in BFM-treated children with precursor T-cell acute lymphoblastic leukemia. *Haematologica.* Jun 2013;98(6):928-936.
37. Trinquand A, Tanguy-Schmidt A, Ben Abdelali R, et al. Toward a NOTCH1/FBXW7/RAS/PTEN-based oncogenetic risk classification of adult T-cell acute lymphoblastic leukemia: a Group for Research in Adult Acute Lymphoblastic Leukemia study. *J Clin Oncol.* Dec 1 2013;31(34):4333-4342.

38. Larson Gedman A, Chen Q, Kugel Desmoulin S, et al. The impact of NOTCH1, FBW7 and PTEN mutations on prognosis and downstream signaling in pediatric T-cell acute lymphoblastic leukemia: a report from the Children's Oncology Group. *Leukemia*. Aug 2009;23(8):1417-1425.
39. Marculescu R, Le T, Bocskor S, et al. Alternative end-joining in follicular lymphomas' t(14;18) translocation. *Leukemia*. Jan 2002;16(1):120-126.
40. Papaemmanuil E, Rapado I, Li Y, et al. RAG-mediated recombination is the predominant driver of oncogenic rearrangement in ETV6-RUNX1 acute lymphoblastic leukemia. *Nat Genet*. Jan 12 2014.
41. Aplan PD, Lombardi DP, Ginsberg AM, Cossman J, Bertness VL, Kirsch IR. Disruption of the human SCL locus by "illegitimate" V-(D)-J recombinase activity. *Science*. Dec 7 1990;250(4986):1426-1429.
42. Larmonie NS, Dik WA, Meijerink JP, Homminga I, van Dongen JJ, Langerak AW. Breakpoint sites disclose the role of the V(D)J recombination machinery in the formation of T-cell receptor (TCR) and non-TCR associated aberrations in T-cell acute lymphoblastic leukemia. *Haematologica*. Aug 2013;98(8):1173-1184.
43. Radich J, Sala O. The Biology of Adult Acute Lymphoblastic Leukemia. In: Advani AS, Lazarus HM, eds. *Adult Acute Lymphocytic Leukemia*: Humana Press; 2011:25-44.
44. Mullighan CG, Miller CB, Radtke I, et al. BCR-ABL1 lymphoblastic leukaemia is characterized by the deletion of Ikaros. *Nature*. May 1 2008;453(7191):110-114.
45. Iacobucci I, Storlazzi CT, Cilloni D, et al. Identification and molecular characterization of recurrent genomic deletions on 7p12 in the IKZF1 gene in a large cohort of BCR-ABL1-positive acute lymphoblastic leukemia patients: on behalf of Gruppo Italiano Malattie Ematologiche dell'Adulto Acute Leukemia Working Party (GIMEMA AL WP). *Blood*. Sep 3 2009;114(10):2159-2167.
46. Clappier E, Auclerc MF, Rapon J, et al. An intragenic ERG deletion (ERG) is a marker of an oncogenic subtype of B-cell precursor acute lymphoblastic leukemia with a favorable outcome despite frequent IKZF1 deletions. *Leukemia*. Sep 25 2013.
47. Waanders E, Scheijen B, van der Meer LT, et al. The origin and nature of tightly clustered BTG1 deletions in precursor B-cell acute lymphoblastic leukemia support a model of multiclonal evolution. *PLoS Genet*. 2012;8(2):e1002533.
48. Raschke S, Balz V, Efferth T, Schulz WA, Florl AR. Homozygous deletions of CDKN2A caused by alternative mechanisms in various human cancer cell lines. *Genes Chromosomes Cancer*. Jan 2005;42(1):58-67.
49. Novara F, Beri S, Bernardo ME, et al. Different molecular mechanisms causing 9p21 deletions in acute lymphoblastic leukemia of childhood. *Hum Genet*. Oct 2009;126(4):511-520.
50. Kitagawa Y, Inoue K, Sasaki S, et al. Prevalent involvement of illegitimate V(D)J recombination in chromosome 9p21 deletions in lymphoid leukemia. *J Biol Chem*. Nov 29 2002;277(48):46289-46297.
51. Ashworth TD, Pear WS, Chiang MY, et al. Deletion-based mechanisms of Notch1 activation in T-ALL: key roles for RAG recombinase and a conserved internal translational start site in Notch1. *Blood*. Dec 16 2010;116(25):5455-5464.
52. Tsuji H, Ishii-Ohba H, Katsube T, et al. Involvement of illegitimate V(D)J recombination or microhomology-mediated nonhomologous end-joining in the formation of intragenic deletions of the Notch1 gene in mouse thymic lymphomas. *Cancer Res*. Dec 15 2004;64(24):8882-8890.
53. Sakata J, Inoue J, Ohi H, et al. Involvement of V(D)J recombinase in the generation of intragenic deletions in the Rit1/Bcl11b tumor suppressor gene in gamma-ray-induced thymic lymphomas and in normal thymus of the mouse. *Carcinogenesis*. Jun 2004;25(6):1069-1075.

54. Grawunder U, Harfst E. How to make ends meet in V(D)J recombination. *Curr Opin Immunol.* Apr 2001;13(2):186-194.
55. Lee J, Desiderio S. Cyclin A/CDK2 regulates V(D)J recombination by coordinating RAG-2 accumulation and DNA repair. *Immunity.* Dec 1999;11(6):771-781.
56. Baer R, Forster A, Rabbitts TH. The mechanism of chromosome 14 inversion in a human T cell lymphoma. *Cell.* Jul 3 1987;50(1):97-105.
57. Arnal SM, Holub AJ, Salus SS, Roth DB. Non-consensus heptamer sequences destabilize the RAG post-cleavage complex, making ends available to alternative DNA repair pathways. *Nucleic Acids Res.* May 2010;38(9):2944-2954.
58. Lewis SM, Hesse JE, Mizuuchi K, Gellert M. Novel strand exchanges in V(D)J recombination. *Cell.* Dec 23 1988;55(6):1099-1107.
59. Elliott JF, Rock EP, Patten PA, Davis MM, Chien YH. The adult T-cell receptor delta-chain is diverse and distinct from that of fetal thymocytes. *Nature.* Feb 18 1988;331(6157):627-631.
60. Davila M, Liu F, Cowell LG, et al. Multiple, conserved cryptic recombination signals in VH gene segments: detection of cleavage products only in pro B cells. *J Exp Med.* Dec 24 2007;204(13):3195-3208.
61. Anderson K, Lutz C, van Delft FW, et al. Genetic variegation of clonal architecture and propagating cells in leukaemia. *Nature.* Jan 20 2011;469(7330):356-361.
62. Mullighan CG, Phillips LA, Su X, et al. Genomic analysis of the clonal origins of relapsed acute lymphoblastic leukemia. *Science.* Nov 28 2008;322(5906):1377-1380.
63. Yang JJ, Bhojwani D, Yang W, et al. Genome-wide copy number profiling reveals molecular evolution from diagnosis to relapse in childhood acute lymphoblastic leukemia. *Blood.* Nov 15 2008;112(10):4178-4183.
64. Kuster L, Grausenburger R, Fuka G, et al. ETV6/RUNX1-positive relapses evolve from an ancestral clone and frequently acquire deletions of genes implicated in glucocorticoid signaling. *Blood.* Mar 3 2011;117(9):2658-2667.
65. Marculescu R, Vanura K, Le T, Simon P, Jager U, Nadel B. Distinct t(7;9)(q34;q32) breakpoints in healthy individuals and individuals with T-ALL. *Nat Genet.* Mar 2003;33(3):342-344.

Legends of Figures

Figure 1. Identification of *PTEN* microdeletions in T-ALL patients. (A) MLPA electropherograms of normal reference DNA and representative examples of T-ALL patients with heterozygous or homozygous *PTEN* microdeletions affecting exons 2-3 or a heterozygous deletion of exons 4-5. Fluorescence intensities of amplified PCR products for specific *PTEN* exons are shown. PCR product sizes are shown at the top. Each arrow points to a homo- or heterozygously deleted exon. (B) Array-CGH plot exhibiting the homozygous *PTEN* exon 2-3 microdeletion in one T-ALL patient sample.

Figure 2. Breakpoints of *PTEN* microdeletions. (A) Schematic representation of the *PTEN* gene. Missense mutations are represented by open triangles above the exons, whereas a silent mutation is presented as a filled grey triangle as shown before¹⁸. Nonsense insertion/deletion mutations are indicated by a filled black triangle. Left or right-pointing open triangles in introns 1, 3 and 5 represent cRSSs. (B) Sequences of cloned intron 1-3 type I and type II breakpoints and the intron 3-5 type III breakpoint for T-ALL patients with *PTEN* microdeletions. Cryptic recombination signal sequences (cRSS) are indicated by a box with the canonic CAC trinucleotide sequences or the corresponding GTG nucleotides in heptamer sequences indicated in bold and underlined. Insertion of non-template, random nucleotides are shown in bold. (C) Examples of sequence traces of cDNA resulting from type I, II and III microdeletions. (D) Involvement of specific cRSSs in illegitimate RAG-mediated recombination events resulting in types I and II microdeletions (signal joint) and aberrant *PTEN* splice variant or the type III with the aberrant exon 4-5 microdeletion *PTEN* transcript (coding joint).

Figure 3. Intronic *PTEN* cryptic RSS mediate RAG recombination events. (A) Upper panel: Linear representation of the GFPi reporter construct that results in the inversion of GFP coding sequence during RAG-mediated recombination, and consequent GFP expression. The inverted GFP sequence (light green box) is flanked by a proximal 12-spacer RSS (light grey triangle) and a distal 23-spacer RSS (dark grey triangle) followed by the IRES-RFP as transfection control reporter (red box). GFP positivity is a measure for recombination potential. Lower panel: Control *in vitro* RAG recombination assay; flow cytometry analysis of HEK293T cells transiently transfected with either an irrelevant, mock vector (in the absence of RAG1/2 expression vectors; negative control) or the GFPi-reporter construct containing the consensus 12- and 23-RSS in the presence of RAG1/2 expression vectors.³⁰ The flow cytometry plots show the expression of GFP and RFP within gated live cells defined by FSC

and SSC parameters (not shown) and the values represent the percentage of each cell population in the quadrants. The gate used to discriminate RFP-positive from RFP-negative cells is depicted by a red square and used for the contour plot analysis. The efficiency of recombination is indicated as the percentage of GFP-positive (recombination positive) cells within the RFP-positive (transfected) population. (B) Flow cytometry analysis of HEK293T cells transiently transfected with the GFPi variant constructs containing specific 12-spacer cRSS (LMO2, SCL/TAL1, *PTEN*-cRRS1 or the J β 2.2-RSS) site combined with the consensus 23-spacer RSS (left panel). The GFPi variant constructs containing the consensus 12-RSS³⁰ was combined with 23-spacer *PTEN* cRSS2, cRSS3, cRSS4 or the control SCL/TAL1 23-spacer cRSS version (right panel). The human LMO2 12-spacer cRSS and the mouse J β 2-2 *bona fide* RSS were used to establish the range of recombination activities for low-efficiency RSSs as measured by the GFPi reporter assay. The 12- and 23-spacer versions of the human SCL/TAL1 cRSS were used to define the lower limit of detection of cRSS function in this reporter assay. Average percentage \pm SD of GFP⁺ cells in the RFP⁺ population are derived from 4-5 independent experiments. (C) Recombination index was determined by normalizing the recombination efficiencies of each indicated reporter to that of GFPi Con12/23 and recombination efficiencies were calculated subtracting the GFP background of each respective unrecombined control. Values represent the mean \pm SEM of 3 independent experiments with 3 replicates per condition; * $p < 0.05$; ** $p < 0.01$ and *** $p < 0.0001$.

Figure 4. *PTEN* microdeletions in xenotransplants of a T-ALL primary patient sample and in human thymocytes from healthy individuals. (A) Schematic representation of the xenotransplantation strategy. Several months post-transplant of the patient's (#24) leukemic cells into immunodeficient NSG mice (n=9), we collected cells from bone marrow (BM), thymus (Thy), spleen (Spl) and liver. Primary (X1) and secondary (X2) xenotransplanted material was then analyzed for the presence of any of the three different *PTEN* microdeletions. (B) Breakpoint sequences of *PTEN* microdeletions as detected in samples from primary (X1) and secondary (X2) xenotransplanted mice. Canonic CAC trinucleotide sequences or the corresponding GTG nucleotides in heptamer sequences are indicated in bold and underlined. (C) Sequences of the breakpoints for *PTEN* type I microdeletions as identified in thymocytes of healthy individuals (H-Thy1 – H-Thy3).

Figure 5. T-ALL patients lacking *PTEN/AKT* mutations and NOTCH-activating mutations have a good outcome. Relapse free survival curves (RFS) for T-ALL patients treated on (A)

Dutch DCOG ALL7/8 or 9 protocols or (B) German COALL-97/03 protocols. Green line; NOTCH-activating mutations including mutations in *NOTCH* and *FBXW7*, red line; PTEN-inactivating or AKT-activating mutations, blue line; NOTCH-activating mutations and PTEN-inactivating or AKT-activating mutations combined, black line; wild type for *NOTCH/FBXW7* and *PTEN/AKT*. Tick-marks in figures refer to patients that are lost from further follow-up. The numbers of patients included at various time points in these studies have been shown.

Table 1. *PTEN* genetics of T-ALL patients

| Patient | PTEN mutation | | PTEN Deletion | PTEN Deletion | Genomic | Aberrant | PTEN | Cytogenetic | NOTCH/FBXW7 |
|--|-------------------|-----------|-----------------------|-------------------|--|----------------|---------|-------------|-------------|
| | Allele A | Allele B | (FISH/Array-CGH) | (MLPA) | breakpoint PCR | transcript | protein | aberration | mutation |
| Clonal inactivation of 2 alleles | | | | | | | | | |
| 1 | R129G | T231fsX24 | - | - | - | ND | Mutant | LMO3 | - |
| 2 | F144fsX37 | R232fsX23 | Subcl Del | - | - | ND | Absent | LMO2 | PEST |
| 3 | P246fsX11 | - | Het Del | Het Del exons 2-9 | - | ND | Absent | SIL-TAL1 | - |
| 4 | D235fsX9 | P245fsX12 | ND | ND | - | ND | Absent | CALM-AF10 | PEST |
| 5 | R232fsX13 | Q244fsX8 | - | - | - | ND | Absent | Unknown | - |
| 6 | R233fsX10 | P245fsX9 | - | - | - | ND | ND | Unknown | FBXW7 |
| 7 | R232fsX10 | P243fsX18 | - | - | - | ND | ND | TLX1 | - |
| 8 | R232fsX10 | P245fsX14 | - | - | - | ND | Absent | LMO2 | FBXW7 |
| 9 | L180fsX2 | I305fsX7 | - | - | Intron 1-3 Del (type II; Subcl) | ND | Absent | NKX2-5 | HD |
| 10 | C104fsX2 | K236fsX5 | - | - | - | ND | Absent | SIL-TAL1 | - |
| 11 | P245fsX3 | - | ND | Het Del exons 2-3 | Intron 1-3 Del (type II, Clonal) | in1/2-ex4 | ND | MYC | - |
| 12 | - | - | Hom Del/FISH negative | Hom Del exons 2-3 | 2 Intron 1-3 Del variants (type I, Clonal) | in1/2-ex4 | Absent | SIL-TAL1 | - |
| Clonal inactivation of 1 allele | | | | | | | | | |
| 13 | R129fsX4/P245fsX3 | - | Subcl Del | Het Del exons 1-9 | - | ND | ND | Unknown | - |
| 14 | R232(STOP) | - | - | - | - | - | ND | SIL-TAL1 | - |
| 15 | C249fsX10 | - | - | - | Intron 1-3 Del (type I; Subcl) | - | Absent | Unknown | - |
| 16 | T231fsX14 | - | - | - | - | - | Absent | NKX2-1 | HD/FBXW7 |
| 17 | T276A | - | - | - | Intron 1-3 Del (type I; Subcl) | - | Absent | Unknown | - |
| 18 | - | - | Het Del | Het Del exons 1-9 | - | WT | Absent | Unknown | PEST |
| 19 | - | - | Het Del | Het Del exons 3-9 | Intron 1-3 Del (type I; Subcl) | Altered | Absent | SIL-TAL1 | - |
| 20 | - | - | - | Het Del exons 4-5 | Intron 3-5 Del (type III, Clonal) / Intron 1-3 Del (type I; Subcl) | ex3-ex6 splice | Absent | SIL-TAL1 | - |
| 21 | - | - | ND | Het Del exons 1-3 | Intron 1-3 Del (type I, Clonal) | in1/2-ex4 | ND | SIL-TAL1 | HD |
| 22 | - | - | Het Del | Het Del exons 2-9 | - | ND | ND | Unknown | - |
| Subclonal inactivating event only | | | | | | | | | |
| 23 | R232fsX (Subcl) | - | - | - | - | - | ND | SIL-TAL1 | - |
| 24 | - | - | - | - | Intron 1-3 Del (type I; Subcl) | ND | ND | SIL-TAL1 | HD/PEST |
| 25 | - | - | ND | - | Intron 1-3 Del (type I; Subcl) | ND | ND | SIL-TAL1 | - |
| 26 | - | - | - | - | Intron 1-3 Del (type I; Subcl) | ND | ND | Unknown | - |
| Others | | | | | | | | | |
| 27 | - | - | - | - | - | - | Absent | TAL2 | FBXW7 |
| 28 | - | - | - | - | - | - | Absent | TAL1 | HD |

PTEN frameshift mutations are indicated with the number of encoded amino acids in the alternative reading frame. *PTEN* deletion status was determined by fluorescent *in situ* hybridization (FISH), array-comparative genomic hybridization (array-CGH) and/or multiplex ligation-dependent probe amplification (MLPA). Introns harboring the genomic breakpoints of *PTEN* deletions and/or microdeletions have been indicated. Exons for alternative spliced *PTEN* transcripts are indicated. Abbreviations: Del, deletion; FS, frameshift; HD, *NOTCH1* heterodimerization domain mutation; Het, heterozygous; Hom, homozygous; PEST, *NOTCH1* mutation in the proline, glutamine, serine and threonine-rich C-terminal region; Subcl, subclonal.

Table 2. NOTCH1-activating and PTEN/AKT mutations predict for poor outcome in pediatric T-ALL treated on DCOG ALL7/8/9 or COALL-97/03 protocols.

| | Univariate analyses using Cox regression model | | | |
|-------------------------------|--|--------------|-------------------------|--------------|
| | <i>n</i> | Hazard ratio | 95% confidence interval | <i>P</i> |
| Male gender | 146 | 3.278 | 1.267-8.486 | 0.014 |
| <i>TLX3</i> | 146 | 2.044 | 1- 4.175 | 0.05 |
| <i>NOTCH1/FBXW7</i> | 141 | 2.077 | 0.946-4.560 | 0.068 |
| § <i>PTEN/AKT</i> aberrations | 142 | 1.675 | 0.787-3.567 | 0.18 |
| | Multivariate analyses using Cox regression model | | | |
| | <i>n</i> | Hazard ratio | 95% confidence interval | <i>P</i> |
| Male gender | 141 | 2.910 | 1.117 – 7.577 | 0.029 |
| <i>TLX3</i> | 141 | 2.018 | 0.921 – 4.424 | 0.079 |
| <i>NOTCH</i> -activating | 141 | 2.588 | 1.083 – 6.183 | 0.032 |
| § <i>PTEN/AKT</i> aberrations | 141 | 3.407 | 1.254-7.400 | 0.014 |

Univariate and multivariate Cox regression analyses stratified for DCOG or COALL treatment protocols using relapse-free survival for various parameters that were significantly associated with poor relapse-free survival (see supplementary table S7). §Includes T-ALL patients that do not express PTEN protein while lacking *PTEN* aberrations, but does not include patient samples with *PTEN* aberrations only on the subclonal level.

Figure 1

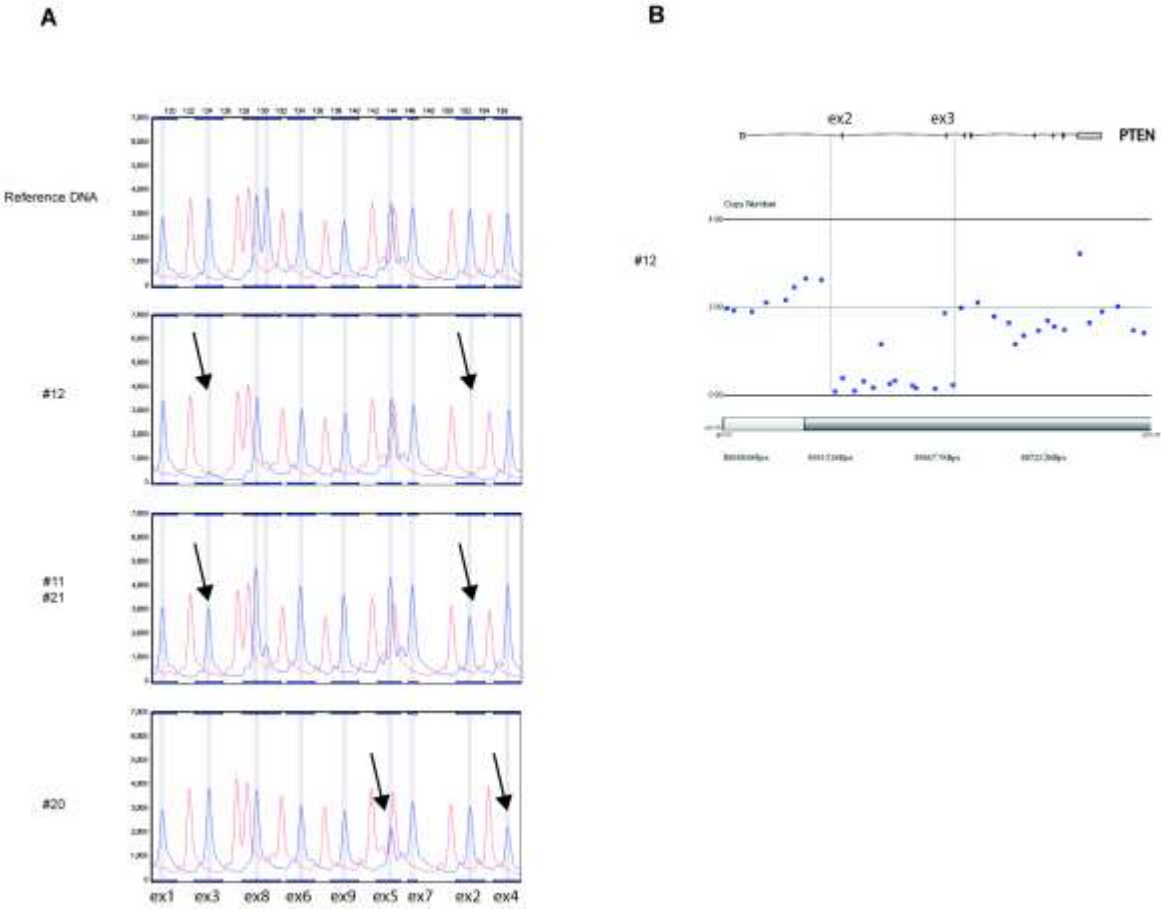
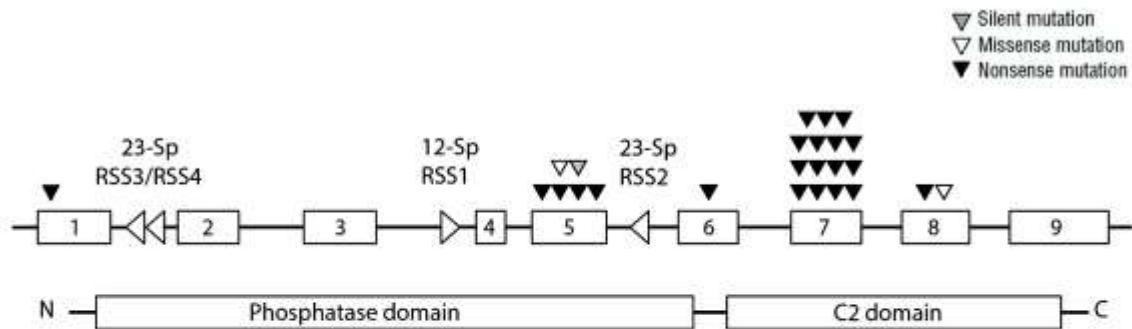


Figure 2

A



B

Type-I micro-deletions

| Patient | brk intr 1 | n-bases | brk intr 3 | Clonality status |
|---------|---|--------------------------|--|------------------|
| WT | TCAT <u>ATATTTTGT (23) TAGGGTG</u> <u>cRSS3</u> | NNNNNNNN (65Kb) NNNNNNNN | <u>CACAGAT (12) ACATAAACA</u> <u>cRSS1</u> | CCCC |
| 21 | TCAT <u>ATATTTTGT (23) TAGGGTG</u> | | CACAGAT (12) ACATAAACA | CCCC Clonal |
| 12 | TCAT <u>ATATTTTGT (23) TAGGGTG</u> | | CACAGAT (12) ACATAAACA | CCCC Clonal |
| 12 | TCAT <u>ATATTTTGT (23) TAGGGTG</u> | GACTGAACCTCT | ---GAT (12) ACATAAACA | CCCC Clonal |
| 24 | TCAT <u>ATATTTTGT (23) TAGGGTG</u> | TAGGGGAG | CACAGAT (12) ACATAAACA | CCCC Subclonal |
| 25 | TCAT <u>ATATTTTGT (23) TAGGGTG</u> | ATCCC | CACAGAT (12) ACATAAACA | CCCC Subclonal |
| 19 | TCAT <u>ATATTTTGT (23) TAGGGTG</u> | GTCCCA | CACAGAT (12) ACATAAACA | CCCC Subclonal |
| 15 | TCAT <u>ATATTTTGT (23) TAGGGTG</u> | GAT | CACAGAT (12) ACATAAACA | CCCC Subclonal |
| 20 | TCAT <u>ATATTTTGT (23) TAGGGTG</u> | | CACAGAT (12) ACATAAACA | CCCC Subclonal |
| 17 | TCAT <u>ATATTTTGT (23) TAGGGTG</u> | | CACAGAT (12) ACATAAACA | CCCC Subclonal |
| 26 | TCAT <u>ATATTTTGT (23) TAGGGT-</u> | TCA | CACAGAT (12) ACATAAACA | CCCC Subclonal |

Type-II micro-deletions

| Patient | brk intr 1 | n-bases | brk intr 3 | Clonality status |
|---------|---|--------------------------|--|------------------|
| WT | TCCA <u>GCTTTAGAT (23) TACAGTG</u> <u>cRSS4</u> | NNNNNNNN (65Kb) NNNNNNNN | <u>CACAGAT (12) ACATAAACA</u> <u>cRSS1</u> | CCCC |
| 11 | TCCA <u>GCTTTAGAT (23) TACAGTG</u> | GCCCTCG | ---AGAT (12) ACATAAACA | CCCC Clonal |
| 9 | TCCA <u>GCTTTAGAT (23) TACAGTG</u> | GGTTT | CACAGAT (12) ACATAAACA | CCCC Subclonal |

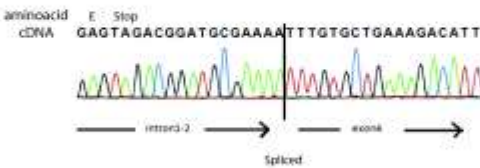
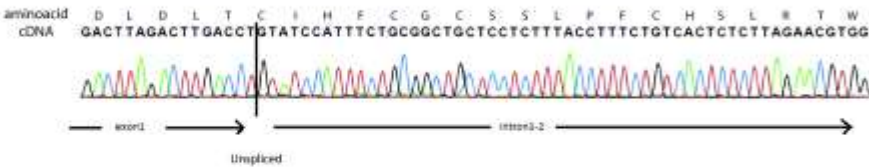
Type-III micro-deletions

| Patient | brk intr 3 | n-bases | brk intr 5 | Clonality status |
|---------|---|----------------------|--|-------------------|
| WT | GTAG <u>CACAGAT (12) ACATAAACA</u> <u>cRSS1</u> | NNNNNN (11Kb) NNNNNN | <u>TGTTCTCAA (23) TACTGTG</u> <u>cRSS2</u> | TAACACCCTAACTTTAG |
| 20 | GTAG ----- | AATTCGGGTCTGCCAG | ----- | TTAG Clonal |

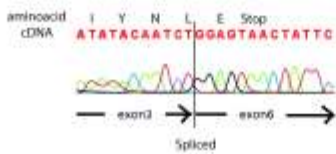
Figure 2

C

cDNA sequence for type I/II micro-deletions in patients #11, #12 and #21

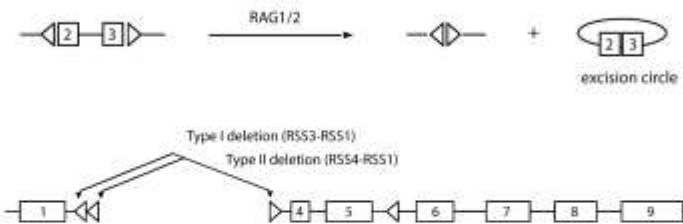


cDNA sequence for the type III micro-deletion in patient #20



D

Signal joint formation in exon 2-3 micro-deletions



Coding joint formation in exon 4-5 micro-deletion

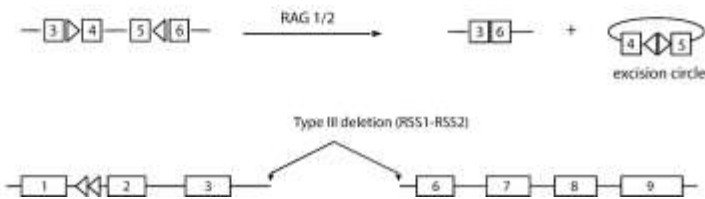


Figure 3

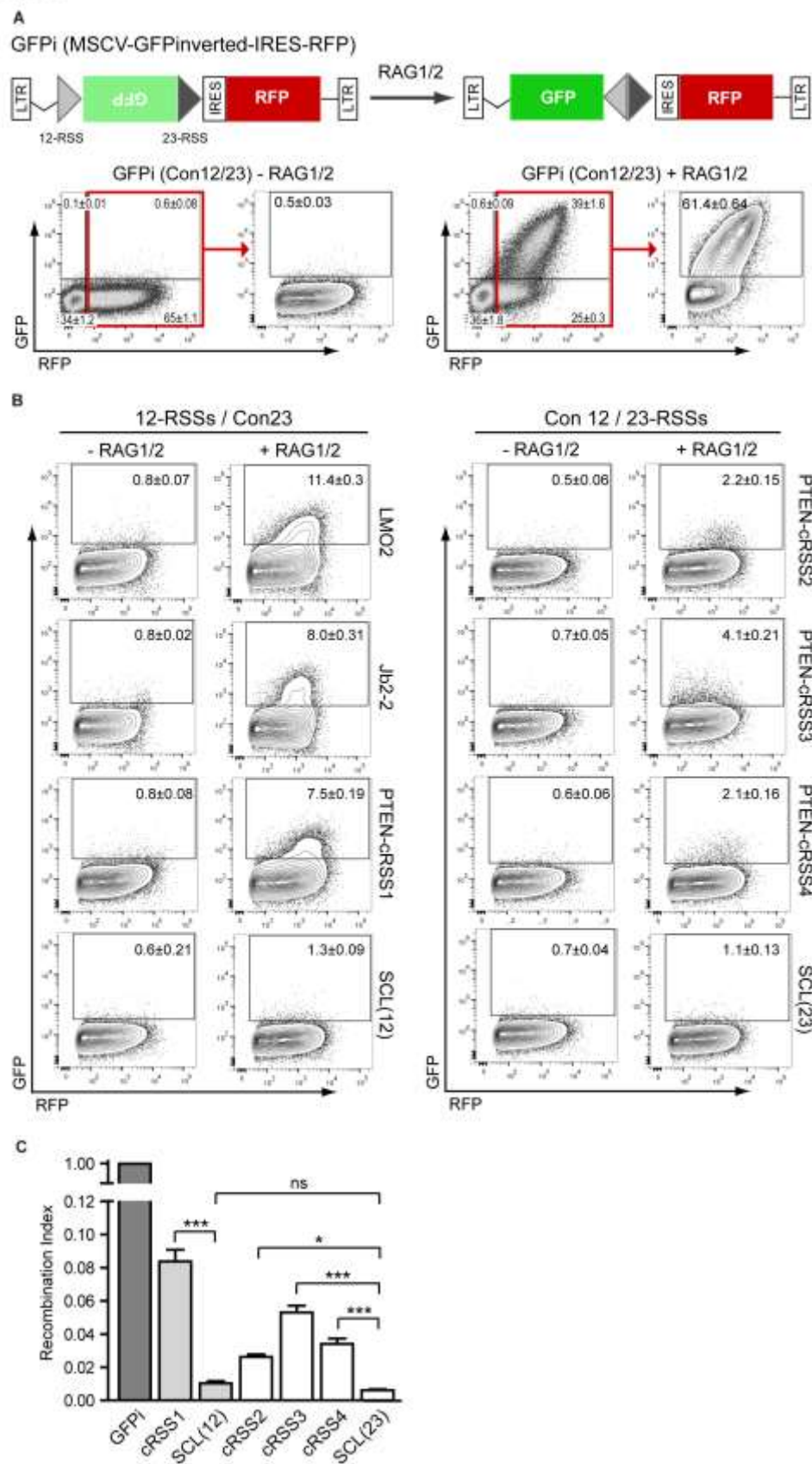
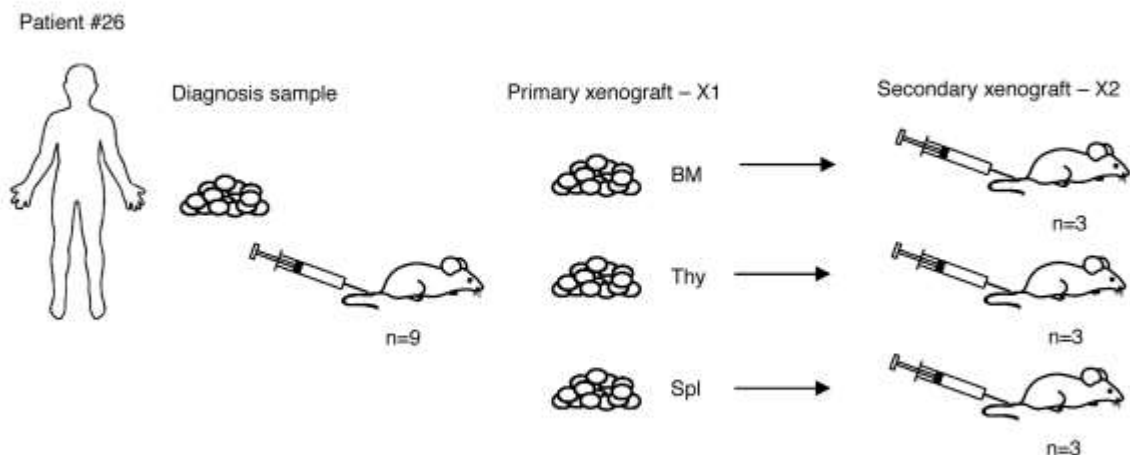


Figure 4

A



B

Type-I micro-deletions

| Sample | brk intr 1 | n-bases | brk intr 3 | Clonality status |
|------------------|----------------------------------|-----------------|--------------------------------------|------------------|
| cRSS3 | | | | |
| WT | TCAT <u>ATATTTTGT(23)TAGGGTG</u> | NNNN(65Kb)NNNN | cRSS1 <u>CACAGAT(12)ACATAAACA</u> | CCCC |
| 24 | TCAT <u>ATATTTTGT(23)TAGGGTG</u> | TAGGGGAG | <u>CACAGAT(12)ACATAAACA</u> | CCCC Subclonal |
| X1-24 thymus-1 | TCAT <u>ATATTTTGT(23)TAGGGTG</u> | TAGGGGAG | <u>CACAGAT(12)ACATAAACA</u> | CCCC Subclonal |
| X2(spl)-24 liver | TCAT <u>ATATTTTGT(23)TAGGGTG</u> | | <u>CACAGAT(12)ACATAAACA</u> | CCCC Subclonal |

Type-II micro-deletions

| Sample | brk intr 1 | n-bases | brk intr 3 | Clonality status |
|----------------|----------------------------------|--------------------|--------------------------------------|------------------|
| cRSS4 | | | | |
| WT | TCCA <u>GCTTTAGAT(23)TACAGTG</u> | NNNNNN(65Kb)NNNNNN | cRSS1 <u>CACAGAT(12)ACATAAACA</u> | CCCC |
| X1-24 thymus-2 | TCCA <u>GCTTTAGAT(23)TACAGTG</u> | TGCC | <u>CACAGAT(12)ACATAAACA</u> | CCCC Subclonal |

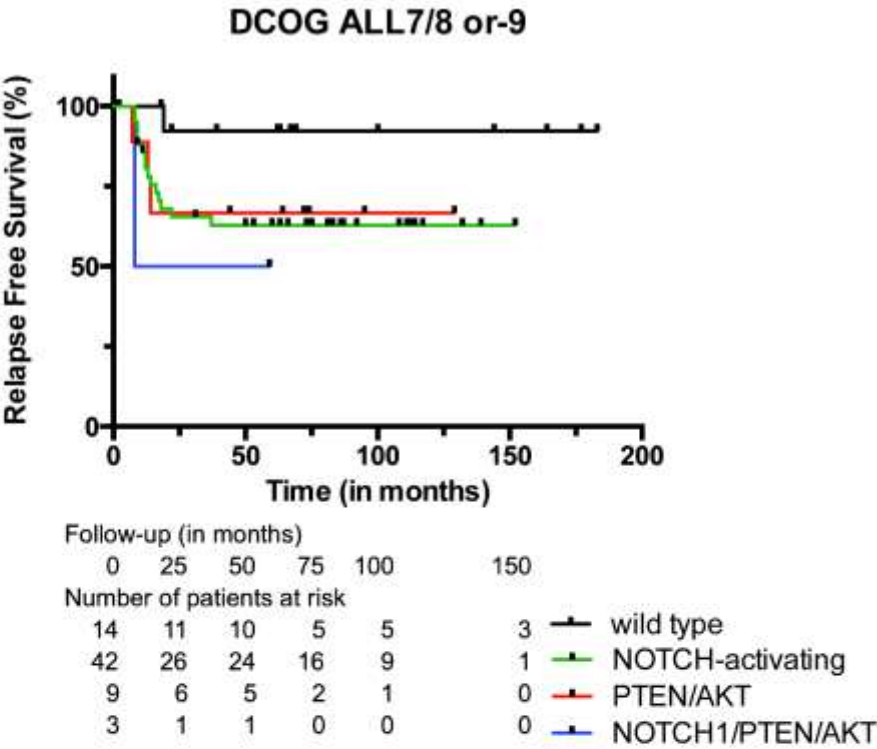
C

Type-I micro-deletions

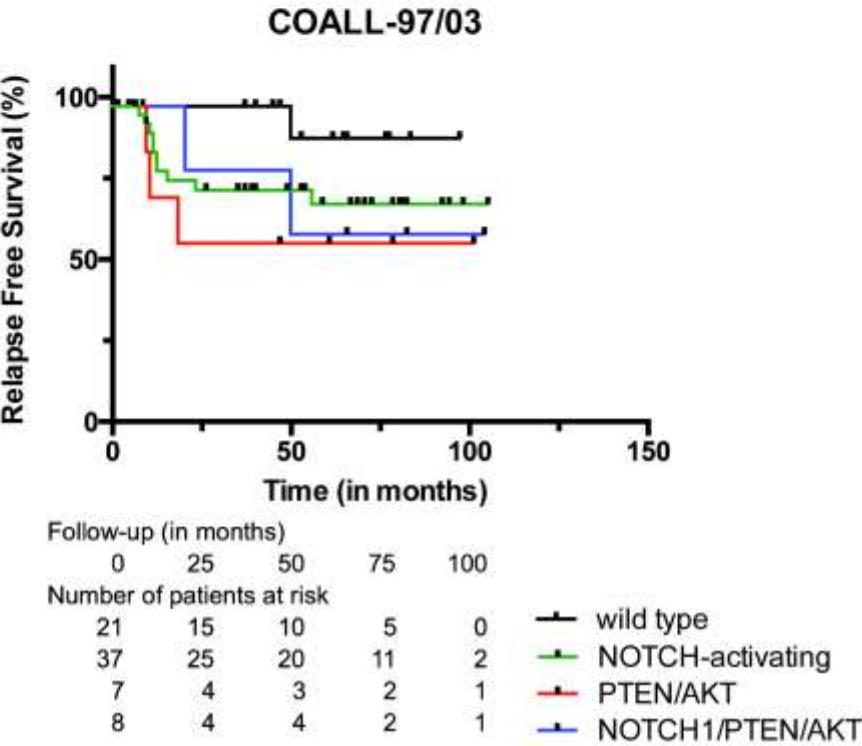
| Sample | brk intr 1 | n-base | brk intr 3 | Clonality status |
|--------|----------------------------------|------------------------|--------------------------------------|------------------|
| cRSS3 | | | | |
| WT | TCAT <u>ATATTTTGT(23)TAGGGTG</u> | NNNNNNNN(65Kb)NNNNNNNN | cRSS1 <u>CACAGAT(12)ACATAAACA</u> | CCCC |
| H-Thy1 | TCAT <u>ATATTTTGT(23)TAGGGTG</u> | | <u>CACAGAT(12)ACATAAACA</u> | CCCC Subclonal |
| H-Thy2 | TCAT <u>ATATTTTGT(23)TAGGGTG</u> | GGGTAAG | <u>---AGAT(12)ACATAAACA</u> | CCCC Subclonal |
| H-Thy3 | TCAT <u>ATATTTTGT(23)TAGGGTG</u> | GCCTCATCAAACC | <u>CACAGAT(12)ACATAAACA</u> | CCCC Subclonal |

Figure 5

A



B



Chapter 4

Lentiviral gene transfer into human and murine hematopoietic stem cells: size matters

Kirsten Canté-Barrett¹, **Rui D. Mendes**¹, Willem K. Smits¹, Yvette M. van Helsdingen-van Wijk¹, Rob Pieters², and Jules P.P. Meijerink¹

¹*Department of Pediatric Oncology/Hematology, Erasmus MC Rotterdam-Sophia Children's Hospital, Rotterdam, the Netherlands*

²*Princess Máxima Center of Pediatric Oncology, Utrecht, the Netherlands*

BMC Res Notes. 2016 Jun 16;9:312

Abstract

Contemporary biomedical research increasingly depends on techniques to induce or to inhibit expression of genes in hematopoietic stem cells (HSCs) or other primary cells to assess their roles on cellular processes including differentiation, apoptosis, and migration. Surprisingly little information is available to optimize lentiviral transduction of HSCs. We have therefore carefully optimized transduction of murine and human HSCs by optimizing vector design, serum-free virus production, and virus quantitation. We conclude that the viral RNA length, even in relatively small vectors, is an important factor affecting the lentiviral gene transfer on the level of both the virus production and the cellular transduction efficiency. Efficient transfer of large gene sequences into difficult-to-transduce primary cells will benefit from reducing the lentiviral construct size.

Introduction

Human immunodeficiency virus-based lentiviral gene transfer has been embraced in contemporary laboratory practice as an efficient procedure to shuttle gene-encoding RNA molecules into target cells, where they are reverse-transcribed and integrated into the host genome. Third-generation lentiviral vector systems have proven to be safe methods in gene therapy with very low risks of ongoing integrations in the host genome or generation of replication-competent viral particles.[1, 2] Lentiviruses infect both dividing and non-dividing cells, making them ideally suited to transduce human and murine hematopoietic stem cells (HSCs).[3-6] Many fields of research often require lentiviral constructs that drive gene expression from a promoter as well as a fluorescent reporter expressed from an internal ribosomal entry site or secondary promoter. The virus production (tested for vectors that encode viral RNA ranging from 4-7.5 kb in length)[7] and efficiency to transduce adherent cell lines seems dependent on the size of the lentiviral vector that encodes for the viral RNA. Vectors with viral RNA ranging from 5-9 kb generally tested decent, whereas those ranging from 10-18 kb transduced very poorly.[8] However, the nature of the target cell (cell line or primary cells) is crucial and most of the studies published to date have not addressed the transduction efficiency of primary cells. Additionally, it is important to control the production process of lentiviral particles (in HEK293T cells), and the quantitation method to determine the number of competent viral particles produced. Since little information is available regarding these variables and their effects on lentiviral transduction, we carefully optimized lentiviral transfer by optimizing vector design, serum-free virus production and quantitation, as well as by optimizing transduction of murine and human HSCs.

Methods

Cloning

For fast, highly flexible and reliable cloning purposes, we have adapted the third-generation self-inactivating lentiviral LeGO-iC2 vector[9] into a Gateway compatible destination vector (LeGO-DEST) by replacing the *Apal/PciI* 2kb insert with the Gateway *ccdB* cassette (Life Technologies). We assembled lentiviral expression constructs in which we cloned a promoter, a gene, a *Thosea asigna* virus 2A (T2A) element,[10] and the blue fluorescent protein reporter (mTagBFP). These, in combination with other lentiviral elements, are flanked by long terminal repeats and encode for the viral RNA. Upon translation, the gene and BFP moieties are efficiently separated by T2A cleavage in target cells (Figure 1).

Cell line

The human T-cell leukemia cell line JURKAT (DSMZ, #ACC-282) identity was confirmed by DNA fingerprinting and cells were regularly tested for mycoplasma contamination.

Virus production, concentration and quantification

We optimized HEK293T transfection in DMEM supplemented with 10% serum using XtremeGENE HP DNA Transfection Reagent (Roche, #06 366 236 001) to produce vesicular stomatitis virus-G pseudotyped virus particles without the addition of serum (low-serum Opti-MEM I with Glutamax: Life Technologies, #51985-026), in batches of 40 ml (harvest twice, with 24-hour intervals, from two confluent 14-cm dishes starting two days after transfection, see Supplementary Protocols). Low serum levels minimize the risk of premature differentiation of HSCs that are subjected to lentiviral transduction. Furthermore, fetal bovine serum can influence transduction,[11] and we found that increasing amounts of serum (ranging from 0% to 10%) negatively affects transduction rates of human HSCs, perhaps due to the aggregation of virus particles in the presence of serum proteins (data not shown). In relation to the quantitation of intact viral particles, the commonly used p24 protein ELISA method overestimates the number of functional viruses by the detection of incomplete, transduction-deficient viral particles as well as soluble p24 protein in the production medium.[12] Quantitative RT-PCR of encapsulated viral RNA particles is a valuable alternative. We therefore calculate the number of viral particles that we produce by quantification of viral RNA copies using RT-qPCR (two RNA copies per viral particle, see Supplementary Protocols). RT-qPCR is performed using primers flanking the cPPT region: 5'-AGGTGGAGAGAGAGACAGAGAC-3' and 5'-CTCTGCTGTCCCTGTAATAAAC-3'

Human CD34⁺ HSC transduction

Human CD34⁺ HSCs were positively selected from umbilical cordblood (Miltenyi Biotec, #130-100-453) and stimulated for 16-20 hours at a concentration of 1×10^6 cells/ml in X-VIVO 10 (Lonza, #BE04-743Q) supplemented with 50 ng/ml rhSCF (R&D, #255-SC), 20 ng/ml rhTPO (R&D, #288-TP/CF) and 50 ng/ml rhFlt3L (Miltenyi Biotec, #130-093-855). Prior to transduction, add protamine sulfate (Sigma, #P4020-1G) to the cells to a final concentration of 4 μ g/ml and pipet the concentrated virus (IVSS VIVASPIN 20 centrifugation concentration columns, Sartorius AG, Sigma-Aldrich, #Z614653-48EA) into a 50 μ g/ml retronectin (r-Fibronectin CH-296: TaKaRa, #T100A)-coated 96-well plate (Falcon, #351172). Add HSCs on top of the virus to a final volume of 200 μ l/well and mix by gently tapping the plate. Spinoculation: centrifuge the cells in the virus-containing medium at 1800 rpm, 32°C for 1 hour. Incubate the transduced cells at 37°C, 5%CO₂ for 24 hours before further use.

Results

We set out to relate the length of the viral RNA to the efficiency of virus production as well as to the potency of these viral batches to transduce the T-cell acute lymphoblastic leukemia line JURKAT. Our optimized lentiviral particle production consistently leads to near equal yields for consecutive batches produced from the same lentiviral construct (average variation of 4.3 fold (range 1.3-12 fold) for repetitive viral batches from 10 different constructs). We observed an inverse exponential correlation between the length of the viral RNA encoded by the construct and the number of viral particles produced by HEK293T cells (Figure 2). RT-qPCR in combination with functional titration experiments on a cell line reliably determines the number of viral particles required to efficiently transduce target cells. We investigated transduction efficiencies for virus batches of three different viral RNA lengths: 4215, 5190 and 5750 bases. For each construct, three independent viral batches were produced. Transduction experiments using serial dilutions of each viral batch were performed on JURKAT cells. Four to six days after transduction, the percentage of BFP-positive (transduced) cells was determined (Figure 3A, representative curves for three independently produced virus batches of each construct). While the intensity of BFP signal was highest for the virus with the shortest viral RNA and lower for longer constructs, in each case the transduced population was easily distinguished from non-transduced cells (inset in Figure 3A). The gene of interest encodes a protein with the C-terminal T2A that separates it from the BFP (Figure 3B). The percentage of transduced JURKAT cells remained unchanged for over three weeks of culture (not shown), reflecting stable integration of the viral genome in the host DNA. All independently produced batches from the same lentiviral

construct were virtually equally efficient to transduce JURKAT cells, indicating that the optimized production, quantitation and transduction procedures are very robust and reproducible. The transduction efficiency was highest for viruses with the shortest viral RNA sequence, and became progressively lower with increasing length. This is also true for all other virus batches produced with viral RNA lengths varying from 4 to 9 kb. In general, more viral particles with longer viral RNAs seem to be required to achieve equal transduction percentages in target cells compared to viruses with shorter viral RNAs.

We then tested the efficiency of these viral batches to transduce primary human CD34⁺ HSCs and murine 'lineage-negative' bone marrow (Lin⁻ BM) cells. To achieve comparable transduction rates for JURKAT, human HSCs and mouse Lin⁻ BM, a thousand-fold more virus particles of the 4215 bp vector are required for the primary cells than JURKAT (Figure 3C). This difference becomes even higher (up to 10⁴) for viruses with larger viral RNAs (5190 and 5750 bases). Thus, the viral RNA size moderately, albeit significantly, affects the transduction rate of a cell line such as JURKAT, but greatly influences transduction efficiencies of freshly isolated human HSCs and murine Lin⁻ BM cells. Transduced human HSCs cultured on stromal support cells retained an equal percentage of BFP-positive cells over the course of three weeks, indicating stable integration (data not shown). To investigate whether specific sequences can influence transduction efficiency, we generated new vectors harboring double or triple BFP sequences (5105 and 5995 bp, respectively; Figure 1). These vectors only differ in size from the single BFP construct, but contain the same sequence. Also for these viruses, the transduction efficiency in JURKAT was highest for the shortest vector and reduced with increased length (Figure 4).

Discussion

We conclude that while larger viral RNA size negatively affects both virus production and transduction of target cells, other factors can also influence the transduction efficiency (e.g. sequence). This is evident from the observation that the 5750 bp vector (containing the 1535 bp gene) revealed lower transduction efficiency than the slightly larger triple BFP vector (5995 bp). The transduction efficiency of human or mouse stem cells decreases tremendously for viruses with viral RNAs approaching 6 kb or larger. Lentiviral vectors encoding smaller viral RNA sequences perform better and even a reduction of merely 600 bp (5750 versus 5190 bp) already improves transduction efficiency by more than three-fold (Figure 3). To produce more efficient lentiviruses, reducing the viral RNA backbone size by removal of non-essential sequences may be effective. Codon optimization alters the gene sequence without affecting the protein

sequence and may also increase transduction efficiency. Additionally, the development of smaller reporter genes or complete removal of the reporter may further enhance the transduction efficiency. In the absence of a fluorescent reporter, integrated lentiviral constructs into the host genome or expression of lentiviral transgene mRNA can be quantified by qPCR[13, 14] or RT-qPCR,[15] respectively. In conclusion, size reduction of lentiviral constructs will facilitate efficient transfer of large gene sequences into difficult-to-transduce primary cells, and will be most helpful in many fields of (basic) research.

Ethics approval and consent to participate

The use of human umbilical cord materials from healthy individuals was approved by the Medical Ethical Review Board of the Erasmus MC Rotterdam (MEC-2009-430) and in accordance with the Declaration of Helsinki. C57Bl/6 mice were housed under specific pathogen free conditions at the animal facility of Erasmus MC according to institutional guidelines. The use of murine bone marrow for the experiments has been approved by the Erasmus MC committee for animal welfare (DEC #103-12-02) and is in compliance with Dutch legislation.

Acknowledgements

We thank Karin Pike-Overzet for helpful discussions. This work was supported by the Children Cancer Free Foundation (Stichting Kinderen Kankervrij); grants KiKa-2008-29 (RDM, WKS, and KC-B), KiKa-2013-116 (KC-B), and KiKa-2014-141 (YMH-W).

Authorship Contributions

KC-B, RDM, WKS, and YMH-W performed experiments and/or analyzed data; KC-B and JPPM designed study and wrote manuscript; RP and JPPM supervised the study.

Conflicts of Interest Disclosure and Supplementary Information

The authors declare no conflict of interest. Supplementary Protocols (.docx file) include protocols for virus production and concentration, viral RNA isolation and RT-qPCR, and human CD34⁺ HSC transduction.

References

1. Debyser Z. Biosafety of Lentiviral Vectors. *Current gene therapy*. 2003; 3:6.
2. Dull T, Zufferey R, Kelly M, Mandel RJ, Nguyen M, Trono D, Naldini L. A Third-Generation Lentivirus Vector with a Conditional Packaging System. *Journal of virology*. 1998; 72:11.
3. Naldini L, Blomer U, Gallay P, Ory D, Mulligan R, Gage FH, Verma IM, Trono D. In Vivo Gene Delivery and Stable Transduction of Nondividing Cells by a Lentiviral Vector. *Science*. 1996; 272:5259.
4. Miyoshi H, Smith KA, Mosier DE, Verma IM, Torbett BE. Transduction of Human Cd34+ Cells That Mediate Long-Term Engraftment of Nod/Scid Mice by Hiv Vectors. *Science*. 1999; 283:5402.
5. Naldini L. Lentiviruses as Gene Transfer Agents for Delivery to Non-Dividing Cells. *Current opinion in biotechnology*. 1998; 9:5.
6. Hong Y, Lee K, Yu SS, Kim S, Kim JG, Shin HY, Kim S. Factors Affecting Retrovirus-Mediated Gene Transfer to Human Cd34+ Cells. *The journal of gene medicine*. 2004; 6:7.
7. al Yacoub N, Romanowska M, Haritonova N, Foerster J. Optimized Production and Concentration of Lentiviral Vectors Containing Large Inserts. *The journal of gene medicine*. 2007; 9:7.
8. Kumar M, Keller B, Makalou N, Sutton RE. Systematic Determination of the Packaging Limit of Lentiviral Vectors. *Human gene therapy*. 2001; 12:15.
9. Weber K, Bartsch U, Stocking C, Fehse B. A Multicolor Panel of Novel Lentiviral "Gene Ontology" (Lego) Vectors for Functional Gene Analysis. *Molecular therapy : the journal of the American Society of Gene Therapy*. 2008; 16:4.
10. Szymczak AL, Vignali DA. Development of 2a Peptide-Based Strategies in the Design of Multicistronic Vectors. *Expert opinion on biological therapy*. 2005; 5:5.
11. Denning W, Das S, Guo S, Xu J, Kappes JC, Hel Z. Optimization of the Transductional Efficiency of Lentiviral Vectors: Effect of Sera and Polycations. *Molecular biotechnology*. 2013; 53:3.
12. Geraerts M, Willems S, Baekelandt V, Debyser Z, Gijssels R. Comparison of Lentiviral Vector Titration Methods. *BMC biotechnology*. 2006; 6.
13. Barczak W, Suchorska W, Rubis B, Kulcenty K. Universal Real-Time Pcr-Based Assay for Lentiviral Titration. *Molecular biotechnology*. 2014.
14. Christodoulou I, Patsali P, Stephanou C, Antoniou M, Kleanthous M, Lederer CW. Measurement of Lentiviral Vector Titre and Copy Number by Cross-Species Duplex Quantitative Pcr. *Gene therapy*. 2016; 23:1.
15. Lizée G, Aerts JL, Gonzales MI, Chinnasamy N, Morgan RA, Topalian SL. Real-Time Quantitative Reverse Transcriptase-Polymerase Chain Reaction as a Method for Determining Lentiviral Vector Titers and Measuring Transgene Expression. *Human gene therapy*. 2003; 14:6.

Figure legends

Figure 1. Schematic representation of the elements cloned into our Gateway compatible LeGO-DEST vector. The different elements and their sizes (bp) that were cloned into our Gateway compatible LeGO-DEST vector to generate lentiviral expression vectors of different sizes between SIN-LTRs: 4215, 5190, and 5750 bp. SFFV: Spleen Focus Forming Virus promoter, T2A: *Thosea asigna* virus 2A element, BFP: blue fluorescent protein. Grey boxes indicate lentiviral elements in LeGO-DEST.

Figure 2. Relationship between the viral RNA length and production of lentiviral particles. The quantity of lentiviral particles that are produced (calculated using RT-qPCR) plotted as a function of the viral RNA length (kb). Each point in the plot represents an individual virus batch, each time produced using the pLEGO-DEST backbone and the same transfection method and producer cell line (see S1 Protocol). The viral RNA length is the length of the sequence flanked by the 5' and 3' SIN-LTRs in the lentiviral expression vector (see Figure 1).

Figure 3. Transduction efficiency of JURKAT, CD34⁺ human HSCs, and murine Lin⁻ BM with lentiviral vectors of varying sizes. Transduction efficiency of different cell types as a function of the amount of virus particles, measured 4-6 days after transduction and expressed as the BFP⁺ percentage of the viable cells. Lentiviruses with three different viral RNA sizes (length between SIN-LTRs) are compared: 4215 (solid black lines), 5190 (dark-grey dashed lines), and 5750 bp (light-grey dashed lines). A. Triplicate transductions \pm SD of JURKAT cells using serial dilutions of each lentivirus, representative of one of three individually produced virus batches. *Inset*: example flow cytometry plot displaying the BFP⁺ transduced fraction. B. Western blot probed with anti-2A antibody. Total lysates expressing the T2A-tagged proteins of the 4215 bp (BFP only, no T2A), 5190 and 5750 bp constructs, separated from BFP. *: non-specific signals. C. Triplicate transductions \pm SD of human CD34⁺ HSCs (solid triangles) and murine Lin⁻ BM (solid circles) with the three lentiviruses.

Figure 4. Transduction efficiency of JURKAT with independently generated and different lentiviral vectors of varying sizes. A. Transduction efficiency of JURKAT cells as a function of the amount of virus particles, measured 4-6 days after transduction and expressed as the BFP⁺ percentage of the viable cells. Lentiviruses with three different viral RNA sizes (length between SIN-LTRs) are compared: 4215 (1xBFP; black), 5105 (2xBFP; blue), and 5995 bp (3xBFP; red). Triplicate transductions \pm SD using serial dilutions of each lentivirus, representative of one of

three individually produced virus batches. B. Average transduction percentage of the different viruses, measured at the number of virus particles that yielded 75% in the '1xBFP' (4215 bp) vector (indicated by dashed vertical lines in Figure 3 and 4A)

Figure1

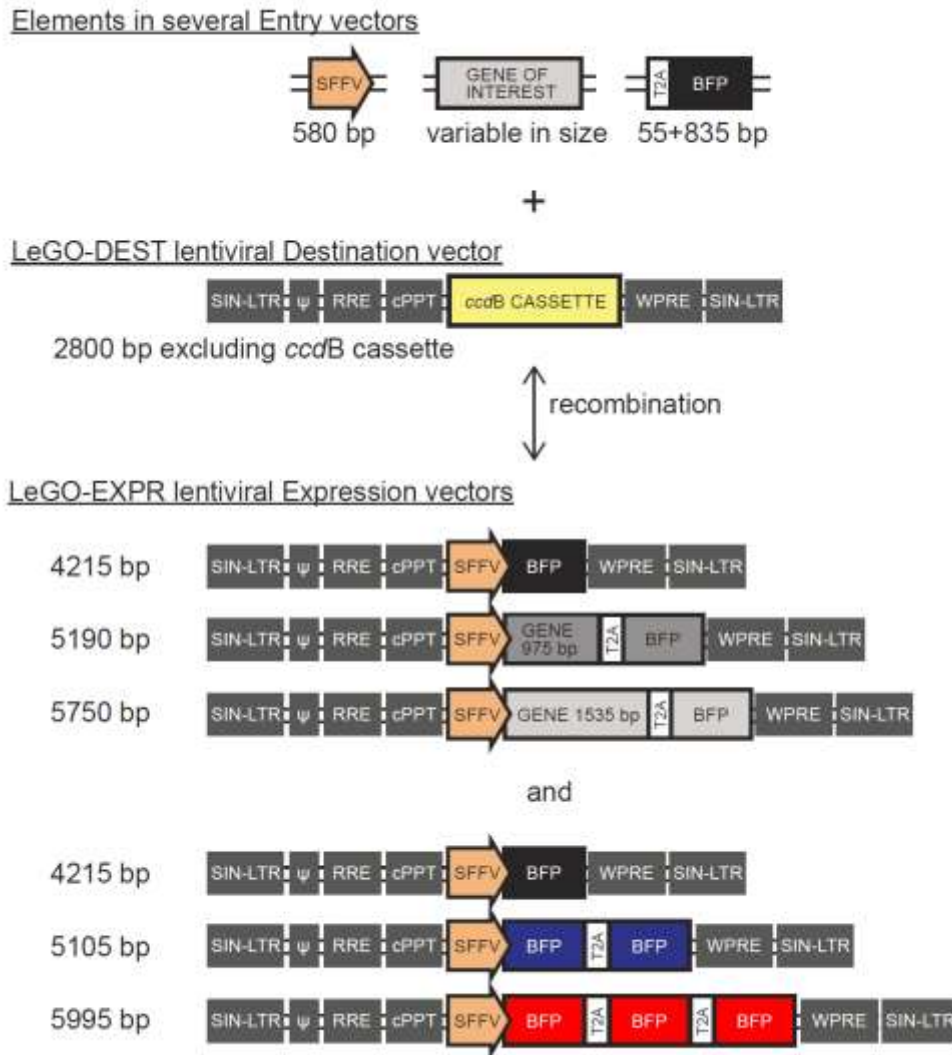


Figure 2

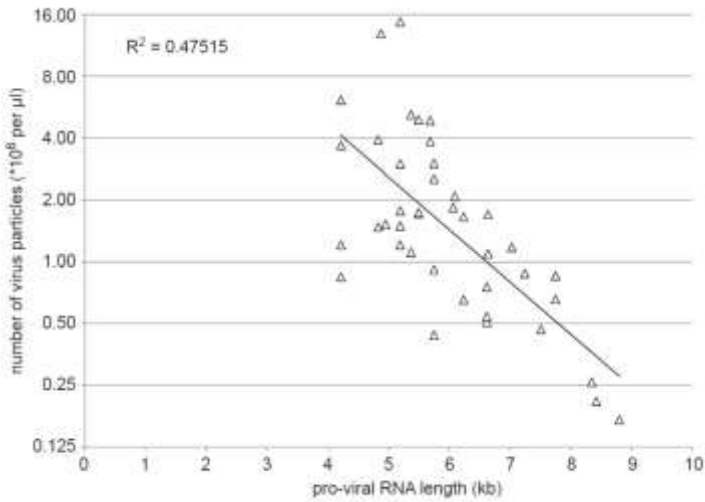


Figure 3

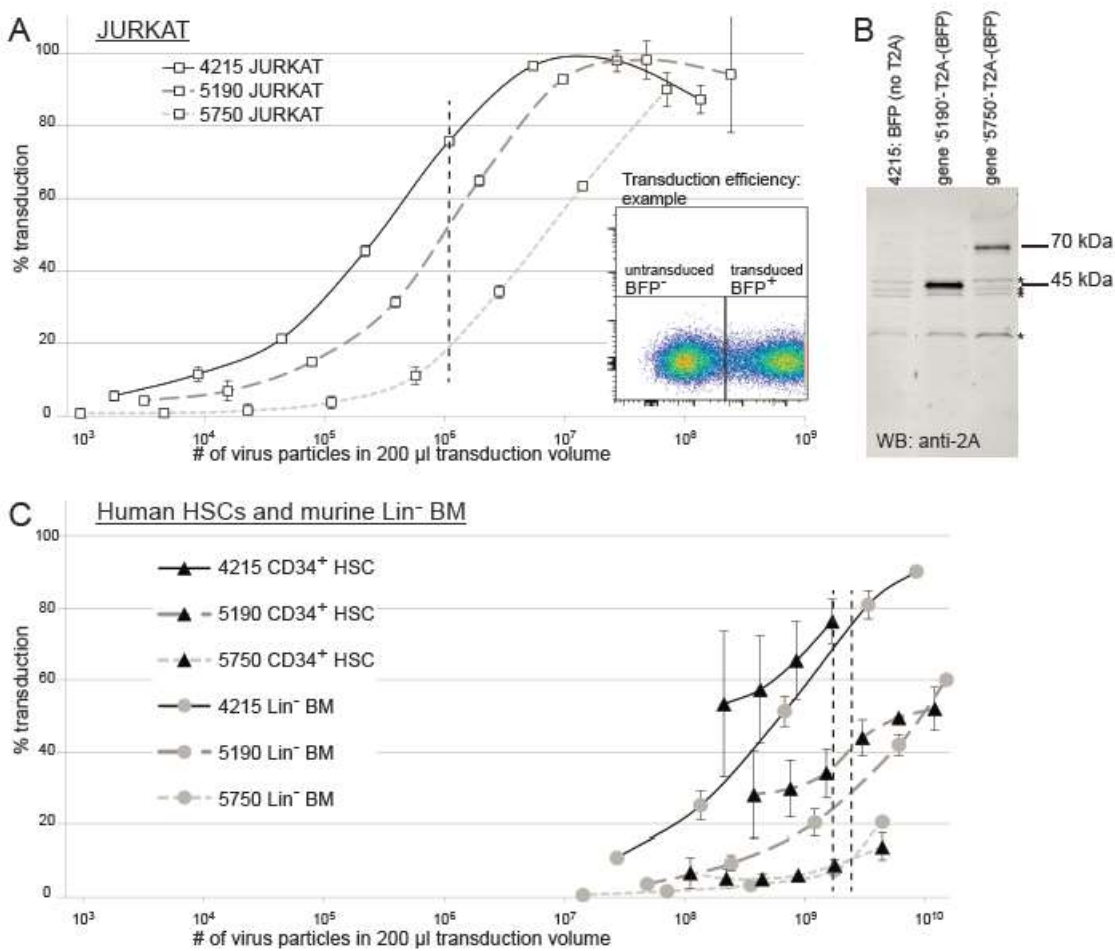
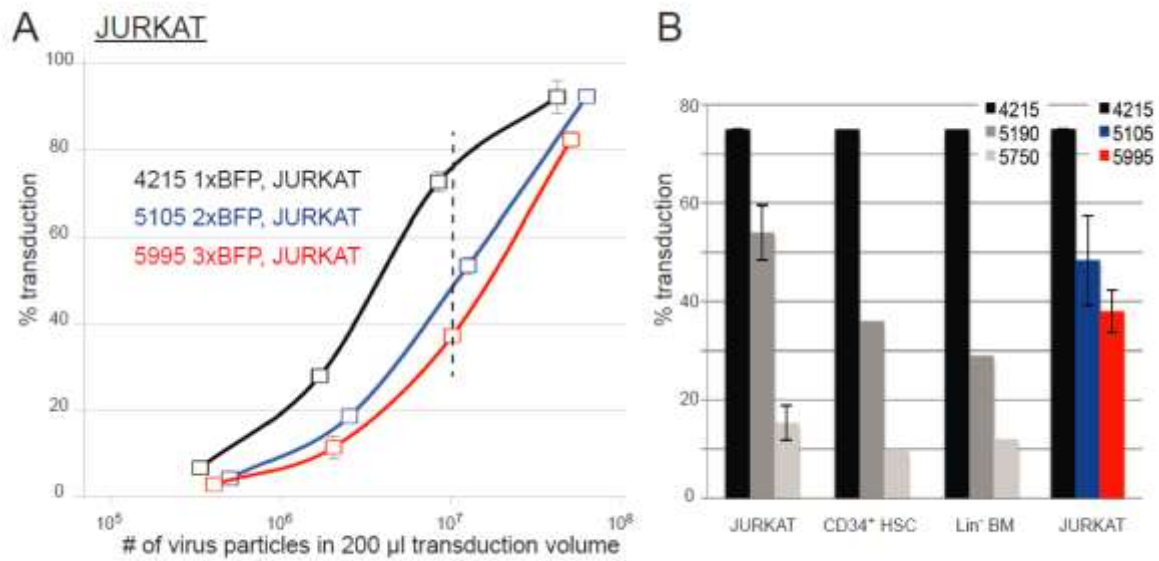


Figure 4



Chapter 5

Downregulation of CD44 functionally defines human T-cell lineage commitment

Rui D. Mendes^{1,5}, Kirsten Canté-Barrett^{1,5}, Yunlei Li¹, Eric Vroegindewij¹, Karin Pike-Overzet², Tamara Wabeke³, Anton W. Langerak³, Rob Pieters^{1,4}, Frank J.T. Staal², and Jules P.P. Meijerink¹

¹Department of Pediatric Oncology/Hematology, Erasmus Medical Center-Sophia Children's Hospital, Rotterdam, The Netherlands; ²Department of Immunohematology and Blood Transfusion, Leiden University Medical Center, Leiden, The Netherlands; ³Department of Immunology, Erasmus Medical Center, Rotterdam, The Netherlands; ⁴Princess Máxima Center for Pediatric Oncology, Utrecht, The Netherlands.

⁵Co-first author

Manuscript submitted

Abstract

Human T-cell development is less well studied than its murine counterpart due to the lack of genetic tools and the difficulty of obtaining cells and tissues. Here we report the transcriptional landscape of 11 immature, consecutive human T-cell developmental stages. The changes in gene expression of cultured stem cells on OP9-DL1 match those of *ex vivo* isolated human thymocytes. These analyses led us to define evolutionary conserved gene signatures that represent pre- and post- $\alpha\beta$ T-cell commitment stages. We found that loss of CD44 marks human T-cell commitment in early CD7⁺CD5⁺CD45^{dim} cells, before the acquisition of CD1a surface expression. The CD44⁻CD1a⁺ post-committed thymocytes have initiated in frame TCR rearrangements and have completely lost the capacity to develop into myeloid, B- and NK-cells, unlike uncommitted CD44⁺CD1a⁻ thymocytes. Therefore, loss of CD44 represents a previously unrecognized stage that defines the earliest committed T-cell population in the human thymus.

Introduction

T-cell development in the thymus is a complex process that depends on sequential transcriptional and epigenetic changes leading to T-cell lineage commitment while suppressing alternative cell fates.(Rothenberg et al., 2008; Zhang et al., 2012) T-cell development has been extensively studied in mice, while human T-cell development is less well defined due to limited availability of human thymus material and inherent genetic diversity. Research has focused on mimicking the thymus environment using *in vitro* differentiation cultures starting with hematopoietic stem cells (HSCs) derived from human cord blood or bone marrow. Historically, human-mouse hybrid fetal thymic organ cultures (FTOC) have been used to functionally define the various stages of human T-cell development.(Plum et al., 1994) Later, OP9-DL1 co-cultures have been proven more useful to study T-cell differentiation.(Schmitt and Zuniga-Pflucker, 2002) The expression of the Delta-like 1 ligand on bone marrow-derived stromal cells from *op/op* M-CSF deficient mice(Yoshida et al., 1990) induces NOTCH signaling in target cells of the hematopoietic lineage. NOTCH signaling promotes T-cell differentiation while inhibiting B-cell differentiation.(Radtke et al., 2004) The differentiation of HSCs in this co-culture recapitulates human *in vivo* T-cell development as measured by the successive acquisition of CD7, CD5, CD1a, and CD4+CD8 surface markers.(Awong et al., 2009; La Motte-Mohs et al., 2005)

During development in the thymus, early T-cell precursors migrate within the cortex from which the positively selected CD4 and CD8 double positive (DP) thymocytes migrate to the

medulla. As a result, developing thymocytes encounter specific signals at specific locations in the thymus.(Halkias et al., 2014) The OP9-DL1 *in vitro* co-culture lacks the proper thymus architecture required for T-cell development. Hence, it supports the development of early T-cell development until the DP stage, with few cells reaching the CD8 single-positive (SP) stage. These few CD8SP cells have most likely been subject to positive selection based on MHC class I expression on OP9-DL1 cells.(Dervovic and Zuniga-Pflucker, 2010) Despite the presence of these rare SP cells it is generally believed that the absence of proper antigen-presenting cells and a thymus environment impair complete positive and negative selection processes in OP9-DL1 co-cultures.

Early T-cell progenitors (ETPs) are pre-committed, multi-potent thymocytes that retain the ability to develop into hematopoietic cells other than the T-cell lineage including NK-cells, B-cells, and cells of the myeloid and erythroid lineages.(Bell and Bhandoola, 2008; Bhandoola et al., 2007; Weerkamp et al., 2006) Fully-committed thymocytes have lost multi-potency and have undergone RAG1/2-mediated T-cell receptor (TCR) alpha and beta chain rearrangements. Human CD4/8 double-negative (DN) thymocytes acquire CD1a at the proliferation stage when cells are committed to the T-cell fate.(Dik et al., 2005) The human hematopoietic stem cell marker CD34 is gradually lost throughout development but it is a poor marker for 'stemness' of pre-committed thymocytes as it remains expressed (albeit dimly) on most CD1a⁺ T-cell committed thymocytes.(Dik et al., 2005; Weerkamp et al., 2006; Weerkamp et al., 2005) Until now, upregulation of CD1a has been used to define human T-lineage commitment,(Weerkamp et al., 2006) but the human DN thymocyte maturation stages and the exact T-cell commitment point had not been clearly defined.

In this study, we have generated gene expression profiles of sequential human T-cell differentiation stages, derived from umbilical cord blood stem/progenitor cells that have full multi-lineage differentiation potential and that were cultured on OP9-DL1 stromal cells. Comparison between these *in vitro*-derived signatures and the gene signatures from normal murine and human early T-cell development stages in the thymus reveals strong conservation of pre- and post-T-cell commitment transcriptional profiles. From this analysis, we found that loss of human CD44 expression at the surface membrane of early DN thymocytes marks the loss of alternative cell fate decisions, initiation of *TCRB* recombination, and T-cell commitment.

Results

Consecutive stages of human *in vitro* T-cell differentiation represent two major gene signatures

CD34⁺ HSCs isolated from human umbilical cord blood were cultured on OP9-DL1 stromal cells. At various time points during a 27-day co-culture (performed in biological triplicate), we sorted seven distinct T-cell populations as progeny from pooled CD34⁺ HSCs from multiple cord blood donors (Fig. 1A). We performed microarray gene expression analysis on a total of 26 OP9-DL1-generated T-cell populations and three CD34⁺ HSC starting pools. Principal component analysis (PCA) revealed three distinct clusters, one of which consists of the CD34⁺ stem/progenitor starting populations (Fig. S1A). After filtering for probes with high variance (top 5%) and high log₂ expression (>8), the expression profile was trimmed to 1387 genes (2179 probesets) that retain the three PCA clusters (Fig. 1B, Table S1). One cluster represents the gene expression profile of CD34⁺ cells. The other two developmental clusters characterize ‘early’ T-cell progenitor (ETP) populations (including the CD7⁺ and the day 7 and day 12 CD7⁺CD5⁺ sorted populations) and ‘late’ T-cell differentiation populations (including the day 18 CD7⁺CD5⁺, CD7⁺CD5⁺CD1a⁺, CD4/8 DP, and SP sorted populations). The ‘early’ and ‘late’ clusters represent a major change in transcriptional programs (Fig. 1B,C). Of note, the CD7⁺CD5⁺ population (in orange) is divided over both clusters (Fig. 1B); those fractions that were sorted on day 7 and day 12 of the co-culture cluster with the early CD7⁺ population, whereas those sorted on day 18 fall in the ‘late’ cluster. Therefore, this CD7⁺CD5⁺CD1a⁻ population is not homogeneous and represents two distinct differentiation stages that dramatically differ in their transcriptional programs.

The 1387 genes of our expression profile harbor 16 different gene signatures of co-expressed genes (Fig. 1C,D, Table S1). Based on gene set enrichment analysis (GSEA) we identified five signatures that are differentially expressed between the early and late T-cell differentiation populations (Figure 1E). Signatures #7 and #8 are enriched for genes that are highly expressed in the early T-cell progenitor stages, and downregulated at later T-cell differentiation stages. Signatures #2, #15 and #16 are enriched for genes that become expressed at later T-cell differentiation stages. These five gene signatures—comprising 547 genes in total—faithfully distinguish the early and late T-cell differentiation populations (Fig. 2A, Fig. S1B, Table S1). Early T-cell differentiation signatures #7 and #8 include transcription (co)factors that are important for HSC/ETP maintenance and/or lineage determination, such as *MEIS1*, *LMO2*, *MEF2C*, *LYL1*, and *HHEX*. (Goodings et al., 2015; Yui and Rothenberg, 2014; Zohren et al., 2012) Genes encoding transcription factors that are responsible for T-cell identity (including *GATA3*, *TCF7*, *BCL11B*) and TCR rearrangements (*RAG1/2*) are present in late signatures #2, #15 or #16. Therefore, these five signatures contain essential genes that are associated with transcriptional programs required for stem/progenitor identity, T-cell specification and lineage commitment.

The *in vitro* gene expression signatures recapitulate *in vivo* signatures of pre- and post-T-cell committed thymocytes

We then compared our gene expression signatures to those from murine and human *ex vivo* sorted thymocyte subsets.(Dik et al., 2005; Mingueneau et al., 2013; Yui and Rothenberg, 2014) We performed GSEA to determine the enrichment of our five gene signatures in the dataset of consecutive T-cell development populations from murine thymi, generated by the Immunological Genome Consortium.(Mingueneau et al., 2013) As a result, we demonstrate that our “early” signatures #7 and #8 are significantly enriched for genes expressed in pre-committed murine subsets ranging from long-term HSCs to DN1-DN2a stages (Fig. 2B; false discovery rate (FDR)-corrected p -value is $q < 0.001$). Known pre-commitment genes such as *HHEX*, *KIT*, *MEF2C*, *LYL1* are among the top-10 genes in our early signatures (Fig. 2B: left panel, Table S2). Our “late” T-cell development gene signatures #2, #15, and #16 are significantly enriched for genes expressed in post-committed murine thymocyte subsets (DN2b and later subsets) and include *BCL11B*, *TCF7*, *LEF1*, *GATA3* and various molecules associated with TCR signaling (FDR $q = 0.028$, Fig. 2B: right panel, Table S2). Next, we applied our *in vitro* gene signatures onto the gene expression data of human *ex vivo* sorted thymocyte populations.(Dik et al., 2005) Interestingly, similar maturation stages of *in vitro* and *ex vivo* populations group together in hierarchical clustering (not shown) and our gene signatures that discriminate pre- and post-commitment also divide the *ex vivo* sorted human thymocytes into pre-committed (the CD1a⁻ populations(Dik et al., 2005)) and post-committed populations (CD1a⁺ and beyond; Fig. 2C and Table S3).(Dik et al., 2005) These results show that the pre- and post-commitment programs of murine and human *in vivo* T-cell development are highly conserved in T-cells cultured from HSCs on OP9-DL1 stromal support.

Loss of CD44 surface expression marks T-cell commitment in early human thymocytes

Further examination of the genes expressed in the *ex vivo* CD34⁺CD1a⁻ human thymocyte sorted populations(Dik et al., 2005) revealed that certain post-commitment genes (including *PTCRA*, *RAG1*, *RAG2* and *DNTT* (known as *TDT*)) are already expressed, suggesting that thymocytes prior to their expression of CD1a have initiated *TCRB* rearrangement and committed to the $\alpha\beta$ T-lineage (Fig. 2C, Table S3). From our pre- and post-commitment gene signatures, we searched for surface antigens that are differentially expressed between pre and post commitment. In doing so, we observed that expression of *CD44*, *CD33*, *CD34*, *CD63*, *CD69*, and *CD300LF* were reduced before the CD1a⁺ stage (Fig. 1D). In order to investigate whether

loss of CD44 protein marks human T-cell commitment, we examined CD44 surface expression on differentiating T-cells in the OP9-DL1 system and observed that CD44 is lost from CD5⁺ T-cells before they express CD1a (Fig. 3A). We sorted subsets at days 12 and 18 from two independent OP9-DL1 co-cultures with the inclusion of CD44 as an additional sorting marker, and were able to discriminate CD5⁺CD44⁺CD1a⁻ (E1), CD5⁺CD44⁻CD1a⁻ (E2) and CD5⁺CD44⁻CD1a⁺ (F) populations (Fig. 3A). Microfluidic cards-based RT-qPCR analysis of 47 pre- or post-commitment genes strongly supported a pre-commitment profile in sorted populations A through E1, and a T-lineage commitment profile in populations E2 and F (Fig. 3B). Therefore, T-cell commitment of differentiating T-cells cultured on OP9-DL1 is associated with loss of CD44 before the acquisition of CD1a surface expression.

Next, we asked whether CD44 represents a pre-commitment marker in *ex vivo* human thymocytes. Thymi from four independent immunologically healthy donors were used to sort CD7⁺CD5⁺CD45^{dim} immature thymocytes (following depletion of CD3, CD4, CD8, and CD19 cells). Within the CD7⁺CD5⁺CD45^{dim} cells, we sorted three populations: CD44⁺CD1a⁻ (population I, equivalent to the OP9-DL1 co-cultured population E1), CD44⁻CD1a⁻ (II, equivalent to E2), and CD44⁻CD1a⁺ (III, equivalent to F) (Fig. 3C). CD44 becomes highly re-expressed on mature thymocytes, requiring depletion of the large bulk of mature cells that constitutes approximately 99% of all thymocytes. CD44 expression on early, pre-committed thymocytes is dim in comparison to CD44 expression on mature T-cells, (Marquez et al., 1995) and correlates with CD45 expression levels (Fig. S2A). CD45^{dim} thymocytes predominantly consist of developmentally early cells including CD4/8 DN, CD4 immature single positive (ISP) and some CD3⁻ DP cells, whereas CD45^{bright} thymocytes mostly consist of CD3⁺ DP, CD4 SP and CD8 SP cells (Fig. S2A). Mature B-, myeloid, and NK-cells are present in very low numbers in the thymus, (Weerkamp et al., 2005) and these mature CD3/4/8 triple-negative cells are also CD45^{bright}CD44^{bright} (Fig. S2B). Therefore, immature (CD34⁺)CD7⁺CD5⁺ thymocytes were sorted from the CD45^{dim} population, and further sorted into sub-populations I-III based on CD44 and CD1a expression. Hierarchical cluster analysis of the gene expression levels of the panel of 47 pre- and post commitment genes identifies population I (CD44⁺) as pre-committed thymocytes, whereas populations II and III (CD44⁻) represent post-committed thymocytes (Fig. 3D). Therefore, loss of CD44 marks the transition to human T-cell commitment *in vivo*, consistent with our *in vitro* data. Of note, expression of the stem-cell marker CD34 gradually decreased in populations I-III but was still present on CD1a⁺ thymocytes (population III, Fig. 3C), in line with previous observations. (Dik et al., 2005; Weerkamp et al., 2006) Therefore, loss of CD34 expression is not directly associated with T-cell commitment.

Loss of CD44 functionally marks human T-cell commitment

$\alpha\beta$ T-cell committed thymocytes have initiated or completed *TCRB* gene rearrangements. To further validate commitment in human thymocyte populations II and III, we analyzed D β -J β and V β -J β recombination events using GeneScan analysis (van Dongen et al., 2003). Analysis of sorted CD44⁺ uncommitted thymocytes (population I) did not reveal detectable rearrangements (Fig. 4A). In contrast, incomplete D β -J β and V β -J β recombination events were identified in the earliest population of post-committed CD44⁺CD1a⁺ thymocytes (II) with a more complete spectrum of full recombinations present in CD44⁺CD1a⁺ thymocytes (III) (Fig. 4A). To provide functional evidence that loss of CD44 marks T-cell commitment, we investigated the intrinsic potential of alternative cell fate decisions in the three sorted human thymocyte populations I-III. We co-cultured thymocyte populations I (pre-committed), II and III (post-committed) on OP9-DL1, OP9-GFP (control), and a mixture of OP9-DL1/OP9-GFP stromal cells. These cultures were performed in the presence of cytokines that support differentiation into various hematopoietic lineages. After seven days of co-culture, microscopic inspection already revealed the inability of committed thymocyte populations II and III to proliferate in co-cultures with OP9-GFP stromal cells that do not support T-cell differentiation (Fig. S3). After ten days of co-culture under various stromal support conditions we determined the total numbers of T-, NK-, B-, and myeloid cells based on lineage markers CD5 and CD1a (T), CD56 and CD94 (NK), CD19 (B), and CD33 (myeloid) (Fig. 4B, C). The pre-committed thymocytes (population I) proliferated well and differentiated into cells of all four lineages on OP9-DL1 support. This pre-commitment population also differentiated into B-, myeloid and NK-cells—but not T-cells—when cultured on OP9 control cells. In contrast, committed thymocyte populations II and III gave rise to CD5⁺CD1a⁺ T-cells only, indicating that these populations have lost the potential for alternative cell fates (Fig. 4C). The fact that both populations II (CD44⁺CD1a⁺) and III (CD44⁺CD1a⁺) fully adopted the T-cell fate illustrates that loss of CD44—rather than the acquisition of CD1a—is the first surface marker that identifies post-committed human thymocytes.

Discussion

The early stages of human T-cell development have been investigated (Spits, 2002) without or with gene expression profiling, (Dik et al., 2005) using FTOC cultures, (Pike-Overzet et al., 2007; Plum et al., 1994; Res et al., 1996) and more recently co-cultures on OP9-DL1. (Van Coppenolle et al., 2009) During the last decade, the OP9-DL1 *in vitro* co-culture system has been used widely to study T-cell development. (de Pooter and Zuniga-Pflucker, 2007) The use of human

hematopoietic progenitors in this culture system offers a relatively easy and efficient tool to translate studies from *in vivo* mouse models to the human situation.(Van Coppenolle et al., 2009) Sequential acquisition of T-cell specific surface markers recapitulates T-cell development from bone marrow or cord blood hematopoietic progenitors to the DP stage.(Awong et al., 2009) Here, we have sorted subsets at different stages during early T-cell differentiation and demonstrated conservation of early T-cell transcriptional profiles including pre- and post-T-cell commitment programs. We identified loss of CD44 expression in immature thymocytes as a marker for T-cell commitment that precedes acquisition of CD1a. Our results demonstrate that population II (CD44⁺CD1a⁻) has initiated TCR D β -J β and V β -J β rearrangements and is already completely T-lineage restricted. These results explain why the CD34⁺CD1a⁻ thymocyte population from previous studies—considered to be uncommitted thymocytes—is in fact heterogeneous with mixed CD44 expression.(Awong et al., 2009; Dik et al., 2005) The multi-lineage potential of CD34⁺CD44⁺ human thymocytes has been suggested,(Marquez et al., 1995) but this study places CD44 expression in the context of other T-cell development markers, and more importantly provides the first functional evidence that loss of CD44 marks human T-cell commitment *in vivo*.

Similar to human early T-cell development, the predominant transcriptional change in murine T-cell development—well before the DP stage—is from pre- to post- $\alpha\beta$ T-cell commitment.(Mingueneau et al., 2013; Zhang et al., 2012) In murine thymocytes, this well-characterized transition occurs in the second CD4/8 double-negative stage (DN2; CD44⁺CD25⁺), before the TCR β checkpoint in DN3 (CD44⁺CD25⁺).(Bhandoola et al., 2007; Godfrey et al., 1994) T-cell commitment is defined by the transition from DN2a to DN2b that is marked by the initial reduction of CD44 surface expression.(Mingueneau et al., 2013; Zhang et al., 2012) The fact that the initial reduction of CD44 in the transition from DN2a to DN2b thymocytes also marks T-cell commitment in mice suggests a functional role for CD44 in pre-committed thymocytes that is conserved in humans.

CD44 is an glycosylated adhesion and migration molecule with several isoforms due to alternatively spliced exons in the extracellular domain.(Zoller, 2015) The standard CD44 isoform (CD44s) is expressed on hematopoietic (stem) cells (as well as many other cell types) and is required for HSC niche formation and quiescence during early hematopoietic development. CD44 has gained clinical interest because CD44s and variant isoforms (CD44v) feature ‘stemness’ potential of many cancer types, including leukemia. Currently, multiple anti-CD44-based therapeutic approaches are investigated for their ability to target leukemia- or cancer-initiating cells (LIC/CIC) with this ‘stemness’ potential.(Zoller, 2015) In addition to its expression

on hematopoietic stem/progenitor and ETP cells, CD44 is highly re-expressed on mature lymphocytes (Fig. 5) and functions in the homing to secondary lymphoid organs. In the thymus, CD44 and its cleavage by metalloproteinases also direct the migration of mature thymocytes within the medulla as well as the exit from the medulla.(Vivinus-Nebot et al., 2004) The rare population of uncommitted CD44⁺ ETPs is present in subcapsular clusters in the cortex of the thymus, suggesting that CD44 also directs thymocyte precursors from the bone marrow to the thymus, entry into the thymus, and potentially intrathymic migration during early T-cell development.(de la Hera et al., 1989; Horst et al., 1990)

By adding CD44, we have expanded the immunophenotypes of the different phases of early T-cell development in the human thymus (Fig. 5). CD44^{dim}CD45^{dim} human thymocytes reflect pre-committed ETP cells that are mainly characterized by the expression of transcription factors from alternative blood lineages, multi-lineage potential, and the absence of *TCRB* rearrangements. Loss of CD44 marks T-cell commitment that coincides with the upregulation of T-cell specific transcription factors (e.g. *GATA3*, *BCL11B*, *TCF7*), loss of alternative cell fate potential, and the initiation of *TCRB* rearrangements. Next, the thymocytes acquire CD1a and complete their *TCRB* rearrangements. Therefore, combining CD44 and CD1a surface markers in CD45^{dim} thymocytes (in addition to other markers) is important in the characterization of pre- and post-committed human thymocytes and will help to better understand human T-cell development and T-cell commitment.

Materials and Methods

Human umbilical cord blood and thymus samples

Human umbilical cord blood (UCB) samples were obtained from consenting mothers after delivery at local hospitals. Mononuclear cells (MNCs) were isolated by Ficoll-Paque density centrifugation, washed and frozen in 10% dimethyl sulfoxide (DMSO) and 90% fetal bovine serum (FBS) for later use. Thymi were obtained as surgical tissue discards from infants 2-9 months of age undergoing cardiac surgery at Erasmus MC Rotterdam, after informed consent from the parents or legal guardians. The children did not have immunological abnormalities. Thymocytes were isolated by cutting the thymic lobes into small pieces and squeezing them through a metal mesh, and stored at -80°C until further analyses. Informed consents were in accordance with the Institutional Review Board of the Erasmus MC Rotterdam and in accordance with the Declaration of Helsinki.

Isolation of CD34⁺ cells from umbilical cord blood

Frozen MNCs were thawed, washed and labeled with MicroBeads conjugated to the monoclonal mouse anti-human CD34 antibody according to the manufacturer's procedure (Miltenyi Biotec). The magnetic separation was performed twice using two MACS Columns (Miltenyi Biotec), consistently reaching a CD34⁺ purity of approximately 95%.

Antibodies

Antibodies (with clone identification) used for flow cytometry: CD1a (HI149), CD3 (UCHT1), CD7 (M-T701), CD10 (HI10a), CD11b (ICRF44), CD13 (WM15), CD14 (MΦP9), CD20 (L27), CD22 (HIB22), CD33 (WM53), CD34 (8G12) (BD Biosciences), CD3 (BW264/56), CD4 (VIT4), CD5 (UCHT2), CD8 (BW135/80), CD19 (LT19), CD33 (AC104.3E3), CD34 (AC136), CD44 (DB105), CD45 (5B1), CD56 (REA196), CD94 (REA113) (Miltenyi Biotec).

OP9-DL1 co-cultures

OP9-DL1 co-cultures were performed according to the original protocol.(Holmes and Zuniga-Pflucker, 2009) Briefly, CD34⁺ human HSC isolated from UCB were plated into 100mm or 145mm culture dishes (Greiner Bio-One B.V.), which were seeded with a monolayer of OP9-DL1 cells beforehand. The α -minimal essential medium (α -MEM) (Life Technologies) used for culture was supplemented with 20% fetal bovine serum (FBS) (Integro B.V.) plus 10 U/mL Penicillin, 10 μ g/mL of streptomycin, and 0.025 μ g/mL fungizone (PSF) (Life Technologies). Recombinant human stem cell factor (SCF) (10 ng/mL) (R&D systems), FLT3L (5 ng/mL) (Miltenyi Biotec), and interleukin-7 (IL7) (2 ng/mL) (Miltenyi Biotec) were added at the initiation of coculture and every 2-3 days during transfer of HSC-derived cells onto new OP9-DL1 plated cells. Co-cultures were disaggregated by vigorous pipetting and passaged through a 40- μ m filter to reduce stromal cell line aggregates and eliminate contaminating OP9-DL1 cells.

Multi-lineage differentiation experiments were performed with the following modifications: recombinant human SCF (10 ng/ml), FLT3L (5 ng/ml), IL7 (5 ng/ml), and IL15 (10 ng/ml) were added every 3-4 days to half of the medium that was refreshed without disturbing the cells. A continuous 10-day OP9- co-culture was initiated with 1000 OP9 (-DL1, -GFP, or 1:1 mix) cells/well plated in 96-well plates one day before adding 4000/well of the sorted thymocytes.

OP9-GFP and OP9-DL1 cell lines were kindly given to us by Dr. Zúñiga-Pflücker; the identity of the cell lines was confirmed by DNA fingerprinting and cells were regularly tested for mycoplasma contamination.

TCRB rearrangements

In-frame D β -J β and V β -J β gene rearrangements were determined using the BIOMED-2 multiplex PCR and visualized using GeneScan.(van Dongen et al., 2003) Primers for the J β 1 cluster were hexachloro-6-carboxy-fluorescein (HEX)-labeled (blue traces), and primers for the J β 2 cluster were 6-carboxy-fluorescein (FAM)-labeled (green traces).

Cell sorting and RNA preparation

Throughout the differentiation of HSC towards the T-cell lineage, we sorted T-cell fractions at various time points (days 0, 7, 12, 18, and 27) based on the expression of cell surface markers including CD45, CD34, CD7, CD5, CD1a, CD4 and CD8. Cell sorting was performed using the FACSARIA II (BD Biosciences), flow-cytometry was performed using the MACS Quant (Miltenyi Biotec), and data analysis was performed using the FlowJo software (TreeStar Inc.). Sorted populations were collected, washed twice in PBS, lysed in buffer RLT (Qiagen) plus 1/100 β -mercaptoethanol and stored at -80°C. RNA was isolated using RNeasy Micro Kit according to the manufacturer's protocol (Qiagen).

Affymetrix gene expression arrays

The quality control, RNA labeling, hybridization and data extraction were performed at ServiceXS B.V. (Leiden, the Netherlands). RNA concentration was measured using the Nanodrop ND-1000 spectrophotometer (Nanodrop Technologies). The RNA quality and integrity was determined using Lab-on-Chip analysis on the Agilent 2100 Bioanalyzer (Agilent Technologies, Inc.) and/or on the Shimadzu MultiNA RNA analysis chips (Shimadzu Corporation). Biotinylated cRNA was prepared using the Affymetrix 3' IVT Express Kit (Affymetrix) according to the manufacturer's specifications with an input of 100ng total RNA. The quality of the cRNA was assessed using the Shimadzu MultiNA in order to confirm if the average fragment size was according to Affymetrix' specifications. Per sample, 7.5 μ g cRNA of the obtained biotinylated cRNA samples was fragmented and hybridized in a final concentration of 0.0375 μ g/ μ l on the Affymetrix HT HG U133+ PM (Affymetrix). After an automated process of washing and staining by the GeneTitan machine (Affymetrix) using the Affymetrix HWS Kit for GeneTitan (part nr. 901530), absolute values of expression were calculated from the scanned array using the Affymetrix Command Console v3.2 software.

The data discussed in this publication have been deposited in NCBI's Gene Expression Omnibus(Edgar et al., 2002) and are accessible through GEO Series accession number GSE79379 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE79379>).

Taqman array microfluidic cards

Reverse-transcriptase (RT) reactions (Promega) were performed to convert mRNA into cDNA. The real-time PCR (qPCR) reactions were performed by Taqman Custom arrays that were pre-designed according to the manufacturer instructions (Applied Biosystems). TaqMan gene expression assay targets were based on selected primers that were pre-loaded into each of the wells of a 384-well Taqman array card. TaqMan Gene Expression Assays consist of a pair of unlabeled PCR primers and a TaqMan probe with a FAM or VIC dye label on the 5' end and minor groove binder (MGB) and nonfluorescent quencher (NFQ) on the 3' end. For the reaction, cDNA samples were diluted and mixed in 1:1 ratio with Taqman Fast Universal PCR Master Mix (Applied Biosystems). Relative levels of gene expression are determined from the fluorescence data generated during PCR using the 7900HT Fast Real-Time PCR System Relative Quantitation software (Applied Biosystems).

Bioinformatics

Affymetrix gene expression CEL files were processed using Partek® Genomics Suite® 6.6. Robust Multi-array Average (RMA) (Irizarry et al., 2003) was applied prior to further analysis. In brief, the intensity levels were quantile normalized after background correction. Probeset-level expression values were summarized by median polish approach and eventually \log_2 transformed. Probesets were filtered for high variance (top 5%) and \log_2 expression values (>8), which resulted in 2179 probesets. Using the annotation file of the array, these probesets were then summarized into 1387 genes by taking the median of all probesets across a gene. Principle component analysis (PCA) and hierarchical clustering were done using Partek® Genomics Suite® 6.6. \log_2 expression values were standardized into z-scores prior to the clustering analysis. Pearson dissimilarity and Euclidean distance were used as distance metric. Gene set enrichment analysis (GSEA) (Subramanian et al., 2005) was performed using GSEA software downloaded from Broad Institute (<http://www.broadinstitute.org/gsea>). For the OP9-DL1 *in vitro* dataset, the GSEA input gene list consisted of 1387 genes that were grouped into the 16 gene signatures identified by the clustering analysis. The T-cell populations were denoted as early or late T-cell program according to their gene expression profile similarities during T-cell differentiation. Gene sets with FDR $q \leq 0.25$ were considered significantly correlated to the class distinction. (Subramanian et al., 2005) The murine *in vivo* microarray data was obtained through access to the Immunological Genome Project data, GEO accession code: GSE15907. (Mingueneau et al., 2013) For analysis of this *in vivo* dataset, the GSEA input gene list consisted of 547 genes enclosed in the 5 signatures that form the T-cell development gene

expression signature. The human *in vivo* microarray data was obtained at <http://www.ebi.ac.uk/miamexpress>, MIAME accession no. E-MEXP-337. Hierarchical clustering on the human thymocyte populations was performed using only 399 out of 547 genes from the T-cell development gene signature as the remaining 148 genes were not profiled in the human dataset.

Relative levels of gene expression data generated by the Taqman array microfluidic cards were processed using Partek® Genomics Suite® 6.6. The sorted *in vitro* populations (A-F) were subject to ANOVA analysis to remove day and pool effects prior to the hierarchical clustering. Standardized z-scores of the expression values in 47 selected T-cell development genes were clustered using Pearson dissimilarity measure and average linkage. These 47 genes are a subset of the 1387 genes with high variation and expression levels in our *in vitro* T-cell differentiation dataset. The same genes and clustering method were used for the human thymocyte populations (I-III) dataset. For this dataset, the expression values obtained by the microfluidic cards were subject to quantile normalization before further analysis.

Author Contributions

RDM, KC-B, and JPPM designed the study and wrote the manuscript; RDM and KC-B performed the experiments, YL analyzed the bioinformatics data, EV flow-sorted the different cell populations, KP-O and FJTS collected and provided the human thymocytes, and TW and AWL performed V-D-J rearrangement experiments. JPPM and RP supervised the study. All authors read and approved the paper, and declare no competing financial interests.

Acknowledgements

We thank Tom Taghon for critically reviewing the manuscript. This study was supported by the Childhood Cancer Foundation (*Stichting Kinderen Kankervrij*, KiKa) grants KiKa2008-29, KiKa2013-116, KiKa2014-141 (RDM and KC-B), KiKa2011-82 (YL), the Dutch Cancer Foundation KWF2010-4691 (EV), and ZonMW E-RARE (grant no.40-41900-98-020 (KPO and FJTS)).

References

- Awong, G., E. Herer, C.D. Surh, J.E. Dick, R.N. La Motte-Mohs, and J.C. Zuniga-Pflucker. 2009. Characterization in vitro and engraftment potential in vivo of human progenitor T cells generated from hematopoietic stem cells. *Blood* 114:972-982.
- Bell, J.J., and A. Bhandoola. 2008. The earliest thymic progenitors for T cells possess myeloid lineage potential. *Nature* 452:764-767.
- Bhandoola, A., H. von Boehmer, H.T. Petrie, and J.C. Zuniga-Pflucker. 2007. Commitment and developmental potential of extrathymic and intrathymic T cell precursors: plenty to choose from. *Immunity* 26:678-689.
- de la Hera, A., A. Acevedo, W. Marston, and F. Sanchez-Madrid. 1989. Function of CD44(Pgp-1) homing receptor in human T cell precursors. *International immunology* 1:598-604.
- de Pooter, R., and J.C. Zuniga-Pflucker. 2007. T-cell potential and development in vitro: the OP9-DL1 approach. *Curr Opin Immunol* 19:163-168.
- Dervovic, D., and J.C. Zuniga-Pflucker. 2010. Positive selection of T cells, an in vitro view. *Semin Immunol* 22:276-286.
- Dik, W.A., K. Pike-Overzet, F. Weerkamp, D. de Ridder, E.F. de Haas, M.R. Baert, P. van der Spek, E.E. Koster, M.J. Reinders, J.J. van Dongen, A.W. Langerak, and F.J. Staal. 2005. New insights on human T cell development by quantitative T cell receptor gene rearrangement studies and gene expression profiling. *J Exp Med* 201:1715-1723.
- Edgar, R., M. Domrachev, and A.E. Lash. 2002. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res* 30:207-210.
- Godfrey, D.I., J. Kennedy, P. Mombaerts, S. Tonegawa, and A. Zlotnik. 1994. Onset of TCR-beta gene rearrangement and role of TCR-beta expression during CD3-CD4-CD8-thymocyte differentiation. *J Immunol* 152:4783-4792.
- Goodings, C., E. Smith, E. Mathias, N. Elliott, S.M. Cleveland, R.M. Tripathi, J.H. Layer, X. Chen, Y. Guo, Y. Shyr, R. Hamid, Y. Du, and U.P. Dave. 2015. Hhex is Required at Multiple Stages of Adult Hematopoietic Stem and Progenitor Cell Differentiation. *Stem cells* 33:2628-2641.
- Halkias, J., H.J. Melichar, K.T. Taylor, and E.A. Robey. 2014. Tracking migration during human T cell development. *Cellular and molecular life sciences : CMLS* 71:3101-3117.
- Holmes, R., and J.C. Zuniga-Pflucker. 2009. The OP9-DL1 system: generation of T-lymphocytes from embryonic or hematopoietic stem cells in vitro. *Cold Spring Harb Protoc* 2009:pdb prot5156.
- Horst, E., C.J. Meijer, T. Radaskiewicz, J.J. van Dongen, R. Pieters, C.G. Figdor, A. Hooftman, and S.T. Pals. 1990. Expression of a human homing receptor (CD44) in lymphoid malignancies and related stages of lymphoid development. *Leukemia* 4:383-389.
- Irizarry, R.A., B. Hobbs, F. Collin, Y.D. Beazer-Barclay, K.J. Antonellis, U. Scherf, and T.P. Speed. 2003. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* 4:249-264.
- La Motte-Mohs, R.N., E. Herer, and J.C. Zuniga-Pflucker. 2005. Induction of T-cell development from human cord blood hematopoietic stem cells by Delta-like 1 in vitro. *Blood* 105:1431-1439.
- Marquez, C., C. Trigueros, E. Fernandez, and M.L. Toribio. 1995. The development of T and non-T cell lineages from CD34+ human thymic precursors can be traced by the differential expression of CD44. *J Exp Med* 181:475-483.
- Mingueneau, M., T. Kreslavsky, D. Gray, T. Heng, R. Cruse, J. Ericson, S. Bendall, M.H. Spitzer, G.P. Nolan, K. Kobayashi, H. von Boehmer, D. Mathis, C. Benoist, C. Immunological Genome, A.J. Best, J. Knell, A. Goldrath, V. Joic, D. Koller, T. Shay, A. Regev, N.

- Cohen, P., Brennan, M., Brenner, F., Kim, T., Nageswara Rao, A., Wagers, T., Heng, J., Ericson, K., Rothamel, A., Ortiz-Lopez, D., Mathis, C., Benoist, N.A., Bezman, J.C., Sun, G., Min-Oo, C.C., Kim, L.L., Lanier, J., Miller, B., Brown, M., Merad, E.L., Gautier, C., Jakubzick, G.J., Randolph, P., Monach, D.A., Blair, M.L., Dustin, S.A., Shinton, R.R., Hardy, D., Laidlaw, J., Collins, R., Gazit, D.J., Rossi, N., Malhotra, K., Sylvia, J., Kang, T., Kreslavsky, A., Fletcher, K., Elpek, A., Bellemare-Pelletier, D., Malhotra, and S. Turley. 2013. The transcriptional landscape of alphabeta T cell differentiation. *Nat Immunol* 14:619-632.
- Pike-Overzet, K., D. de Ridder, F. Weerkamp, M.R. Baert, M.M. Verstegen, M.H. Brugman, S.J. Howe, M.J. Reinders, A.J. Thrasher, G. Wagemaker, J.J. van Dongen, and F.J. Staal. 2007. Ectopic retroviral expression of LMO2, but not IL2Rgamma, blocks human T-cell development from CD34+ cells: implications for leukemogenesis in gene therapy. *Leukemia* 21:754-763.
- Plum, J., M. De Smedt, M.P. Defresne, G. Leclercq, and B. Vandekerckhove. 1994. Human CD34+ fetal liver stem cells differentiate to T cells in a mouse thymic microenvironment. *Blood* 84:1587-1593.
- Radtke, F., A. Wilson, and H.R. MacDonald. 2004. Notch signaling in T- and B-cell development. *Curr Opin Immunol* 16:174-179.
- Res, P., E. Martinez-Caceres, A. Cristina Jaleco, F. Staal, E. Noteboom, K. Weijer, and H. Spits. 1996. CD34+CD38dim cells in the human thymus can differentiate into T, natural killer, and dendritic cells but are distinct from pluripotent stem cells. *Blood* 87:5196-5206.
- Rothenberg, E.V., J.E. Moore, and M.A. Yui. 2008. Launching the T-cell-lineage developmental programme. *Nat Rev Immunol* 8:9-21.
- Schmitt, T.M., and J.C. Zuniga-Pflucker. 2002. Induction of T cell development from hematopoietic progenitor cells by delta-like-1 in vitro. *Immunity* 17:749-756.
- Spits, H. 2002. Development of alphabeta T cells in the human thymus. *Nat Rev Immunol* 2:760-772.
- Subramanian, A., P. Tamayo, V.K. Mootha, S. Mukherjee, B.L. Ebert, M.A. Gillette, A. Paulovich, S.L. Pomeroy, T.R. Golub, E.S. Lander, and J.P. Mesirov. 2005. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 102:15545-15550.
- Van Coppennolle, S., G. Verstichel, F. Timmermans, I. Velghe, D. Vermijlen, M. De Smedt, G. Leclercq, J. Plum, T. Taghon, B. Vandekerckhove, and T. Kerre. 2009. Functionally mature CD4 and CD8 TCRalphabeta cells are generated in OP9-DL1 cultures from human CD34+ hematopoietic cells. *J Immunol* 183:4859-4870.
- van Dongen, J.J., A.W. Langerak, M. Bruggemann, P.A. Evans, M. Hummel, F.L. Lavender, E. Delabesse, F. Davi, E. Schuurin, R. Garcia-Sanz, J.H. van Krieken, J. Droese, D. Gonzalez, C. Bastard, H.E. White, M. Spaargaren, M. Gonzalez, A. Parreira, J.L. Smith, G.J. Morgan, M. Kneba, and E.A. Macintyre. 2003. Design and standardization of PCR primers and protocols for detection of clonal immunoglobulin and T-cell receptor gene recombinations in suspect lymphoproliferations: report of the BIOMED-2 Concerted Action BMH4-CT98-3936. *Leukemia* 17:2257-2317.
- Vivinus-Nebot, M., P. Rousselle, J.P. Breittmayer, C. Cenciarini, S. Berrih-Aknin, S. Spong, P. Nokelainen, F. Cottrez, M.P. Marinkovich, and A. Bernard. 2004. Mature human thymocytes migrate on laminin-5 with activation of metalloproteinase-14 and cleavage of CD44. *J Immunol* 172:1397-1406.
- Weerkamp, F., M.R. Baert, M.H. Brugman, W.A. Dik, E.F. de Haas, T.P. Visser, C.J. de Groot, G. Wagemaker, J.J. van Dongen, and F.J. Staal. 2006. Human thymus contains multipotent progenitors with T/B lymphoid, myeloid, and erythroid lineage potential. *Blood* 107:3131-3137.

- Weerkamp, F., E.F. de Haas, B.A. Naber, W.M. Comans-Bitter, A.J. Bogers, J.J. van Dongen, and F.J. Staal. 2005. Age-related changes in the cellular composition of the thymus in children. *The Journal of allergy and clinical immunology* 115:834-840.
- Yoshida, H., S. Hayashi, T. Kunisada, M. Ogawa, S. Nishikawa, H. Okamura, T. Sudo, L.D. Shultz, and S. Nishikawa. 1990. The murine mutation osteopetrosis is in the coding region of the macrophage colony stimulating factor gene. *Nature* 345:442-444.
- Yui, M.A., and E.V. Rothenberg. 2014. Developmental gene networks: a triathlon on the course to T cell identity. *Nat Rev Immunol* 14:529-545.
- Zhang, J.A., A. Mortazavi, B.A. Williams, B.J. Wold, and E.V. Rothenberg. 2012. Dynamic transformations of genome-wide epigenetic marking and transcriptional control establish T cell identity. *Cell* 149:467-482.
- Zohren, F., G.P. Souroullas, M. Luo, U. Gerdemann, M.R. Imperato, N.K. Wilson, B. Gottgens, G.L. Lukov, and M.A. Goodell. 2012. The transcription factor Lyl-1 regulates lymphoid specification and the maintenance of early T lineage progenitors. *Nat Immunol* 13:761-769.
- Zoller, M. 2015. CD44, Hyaluronan, the Hematopoietic Stem Cell, and Leukemia-Initiating Cells. *Front Immunol* 6:235.

Figure Legends

Figure 1. Consecutive stages of human early *in vitro* T-cell differentiation represent two major gene signatures. (A) Schematic representation of the OP9-DL1 co-culture with the sorting strategy of consecutive T-cell differentiation stages. The table displays the sorted populations (and the replicates) based on surface markers (CD45, CD34, CD7, CD5, CD1a, CD4, CD8) and number of days in co-culture. All populations were also gated for CD45⁺GFP⁻ to exclude OP9-DL1 cells that are GFP positive. (B) Principle component analysis of 29 samples based on the 2179 probesets with high variance (top 5%) and log₂ expression values (>8). These probesets were summarized into 1387 genes by taking the median of all probesets across a gene. (C) Hierarchical clustering analysis using 1387 genes. Pearson dissimilarity measure and average linkage were applied to cluster the samples, and Euclidean distance and Ward's method were applied to define 16 distinct gene signatures with similar expression patterns. (D) Average expression pattern of genes in the 16 gene signatures. z-scores are on the y-axis and 29 populations on the x-axis, ordered according to Fig. 1C. The grey area represents 1 SD. (E) Gene set enrichment analysis (GSEA) on the *in vitro* T-cell differentiation dataset to select gene signatures that are significantly enriched in the early or late T-cell program (marked in green and bold). The Normalized Enrichment Score (NES) and False Discovery Rate (FDR) *q*-value for each gene signature are given on top of each sub-figure. Populations belonging to the early/late T-cell program were determined according to Fig. 1C. Gene sets with FDR $q \leq 0.25$ were considered significantly correlated to the class distinction. (Subramanian et al., 2005)

Figure 2. The *in vitro* gene expression signatures recapitulate *in vivo* signatures of pre- and post-T-cell committed thymocytes. (A) Hierarchical clustering analysis (Pearson average method) of the different sorted populations based on the 547 genes that discriminate early (characterized by gene signatures #7 and #8) from late (characterized by gene signatures #15, #16 and #2) T-cell programs. The clustering obtained is similar to Fig.1C where 1387 genes were used. (B) Gene set enrichment analysis of our *in vitro* gene signature on the gene expression dataset of consecutive T-cell development populations from murine thymi. The first and second sub-figures from the left are the analysis results using gene signatures #7 and #8. The Normalized Enrichment Score (NES), Nominal (NOM) *p*-value, and False Discovery Rate (FDR) *q*-value indicate that these two gene signatures correlate with the pre-commitment stages. Also shown in the first sub-figure is the list of top 10 input genes with the highest signal-to-noise ratios in this dataset. The third and fourth sub-figures are the same analysis for gene signatures #2, #15, and

#16, which correlate with the post-commitment stages. (C) Hierarchical clustering (Euclidean distance and Ward's method) on human thymocyte populations using 399 out of the 547 *in vitro* pre-/post-commitment signature genes (the remaining 148 genes were not profiled because they were not on the HG-U133A array used in the dataset by Dik *et al.*, 2005). The dendrogram on the left shows the clustering of different sorted human thymocytes, separating the pre- and post-committed populations.

Figure 3. Loss of CD44 surface expression marks T-cell commitment in early human thymocytes. (A) Sorting strategy for isolating T-cell differentiation stages A-F from duplicate (n=2) 12- and 18-day co-cultures started with pooled CD34⁺ HSCs, based on surface markers CD45, CD34, CD7, CD5, CD1a, and CD44. (B) Hierarchical clustering (Pearson average) of sorted populations A-F using 47 selected T-cell development genes. Expression values were obtained by Taqman array microfluidic cards, followed by ANOVA to remove day and pool effects. (C) Sorting strategy for isolating T-cell developmental stages I-III from four independent donors (n=4), based on CD45, CD7, CD5, CD1a, and CD44 after pre-depletion of CD3⁻, CD4⁻, CD8⁻, and CD19-expressing thymocytes. (D) Hierarchical clustering (Pearson average) of sorted populations I-III using 47 selected T-cell development genes. Expression values were obtained by Taqman array microfluidic cards, followed by quantile normalization.

Figure 4. Loss of CD44 functionally marks human T-cell commitment. (A) GeneScan visualization of in-frame D β -J β (left column) and V β -J β (middle and right columns) gene rearrangements, determined using the BIOMED-2 multiplex PCR.(van Dongen *et al.*, 2003) One representative example of three independent donors is shown. Healthy peripheral blood mononuclear cells were used as a positive, polyclonal control (top row); populations I-III are indicated (rows 2-4). Primers for the J β 1 cluster were hexachloro-6-carboxy-fluorescein (HEX)-labeled (blue traces), and primers for the J β 2 cluster were 6-carboxy-fluorescein (FAM)-labeled (green traces). Red traces: internal size markers. (B) Flow cytometry analysis of sorted human thymocyte populations I-III after 10 days of co-culture on OP9-DL1:GFP mixed cells. Differentiation into various hematopoietic lineages is defined as follows: T-lineage (blue): CD5⁺CD1a⁺, NK-lineage (red): CD56⁺CD94⁺, B-lineage (yellow): CD19⁺, myeloid lineage (green): CD33⁺. During differentiation, all populations remain CD45⁺ and CD7⁺. (C) Absolute cell numbers \pm SD of each differentiated and proliferated hematopoietic lineage from a representative of three experiments, after 10 days of co-culture on a layer of OP9-DL1, a 1:1 mix, or OP9-GFP

(ctrl.) cells. The number of cells from populations I-III used to initiate the co-cultures on day 0 was 4,000 cells/well, indicated by the dashed horizontal line.

Figure 5. Schematic comparison of human and mouse early T-cell development in the thymus. Bone marrow-derived lymphoid progenitors enter the thymus (depicted as a blue oval) and exhibit dim CD44 expression. T-cell commitment is marked by the downregulation of CD44. As cells mature, they regain CD44 expression.

Figure 1. Consecutive stages of human early *in vitro* T-cell differentiation represent two major gene signatures.

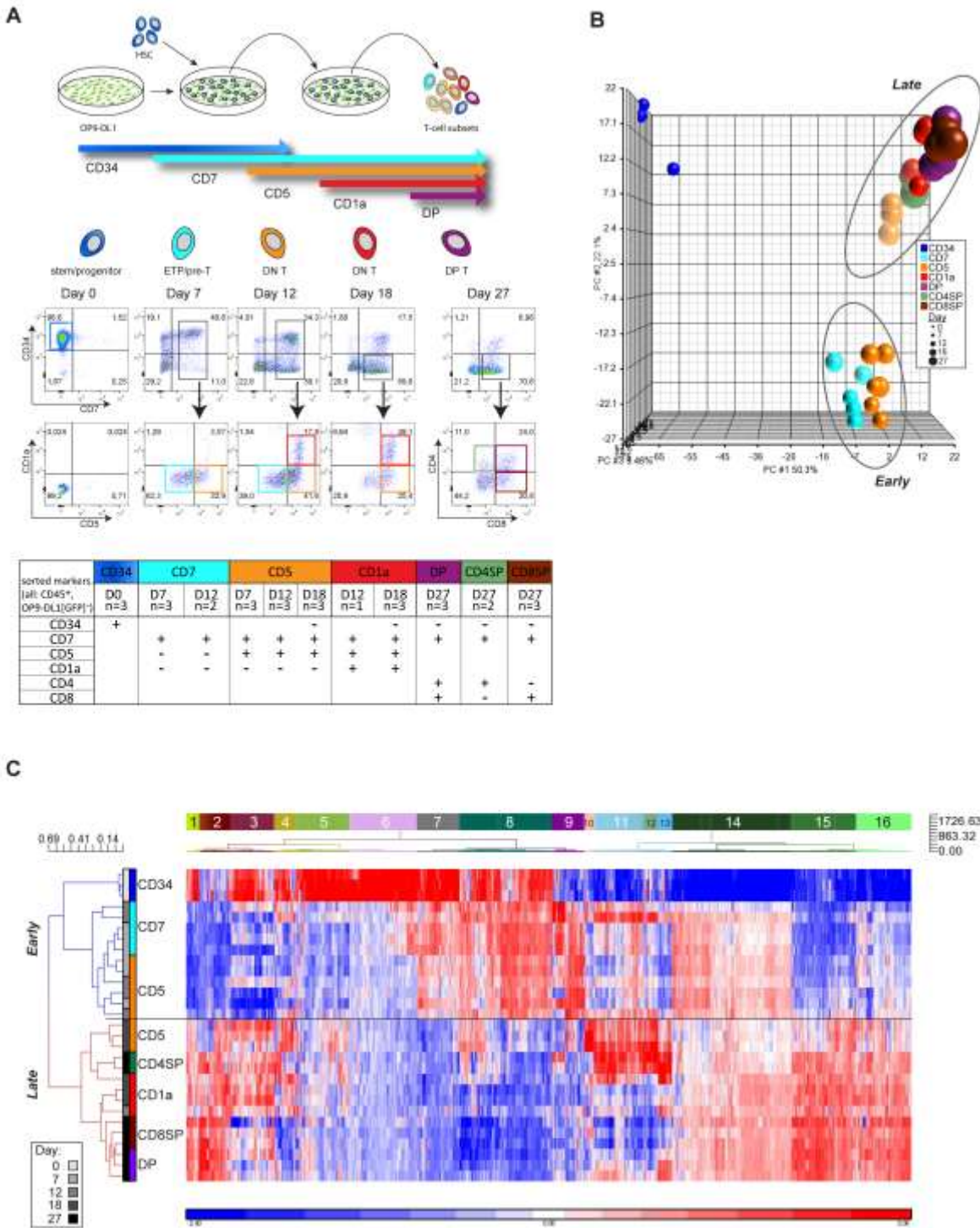
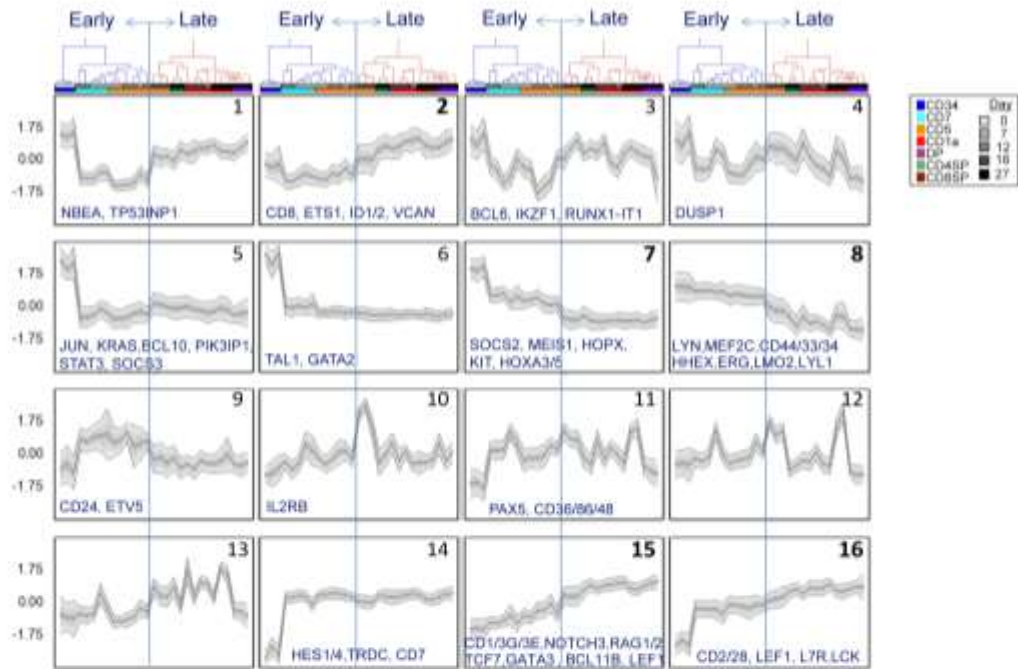


Figure 1. (continued)

D



E

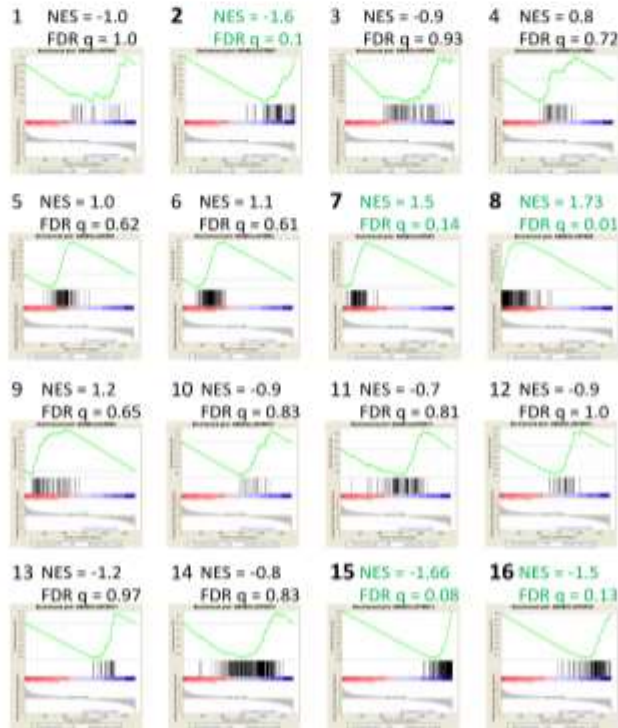


Figure 2. The *in vitro* gene expression signatures recapitulate *in vivo* signatures of pre- and post- T-cell committed thymocytes.

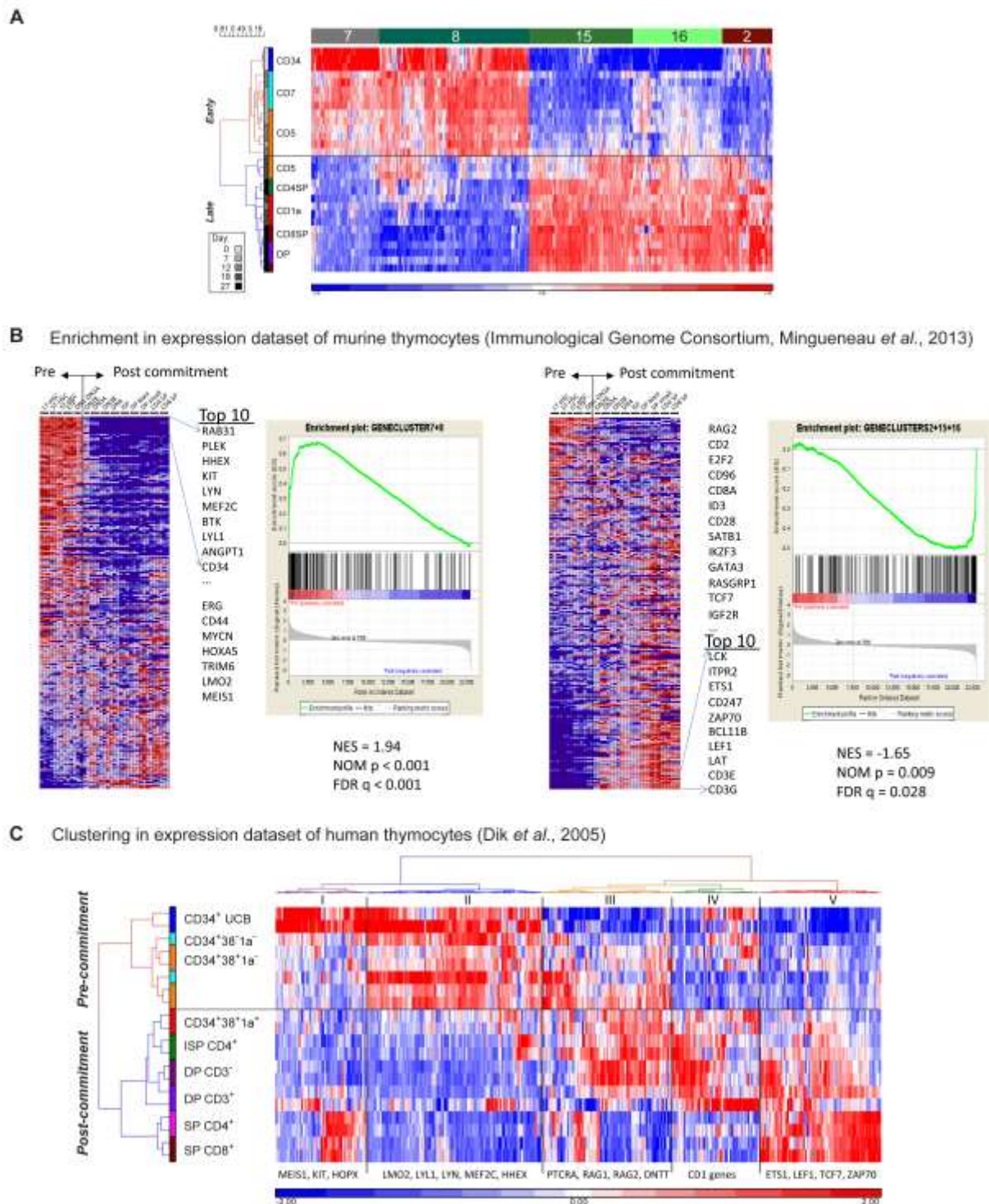
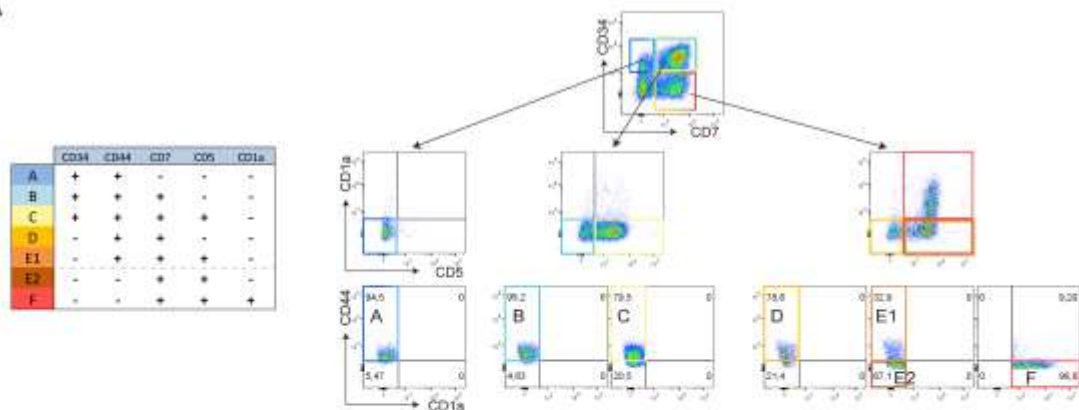
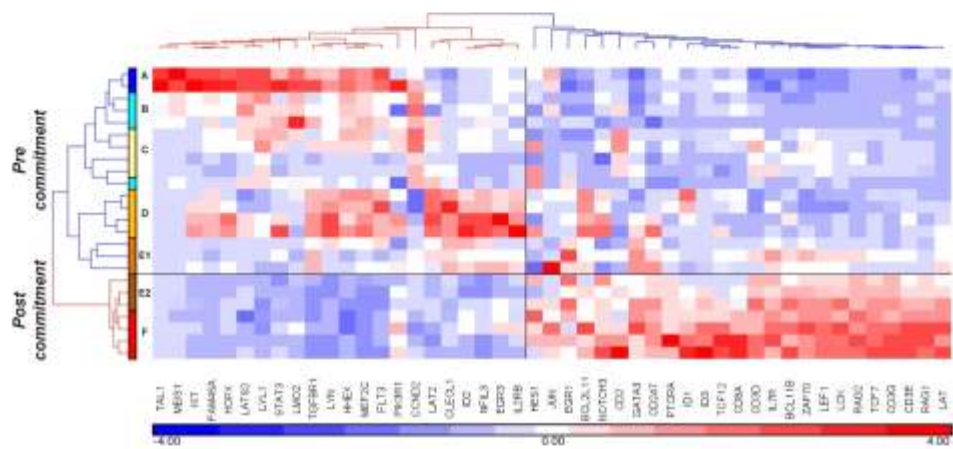


Figure 3. Loss of CD44 surface expression marks T-cell commitment in early human thymocytes.

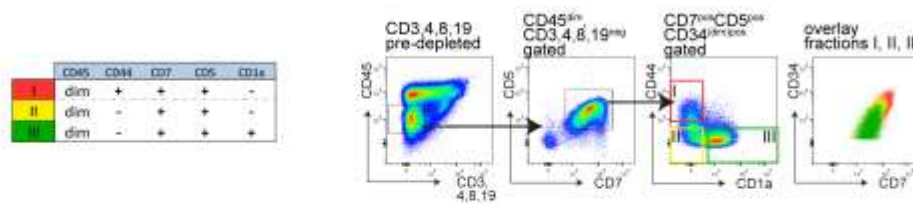
A



B



C



D

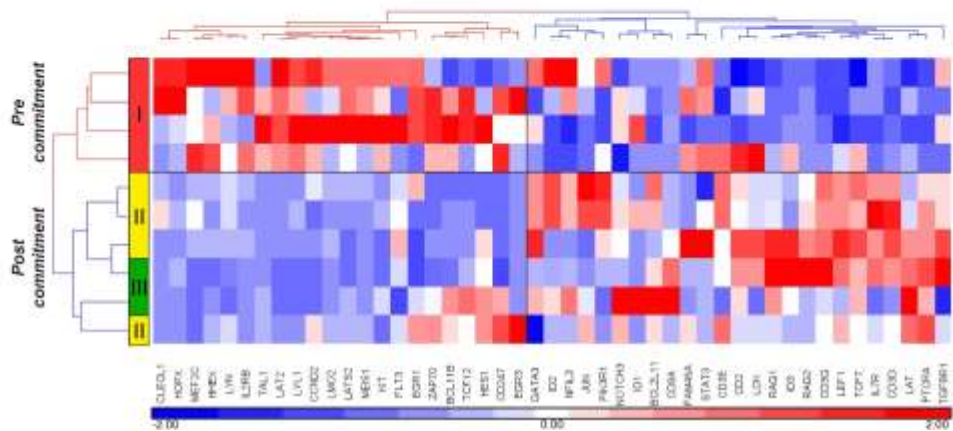


Figure 4. Loss of CD44 functionally marks human T-cell commitment.

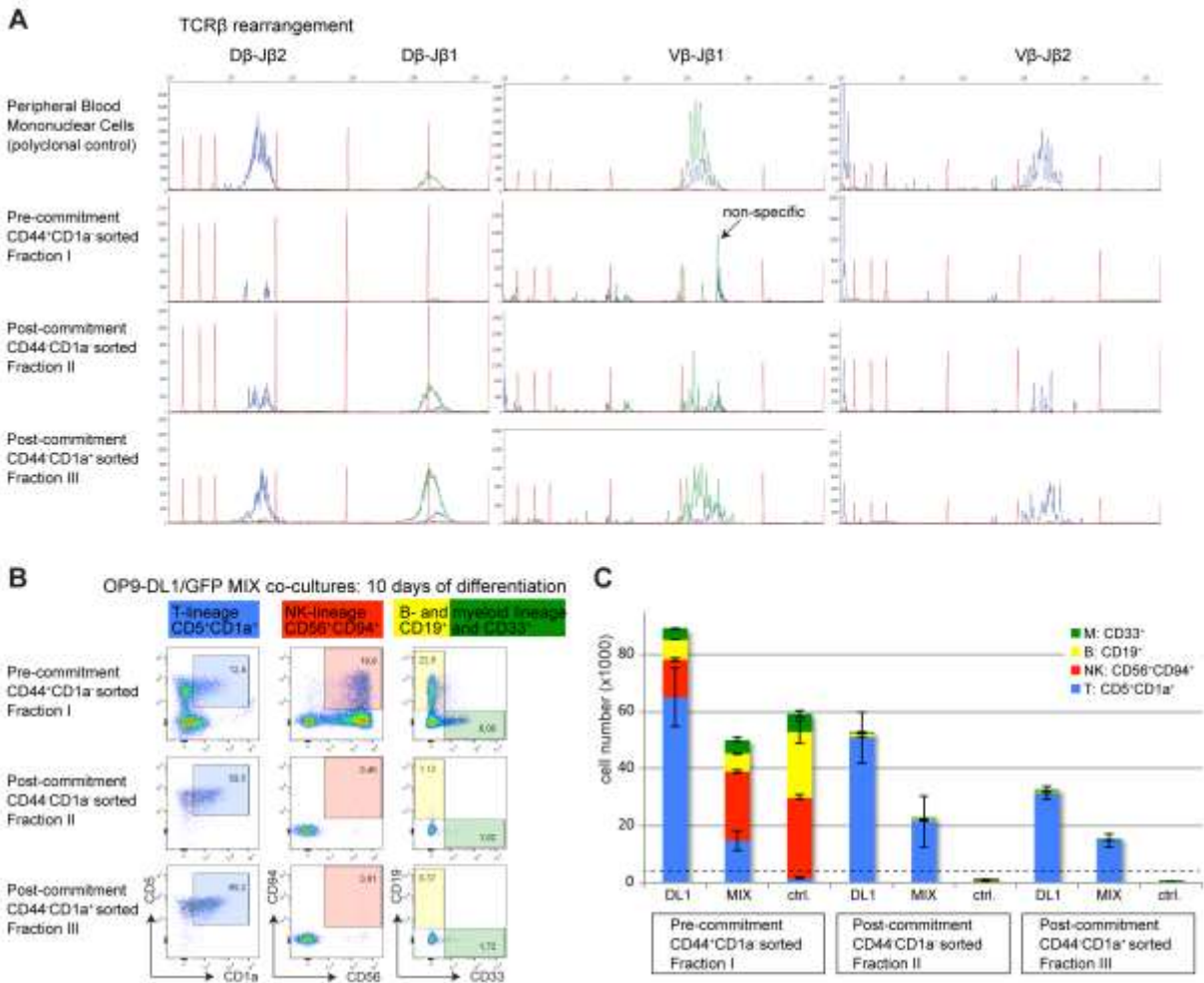
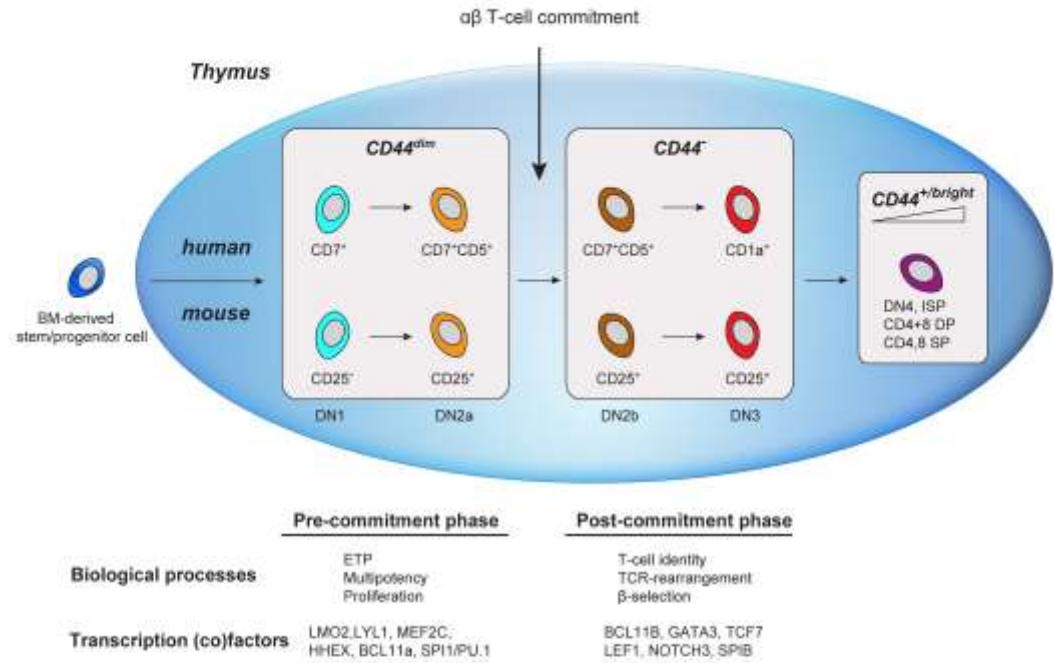


Figure 5. Schematic comparison of human and mouse early T-cell development in the thymus



Chapter 6

T-Cell Acute Lymphoblastic Leukemia Oncogenes hijack T-Cell Developmental Programs for Leukemogenesis: the arrest of Early T-cell programs by MEF2C, LYL1 or LMO2

Rui D. Mendes¹, Kirsten Canté-Barrett¹, Yunlei Li¹, Wilco Smits¹, Rob Pieters^{1,2} and Jules P.P. Meijerink¹

¹Department of Pediatric Oncology/Hematology, Erasmus Medical Center-Sophia Children's Hospital, Rotterdam, The Netherlands; ²Princess Máxima Center of Pediatric Oncology, Utrecht, The Netherlands.

Manuscript in preparation

Abstract

Pediatric T-cell acute lymphoblastic leukemia (T-ALL) is a cancer of developing T-cells in the thymus. The T-ALL subtypes are characterized by specific chromosomal rearrangements driving the ectopic expression of oncogenes that result in arrest of leukemic cells at specific T-cell developmental stages. Here, we show that T-ALL subtypes are strongly associated with the expression of particular T-cell developmental programs. Immature/early T-cell precursor (ETP)-ALL and TLX subtypes of T-ALL express a pre-commitment program whereas the proliferative and TALLMO subtypes express a post-commitment program. This suggests that T-ALL oncogenes require the expression of particular T-cell transcriptional programs to exert an oncogenic function, and therefore may hijack T-cell developmental programs by facilitating a developmental arrest. That is further supported by functional *in vitro* studies where we investigate the impact of typical ETP-ALL (proto-)oncogenes, including MEF2C, LYL1 or LMO2 on early T-cell development. Umbilical cord blood stem cells transduced with lentiviral expression constructs that constitutively drive MEF2C, LYL1 or LMO2 were cultured on OP9-DL1 cells for differentiation towards the T-cell lineage. Expression of MEF2C or its putative co-factors (LYL1 or LMO2) blocks T-cell development at ETP-stages, in line with their roles in immature T-ALL. Contrarily, expression of the oncogenes in T-cells beyond the ETP-stages does not lead to T-cell development arrest, indicating that their leukemogenic potential is dependent on ETP-stage and expression of pre-commitment transcriptional networks. Thus, we conclude that the specific interaction of an aberrant oncogenic transcription factor with specific transcriptional network in a developing T-cell is an important requirement for oncogenesis.

Introduction

T-cell development in the thymus is a complex process that depends on sequential transcriptional and epigenetic events that promote T-cell lineage commitment while simultaneously suppressing alternative cell fates.[1, 2] During this process, inappropriate activation of specific oncogenic transcription factors arrest thymocytes at specific T-cell developmental stages.[3] The ectopic expression of these oncogenes results from specific chromosomal rearrangements that place T-ALL oncogenes under the control of regulatory elements, such as enhancers of T-cell receptor or *BCL11B* genes, promoters from other genes, or occasionally result in oncogenic fusion proteins. Based on unsupervised gene cluster analysis we previously identified four T-ALL subtypes with specific gene expression signatures that are denoted as immature, TLX, proliferative or TALLMO clusters.[4]

Gene expression signature and clustering analysis of the immature T-ALL cluster highly overlap with the previously described high-risk early T-cell precursor (ETP)-ALL,[5] and represent the same disease entity.[6] The immature/ETP-ALL subtype is characterized by the overexpression of stem cell associated transcription factors that are normally extinguished during early T-cell development. These include the high expression of the MEF2C transcription factor,[4] and the expression of LMO2 and LYL1 [7, 8] that form a macromolecular transcriptional complex.[9] We have previously demonstrated that the immature/ETP-ALL subtype exhibits rearrangements of MEF2C and translocations involving transcription factors that target MEF2C (e.g. NKX2.5, SPI1) or MEF2C-associated cofactors (e.g. NCOA2),[4] and we showed that MEF2C regulates transcription factors as LYL1, LMO2 and HHEX.[4] This implies that those genes may be involved in MEF2C-deregulated T-ALL. The ectopic expression of LMO2 and HHEX are associated with a hematopoietic stem cell (HSC)-specific transcriptional program that is responsible for an arrest and self-renewal of preleukemic thymocytes, which impedes the new entry of seeding T-cell precursors in the thymus.[10, 11] Recently, this disruption in the turnover of cells in thymus was described to promote T-ALL.[12] However, these transcription factors are expressed in normal T-cell precursors that continuously seed the thymus, demonstrating that their expression is not exclusive for T-ALL.[13, 14] Normally, these transcription factors can exert roles in proliferation, survival and self-renewal of hematopoietic cells, and their expression is retained at the early phases of T-cell development until T-cell lineage commitment. [13] (R.D. Mendes and K. Cante-Barrett et al. unpublished data)

In our previous study, we used the OP9-DL1 *in vitro* culture system[15] to investigate the transcriptional landscape of differentiating human T-cells from umbilical cord-blood isolated HSCs. (Chapter 5, R.D. Mendes and K. Cante-Barrett et al. unpublished data) The changes in gene expression of five subsequent development stages revealed important early and late T-cell programs that are conserved in mouse and human T-cell development *in vivo*, and represent pre-commitment and post-commitment programs, respectively. Importantly, we found that loss of CD44 functionally defines human T-cell lineage commitment in ETP cells, before the acquisition of CD1a surface expression.

In the present study, we have investigated the effect of ectopic expression of MEF2C, LYL1 or LMO2 transcription factors in T-cell development to better understand how aberrant expression of these (proto-)oncogenes can contribute to differentiation arrest during early T-cell development. Thus, we ectopically expressed MEF2C, LYL1 or LMO2 in HSCs and then cultured these cells in the OP9-DL1 *in vitro* system. These proteins each result in the

developmental arrest within the ETP-stage and may therefore play an important role in the pathogenesis of ETP-ALL.

Materials and methods

Patients and gene expression profiles

A total of 117 primary pediatric T-ALL patient samples enrolled in the Dutch Childhood Oncology Group (DCOG) protocols or the German Cooperative Study Group for Childhood Acute Lymphoblastic Leukemia study (COALL) were included in this study. Informed consents were in accordance with the institutional review boards of the Erasmus MC, Rotterdam, and the Declaration of Helsinki. The microarray datasets are available at <http://www.ncbi.nlm.nih.gov/geo/> under accession number GSE26713.

Lentiviral expression vectors

The third-generation self-inactivating lentiviral LeGO-iC2 vector[16] was adapted into a Gateway compatible destination vector (LeGO-DEST) by replacing the *Apal/Pcil* 2kb insert with the Gateway *ccdB* cassette (Life Technologies, Bleiswijk, the Netherlands). The lentiviral expression constructs were assembled by cloning the Spleen Focus Forming Virus (SFFV) promoter, the *MEF2C*, *LYL1* or *LMO2* oncogenes flanked by a *Thosea asigna* virus 2A (T2A) element [17] and the blue fluorescent protein reporter (mTagBFP). These components in combination with other lentiviral elements are flanked by long terminal repeats, encoding for the proviral RNA. The 3 lentiviral expression constructs encoding the oncogenes *LMO2*, *LYL1* or *MEF2C* exhibit proviral RNA lengths of 4830 bp, 5190 bp and 5750 bp, respectively. The control vector is solely composed of the SFFV promoter followed by the BFP reporter and contains 4215 bp.

Virus Production

Virus productions were performed according to the protocol described in Chapter 4. Briefly, the different virus batches were produced in serum-free Opti-MEM I with Glutamax (Life Technologies) and were concentrated using VIVASPIN 20 centrifugation concentration columns (Sartorius, Goettingen, Germany). The number of viral particles was measured by quantification of proviral RNA copies using RT-qPCR.

Western-blot

Lentiviral vectors were functionally tested by western-blot (WB), using transduced Jurkat and Loucy T-ALL cell lines. Transduced cells were propagated and monitored for the expression of BFP protein. Cells were lysed in RIPA buffer and 40 µg of total protein was loaded in lanes of Any KD Mini-PROTEAN TGX Precast Protein Gels (Biorad, Veenendaal, the Netherlands). The primary antibodies (abs) with reactivity against the human proteins were used according to the manufacturer's protocol; and include the anti-2A Peptide (ABS31) polyclonal Rabbit Ab (Millipore BV, Amsterdam-Zuidoost, the Netherlands), the anti-LMO2 polyclonal Goat Ab (R&D systems), the anti-LYL1 polyclonal Rabbit Ab (Thermo Fisher Scientific, Waltham, MA, USA), and the anti-MEF2C (D80C1) XP monoclonal Rabbit Ab (Cell Signaling, Danvers, MA, USA). Incubations with the primary antibodies were followed by incubations with secondary IRDye 680RD goat anti-rabbit or 680LT donkey anti-goat antibodies (LI-COR Biosciences, Lincoln, NE, USA), and detected with the Odyssey imaging system (LI-COR).

Human umbilical cord blood

Human umbilical cord blood (UCB) samples were obtained by syringe extraction and collected in a blood-pack unit containing citrate phosphate dextrose anticoagulant from mothers after delivery in local hospitals. Mononuclear cells (MNCs) were isolated by Ficoll-Paque density centrifugation, washed and frozen in 10% dimethyl sulfoxide (DMSO) and 90% fetal bovine serum (FBS) for later use. Informed consents were in accordance with the Institutional Review Boards of the Erasmus MC (Rotterdam, the Netherlands) and the Declaration of Helsinki.

Isolation of CD34⁺ cells from umbilical cord blood

Frozen MNCs from UCB samples were thawed, washed and labeled with MicroBeads conjugated to the monoclonal mouse anti-human CD34 antibody according to the manufacturer's protocol (Miltenyi Biotec, Bergisch Gladbach, Germany). Magnetic separation was performed twice using two MACS Columns (Miltenyi Biotec), consistently reaching a CD34⁺ purity of approximately 95%. CD34⁺ HSCs were resuspended to concentration of 0.5×10^6 cell/ml in pre-stimulation medium consisting of serum-free X-VIVO 10 medium (Lonza, Verviers, Belgium) supplemented with 50 ng/ml rhSCF (R&D systems, Abingdon, UK), 20 ng/ml rhTPO (R&D systems) and 50 ng/ml rhFlt3L (Miltenyi Biotec). Cells were incubated for 16-20 hours at 37°C and 5% CO₂.

Transduction of Human CD34⁺ cells

Cells were transduced immediately after 16-20 hours of pre-stimulation. Transduction of cells was performed according to the protocol described in Chapter 4. Briefly, 96-well nontreated surface plates (BD Biosciences, Breda, The Netherlands) were coated with retronectin (Takara, Saint-Germain-en-Laye, France), and 1.0×10^5 cells were mixed with 10^9 - 10^{10} virus particles in a final volume of 200 μ l/well, followed by spinoculation. Transduced cells were incubated for 24 hours at 37°C, 5%CO₂.

OP9-DL1 co-cultures

OP9-DL1 co-cultures were performed according to the original protocol.[18] Briefly, transduced HSC were plated into 6-well cell culture plates (Greiner Bio-One B.V., Alphen aan den Rijn, Netherlands), which were seeded the day before with 7×10^4 OP9-DL1 cells to reach 70-80% confluence. The α -minimal essential medium (α -MEM) (Life Technologies) used for culture was supplemented with 20% fetal bovine serum (FBS) (Integro b.v., Dieren, the Netherlands) plus 10 U/mL Penicillin, 10 μ g/mL of streptomycin, and 0.025 μ g/mL fungizone (PSF) (Life technologies). Recombinant human stem cell factor (SCF) (10 ng/mL) (R&D systems, Abingdon, UK), FLT3L (5 ng/mL) (Miltenyi Biotec), and interleukin-7 (IL7) (2 ng/mL) (Miltenyi Biotec) were added at the initiation of coculture and every 2-3 days during transfer of HSC-derived cells onto fresh plates that had been seeded with OP9-DL1 cells beforehand. Cocultures were disaggregated by vigorous pipetting and passaged through a 40- μ m filter to reduce stromal cell line aggregates.

Flow-cytometry

T-cell differentiation was followed-up by flow-cytometry using the MACS Quant (Miltenyi Biotec). The different stages of differentiation were based on the expression of cell surface markers including the following antibodies (with clone identification): CD1a (HI149), CD7 (M-T701), CD13 (WM15), CD33 (WM53), CD34 (8G12) (BD Biosciences), CD4 (VIT4), CD5 (UCHT2), CD8 (BW135/80), CD19 (LT19), CD33 (AC104.3E3), CD34 (AC136), CD44 (DB105), CD45 (5B1) (Miltenyi Biotec). The data analysis was performed using the FlowJo software (TreeStar Inc., Ashland, OR).

Results

T-ALL subtypes are associated with pre-/post- commitment developmental programs

In our previous study, we performed gene expression arrays of consecutive T-cell developmental stages from early thymus progenitors (ETP) to CD4⁺CD8⁺ double-positive (DP) T-cells (Supplemental Figure S1a). Using bioinformatics tools we obtained a gene signature of

547 genes, which is composed of differentially expressed genes across T-cell differentiation (Supplemental Figure S1b). Importantly, the changes in gene expression of the subsequent stages matched those of human and murine *in vivo* thymocyte subsets, and consequently two distinct T-cell differentiation programs named “early” and “late” were correlated with *in vivo* pre- and post- $\alpha\beta$ T-cell commitment profiles, respectively (Supplemental Figure S1b).

Following that work, here we examine whether this gene signature can distinguish T-ALL patient subtypes, since different T-ALL subtypes exhibit a specific developmental arrest that is driven by the ectopic expression of specific oncogenic rearrangements. Thus, the gene signature of 547 genes was applied to the gene expression data obtained from 117 pediatric T-ALL samples. Interestingly, the PCA mapping showed clustering of the 117 samples into four main clusters that closely resembled the four clusters as previously obtained by unsupervised cluster analysis,[4] i.e. the immature/ETP-ALL, TLX, proliferative and TALLMO clusters (Figure 1a). In addition, analysing T-ALL samples together with the T-cell sorted fractions obtained using the OP9-DL1 *in vitro* system, revealed that UCB-CD34⁺ stem cells co-cluster with immature/ETP-ALL samples. Furthermore, immature/ETP-ALL and TLX cases, which frequently exhibit an immature or $\gamma\delta$ -lineage associated phenotype,[8] both express a pre-commitment gene signature as they co-cluster with “early”/pre-commitment T-cell stages. Proliferative and TALLMO cases express a post-committed gene signature and co-cluster with “late”/post-commitment T-cell stages (Figure 1b). Yet, a few samples are apparently misclustered (arrows), fact that can be possibly explained by the translocation partners observed on their rearrangements. For instance, translocations or fusions with *Bcl11b*, which is a haplo-insufficient tumor suppressor, could impair TALLMO rearranged leukemic cells to pass beyond the T-cell commitment point, whereas TLX3 translocations with *TCR- β* or *TCR- δ* loci among normal, unrearranged *Bcl11b* loci may allow cells to pass T-cell commitment and $\alpha\beta$ -lineage development. (work in progress)

Instead of the 547 genes signature, which is purely based on differentially expressed genes during normal T-cell development, PCA mapping was also performed using the 435 probeset (398 genes) gene signature (Figure 1c), which previously separated T-ALL patient samples into the 4 T-ALL clusters by using an unsupervised approach.[4] As expected, the samples could be separated into the 4 T-ALL clusters (immature/ETP-ALL, TLX, proliferative and the TALLMO), showing a close overlap with the clustering that distinguished pre- and post-commitment T-cell programs (Figure 1a). This suggests that the 398-gene signature could already contain differentially expressed T-cell developmental genes among the T-ALL clusters. We then examined the coinciding genes between the 547 T-cell development gene signature

and the 398 gene set (435 probes) from our previous unsupervised T-ALL cluster analysis, and identified 107 overlapping genes that are differentially expressed during T-cell development as well as among the 4 clusters of T-ALL. Hierarchical clustering of the T-ALL samples based on these 107 genes revealed four main gene groups (I-IV) that are closely related to the four T-ALL clusters (Figure 1d and Supplemental Table S1). For instance, immature/ETP-ALL is associated with high expression of genes known to be expressed during the ETP-stage, such as LMO2 and HHEX (Group I); the TLX cluster that comprises HOXA-activated and TLX3-rearranged T-ALL cases expresses NOTCH3, a relevant gene for gamma-delta T cell differentiation (Group II)[19]; the proliferative cluster exhibits genes as DNTT (also known as TDT) that may reflect the initiation of T-cell receptor recombination events, and CD1A, which is a hallmark of this cluster (Group III), and finally the TALLMO group contained genes of more mature T-cells such as LEF1 and CD8 (Group IV). Thus, the gene signature of 435 probesets that we established to cluster different T-ALL samples is highly enriched with genes from the T-cell developmental signature (Supplemental Figure S1c), indicating that T-ALL oncogenes cooperate with T-cell developmental programs for leukemogenesis. This may imply that oncogenes can only exert their leukemogenic potential and promote differentiation arrest when ectopically expressed during specific T-cell developmental programs (Figure 1e).

Validation of lentiviral vectors expressing MEF2C, LYL1, or LMO2 in Jurkat cells.

We further studied this in OP9-DL1 co-cultures analyzing the effect of various genes that are highly expressed in immature/ETP-ALL patients (LYL, LMO2 and MEF2C) on normal T-cell development. For this, we designed and produced lentiviral vectors that allow the overexpression of MEF2C, LYL1 or LMO2 in differentiating HSCs *in vitro* (Figure 2a). Following virus production and titration of the lentiviral oncogene vectors or BFP-control vector we checked expression of oncogenes upon transduction of T-ALL cell lines, Jurkat (Figure 2b) and Loucy cells (Supplemental Figure S2). The oncogenes and BFP moieties are efficiently expressed in target cells and therefore transduction efficiencies can be measured by the percentage of BFP-positive cells (Figure 2b, upper panel). The percentage of transduced Jurkat cells (BFP-positive) remained unchanged (>95%) for over three weeks of culture, reflecting stable integration of the viral genome in the host DNA (data not shown). The only exception occurred to cells transduced with the MEF2C lentiviral vector as the percentage of BFP-positive cells gradually decreased, possibly reflecting toxic effects of MEF2C overexpression. To confirm the ectopic expression of the oncogenes on transduced cells, MEF2C, LYL1 and LMO2 were identified by western-blot analysis using T2A or specific antibodies (Figure 2b and Supplemental Figure S2).

Constitutive Expression of MEF2C, LYL1, or LMO2 impairs T-cell development

To study the effect of MEF2C, LYL1 or LMO2 on normal T-cell development, we transduced umbilical cord blood-derived HSCs with each lentiviral construct, in at least 3 independent experiments. All lentiviral batches transduced UCB-HSCs with 40-60% transduction efficiencies, as measured by the percentage of BFP-positive cells (data not shown). The MEF2C construct provided low transduction efficiency that barely exceeded 10% transduction rates. The effect of constitutive expression of each of these genes in differentiating HSCs was analyzed upon culture of transduced HSCs on the OP9-DL1 *in vitro* system (Figure 3a).

The dynamics of the co-cultures was evaluated by detection of CD34 and CD7 expression (Figure 3b and Figure 3c). HSCs are CD34⁺CD7⁻, CD34⁻CD7⁻ are non T-cell populations (including myeloid-, B- and NK-cells), and CD34⁺CD7⁺ and CD34⁻CD7⁺ cells represent T-cell populations (Figure 3b). The BFP reporter allows for flow-cytometry gating and analysis of transduced HSCs (Figure 3c, two right graphs with blue dots), whereas BFP-negative cells from non-transduced HSCs can be monitored as a control in the same culture (two left graphs with pink dots). BFP-negative cells from the transduction experiments with each expression construct and the BFP-positive cells transduced with the BFP control vector normally differentiate into the T-cell lineage (top graphs). However, expression of each of the three ETP genes LMO2, LYL1 or MEF2C results in a reduced number of T-cells. This was not due to a defective T-cell development caused by a culture defect, as untransduced HSCs (BFP⁻) in the same experiment efficiently generated T-cells.

MEF2C not only blocked T-cell development, but also blocked differentiation of HSCs into alternative B-, myeloid or NK- lineages. (Figure 3c) Also, expression of LYL1 or LMO2 reduced the T-cell differentiation potential of HSCs, while the potential to develop into alternative cell fates remains unchanged. As a consequence, the number of untransduced (BFP⁻) T-cells and other cell lineages is higher compared to the BFP control due to the extra availability of space and cytokines in the absence of surviving transduced cells.

Notably, ectopic expression of MEF2C, LYL1 or LMO2 in HSCs arrested developing cells at early T-cell development (Figure 3d). The expression of MEF2C leads to an immediate block in T-cell differentiation and cells fail to acquire CD7 as the first T-cell marker (Day 5). Instead, with increased passage cells gradually lose CD34 expression and eventually disappear from the co-culture due to apoptosis. Overexpression of LYL1 or LMO2 allows the cells to acquire CD7, although fewer BFP⁺, CD7⁺ cells are obtained compared to the control (Day 9 and 12, respectively), and there is also a lower percentage of CD7⁺ cells losing CD34 as a parameter of

T-cell maturation. Thus, constitutive expression of MEF2C, LYL1, or LMO2 impairs early T-cell development *in vitro*, comparable to what is seen in ETP-ALL patients.

Constitutive expression of MEF2C, LYL1, or LMO2 is dependent on ETP-transcriptional programs to promote differentiation block before T-cell commitment

To distinguish and characterize the differentiation arrest caused by each oncogene, we measured five immunophenotype markers (CD34, CD44, CD7, CD5, CD1a) that can discriminate consecutive T-cell developmental stages (A-F), from HSC till the CD1⁺ stage (Figure 4a). Those populations together with the non T-cell (M) population represent >85% of the total of cells detected in each well by flow-cytometry, while <15% of cells remain uncharacterized as these fall outside the defined flow-cytometry gates. The T-cell lineage commitment occurs in the CD7⁺CD5⁺ stage when cells lose the expression of the CD44 surface marker (Chapter 5). Figure 4b shows the total number of BFP-positive cells per fraction, representing the emergence of consecutive T-cell developmental stages over different days of co-culture. Cells transduced with the BFP control vector at day 2 of co-culture remain mostly as hematopoietic progenitors (CD34⁺CD44⁺CD7⁻), while at days 5 and 8 a substantial number of early T-cells are observed based on the expression of CD7 and CD5 (fractions B and C, respectively). Besides the presence of early T-cell populations, day 12 is characterized by further differentiated T-cell subsets that have lost CD34 expression (D-E2), including cells that have undergone T-cell commitment (E2) as characterized by loss of CD44. Populations expressing CD1a appear at day 18 of co-culture. Importantly, the dynamics of the populations (A-F) during the co-culture is identical between HSCs that are transduced with the BFP control vector (BFP-positive) in comparison to cells that have not been transduced (BFP-negative) (Supplemental Figure S3a).

The expression of MEF2C leads to an immediate differentiation block of the transduced (BFP⁺) HSCs, as practically there are no cells that have differentiated in either the T-cell or non-T-cell lineages (Figure 4b). The expression of LYL1 facilitates a partial differentiation block at the most immature CD7⁺ ETP stage that impairs further differentiation into CD7⁺CD5⁺ cells (fraction C). Ectopic expression of LMO2 facilitates T-cell development up till the CD7⁺CD5⁺ ETP-stage, where cells are partially blocked. Few cells may escape this block and further lose CD34 and CD44, and acquire CD1a (D-F). Overall, each of these oncogenes facilitate a T-cell developmental arrest of differentiating T-cells before the T-cell commitment checkpoint, consistent with the early T-cell development block of leukemic ETP-ALL cells (Figure 4c). This indicates that the oncogenes may only exert their ability to block differentiation in the context of

appropriate T-cell development. In order to prove this point, we investigated whether MEF2C, LYL1, or LMO2 might block T-cell differentiation in further differentiated T-cells. We therefore transduced early CD7⁺ T-cells, and studied their differentiation potential on OP9-DL1 stromal cells. Remarkably, when CD7⁺-sorted cells were transduced (day 6 of co-culture) with the MEF2C or LYL1 lentiviral vector, those cells were able to differentiate normally into the various T-cell stages (Supplemental Figure S3b). This indicates that the differentiation block caused by overexpression of MEF2C or LYL1 is dependent on the presence of a very premature T-cell developmental program, before the expression of CD7. In contrast, sorted CD7⁺ cells transduced with LMO2 lentiviral vector still exhibited a partial block at the CD7⁺CD5⁺ ETP-stage (Supplemental Figure S3b). Overall, constitutive expression of MEF2C, LYL1 or LMO2 *in vitro* results in a premature differentiation arrest of developing T-cells, but at different ETP-stages. Further, we also demonstrate that MEF2C, LYL1 or LMO2 are dependent on those early T-cell developmental programs to exert their leukemogenic effect.

Discussion

In this study we reveal that T-cell development gene signatures provide a similar rearrangement-associated clustering of T-ALL subtypes as previously obtained by unsupervised cluster analysis. We further demonstrate that T-ALL oncogenes are only able to exert an oncogenic function in the context of expression of particular T-cell transcriptional programs. In addition, the observation that the previously described T-ALL signature was highly enriched for important T-cell development regulators indicated that the T-ALL oncogenes highly cooperate with T-cell developmental programs for leukemogenesis. This was further supported by the OP9-DL1 *in vitro* studies, where MEF2C, LYL1 or LMO2 were dependent on the transcriptional networks from the initial stages of T-cell development to induce a differentiation arrest characteristic of ETP-ALL. In addition, we showed that the TLX subtype that is mostly associated with $\gamma\delta$ -lineage express a pre-commitment transcriptional program of T-cell development. This is in line with our own patient data, as most TLX3-rearranged cases have translocations where BCL11B is involved, and are therefore haploinsufficient for the essential role of BCL11B in T-cell lineage commitment and maintenance of T-cell identity.[20, 21] Of note, the OP9-DL1 system clearly favors the differentiation of $\alpha\beta$ -T cells, thus to specifically comprehend the transcriptional program of $\gamma\delta$ -lineage a similar *in vitro* system needs to be established that will support $\gamma\delta$ -T-cell lineage development.

The transcription factors MEF2C, LYL1 and LMO2 play central roles in normal early T-cell development and consequently abnormal regulation of MEF2C can promote

leukemogenesis, resulting in ETP-ALL. We have previously demonstrated that LYL1 and LMO2 are regulated by MEF2C, and showed that MEF2C can directly bind to the promoter regions of LMO2. These observations support the oncogenic roles of LYL1 and LMO2 in MEF2C-deregulated ETP-ALL.[4] Thus, it is crucial to understand both the oncogenic function of each protein individually and their involvement in major transcriptional complexes by identification of protein-protein interactions. ETP-ALL cells usually express myeloid surface markers (CD13 and/or CD33) and have lower expression of T-cell specific genes compared to other types of T-ALL.[22] This may be derived from the fact that the expression of MEF2C is higher in myeloid and B-cells compared to T-cells and may therefore favor the differentiation of other non T-cell lineages, and also ectopic expression of MEF2C has been linked to MLL-rearranged acute myeloid leukemia.[14] Consequently, MEF2C-transduced HSCs in the OP9-DL1 *in vitro* system blocks T-cell development at very early stage when cells are still CD34⁺, and exhibit high expression of CD13 and CD33 in cells that attempt to differentiate into non T-cell lineages (such as B-, myeloid or NK-cells). Possibly, T-cells arrested at this immature/ETP-ALL stage do not survive *in vitro*, as these cells may not be sustained with enough proliferation and survival signals, which may be supported by IL7R[23] or RAS mutations in ETP-ALL. (Y. Li and J.G. Buijs-Gladdines et al. unpublished data). Importantly, the MEF2C-driven oncogenic mechanism that leads to differentiation block in ETP-cells is dependent on a very early T-cell developmental stage, as MEF2C overexpression at later stages did not provoke developmental arrest *in vitro*.

Furthermore, constitutive expression of LYL1 or LMO2 in HSC is also incompatible with development into and beyond the T-cell commitment checkpoint and these proteins could therefore play important roles in the pathogenesis of ETP-ALL. The partial block observed at the constitutive expression of LYL1 or LMO2 may be related to “escaper cells” that have a lower expression of those oncogenes or due to continuous NOTCH1-activation promoted by the DL1-ligand expressed on the membrane of OP9-DL1 cells that enforces T-cell differentiation. Similarly, a study using fetal thymus organ cultures showed that human CD34⁺ cells with ectopic expression of LMO2 result in an incomplete developmental block, exhibiting a lower percentage of CD1⁺ cells compared to untransduced cells.[24] However, LMO2 is an oncogene for TALLMO patients that have arrested at more advanced $\alpha\beta$ -lineage T-cells, therefore ectopic expression of LMO2 may facilitate multiple stages of T-cell developmental arrest. In line with this, a recent study with humanized NSG mice transplanted with human CD34⁺ stem/progenitor cells that were genetically modified to overexpress LMO2, caused 3 aberrant effects on human T-cell development *in vivo*. [25] Specifically, it was observed a block at the DN/ ISP stage, an accumulation of CD4⁺CD8⁺ DP CD3⁻ cells, and an altered CD8/CD4 ratio with enhanced

peripheral T lymphocytes. These results highlight the limitations of the OP9-DL1 *in vitro* system in developing T-cells into more mature stages, and demonstrate the strength of complementing this kind of studies with humanized mouse model. Furthermore, CD2-LMO2 mice give rise to highly penetrant T-ALL by two distinct patterns of gene expression: with ETP-ALL features such as high expression of HHEX, LYL1 and MEF2C [10] or alternatively, with Notch1 target gene activation. [26] These mice studies point to a synergy between co-expression of LMO2, LYL1 or HHEX with ectopic expression of MEF2C in ETP-ALL. For instance, HHEX and LYL1 seem to tightly collaborate with MEF2C as the incidence of T-ALLs in those mice is strongly reduced in the absence of each transcription factor. [9, 26] Therefore, we hypothesize that MEF2C synergizes with LMO2 and LYL1 in the dysregulation of important signaling pathways, resulting in ETP-ALL.

Overall, we conclude that T-ALL oncogenes can only promote leukemogenesis if they are activated at specific developmental stages where they can perturb normal T-cell development. This hypothesis is supported by cytogenetic analysis of T-ALL patients, as specific translocations (e.g. Bcl11b or TCR rearrangements) only occur at developmental contexts when the chromatin of those loci is open. The developmental arrest allows the rapid acquisition of additional genetic mutations, so called type-B mutations that will give proliferative advantage to those pre-leukemic cells and ultimately lead to T-ALL. The *in vitro* system approach is a valuable tool to investigate the mechanisms that promote cellular arrest and it may help to understand how specific oncogenes and genetic lesions cooperate during T-cell leukemogenesis and disease progression.

Acknowledgements

R.D.M. and K.C.-B. were financed by the Children Cancer Free Foundation (Stichting Kinderen Kankervrij (KiKa 2007-12, KiKa 2008-29 and KiKa 2013-116)).

Contribution

R.D.M., K.C.-B. and J.P.P.M. designed the study; R.D.M., K.C.-B. and W.K.S. cloned the lentivirus expression constructs and produced the lentiviral constructs. R.D.M. performed the OP9-DL1 experiments and wrote the manuscript. R.D.M., K.C.-B., R.P. and J.P.P.M. analyzed and interpreted data; and all authors read, revised, and approved the paper.

References

1. Rothenberg, E.V., J.E. Moore, and M.A. Yui, *Launching the T-cell-lineage developmental programme*. Nat Rev Immunol, 2008. **8**(1): p. 9-21.
2. Zhang, J.A., et al., *Dynamic transformations of genome-wide epigenetic marking and transcriptional control establish T cell identity*. Cell, 2012. **149**(2): p. 467-82.
3. Meijerink, J.P., *Genetic rearrangements in relation to immunophenotype and outcome in T-cell acute lymphoblastic leukaemia*. Best Pract Res Clin Haematol, 2010. **23**(3): p. 307-18.
4. Homminga, I., et al., *Integrated transcript and genome analyses reveal NKX2-1 and MEF2C as potential oncogenes in T cell acute lymphoblastic leukemia*. Cancer Cell, 2011. **19**(4): p. 484-97.
5. Coustan-Smith, E., et al., *Early T-cell precursor leukaemia: a subtype of very high-risk acute lymphoblastic leukaemia*. Lancet Oncol, 2009. **10**(2): p. 147-56.
6. Zuurbier, L., et al., *Immature/early T-cell precursor acute lymphoblastic leukemia patients treated on the COALL-97 protocol are not associated with poor outcome* Submitted for publication, 2012.
7. Ferrando, A.A., et al., *Gene expression signatures define novel oncogenic pathways in T cell acute lymphoblastic leukemia*. Cancer Cell, 2002. **1**(1): p. 75-87.
8. van Grotel, M., et al., *Prognostic significance of molecular-cytogenetic abnormalities in pediatric T-ALL is not explained by immunophenotypic differences*. Leukemia, 2008. **22**(1): p. 124-31.
9. McCormack, M.P., et al., *Requirement for Lyl1 in a model of Lmo2-driven early T-cell precursor ALL*. Blood, 2013. **122**(12): p. 2093-103.
10. McCormack, M.P., et al., *The Lmo2 oncogene initiates leukemia in mice by inducing thymocyte self-renewal*. Science, 2010. **327**(5967): p. 879-83.
11. Shields, B.J., et al., *Hhex regulates Kit to promote radioresistance of self-renewing thymocytes in Lmo2-transgenic mice*. Leukemia, 2015. **29**(4): p. 927-38.
12. Martins, V.C., et al., *Cell competition is a tumour suppressor mechanism in the thymus*. Nature, 2014. **509**(7501): p. 465-70.
13. Yui, M.A. and E.V. Rothenberg, *Developmental gene networks: a triathlon on the course to T cell identity*. Nat Rev Immunol, 2014. **14**(8): p. 529-45.
14. Cante-Barrett, K., R. Pieters, and J.P. Meijerink, *Myocyte enhancer factor 2C in hematopoiesis and leukemia*. Oncogene, 2014. **33**(4): p. 403-10.
15. Schmitt, T.M. and J.C. Zuniga-Pflucker, *Induction of T cell development from hematopoietic progenitor cells by delta-like-1 in vitro*. Immunity, 2002. **17**(6): p. 749-56.
16. Weber, K., et al., *A multicolor panel of novel lentiviral "gene ontology" (LeGO) vectors for functional gene analysis*. Mol Ther, 2008. **16**(4): p. 698-706.
17. Szymczak, A.L. and D.A. Vignali, *Development of 2A peptide-based strategies in the design of multicistronic vectors*. Expert Opin Biol Ther, 2005. **5**(5): p. 627-38.
18. Holmes, R. and J.C. Zuniga-Pflucker, *The OP9-DL1 system: generation of T-lymphocytes from embryonic or hematopoietic stem cells in vitro*. Cold Spring Harb Protoc, 2009. **2009**(2): p. pdb prot5156.
19. Van de Walle, I., et al., *Specific Notch receptor-ligand interactions control human TCR-alpha/beta/gammadelta development by inducing differential Notch signal strength*. J Exp Med, 2013. **210**(4): p. 683-97.
20. Li, L., M. Leid, and E.V. Rothenberg, *An early T cell lineage commitment checkpoint dependent on the transcription factor Bcl11b*. Science, 2010. **329**(5987): p. 89-93.
21. Li, P., et al., *Reprogramming of T cells to natural killer-like cells upon Bcl11b deletion*. Science, 2010. **329**(5987): p. 85-9.

22. Zhang, J., et al., *The genetic basis of early T-cell precursor acute lymphoblastic leukaemia*. Nature, 2012. **481**(7380): p. 157-63.
23. Zenatti, P.P., et al., *Oncogenic IL7R gain-of-function mutations in childhood T-cell acute lymphoblastic leukemia*. Nat Genet, 2011. **43**(10): p. 932-9.
24. Pike-Overzet, K., et al., *Ectopic retroviral expression of LMO2, but not IL2Rgamma, blocks human T-cell development from CD34+ cells: implications for leukemogenesis in gene therapy*. Leukemia, 2007. **21**(4): p. 754-63.
25. Wiekmeijer, A.S., et al., *Overexpression of LMO2 causes aberrant human T-Cell development in vivo by three potentially distinct cellular mechanisms*. Exp Hematol, 2016. **44**(9): p. 838-849.e9.
26. Smith, S., et al., *LIM domain only-2 (LMO2) induces T-cell leukemia by two distinct pathways*. PLoS One, 2014. **9**(1): p. e85883.

Figure legends

Figure 1. T-ALL subtypes are associated with pre-/post- commitment developmental programs (a) Co-clustering of 117 T-ALL patient samples and *in vitro* T-cell sorted fractions (Chapter 5). The PCA analysis is based on the 547 genes that are differentially expressed across T-cell differentiation. The T-ALL subtypes are divided into the four T-ALL unsupervised clusters: Immature/ETP-ALL, TLX, Proliferative and TALLMO; and the T-cell sorted fractions are divided into CD34⁺ UCB-derived stem/progenitor cells, pre-commitment and post-commitment stages. (b) Hierarchical clustering analysis using the Kendall average method for 236 genes (547-Var50Val50). The dendrogram at the left represents the similarity between the 117 T-ALL patients divided into 4 T-ALL clusters. The dendrogram at the top indicates how related is the expression of the 547 genes throughout the different T-ALL samples. The dashed line represents the T-cell lineage commitment and the two arrows identify two T-ALL samples that are misclustered due to translocation partners (c) PCA mapping of 117 T-ALL patient samples using a gene signature of 435 probesets (398 genes) that was used to distinguish the 4 gene expression-based clusters of T-ALL in Homminga et al. Cancer Cell 2011. (d) Hierarchical clustering analysis (PearsonWards method) using 107 genes that overlap between the 547-gene signature and the T-ALL signature of 435 genes probesets. The dendrogram at the left represents the similarity of expression of 107 genes across the 117 T-ALL patient samples. The dendrogram at the top indicates how related are the 117 T-ALL patients samples and the respective cluster is indicated. (e) Schematic representation of the different subtypes of T-ALL with the corresponding developmental arrest stages provoked by the ectopic expression of the oncogenes. The four T-ALL clusters are also separated into pre-commitment and post-commitment T-cell developmental programs.

Figure 2. Constitutive Expression of MEF2C, LYL1, or LMO2 impairs T-cell development (a) Schematic representation of the elements cloned into our Gateway compatible LeGO-DEST vector. The 3 lentiviral expression constructs coding the oncogenes *LMO2*, *LYL1* or *MEF2C* exhibit 4830 bp, 5190 bp and 5750 bp of proviral RNA lengths, respectively. The BFP control vector composed by the SFFV promoter followed by the BFP reporter contains 4215 bp. SFFV: Spleen Focus Forming Virus promoter, T2A: Thosea asigna virus 2A element, BFP: blue fluorescent protein. Grey boxes indicate lentiviral elements in LeGO-DEST. (b) Top panel: flow-cytometry image demonstrating BFP⁻ and BFP⁺ (transduced) Jurkat cells; bottom panel: Western-blot of lysates from Jurkat cells transduced with MEF2C, LYL1, or LMO2 lentiviral

constructs and the SFFV-BFP control. After cleavage of the proteins, the MEF2C, LYL1, or LMO2 retain the majority of the T2A peptide that can serve as an epitope tag and can be recognized by the anti-2A antibody (left blot). The BFP reporter that follows the T2A sequence is separated and not recognized by anti-2A. Direct antibodies against MEF2C, LYL1, or LMO2 were used to detect the expression of those proteins in transduced Jurkat cells.

Figure 3. Constitutive Expression of MEF2C, LYL1, or LMO2 impairs T-cell development (a) Schematic representation of the OP9-DL1 co-cultures used to differentiate transduced HSCs into consecutive T-cell differentiation stages. Briefly, HSC were transduced with the different lentivirus constructs and cultured on top of OP9-DL1 cells for 18 days until the cells reached the CD1a⁺ stage. (b) Schematic representation of the flow-cytometry analysis used to distinguish the different populations present in the co-cultures. The use of CD34 and CD7 antibodies allows the detection of 4 different populations: HSCs (CD34⁺CD7⁻), non-T-cell populations (CD34⁻CD7⁻), ETPs (CD34⁺CD7⁺) and T-cells (CD34⁻CD7⁺). (c) Transduced cells are illustrated as BFP-positive cells in flow-cytometry analysis. Analysis of transduced cells is depicted as blue dots on the right graphs and analysis of non-transduced cells is depicted as pink dots on the left graphs. The two top-graphs correspond to the number of T-cells (CD7⁺CD34^{+/+}) and the two bottom graphs correspond to the number of non T-lineage cells (CD7⁻CD34⁻) in 3 independent single well experiments for each construct over 18 days of co-culture (n=3). (d) Flow-cytometry plots illustrating distinct T-cell differentiation stages based on the expression of CD34 and CD7 markers at days 5, 9 and 12. Cells transduced with the control vector are on top and cells transduced with lentiviral constructs expressing oncogenes are at the bottom. All populations were also gated for CD45⁺GFP⁻ to exclude OP9-DL1 cells that are GFP positive.

Figure 4. Constitutive expression of MEF2C, LYL1, or LMO2 is dependent on ETP-transcriptional programs to promote differentiation block before T-cell commitment (a) Immunophenotype of 7 T-cell subsets (A-F) measured over time in OP9-DL1 co-cultures based on surface markers CD45, CD34, CD7, CD5, CD1a, and CD44. (b) Total cell numbers of distinct T-cell subsets (A-F, all BFP⁺ gated) upon transduction of HSCs with control-, MEF2C-, LYL1-, or LMO2-lentiviral vectors and differentiation in OP9-DL1 during 18 days. All subsets were also gated on CD45⁺ and GFP⁻ to exclude OP9-DL1 cells. (c) Schematic representation of the differentiation block caused by constitutive expression of MEF2C, LYL1, or LMO2 before T-cell commitment.

Figure 1. T-ALL subtypes are associated with pre-/post- commitment developmental programs.

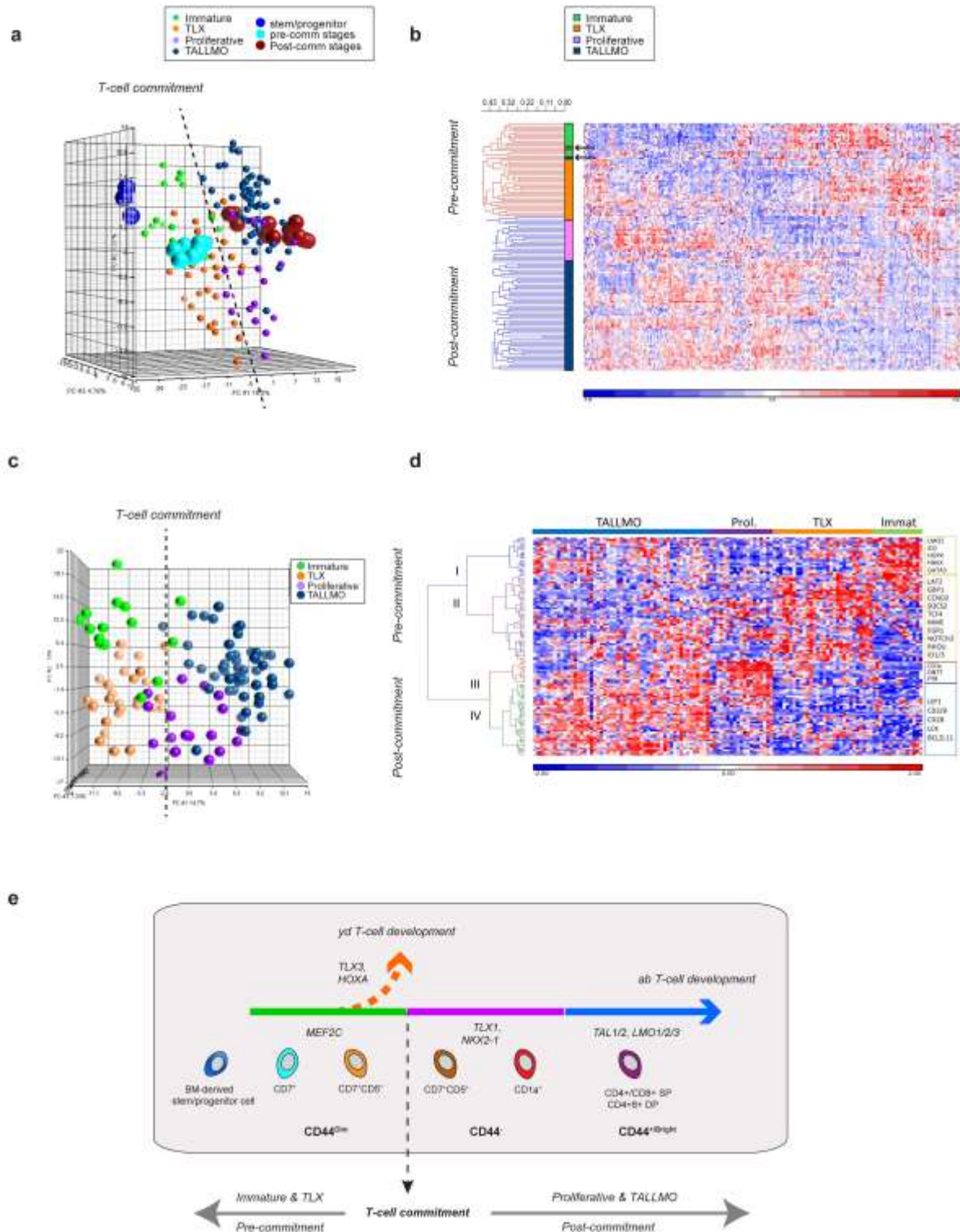
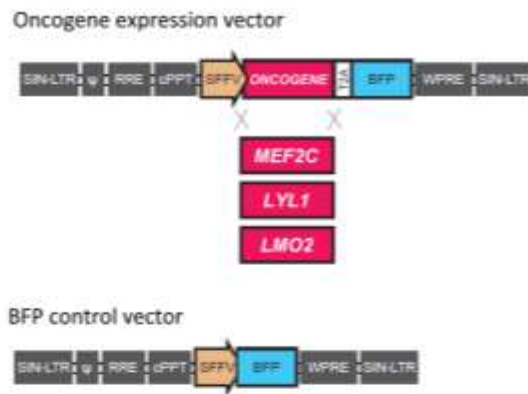


Figure 2

a



b Transduction in Jurkat cell line

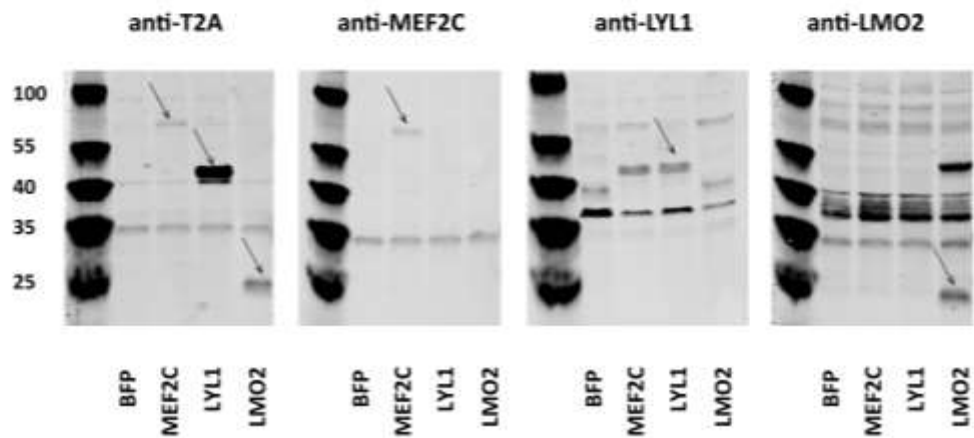
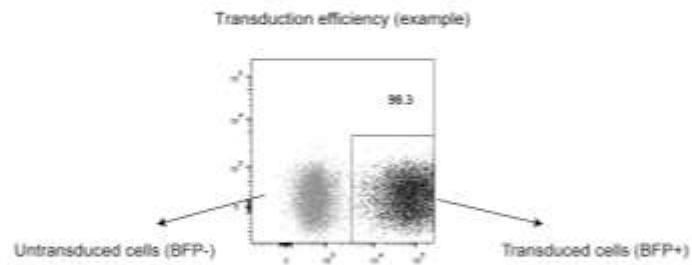


Figure 3

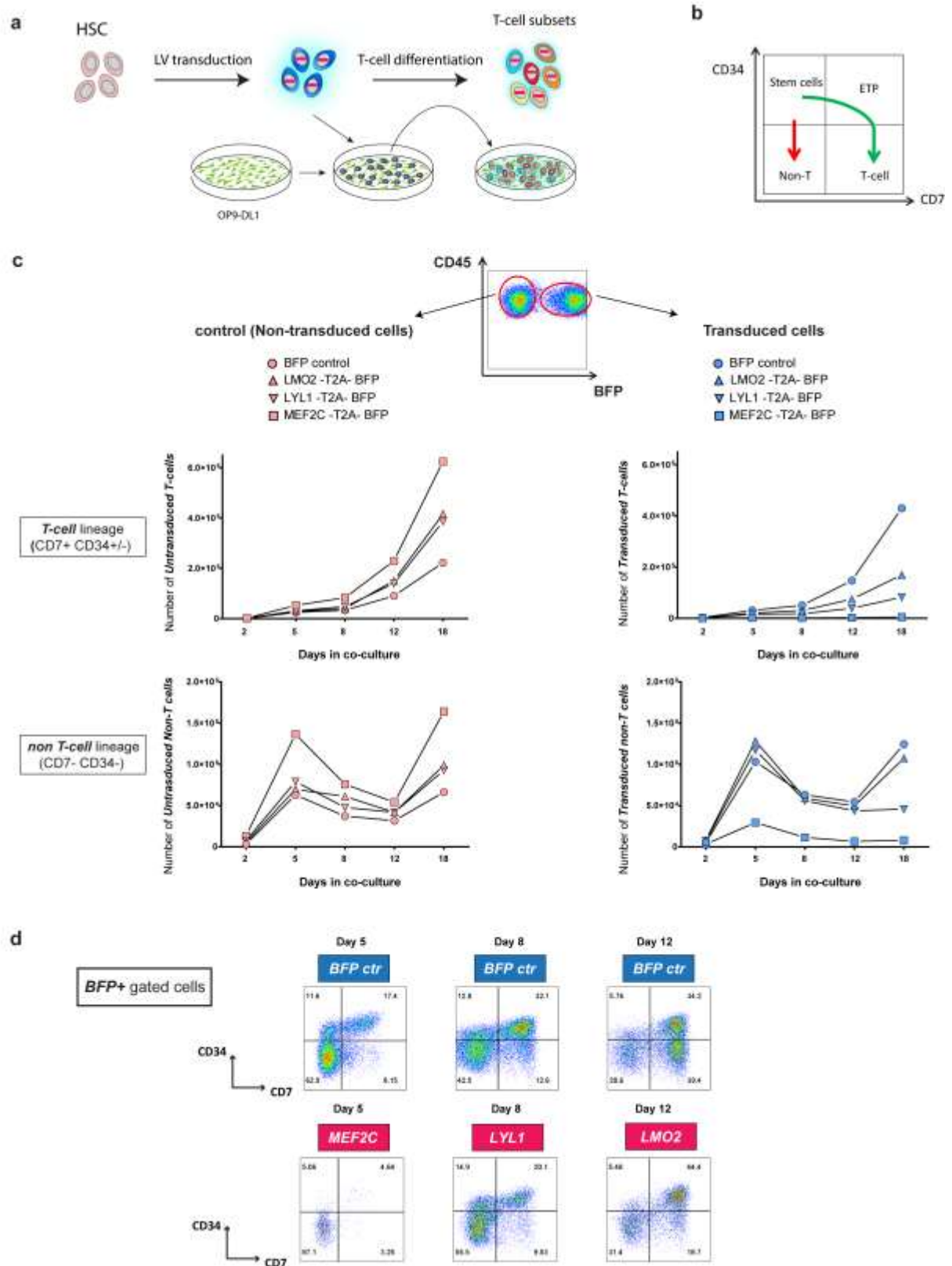
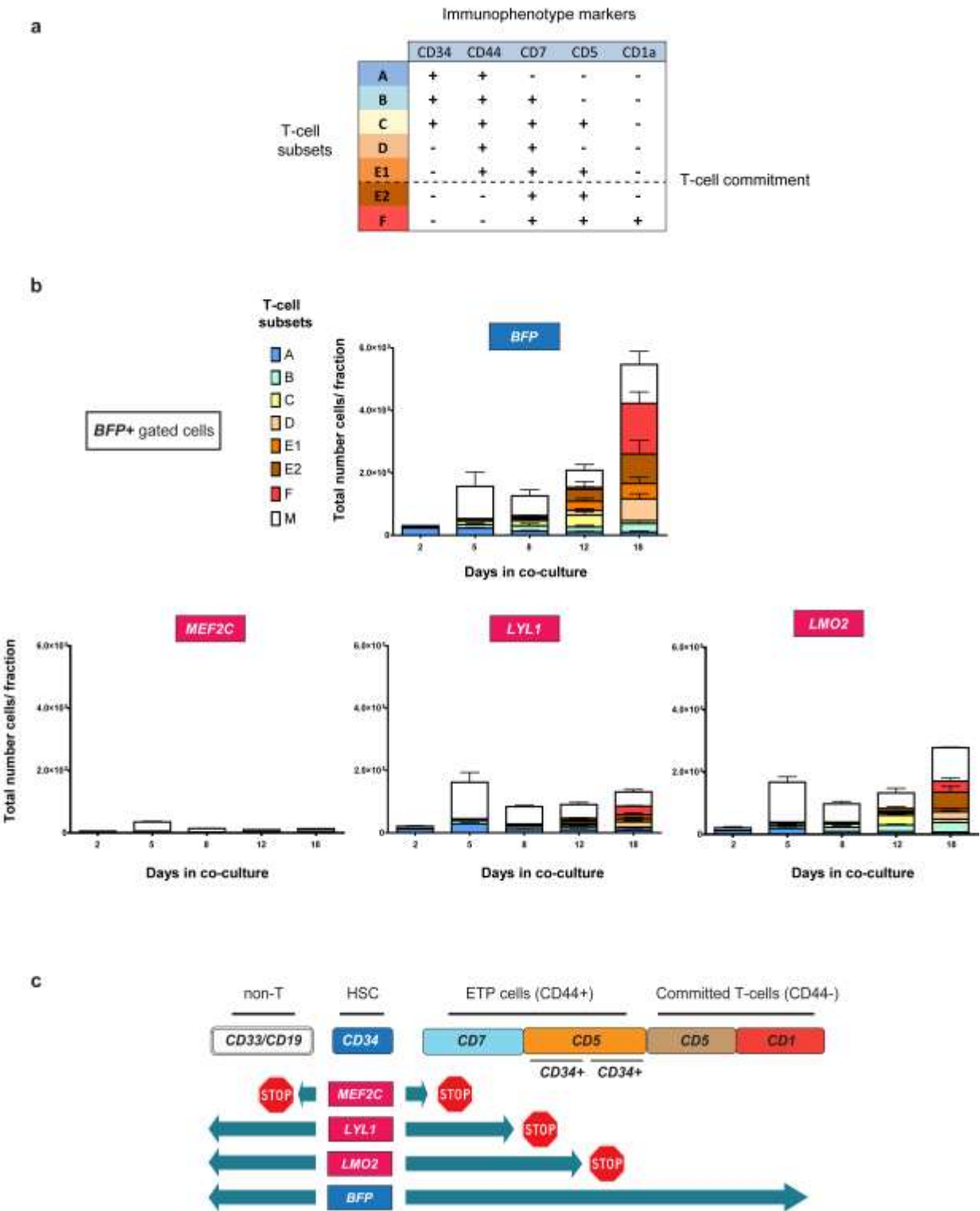


Figure 4



Chapter 7

General discussion & Future perspectives

General discussion and future perspectives

Over the last decades, diverse genomic aberrations were identified and characterized in T-ALL patient samples, resulting in the activation of specific oncogenes or the inactivation of tumor suppressor genes. The overexpression of rearranged oncogenic transcription factors classify distinct genetic T-ALL subgroups that harbor specific gene expression profiles. These aberrations (type-A) are considered the driving oncogenic events, which are associated with arrest at specific stages of T-cell differentiation. In addition, recurrent genetic aberrations that affect cell viability and/or proliferation are denoted as type B hits, and are found in nearly all T-ALL genetic subgroups. Despite the extensive identification of genetic defects driving human T-ALL, multi-agent chemotherapy sometimes combined with stem cell transplantation is the standard treatment for pediatric T-ALL patients. Besides the risk for severe acute toxicities and long-term side effects that are associated with this intensive treatment, still 15% of the patients relapse and die due to therapy resistance. The current challenge is therefore to search for alternative patient-tailored treatment regimens that may reduce toxicity or enhance the cellular response to chemotherapeutic agents. To achieve this we need to understand how those genetic aberrations occur, which oncogenic mechanisms are being activated, and how these events contribute to leukemogenesis, disease progression or relapse. This thesis clearly shows the benefits of using *in vitro* methodologies to understand the impact of oncogenic mechanisms and their contribution to leukemogenesis. These data may therefore contribute to the development of novel targeted therapies.

The impact of type B mutations in T-ALL heterogeneity and clonal evolution

Clonal genetic diversity is a common feature of cancer and is probably, from a Darwinian natural selection perspective, the essential substrate for clonal evolution, disease progression and relapse. T-ALL is a genetically complex type of leukemia, exhibiting a mixture of leukemic subclones carrying different combinations of genetic lesions at diagnosis that are under constant selection pressure during chemotherapeutic treatment. In contrast to type-A mutations, which are considered the driving genetic events present in all cells of the leukemia, type-B mutations may not be clonal and are frequently acquired or lost during disease progression as result of clonal evolution. Studies on clonal evolution are possible nowadays due to advances in Next-generation sequencing (NGS) techniques and the development of bioinformatic tools that can distinguish these mutations at the clonal and subclonal levels. Identifying the full complexity of

subclonal architecture and genetic diversity in T-ALL can be done through deep sequencing of entire tumor cell populations. This approach allows to discriminate subclonal passenger mutations from the oncogenic driver mutations that most likely contribute to disease progression and therapy resistance.¹ Targeted therapy directed against mutant molecules may have limited efficacy if the targets themselves are not initiating lesions, but reflect secondary mutations occurring in leukemia subclones. Because genetic variation may represent a significant obstacle to effective therapy, WGS will certainly support personalized targeted therapy, namely the combination of specific signaling inhibitors. For instance, NOTCH1-activated leukemias should be screened for the presence of subclonal PTEN aberrations – such as *PTEN* microdeletions described in Chapter 3 – since newly acquired *PTEN* mutations may drive resistance to NOTCH1-targeting therapies as discussed in Chapter 2. In this case, combined NOTCH and AKT-directed therapies may prove an effective treatment relapse of T-ALL blasts (Chapter 2). Different lines of research are currently focusing on the understanding of molecular mechanisms that are associated with chemotherapy resistance, as the biological basis for disease relapse in T-ALL is still poorly understood. Recently, whole-exome sequencing of matched diagnosis, remission and relapse DNA samples, identified recurrent mutations that are acquired in relapse samples after chemotherapy, indicating that relapse is not only mediated by minor leukemic subclones that are already present at diagnosis. Specifically, activating mutations in the cytosolic 5'-nucleotidase II gene (*NT5C2*) were found in about 20% of relapsed T-ALL patients and are associated with the outgrowth of 6-mercaptopurine (6-MP)- or 6-thioguanine (6-TG)-resistant clones.^{2,3} *NT5C2* encodes a 5'-nucleotidase enzyme that is responsible for the inactivation of nucleoside-analog chemotherapy drugs routinely used in T-ALL maintenance therapy, including 6-MP and 6-TG.^{4,5} Accordingly, expression of *NT5C2* mutant in ALL lymphoblasts resulted in increased nucleotidase activity *in vitro* conferring resistance to 6-MP and 6-TG.^{2,3} These results demonstrated the impact of increased nucleoside-analog metabolism in disease progression and chemotherapy resistance. Furthermore, increasing evidence is demonstrating the role of leukemic stem cells (LSC),⁶⁻⁸ epigenetic changes^{9,10} and metabolism alterations⁹ in the evolution of resistant leukemic clones. Overall, the identification and comprehension of molecular mechanisms involved in disease progression and therapy-resistance is critical for the development of new strategies that may prevent the acquisition or outgrowth of treatment resistant subclones.

Genetic abnormalities in normal healthy tissues

In our studies in Chapter 3 we reveal that *PTEN* microdeletions can also occur in the thymus of normal healthy thymocytes. This observation together with other studies^{11,12} show that healthy human thymocytes are also prone to illegitimate RAG-mediated recombination events as observed in T-ALL blasts, resulting in the inactivation of tumor suppressor genes or activation of oncogenes. However, in healthy thymocytes, those mutations may occur as single events or in cells that lack essential collaborating mutations to become a pre-leukemic cell. As proposed by Mel Greaves¹³ this phenomena corroborates the vulnerability of human tissues to oncogenic processes, which is supported by the high prevalence of “hidden” malignant cancers or pre-malignant carcinomas identified in organs of individuals dying of non-cancer causes. In line with this, the author defends that the process of tumorigenesis is a lottery, since most tumors never emerge all the bottlenecks. As an example, the screening of mutations in archived neonatal blood spots or Guthrie cards showed that many healthy babies carry pre-leukemic translocations such as TEL/AML gene fusion, but never develop leukemias as the secondary hits never occur.^{14,15} Hence, cancer is commonly or ubiquitously initiated but promotion to full malignancy is rare as the emergence of a cancer clone is a relatively inefficient evolutionary process. Accordingly, despite the frequent acquisition of *PTEN* microdeletions in thymocytes, this event by itself is not sufficient to enable increased cell survival or proliferation that leads to development of T-ALL. To become malignant, developing thymocytes with *PTEN* inactivation may require the collaboration with specific oncogenes (type A aberrations), such as TAL or LMO genes, that promote a late developmental arrest and allow the acquisition of additional genetic hits. Late T-cell development stages (post-commitment program) are especially prone to such oncogenic chromosomal translocations, because the “recombination machinery” that is responsible for TCR recombinations (i.e. V(D)J rearrangements) and for these leukemogenic events is especially expressed after T-cell commitment.

Normal T-cell development as a reference to elucidate the role of T-ALL oncogenes in leukemogenesis

In this thesis we have used two different approaches to understand the role of oncogenic transcription factors (Type-A aberrations) in T-ALL. First, we determined the expression of those oncogenes in normal T-cell development (Chapter 5); second, we examined the associations of these oncogenes with particular T-cell developmental programs that may be required for leukemogenesis and accordingly, we tested the enforced expression of ETP-ALL oncogenes during early T-cell development (Chapter 4 and 6). In the first approach (Chapter 5), supervised

cluster analysis of consecutive human T-cell differentiation stages in the OP9-DL1 system demonstrated that *in vitro* development closely resembles the *in vivo* situation both in human and mice. It therefore validates the OP9-DL1 culture system using umbilical cord blood-derived HSCs to study human T-cell development. Gene expression analysis exhibits the pattern of expression of the proto-oncogenes and other transcription factors throughout normal T-cell development, and helps to comprehend the function of expressed oncogenes in the different stages of the developmental process. Although some oncogenes as TLX1, TLX3, NKX2.1 and fusion proteins are not expressed during normal T-cell development, their ectopic expression in malignant thymocytes may interfere with important T-cell development regulators that are expressed during normal T-cell development. For instance, it has been described that TLX1 and TLX3 can interact with ETS1, which is a critical component of the complex that binds and activates the TCR enhancer element (E α), resulting in an enrichment of repressive histone modifications and lack of TCR α gene expression that ultimately leads to a differentiation arrest.¹⁶ Accordingly, in order to distinguish the activation of particular oncogenic mechanisms upon ectopic expression of the oncogenes and to understand how their abnormal regulation can promote leukemogenesis, it is crucial to fully understand normal T-cell development.

Remarkably, we identified the T-cell lineage commitment point in humans to occur during the CD7⁺CD5⁺stage, precisely when cells downregulate the expression of CD44 at the surface membrane. This exactly resembles the DN2a-DN2b transition that has been revealed as T-cell commitment point in mice. Despite a previous study that described a differential expression of CD44 earlier than CD1a expression in human T-cell development,¹⁷ here we place CD44 expression in the context of other T-cell development markers, which allows to distinguish consecutive T-cell developmental stages. More importantly, we functionally demonstrated that loss of CD44 marks human T-cell lineage commitment. This contributed to define pre- and post-commitment programs of T-cell development and helps understanding the association of different oncogenes with these particular programs. Defining the consecutive stages of normal T-cell development is also particularly useful when testing the leukemogenic effects of oncogenic transcription factors *in vitro*, as shown by the developmental arrest at the ETP-stage upon ectopic expression of MEF2C, LYL1 or LMO2. Overall, these results can be potentially interesting for the field of normal human T-cell development and accompanying T-cell (developmental) malignancies.

The oncogenic effect of T-ALL oncogenes depends on specific developmental programs

Furthermore, gene expression analysis of the consecutive T-cell stages obtained with the OP9-DL1 *in vitro* system allowed us to identify a gene expression signature of 1387 genes that cluster human T-ALL samples in four T-cell developmental clusters. Importantly, these clusters are nearly identical to the 4 genetic subgroups (ETP-ALL, TLX, proliferative and TALLMO) that we previously identified using unsupervised hierarchical clustering analysis of 117 pediatric T-ALL samples.¹⁸ Particularly, samples that are classified as ETP-ALL or TLX are characterized by the expression of a pre-commitment profile, while proliferative and TALLMO patient samples exhibited a post-commitment profile. This was expected for ETP-ALL, since it is the most immature T-ALL entity, but more importantly it supports our hypothesis that $\gamma\delta$ -lineage cells – characteristic of the TLX phenotype – diverge from $\alpha\beta$ -development before T-cell commitment. This hypothesis is supported by genetic analysis of T-ALL patient samples: TLX and ETP-ALL patients are frequently characterized by *BCL11B* enhancer-driven translocations to *TLX3* or other genes (*NKX2.5*, *SPI1*, *LMO2*). In contrast, TALLMO and proliferative subtypes frequently have TCR-driven rearrangements, supporting a pathogenic role of RAG1/2 proteins in pre-leukemic T-cells that have passed T-cell commitment.

Remarkably, about half of the genes represented in the gene signatures obtained from *in vitro* T-cell development data overlapped with the T-ALL gene signature composed of 435 probesets that previously led to the identification of the four T-ALL genetic subgroups.¹⁸ T-ALL oncogenic aberrations are therefore closely associated with pre- or post-commitment developmental programs. This supports our hypothesis that oncogenes need to cooperate with those programs for leukemogenesis, and actually may only exert their leukemogenic role in the context of specific T-cell developmental programs. This may help to understand how the oncogenes can block T-cell differentiation, and to what extent they are responsible for driving the expression of genes in these programs. Regarding this, we have physiologically tested the leukemogenic potential of MEF2C, LYL1 or LMO2 using the OP9-DL1 system, providing evidence that all three are able to block T-cell development within the ETP-stage. Contrarily, ectopic expression of these genes in thymocytes that already passed the ETP-stage revealed that these oncogenes did no longer affect the developmental process as cells could normally differentiate into the T-cell lineage. Accordingly, we conclude that solely the ectopic expression of the oncogene is not sufficient for leukemogenesis, but instead it requires a specific transcriptional program characteristic of distinctive T-cell developmental stages to cause a differentiation arrest. Moreover, although MEF2C overexpression was capable of arresting the cells at a very immature stage, those cells eventually disappeared from culture. Probably, those cells require extra signals for survival and proliferation that are usually derived from type B

mutations in T-ALL, such as the IL7R or RAS mutations found in ETP-ALL. Importantly, if we understand the oncogenic mechanisms of type A mutations, we could alleviate the developmental block of leukemic cells by targeting critical downstream molecules that are responsible for differentiation arrest. This may not only impair the acquisition of additional genetic hits, but also prevent the uncontrolled proliferation of cells at that stage.

Moreover, it would be interesting to further investigate the genes that are contained in the T-ALL gene signature (435 probesets), especially after filtering out all genes that are associated with T-cell development. Studying the resultant genes could provide new insights into important oncogenic mechanisms that are specific for each T-ALL subgroup, such as the activation of critical proliferation or survival signaling molecules, e.g. MYC and TP53. This would indicate that there are alternative pathways apart from pre- or post-commitment networks that are prone to become mutated in each T-ALL subtype to promote proliferation and/or survival mechanisms. Overall, this could give a better rationale for the use of specific inhibitors to different T-ALL entities.

Current methodologies to unravel the leukemogenic mechanisms of ETP-ALL

Despite the important insights obtained on the leukemogenic potential of ETP-ALL oncogenes using the OP9-DL1 *in vitro* system, some questions remain to be investigated. First, we would like to examine the combinations of oncogene and type B aberrations that are essential to drive leukemogenesis, and investigate the downstream activation of oncogenic signaling mechanisms. Second, it is important to develop models to test the efficacy of specific pathway inhibitors in T-ALL.

We are currently using *in vitro* techniques in T-ALL cell lines to characterize the transcriptional complex that comprises MEF2C, to investigate molecular interactions between MEF2C and other proteins that participate in an oncogenic transcription complex (e.g. LYL1 and LMO2). This may help to identify new therapeutic targets that could dismantle such a complex. Furthermore, it may lead to a better understanding on how MEF2C drives particular genes at the ETP-stage, which can be additional targets for therapy. In addition, we are working with *in vivo* mouse models that contain conditional deletion (KO) and overexpression (KI) of MEF2C to understand the role of this oncogene in normal and malignant *in vivo* T-cell development, respectively. Gene expression analysis of leukemic or pre-leukemic (potentially arrested) thymocytes (KI) will determine the effects of enforced expression of MEF2C transcriptional complex components or critical downstream targets on T-cell proliferation and differentiation.

Thus, MEF2C-driven leukemias in these mice will form a relevant model to investigate and characterize oncogenic MEF2C transcriptional complexes and its downstream target genes. It can also represent an important *in vivo* model to test the effects of compounds against relevant MEF2C-driven oncogenic pathways.

In addition, we can further explore the OP9-DL1 *in vitro* to test the expression of ETP-ALL oncogenes in combination with frequent ETP-ALL type-B aberrations, such as IL7R or RAS activating mutations. Although, we have shown that MEF2C-, LYL1- or LMO2- overexpression at the very early stages of T-cell development was capable of promoting a differentiation arrest, it would be interesting to transduce arrested cells with other collaborative events such as IL7R or RAS activating mutations to promote cell proliferation and survival. This system could provide new insights into the downstream molecular pathways that are activated to promote leukemogenesis and it may be a useful tool for testing drugs that target (pre-)leukemic cells. Alternatively, cells expressing the combination of oncogenes and type-B mutations could be xenotransplanted into sublethally irradiated mice to further examine the oncogenic potential of different mutations and the resulting malignancies at a cellular and molecular level.

Future research in T-ALL

Despite a detailed understanding of the genetic defects that drive T-ALL, we face a number of interesting challenges. First, we need to better understand how specific T-ALL oncogenes and tumor suppressors cooperate to drive full leukemic transformation. We still lack detailed information whether the co-occurrence of certain mutations may be synergistic in the activation of molecular oncogenic mechanisms that transform normal developing T-cells into (pre-)leukemic blasts. Secondly, the potential contribution of the exact order in which genetic lesions are acquired during disease initiation, progression, and/or maintenance, which underlie disease subclonality and evolutional selection before and after therapy, should also be further investigated. Moreover, the existence of leukemic initiating cells (LICs) in the context of T-ALL is still under debate, and the molecular mechanism that could regulate LIC activity remains poorly characterized. Furthermore, deregulated expression of microRNAs, long non-coding RNAs, enhancer activities, aberrant chromatin remodeling, and/or epigenetic changes in the context of malignant T-cell transformation need to be further investigated.

References

1. Ding L, Ley TJ, Larson DE, et al. Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. *Nature*. 2012;481(7382):506-510.
2. Tzoneva G, Perez-Garcia A, Carpenter Z, et al. Activating mutations in the NT5C2 nucleotidase gene drive chemotherapy resistance in relapsed ALL. *Nat Med*. 2013;19(3):368-371.
3. Meyer JA, Wang J, Hogan LE, et al. Relapse-specific mutations in NT5C2 in childhood acute lymphoblastic leukemia. *Nat Genet*. 2013;45(3):290-294.
4. Pieters R, Huismans DR, Loonen AH, et al. Relation of 5'-nucleotidase and phosphatase activities with immunophenotype, drug resistance and clinical prognosis in childhood leukemia. *Leuk Res*. 1992;16(9):873-880.
5. Pieters R, den Boer ML, Durian M, et al. Relation between age, immunophenotype and in vitro drug resistance in 395 children with acute lymphoblastic leukemia--implications for treatment of infants. *Leukemia*. 1998;12(9):1344-1348.
6. King B, Trimarchi T, Reavie L, et al. The ubiquitin ligase FBXW7 modulates leukemia-initiating cell activity by regulating MYC stability. *Cell*. 2013;153(7):1552-1566.
7. Roderick JE, Tesell J, Shultz LD, et al. c-Myc inhibition prevents leukemia initiation in mice and impairs the growth of relapsed and induction failure pediatric T-ALL cells. *Blood*. 2014;123(7):1040-1050.
8. Schubbert S, Cardenas A, Chen H, et al. Targeting the MYC and PI3K Pathways Eliminates Leukemia-Initiating Cells in T-cell Acute Lymphoblastic Leukemia. *Cancer Res*. 2014;74(23):7048-7059.
9. Peirs S, Van der Meulen J, Van de Walle I, et al. Epigenetics in T-cell acute lymphoblastic leukemia. *Immunol Rev*. 2015;263(1):50-67.
10. Knoechel B, Roderick JE, Williamson KE, et al. An epigenetic mechanism of resistance to targeted therapy in T cell acute lymphoblastic leukemia. *Nat Genet*. 2014;46(4):364-370.
11. Marculescu R, Le T, Simon P, Jaeger U, Nadel B. V(D)J-mediated translocations in lymphoid neoplasms: a functional assessment of genomic instability by cryptic sites. *J Exp Med*. 2002;195(1):85-98.
12. Marculescu R, Vanura K, Le T, Simon P, Jager U, Nadel B. Distinct t(7;9)(q34;q32) breakpoints in healthy individuals and individuals with T-ALL. *Nat Genet*. 2003;33(3):342-344.
13. Greaves M. Darwinian medicine: a case for cancer. *Nat Rev Cancer*. 2007;7(3):213-221.
14. Gale KB, Ford AM, Repp R, et al. Backtracking leukemia to birth: identification of clonotypic gene fusion sequences in neonatal blood spots. *Proc Natl Acad Sci U S A*. 1997;94(25):13950-13954.
15. Wiemels JL, Cazzaniga G, Daniotti M, et al. Prenatal origin of acute lymphoblastic leukaemia in children. *Lancet*. 1999;354(9189):1499-1503.
16. Dadi S, Le Noir S, Payet-Bornet D, et al. TLX homeodomain oncogenes mediate T cell maturation arrest in T-ALL via interaction with ETS1 and suppression of TCRalpha gene expression. *Cancer Cell*. 2012;21(4):563-576.
17. Marquez C, Trigueros C, Fernandez E, Toribio ML. The development of T and non-T cell lineages from CD34+ human thymic precursors can be traced by the differential expression of CD44. *J Exp Med*. 1995;181(2):475-483.
18. Homminga I, Pieters R, Langerak AW, et al. Integrated transcript and genome analyses reveal NKX2-1 and MEF2C as potential oncogenes in T cell acute lymphoblastic leukemia. *Cancer Cell*. 2011;19(4):484-497.

Chapter 8

Summary / Samenvatting

Supplementary Data

Curriculum Vitae

List of publications

Portfolio

Acknowledgements

Summary

This thesis aimed to: 1) identify additional PTEN-inactivating events as novel type-B mutations that may provide survival and proliferation advantage to (pre)-leukemic cells, and understand their clinical impact in T-ALL; 2) to establish an *in vitro* co-culture system (OP9-DL1) that supports T-cell development from human umbilical cord blood-derived hematopoietic stem cells (HSCs), with the intention of defining distinct human T-cell development stages in relation to the underlying transcriptional programs and expression of cell surface markers; 3) to optimize a lentivirus system that supports transduction and stable expression of oncogenes in human and mouse HSCs and 4) to study the impact of ETP-ALL expressed oncogenes (such as MEF2C, LYL1 or LMO2) on human T-cell development using the OP9-DL1 *in vitro* system.

Chapter 2. The tumor suppressor PTEN negatively regulates PI3K-AKT signaling and is often inactivated by mutations (including deletions) in a variety of cancer types, including T-ALL. In chapter 2, we review mutation-associated mechanisms that inactivate PTEN together with other molecular mechanisms that activate AKT and contribute to T-cell leukemogenesis. In addition, we discuss how *Pten* mutations in mouse models affect the efficacy of gamma-secretase inhibitors (GSI) to block NOTCH1 signaling through activation of AKT. Based on these models and on observations in primary diagnostic T-ALL patient samples, we propose that PTEN-deficient cells employ a compensatory mechanism in which PI3K-AKT signaling is dampened over time. As a result of this reduced PI3K-AKT signaling, the level of AKT activation is insufficient to compensate for NOTCH1 inhibition, resulting in responsiveness to GSI. On the other hand, de novo acquired PTEN inactivating events in a NOTCH1-dependent T-ALL result in immediate, strong activation of PI3K-AKT signaling, increased glycolysis and glutaminolysis, and consequently GSI resistance. Due to the central role of PTEN-AKT signaling in T-ALL and in resistance to NOTCH1 inhibition, AKT inhibitors may be a promising addition to current treatment protocols for T-ALL.

Chapter 3. Phosphatase and tensin homolog (PTEN)-inactivating mutations and/or deletions are an independent risk factor for relapse of T-cell acute lymphoblastic leukemia (T-ALL) patients treated on Dutch Childhood Oncology Group or German Cooperative Study Group for Childhood Acute Lymphoblastic Leukemia protocols. Some monoallelic mutated or PTEN wild-type patients lack PTEN protein, implying that additional PTEN inactivation mechanisms exist. We show that

PTEN is inactivated by small deletions affecting only a few exons in 8% of pediatric T-ALL patients. These so-called microdeletions were clonal in 3% and subclonal in 5% of patients. Conserved deletion breakpoints were identified that are flanked by cryptic recombination signal sequences (cRSSs) and frequently have non-template-derived nucleotides inserted in between breakpoints, pointing to an illegitimate RAG recombination activity. Identified cRSSs drive RAG-dependent recombination in a reporter system as efficient as bona fide RSSs elements that flank T-cell receptor locus gene segments. Microdeletions strongly associate with the TALLMO subtype that is characterized by TAL1 or LMO2 rearrangements, in line with the previous association of PTEN non-sense mutations or entire locus deletions. Primary and secondary xenotransplantation of TAL1- rearranged T-ALL allowed development of leukemic subclones that acquired new PTEN microdeletions, indicating that PTEN aberrations may be continuously acquired and a mechanism for subclonal diversity. Remarkably, equivalent microdeletions were also detected in thymocytes of healthy individuals indicating that ongoing RAG activity may actively contribute to the acquisition of preleukemic hits in normal thymocytes that will be effectively eradicated by the immune system.

Chapter 4. Leukemia research increasingly depends on techniques to induce or to inhibit expression of oncogenes or tumor-suppressors in HSCs or primary leukemic cells. Surprisingly little information is available to optimize lentiviral transduction of primary leukemic cells or HSCs. We have therefore carefully optimized transduction of murine and human HSCs using an optimized lentiviral vector design, serum-free virus production of recombinant viruses and quantitation, and optimized delivery (transduction) into primary human or mouse HSCs. We developed a quantitative RT-PCR-method to quantify the encapsulated proviral RNA particles. We have observed an inverse exponential relation between the length of the proviral RNA encoded by the construct and the number of viral particles that are being produced in HEK293T-cells. The length of the proviral RNA also determines the transduction efficiency of target cells. Increased concentrations of serum (ranging from 0% to 10%) during transduction negatively affect the transduction rate of human HSCs. In conclusion, we provide an optimized serum-free lentiviral gene transfer into human and murine hematopoietic stem cells, where the reduction of lentiviral constructs size will facilitate efficient transfer of large oncogene sequences into human or mouse HSCs or primary leukemic cells.

Chapter 5. Human T cell development is less well studied than its murine counterpart due to the lack of genetic tools and the difficulty of obtaining cells and tissues. However, recent

technological advances allow identification of the transcriptional landscape of differentiating human thymocytes. Here we report the gene expression profiles of 11 immature, consecutive T-cell developmental stages that were obtained using the OP9-DL1 *in vitro* culture system. The changes in gene expression of cultured stem cells on OP9-DL1 match those of ex vivo isolated human thymi. These analyses let us to define a gene signature that distinguishes pre- and post- $\alpha\beta$ T-cell commitment. We found that loss of CD44 marks T-cell commitment in early CD7+CD5+ cells, before the acquisition of CD1a surface expression. The CD44-CD1a- post-committed thymocytes have initiated in frame TCR rearrangements and have completely lost the capacity to develop into myeloid, B- and NK-cells, unlike uncommitted CD44+CD1a- thymocytes. Therefore, loss of CD44 represents a previously unrecognized stage that defines the earliest committed T-cell population in the human thymus.

Chapter 6. Pediatric T-ALL is a cancer of developing T-cells in the thymus. The T-ALL subtypes are characterized by specific chromosomal rearrangements driving the ectopic expression of oncogenes that result in arrest of leukemic cells at specific T-cell developmental stages. Here, we show that T-ALL subtypes are strongly associated with the expression of particular T-cell developmental programs. Immature/ETP-ALL and TLX subtypes of T-ALL express a pre-commitment program whereas the proliferative and TALLMO subtypes express a post-commitment program. This suggests that T-ALL oncogenes require the expression of particular T-cell transcriptional programs to be able to exert an oncogenic function, and therefore may hijack T-cell developmental programs by facilitating a developmental arrest. That is further supported by functional *in vitro* studies where we investigate the impact of typical ETP-ALL (proto-)oncogenes, including MEF2C, LYL1 or LMO2 on early T-cell development. Umbilical cord blood stem cells transduced with lentiviral expression constructs that constitutively drive MEF2C, LYL1 or LMO2 were cultured on OP9-DL1 cells for differentiation towards the T-cell lineage. Expression of MEF2C or its putative co-factors (LYL1 or LMO2) blocks T-cell development at ETP-stages, in line with their roles in immature T-ALL. Contrarily, expression of the oncogenes in T-cells beyond the ETP-stages does no longer support T-cell development arrest, indicating that their leukemogenic potential is dependent on ETP, pre-commitment transcriptional networks. Thus, we conclude that the specific interaction of an aberrant oncogenic transcription factor with specific transcriptional network in a developing T-cell is an important requirement for oncogenesis.

Chapter 7. The findings described in the thesis are discussed and put into perspective.

Nederlandse samenvatting

Dit proefschrift heeft de volgende doelstellingen gehad: 1) Het identificeren van bepaalde afwijkingen in het *PTEN* gen, in toevoeging van de al bekende genetische afwijkingen, die als nieuwe type-B mutaties een overlevings- of groeivoordeel zouden kunnen bieden voor (pre-) leukemische cellen. Ook het begrijpen en beschrijven van de klinische impact van *PTEN* afwijkingen in T-cel acute leukemie (T-ALL) patiënten behoorde tot deze doelstelling. 2) Het opzetten van een kweekstelsel (OP9-DL1) waarin bloedstamcellen verkregen uit navelstrengbloed in een petrischaal gedurende enkele weken kunnen uitgroeien tot T-cellen. Het bepalen van genexpressieprofielen van de verschillende T-cel ontwikkelingsstadia aan de hand van specifieke oppervlakte markers heeft het functioneren van het stelsel gevalideerd en ook geleid tot de identificatie van een extra oppervlakte marker. 3) De optimalisatie van een methode waarmee met behulp van lenti-virussen bepaalde oncogenen ingebracht en geactiveerd kunnen worden in bloedstamcellen van menselijke en muizen oorsprong. 4) Het bestuderen van het effect van de activatie van oncogenen die specifiek voorkomen in een vroege vorm van T-ALL (genaamd ETP-ALL) op de menselijke T-cel ontwikkeling in het OP9-DL1 kweekstelsel.

Hoofdstuk 2. De functie van het zogenaamde ‘tumor suppressor gen’ *PTEN* is het remmen van een belangrijk signalerings-pad PI3K-AKT dat na activatie de overleving en groei van cellen stimuleert. In T-ALL, net als in verscheidene andere kanker soorten, komen afwijkingen voor die leiden tot het inactief of afwezig zijn van *PTEN*, en daardoor tot verminderde remming van het PI3K-AKT pad. In hoofdstuk 2 hebben we de literatuur samengevat rondom *PTEN*, PI3K-AKT en een ander belangrijk signalerings-pad dat begint bij de NOTCH1 receptor. We behandelen hoe bepaalde mutaties en andere afwijkingen in deze verschillende paden samen kunnen leiden tot de ontwikkeling van T-ALL, en hoe verschillende remmers aangrijpen op deze mechanismes en dus mogelijk klinisch relevant zijn. We concluderen dat AKT remmers veelbelovend kunnen zijn als toevoeging aan de huidige behandelingsprotocollen voor T-ALL.

Hoofdstuk 3. T-ALL patiënten met *PTEN* afwijkingen die in Nederland en Duitsland zijn behandeld hebben een verhoogd risico op terugval. Hoewel veel van deze afwijkingen verklaard kunnen worden door genetische mutaties of deleties van het *PTEN* gen die resulteren in het

niet-functioneren of afwezig zijn van het PTEN eiwit, zijn er ook cellen van patiënten zonder een dergelijke verklaring voor de afwezigheid van PTEN. We hebben in deze cellen hele kleine, gedeeltelijke deleties in het *PTEN* gen gevonden die ook leiden tot de afwezigheid van PTEN. Deze zogenaamde ‘micro-deleties’ werden met de eerdere technieken over het hoofd gezien, en komen voor in 8% van kinderen met T-ALL. We hebben laten zien dat deze micro-deleties worden veroorzaakt door een herschikkings-proces wat normaliter de herschikking van T- en B-cel receptor genen verzorgt en dus een normaal actief proces is in rijpende T-cellen. Opmerkelijk genoeg hebben we deze *PTEN*-microdeleties ook gevonden in sommige rijpende T-cellen van enkele gezonde kinderen, wat aangeeft dat er meerdere mutaties moeten optreden voordat deze microdeleties een risico kunnen worden.

Hoofdstuk 4. Ten behoeve van ons T-ALL onderzoek naar het effect van de activatie van bepaalde oncogenen introduceren (‘transductie’) we grote DNA moleculen in normale bloedstamcellen met behulp van specifieke lenti-virussen. Omdat deze techniek uiterst inefficiënt is in stamcellen, hebben we het door meerdere aanpassingen verbeterd. We hebben de volgende facetten geoptimaliseerd: ontwerp lenti-virale DNA vector, virus productie, nauwkeurige bepaling hoeveelheid virus, transductie in bloedstamcellen. We hebben gezien dat met toenemende lengte van de vector zowel de virus productie alsook de transductie efficiëntie afneemt. Om toch grote oncogenen te kunnen bestuderen in bloedstamcellen, zullen andere elementen uit de DNA vector verkleind of verwijderd moeten worden om een betere transductie efficiëntie te realiseren.

Hoofdstuk 5. T-cel ontwikkeling/rijping in mensen vindt plaats in de thymus, voornamelijk gedurende de eerste 10 levensjaren. Vergeleken met het onderzoek naar T-cel ontwikkeling in de muis is de menselijke T-cel ontwikkeling weinig bestudeerd, gedeeltelijk omdat primair materiaal van kinderen weinig beschikbaar is. Met de opkomst van geavanceerde kweeksystemen (OP9-DL1) kan ook de menselijke T-cel ontwikkeling nagebootst en bestudeerd worden in een petrischaal. We hebben dit systeem opgezet en gebruikt om genexpressieprofielen van meerdere opeenvolgende menselijke T-cel ontwikkelingsstadia (aan de hand van oppervlakte markers) te bepalen. Uit de analyse van deze profielen blijkt dat ook op het niveau van de genexpressie het kweekstelsel de werkelijke ontwikkeling goed weergeeft. Bovendien heeft de analyse geleid tot een extra oppervlakte marker CD44 die onze genexpressieprofielen van voor en na een belangrijk checkpoint in T-cel ontwikkeling, de zogenaamde ‘T-cel commitment’, scheidt. We hebben getest dat de vroegste cellen in de

thymus die nog CD44 hebben, ook nog kunnen uitrijpen naar andere bloedceltypes; de iets latere 'gecommitteerde' T-cellen hebben geen CD44 en kunnen alleen nog T-cellen worden.

Hoofdstuk 6. T-ALL in kinderen kan, gebaseerd op genexpressieprofielen, grofweg in vier subgroepen worden verdeeld: de meest vroege ETP-ALL, TLX groep, proliferatieve groep, en de TALLMO groep. De TLX en TALLMO groep zijn genoemd naar de belangrijkste oncogenen die de T-ALL veroorzaken: TLX3, TAL1 en LMO2. De proliferatieve groep wordt gekenmerkt door de oncogenen TLX1 en NKX2.1 terwijl de meest vroege ETP-ALL groep gekenmerkt wordt door MEF2C, LYL1 en ook LMO2. We hebben de genexpressieprofielen van voor en na normale T-cel commitment (**hoofdstuk 5**) gebruikt om de T-ALL patiënten opnieuw te klusteren en vonden grofweg dezelfde vier groepen terug. De 'vroege' ETP-ALL en TLX groepen overlappen met het profiel 'voor-commitment', terwijl de 'latere' proliferatieve en TALLMO groepen overlappen met het 'na-commitment' profiel. Dit suggereert dat de T-ALL oncogenen functioneren in de context van normale T-cel ontwikkelingsgenen die ook actief zijn, en dat T-ALL oncogenen het normale ontwikkelingsprogramma blokkeren in specifieke stadia. Gebruikmakend van de expertise ontwikkeld in **hoofdstukken 4 en 6**, hebben we dit nagebootst voor ETP-ALL, waarin de oncogenen MEF2C, LYL1 en LMO2 actief zijn. We hebben deze oncogenen ingebracht in normale bloedstamcellen middels virale transductie, en vervolgens de ontwikkeling gevolgd in het OP9-DL1 kweekstelsel. Deze oncogenen leiden alle drie tot een blokkade vroeg in de ontwikkeling, terwijl ze geen effect hebben als ze pas in latere (na 'commitment') cellen aangezet worden. Deze observatie ondersteunt de hypothese dat de normale T-cel ontwikkeling alleen geblokkeerd wordt door specifieke oncogenen in dát specifieke stadium waarin het normale ontwikkelingsprogramma actief is en het de functie van het oncogen en de leukemie ontwikkeling ondersteunt.

Supplementary Data

Chapter 3

Supplementary Materials and Methods

Multiplex ligation-dependent probe amplification (MLPA) reaction

For MLPA analysis, the SALSA MLPA kit (MRC-Holland, Amsterdam, the Netherlands) was used in combination with the P200-A1 Human DNA reference-1 probemix kit (MRC-Holland). Two synthetic left and right hybridization probe oligonucleotide (LPO and RPO) pairs were designed per exon for all 9 exons of *PTEN*. Probe pairs were designed according to the manufacturer guidelines, and obtained from IDT (Coralville, USA). Probe pairs were combined into a probe-mix with a concentration of 4nM per oligonucleotide (Table S1). Reactions were carried out in a model 2320 thermocycler (Applied Biosystems, Bleiswijk, the Netherlands). MLPA fragment analysis was performed using GeneMarker V1.85 (Softgenetics, State College, PA, USA).

Breakpoint Mapping and PCR-based screening

To map breakpoints for *PTEN* exons 2-3 and 4-5 deletions, multiple forward or reverse primers were designed for *PTEN* introns 1, 3 and 5 (Table S2 and S3). Specific primers (Table S4) were selected for *PTEN* exon 2-3 and 4-5 breakpoint screening. Positive reactions were directly sequenced for both strands or following cloning into the TOPO TA vector (Life Technologies). Sequence reactions were run on the 3130x capillary sequencer (Applied Biosystems) and analyzed using CLC Main Workbench software (Aarhus, Denmark).

Computational detection of putative RAG recombination signal sequences

The human *PTEN* gene (ENSG00000171862) was screened for the presence of cryptic RAG recombination signal sequences (cRSS) using the PERL software algorithms developed by Cowell *et al.*²⁸ Mouse 12- and 23-RSSs (n=356) from all TCR and Ig loci were used for modeling. Both the program and 12-/23-RSSs were available at <http://www.duke.edu/~lqcowell/>. The program computes RSS information content (RIC) scores for 12-spacer RSSs (i.e. 28-bp sequences comprising the heptamer, a 12-bp spacer and the nonamer sequences: RIC12) or 23-spacer RSSs (i.e. 39-bp sequences comprising the heptamer, a 23-bp spacer and the nonamer sequences: RIC23) to predict cryptic RSSs (cRSS) based on a comparison to the RIC score for actual immunoglobulin and T-cell receptor RSS. RSSs that potentially function within the range of actual antigen receptor loci RSSs are predicted to score above the -38,81 as RIC12 and -58,45 as RIC23 thresholds. RSSs that potentially function within the range predicted for degenerated RSSs, i.e. the so-called cryptic RSSs (cRSS), score below these thresholds but above the background (e.g. the mean of RIC scores determined for non-RSS DNA sequences: -60,07 for RIC12 and -77,76 for RIC23) as previously described.¹

Generation of GFPi-*PTEN* cRSS reporter constructs and recombination assay

To measure efficiencies of predicted cRSSs in mediating recombination of GFPi-mRFP reporter constructs, PCR amplified *PTEN* cRSS1-4 or defined RSS control sequences were cloned into this reporter construct and recombination assays were carried out as described,² with minor adaptations (see Supplementary Materials and Methods). Predicted *PTEN* cRSS1-4 or control RSSs were tested in combination with a consensus 12 or 23 RSS (Figure 3B). All RSS or cRSS used in the GFPi-mRFP constructs are summarized in Table S5.

Briefly, HEK293T cells grown in standard culture conditions were seeded at 2×10^5 cells per well in 12-well plates. After 24h, cells were transfected with a transfection mix that contains 5 μ L of Lipofectamine 2000 (Invitrogen, Carlsbad, CA, USA), 300 μ L of Optimem (Gibco, Life Technologies, Paisley, UK), 0.8 μ g of CMV-RAG1,² 0.7 μ g of CMV-RAG2,² and 2.5 μ g of a GFPi-mRFP construct variant. Medium was replaced 16h after transfection. Cells were harvested at 48h following transfection, and analyzed by flow

cytometry for GFP and RFP fluorescence using a FACSAria III (BD Bioscience, Madrid, Spain). Data analysis was performed using the FlowJo software (TreeStar Inc., Ashland, OR).

Xenotransplantation of primary T-ALL patient material

NOD-SCID-IL2R γ ^{null} (NSG) mice were bred and housed under specific pathogen free (SPF) conditions at the Experimental Animal Facility of ErasmusMC. All experiments have been approved by the committee on animal welfare of ErasmusMC and are in compliance with Dutch legislation. Sets of 6–10 week old mice were intravenously injected with $1-5 \times 10^6$ human bone marrow leukemic cells. Upon signs of illness, mice were euthanized and leukemic cells were collected from the different organs. Likewise, primary transplanted leukemic cells from bone marrow, thymus or spleen were retransplanted.

References

1. Cowell LG, Davila M, Kepler TB, Kelsoe G. Identification and utilization of arbitrary correlations in models of recombination signal sequences. *Genome Biol.* 2002;3:RESEARCH0072.
2. Trancoso I, Bonnet M, Gardner R, et al. A Novel Quantitative Fluorescent Reporter Assay for RAG Targets and RAG Activity. *Front Immunol.* 2013;4:110.

Supplementary Tables

[illegible]

Table S2. Primers used for *PTEN* exon 2-3 breakpoint mapping

| Primer sequence Forward/Reverse | Primers position (GRCh37) |
|----------------------------------|------------------------------|
| FW: 5'-CGAGGGGCATCGACTA-3' | Chr:10 89.624.110-89.624.125 |
| FW: 5'-TCTGGGAGACAGATTTCTTTC-3' | Chr:10 89.626.319-89.626.339 |
| FW: 5'-TTGGGTAAACAAACAGTTCAG-3' | Chr:10 89.628.198-89.628.219 |
| FW: 5'-TTGGGCCATGTTAGGATT-3' | Chr:10 89.630.012-89.630.029 |
| FW: 5'-TTTGGGGATGTCATTAAGC-3' | Chr:10 89.631.816-89.631.834 |
| FW: 5'-TTGGCTGATTGTGAAAGTGT-3' | Chr:10 89.633.635-89.633.654 |
| FW: 5'-TTCCCAATGAGGTAAAATTG-3' | Chr:10 89.635.523-89.635.543 |
| FW: 5'-CCAAGCAATGCAGACTTACT-3' | Chr:10 89.637.291-89.637.310 |
| FW: 5'-AGGCTTGAGCAGTTGTAGATAG-3' | Chr:10 89.639.040-89.639.061 |
| FW: 5'-GGCTGCATTAATTGATCTGTA-3' | Chr:10 89.640.922-89.640.943 |
| FW: 5'-CCAGGCAGCTTCAGACTT-3' | Chr:10 89.642.530-89.642.547 |
| FW: 5'-GTTCCCAAATCTAAACAGAGAC-3' | Chr:10 89.644.400-89.644.422 |
| FW: 5'-GCCAGCTGAATTTACTTTAATG-3' | Chr:10 89.646.287-89.646.308 |
| FW: 5'-CCCCCTTTTCTCTCATATGT-3' | Chr:10 89.648.219-89.648.239 |
| FW: 5'-CTGGCTTGTGAAGAGAGAAT-3' | Chr:10 89.649.821-89.649.841 |
| FW: 5'-CTGGGCAACATGATGAAA-3' | Chr:10 89.651.378-89.651.395 |
| FW: 5'-GGGCATTGATTAGGTCA-3' | Chr:10 89.652.427-89.652.444 |
| FW: 5'-TTGGGGTTTGTGTTTGATTT-3' | Chr:10 89.653.486-89.653.505 |
| RV: 5'-TGGGCAGTCCTTTAAACACTC-3' | Chr:10 89.685.808-89.685.789 |
| RV: 5'-CCCCATCTGTAATGAGAAG-3' | Chr:10 89.687.310-89.687.290 |
| RV: 5'-TCCCTACTTTCTCAGGATC-3' | Chr:10 89.689.092-89.689.073 |
| RV: 5'-GGCGGTGTCATATGTCTT-3' | Chr:10 89.690.830-89.690.812 |

Table S3. Primers used for *PTEN* exon 4_5 breakpoint mapping

| Primer sequence Forward/Reverse | Primers position (GRCh37) |
|------------------------------------|------------------------------|
| FW: 5'-CAAAGCATAAAAACCAATTACAAG-3' | Chr:10 89.685.281-89.685.303 |
| FW: 5'-TGCCACAGATTCATGTTACTT-3' | Chr:10 89.685.510-89.685.530 |
| FW: 5'-AGCCGACCATCCACTG-3' | Chr:10 89.687.368-89.687.383 |
| FW: 5'-AATGCTGGCTTTGACTGAA-3' | Chr:10 89.688.668-89.688.686 |
| FW: 5'-GGAAGCACCTGAATTTACAGTA-3' | Chr:10 89.690.506-89.690.527 |
| RV: 5'-GGAGGGAGGAACACAAAGAT-3' | Chr:10 89.693.370-89.693.352 |
| RV: 5'-GAAGGCCAAAGAAAATGATGAA-3' | Chr:10 89.695.389-89.695.369 |
| RV: 5'-CAGGGATAAAGAAGGTAAGAACA-3' | Chr:10 89.697.244-89.697.222 |
| RV: 5'-CGCTTGAACCTCAGGAGGTA-3' | Chr:10 89.699.184-89.699.166 |
| RV: 5'-TCCGACTGCACGTAATTATC-3' | Chr:10 89.701.201-89.701.182 |
| RV: 5'-CTGGATGATTTCACTACTGTG-3' | Chr:10 89.703.130-89.703.110 |
| RV: 5'-CCCCTCTGTCAAACCAAT-3' | Chr:10 89.705.100-89.705.083 |
| RV: 5'-AGGGCTTCTGGTCAACAG-3' | Chr:10 89.706.988-89.706.971 |
| RV: 5'-TTGGGAGGCTCACATGTA-3' | Chr:10 89.708.947-89.708.930 |
| RV: 5'-GGCGCTGGTGATGTGTA-3' | Chr:10 89.710.642-89.710.626 |
| RV: 5'-GCTGCCACGCTTATCACT-3' | Chr:10 89.712.480-89.712.462 |

Table S4. Specific primers to screen for *PTEN* exon 2_3 and 4_5 breakpoints.

| Screening breakpoint <i>PTEN</i> deletion (DNA/cDNA) | Primer sequence Forward/Reverse | Primers position (GRCh37) |
|--|--|--|
| Exon 2_3: DNA | FW: 5'-GCTGCTCCTTTACCTTTC-3' or FW: 5'-GGCTGCTCCTTTACCTT-3' | Chr:10 89.624.321-89.624.340 Chr:10 89.624.320-89.624.338 |
| | RV: 5'-ATCCCAAAAGTTCAAACATC-3' or RV: 5'-AGCAAAATCCTGATCTGAAGTAAT-3' | Chr:10 89.689.878-89.689.858 Chr:10 89.689.797-89.689.819 |
| | FW: 5'-CGAGGGGCATCAGCTA-3' or FW: 5'-CCGTTTTTAGGTTTCAGGTC-3' | Chr:10 89.624.110-89.624.125 Chr:10 89.625.248-89.625.267 |
| | RV: 5'-GGCGGTGTCATAATGTCTT-3' or RV: 5'-AGCAAAATCCTGATCTGAAGTAAT-3' | Chr:10 89.690.830-89.690.812 Chr:10 89.689.797-89.689.819 |
| Exon 2_3: cDNA | | |
| Exon 4_5: DNA | FW: 5'-GAAACCAATGGATTGTAGTTTATT-3' or RV: 5'-TCCGACTGCACGTAATTATC-3' | Chr:10 89.689.486-89.689.509 Chr:10 89.701.201-89.701.182 |
| | | |
| Exon 4-5: cDNA | FW: 5'-AAGCATAAAAACCAATTACAAGAT-3' or RV: 5'-CAGTGCCACTGGTCTATAATC-3' | Chr:10 89.685.283-89.685.305 Chr:10 89.711.961-89.711.941 |

Table S5. RSSs tested with the GFPi reporter construct

| 12-spacer RSS | heptamer | 12-spacer | nonamer | RIC* |
|------------------------------|----------|-------------------------|-----------|--------|
| Consensus 12-sp ² | CACAGTG | CTACAGACTGGA | ACAAAAACC | |
| Human <i>PTEN</i> cRSS1 | CACAGAT | AAATATACTTT | ACATAAACA | -34.2 |
| Human <i>PTEN</i> exon 1 | CAGAGAG | ATCGTAGCAGA | AACAAAAGG | -45.48 |
| Human SCL12 ^{2,3} | CACAGCC | TCGCGATTCT | GTAATITGC | |
| Human LMO2 ^{2,4} | CACAGTA | TTGTCTACCCA | GCATAATT | |
| Mouse Jβ2-2 ^{2,5} | CACAGTC | GTCGAAATGCTG | GCACAACC | |
| Mouse VH87 IgH ⁶ | CACATAT | AGGATCAATCCT | ICAAATCCA | |
| 23-spacer RSS | heptamer | 23-spacer | nonamer | RIC* |
| Consensus 23-sp ² | CACAGTG | CTACAGCTCCACTGTCTACTGGA | ACAAAAACC | |
| Human <i>PTEN</i> cRSS2 | CACAGTA | TTTCACTTCTATGAACTAATTA | TTGAGAACA | -55.58 |
| Human <i>PTEN</i> cRSS3 | CACCTTA | GGTTGAATACACAGAAAGAAAC | ACAAAATAT | -59.78 |
| Human <i>PTEN</i> cRSS4 | CACIGTA | TGAAAAAGCTAACATACCTACA | ATCTAAAGC | -75.59 |
| Human SCL 23 ⁷ | CACAGCC | TCGCGATTCTGTATATTGCGT | AAGGAAAAG | |

*RIC scores are only stated for newly identified cRSS in *PTEN*.

References:

- Cowell LG, Davila M, Kepler TB, Kelsos G. Identification and utilization of arbitrary correlations in models of recombination signal sequences. *Genome Biol.* 2002;3:RESEARCH0072.
- Trancoso I, Bonnet M, Gardner R, et al. A Novel Quantitative Fluorescent Reporter Assay for RAG Targets and RAG Activity. *Front Immunol.* 2013;4:110.
- Raghavan SC, Kirsch IR, Lieber MR. Analysis of the V(D)J Recombination Efficiency at Lymphoid Chromosomal Translocation Breakpoints. *Journal of Biological Chemistry.* 2001;276:29126-29133.
- Cowell LG, Davila M, Yang K, Kepler TB, Kelsos G. Prospective estimation of recombination signal efficiency and identification of functional cryptic signals in the genome by statistical modeling. *J Exp Med.* 2003;197:207-220.
- Davila M, Liu F, Cowell LG, et al. Multiple, conserved cryptic recombination signals in VH gene segments: detection of cleavage products only in pro B cells. *J Exp Med.* 2007;204:3195-3208.

Table S6. Overall clinical and molecular cytogenetics of *PTEN*-mutated patients versus wild-type patients.

| Clinical (n=142) | <i>PTEN</i> mutation/deletion | | <i>p</i> -value |
|---------------------------------------|-------------------------------|------------------------|-----------------------------------|
| | WT | Mut | |
| Gender | | | 0.91 |
| Male (n=97) | 79 | 18 | |
| Female (n=45) | 37 | 8 | |
| Median age (range) | 7.9 (1.3-17.8) | 4.7 (1.1-15.9) | 0.03* |
| Median WBC (range) | 119.2 (2.0-900.0) | 134 (5.0-600.0) | 0.14* |
| Cytogenetics (n=142) | WT n(%) | Mut n(%) | <i>p</i>-value[†] |
| TAL-related[‡] (n=29) | 18 (60%) | 11 (40%) | 0.002 |
| LMO-related[‡] (n=19) | 16 (84%) | 3 (16%) | 1# |
| TLX3+ (n=28) | 28 (100%) | 0 (0%) | 0.002# |
| TLX1+ (n=7) | 6 (86%) | 1 (14%) | 1# |
| HOXA+ (n=13) | 13 (100%) | 0 (0%) | 0.13# |
| Unknown (n=48) | 37 (77%) | 11 (23%) | 0.31 |
| Unsupervised clusters (n=113) | WT n(%) | Mut n(%) | <i>p</i>-value[†] |
| TAL/LMO+ (n=51) | 35 (69%) | 16 (31%) | 0.003# |
| Proliferative (n=19) | 16 (84%) | 3 (16%) | 1# |
| Immature (n=15) | 13 (87%) | 2 (13%) | 0.73# |
| TLX (n=28) | 28 (100%) | 0 (0%) | 0.002# |
| NOTCH1/FBXW7 status (n=141) | WT n(%) | Mut n(%) | <i>p</i>-value[†] |
| wild-type (n=51) | 34 (66%) | 17 (34%) | 0.001 |
| mutant (n=90) | 81 (90%) | 9 (10%) | |

[†], statistical analysis of the prevalence of *PTEN* aberrations within specific genetic T-ALL subgroups, as indicated, compared to other all other T-ALL subgroups combined. Abbreviations: Mut, mutant; WT, wild-type; WBC, white blood cell count. Significant *p*-values are indicated in bold. All *p*-values were calculated by using Pearson's Chi-square test, unless indicated: *Mann-Whitney U test, #Fisher's exact test. [‡]Two T-ALL cases have *TAL2* and *LMO1* or *TAL1* and *LMO2* rearrangements; The TAL/LMO group consists of *TAL1* and/or *LMO2* rearranged patients and patients lacking these aberrations but that have an identical or TAL/LMO-like gene expression profile.

Table S7. 5-years relapse free survival of patient groups based on clinical, (cyto)genetic or biological characteristics

| Clinical (n=146) | DCOG | | COALL | | Overall stratified analysis | |
|--|--------------------|-------------|-------------------|-------------|-----------------------------|--------------|
| | 5-ys RFS (% ± SD) | <i>P</i> | 5-ys RFS (% ± SD) | <i>P</i> | 5-ys RFS (% ± SD) | <i>P</i> |
| male (n=100) vs female (n=46) | 56 ± 7 vs 84 ± 8 | 0.19 | 61 ± 9 vs 91 ± 6 | 0.02 | 65 ± 5 vs 88 ± 5 | 0.009 |
| age ≥ 10 yrs (n=50) vs age < 10 yrs (n=96) | 82 ± 11 vs 74 ± 6 | 0.53 | 79 ± 9 vs 89 ± 8 | 0.51 | 72 ± 7 vs 72 ± 5 | 0.99 |
| WBC *10 ⁹ leukocytes/liter ≥ 50 (n=114) vs WBC < 50 (n=31) | 69 ± 6 vs 80 ± 13 | 0.4 | 68 ± 7 vs 84 ± 10 | 0.11 | 69 ± 5 vs 83 ± 8 | 0.06 |
| Driving oncogenic type A mutations/rearrangements (n=146) | | | | | | |
| <i>TAL1</i> , <i>TAL2</i> or <i>LYL1</i> (n=30) vs other (n=116) | 93 ± 6 vs 65 ± 6 | 0.04 | 66 ± 14 vs 74 ± 7 | 0.49 | 81 ± 8 vs 71 ± 5 | 0.25 |
| <i>LMO1</i> , <i>LMO2</i> or <i>LMO3</i> (n=19) vs other (n=127) | 50 ± 16 vs 75 ± 6 | 0.09 | 75 ± 22 vs 72 ± 6 | 0.67 | 59 ± 13 vs 74 ± 4 | 0.29 |
| <i>TLX3</i> (n=29) vs other (n=117) | 57 ± 12 vs 75 ± 6 | 0.11 | 54 ± 17 vs 76 ± 6 | 0.22 | 56 ± 10 vs 76 ± 4 | 0.04 |
| <i>TLX1</i> (n=8) vs other (n=138) | 80 ± 18 vs 70 ± 6 | 0.54 | 100 vs 72 ± 6 | 0.45 | 86 ± 13 vs 71 ± 4 | 0.36 |
| <i>SHOXA</i> -activated (n=13) vs other (n=133) | 40 ± 22 vs 74 ± 6 | 0.15 | 71 ± 17 vs 72 ± 6 | 0.82 | 58 ± 14 vs 74 ± 4 | 0.26 |
| <i>MEF2C</i> -activated (n=6) vs other (n=140) | 100 vs 71 ± 6 | 0.56 | 80 ± 18 vs 72 ± 6 | 0.80 | 83 ± 15 vs 72 ± 4 | 0.62 |
| <i>NKX2-1/NKX2-2</i> (n=7) vs other (n=139) | 100 vs 70 ± 6 | 0.40 | 100 vs 70 ± 7 | 0.19 | 100 vs 71 ± 4 | 0.12 |
| Unsupervised gene expression clusters (n=117) | | | | | | |
| TAL/LMO cluster (n=53) vs other (n=64) | 73 ± 10 vs 67 ± 10 | 0.70 | 69 ± 10 vs 77 ± 7 | 0.49 | 71 ± 7 vs 73 ± 6 | 0.81 |
| TLX cluster (n=30) vs other (n=87) | 60 ± 14 vs 74 ± 8 | 0.25 | 65 ± 13 vs 77 ± 6 | 0.45 | 63 ± 9 vs 76 ± 5 | 0.18 |
| Proliferative cluster (n=19) vs other (n=98) | 63 ± 15 vs 68 ± 6 | 0.39 | 82 ± 12 vs 73 ± 7 | 0.70 | 82 ± 10 vs 71 ± 5 | 0.39 |
| Immature cluster (n=15) vs other (n=102) | 50 ± 35 vs 70 ± 7 | 0.97 | 91 ± 9 vs 70 ± 7 | 0.18 | 85 ± 10 vs 71 ± 5 | 0.28 |
| Type B mutations | | | | | | |
| <i>PTEN/AKT</i> mutant/no <i>PTEN</i> protein (n=31) vs other (n=111) | 64 ± 15 vs 70 ± 6 | 0.54 | 57 ± 15 vs 76 ± 6 | 0.19 | 60 ± 10 vs 74 ± 4 | 0.17 |
| <i>NOTCH1/FBXW7</i> mutant (n=90) vs other (n=51) | 62 ± 6 vs 82 ± 8 | 0.10 | 68 ± 8 vs 80 ± 9 | 0.33 | 65 ± 5 vs 81 ± 6 | 0.06 |
| <i>PTEN/AKT/NOTCH1/FBXW7</i> mutant/no <i>PTEN</i> protein (n=109) vs other (n=32) | 63 ± 7 vs 92 ± 7 | 0.04 | 66 ± 7 vs 90 ± 10 | 0.06 | 65 ± 5 vs 92 ± 6 | 0.005 |
| WT1 mutant (n=17) vs other (n=129) | 63 ± 17 vs 72 ± 6 | 0.47 | 46 ± 22 vs 76 ± 6 | 0.16 | 57 ± 13 vs 74 ± 4 | 0.14 |
| Del 9p21 (n=116) vs other (n=25) | 70 ± 6 vs 75 ± 13 | 0.66 | 64 ± 6 vs 94 ± 6 | 0.03 | 68 ± 5 vs 86 ± 6 | 0.07 |

Significant log-rank *p*-values for DCOG or COALL cohort analyses are indicated in bold; RFS, Relapse free survival; SD, standard deviation; *P*, *p*-value; WBC, white blood cell count; [†]All patients having one of the specified cytogenetic aberrations are being compared to patients lacking those specific aberrations (indicated as "other"); [‡]Different genetic aberrations have been identified that activate *HOXA* or *MEF2C* oncogenes in specific genetic subtypes of T-ALL patients (Meijerink JP. Best Pract Res Clin Haematol, 2010; Homminga I. et al., Cancer Cell 2011); [‡]Unsupervised clusters as defined in Homminga I. et al., Cancer Cell 2011. The del9p21 status was determined by multiplex ligation-dependent probe amplification (MLPA) analysis for *p15/CDKN2B* and *p16/CDKN2A* loci as designed by MRC-Holland, Amsterdam, the Netherlands. [†]*PTEN* status does not include patient samples having *PTEN* mutations or deletions on the subclonal level.

Chapter 4

Supplementary Protocols: virus production and concentration, viral RNA isolation and RT-qPCR, and human CD34⁺ HSC transduction.

Reagents

X-tremeGENE HP DNA Transfection Reagent: Roche *Cat. #06 366 236 001*

Opti-MEM I with Glutamax: Life Technologies *Cat. #51985-026*

0.45 µm filter: Minisart *Cat. #17598* or equivalent

IVSS VIVASPIN 20 centrifugation concentration column: Sartorius AG, Sigma-Aldrich *Cat. #Z614653-48EA*

Retronectin: r-Fibronectin CH-296: TaKaRa *Cat. #T100A*

Non-tissue culture treated plates: Falcon (96 wp) *Cat. #351172*

X-VIVO 10 medium: Lonza *Cat. #BE04-743Q*

Recombinant human stem cell factor (rhSCF): R&D *Cat. #255-SC*

Recombinant human thrombopoietin (rhTPO): R&D *Cat. #288-TP/CF*

Recombinant human FLT3 ligand (rhFlt3L): Miltenyi Biotec *Cat. #130-093-855*

Protamine Sulfate salt from Salmon grade X: Sigma *Cat. #P4020-1G*

Virus production and concentration (biosafety level 2)

- 1) For a 14 cm dish, plate $\sim 8 \times 10^6$ HEK293T cells in 15 ml DMEM + 10% FCS and grow O/N to near confluency
- 2) Mix lentiviral vector DNA (25 µg), VSV-G (3.7 µg), RRE (GAG/POL; 5 µg), REV (3.1 µg) and serum-free Opti-MEM (4 ml)
- 3) Add X-tremeGENE HP DNA Transfection Reagent (37 µl) to the DNA mix
- 4) Incubate at room temperature for 15 minutes
- 5) In a drop-wise manner: add the transfection mixture to the cells
- 6) Incubate cells at 37°C, 5%CO₂ for 16-20 hours
- 7) Remove DMEM, wash cells once with pre-warmed PBS, and add 11 ml of pre-warmed serum-free Opti-MEM
- 8) Incubate cells at 37°C, 5%CO₂ for 24 hours
- 9) Collect the first batch of virus-containing medium and pass through a 0.45 µm filter
- 10) Concentrate roughly 40-fold (for example: 2 plates: concentrate 20 ml to 0.5 ml) using a centrifugation concentration column according to the manufacturer's recommendations and keep at 4°C
- 11) Add 11 ml of fresh pre-warmed serum-free Opti-MEM to the cells and incubate at 37°C, 5%CO₂ for another 24 hours
- 12) Collect the second batch of virus-medium and repeat steps 9 and 10
- 13) Combine both batches and determine the concentration factor
- 14) Take 2 µl of the concentrated virus for viral RNA isolation in step 17
- 15) Aliquot virus and use fresh or store at -80°C

Viral RNA isolation and RT-qPCR

- 16) To determine the number of viral particles by PCR quantification, make a serial dilution of a viral vector DNA plasmid with a known concentration
- 17) Trizol isolation of viral RNA is performed according to the manufacturer's procedure
- 18) RT-qPCR is performed using primers flanking the cPPT region: 5'-AGGTGGAGAGAGAGACAGAGAC-3' and 5'-CTCTGCTGTCCCTGTAATAAAC-3'

Retronectin-coated plates

- 19) Coat retronectin at a concentration of 50 µg/ml in PBS in non-tissue culture treated plates at 4°C (overnight)
- 20) Before using retronectin-coated plate: remove retronectin, add 2%BSA/PBS, incubate at 37°C for 30 minutes, and wash twice with PBS

Human CD34⁺ HSC O/N stimulation

- 21) Resuspend HSCs to concentration of $0.5 \times 10^6/\text{ml}$ in serum-free X-VIVO 10 supplemented with 50 ng/ml rhSCF, 20 ng/ml rhTPO and 50 ng/ml rhFlt3L
- 22) Incubate the cells at 37°C, 5%CO₂ for 16-20 hours

Human CD34⁺ HSC transduction (96 well-plate) (biosafety level 2)

- 23) Collect HSCs and centrifuge to adjust to a concentration of $1 \times 10^6/\text{ml}$, keeping them in the same medium with cytokines
- 24) Add protamine sulfate to the cells to a final concentration of 4 µg/ml
- 25) Pipet concentrated virus into retronectin-coated (96-well) plate
- 26) Add HSCs on top of the virus to a final volume of 200 µl/well
- 27) Mix by gently tapping the plate
- 28) Spinoculation: centrifuge at 1800 rpm, 32°C for 1 hour
- 29) Incubate the transduced cells at 37°C, 5%CO₂ for 24 hours before further use

Chapter 5

Supplementary Figures

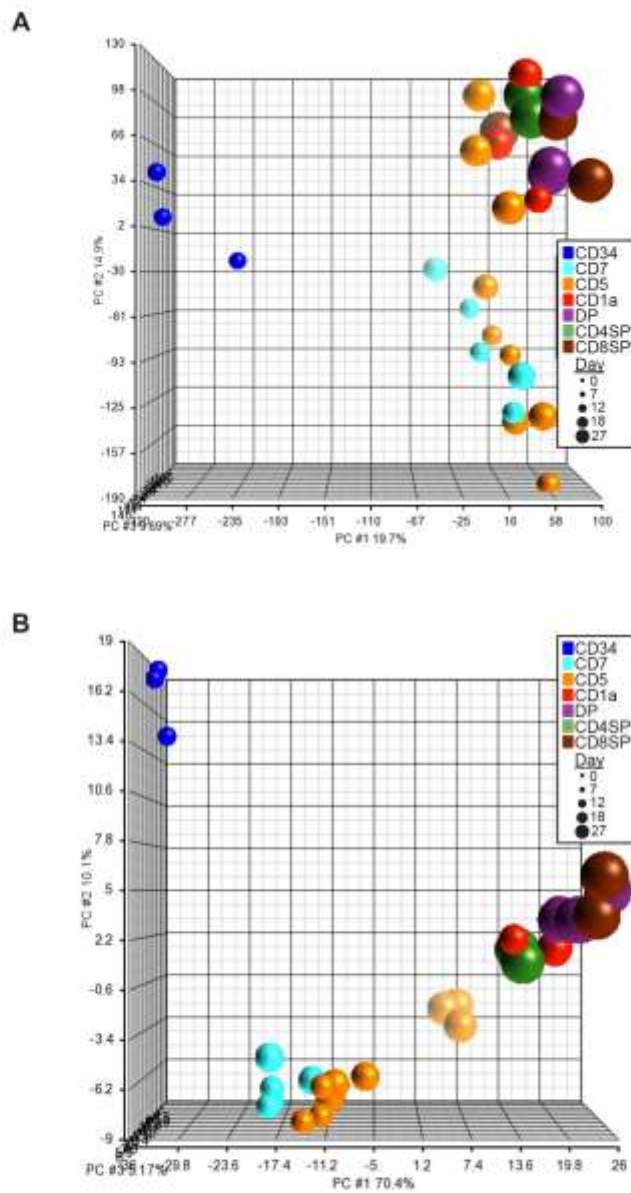
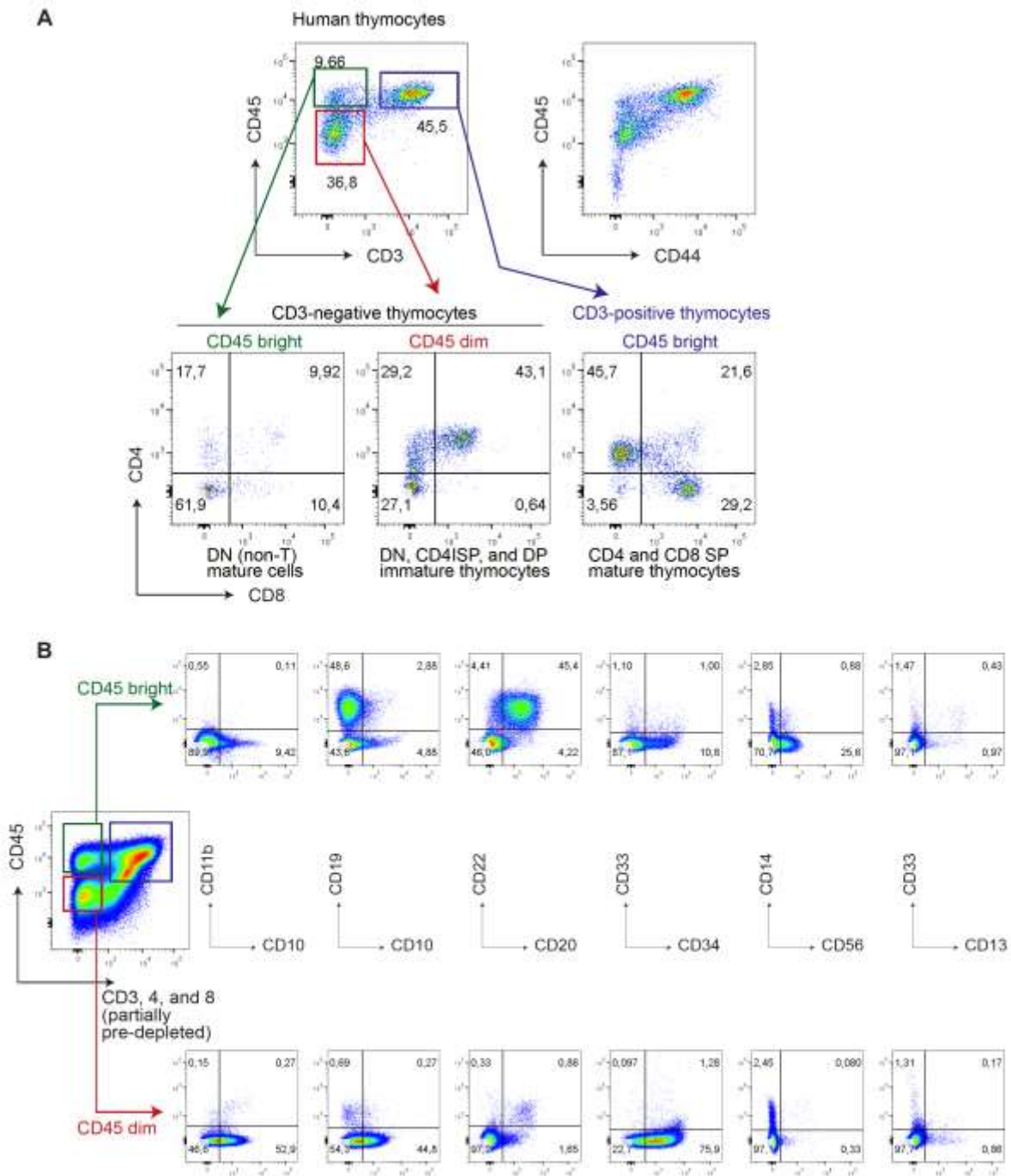
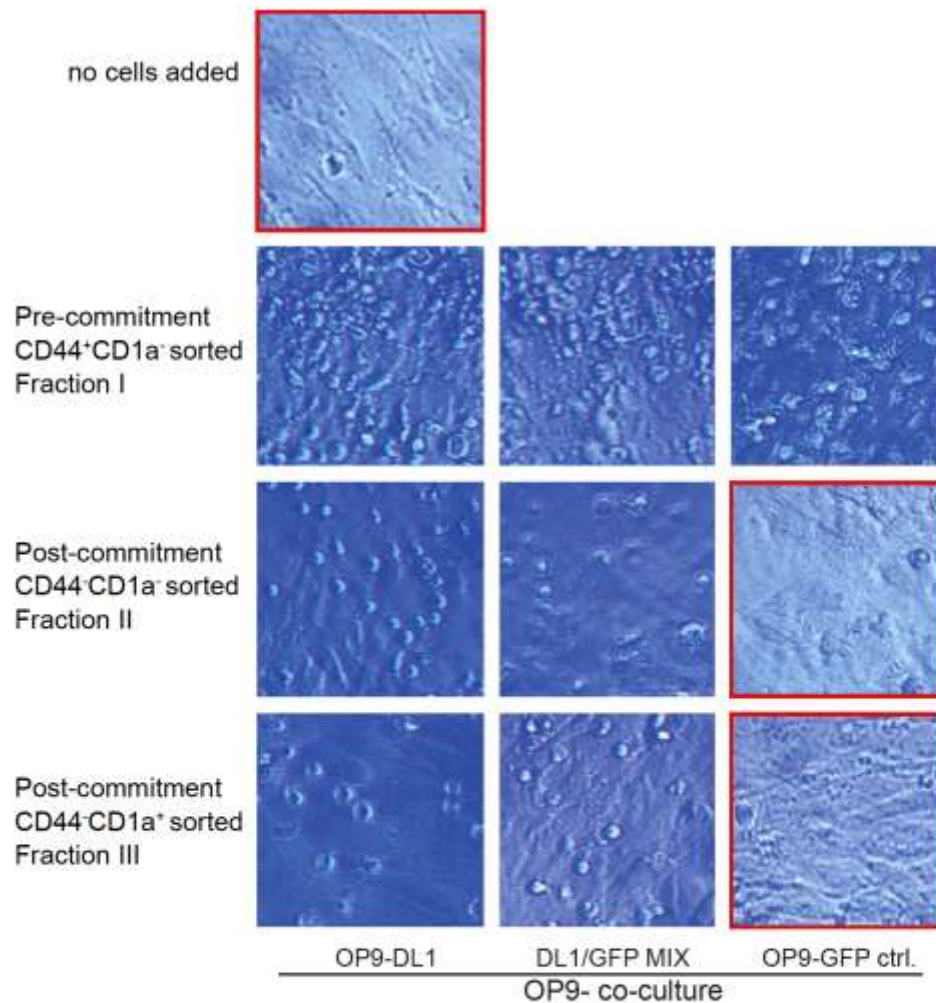


Figure S1. Related to Figure 1. Consecutive stages of *in vitro* T-cell differentiation can be divided into clusters with distinct gene expression. Principal component analysis of gene expression data of 29 sorted populations using all probesets (n=54715) of the array (A) or using 547 genes (gene clusters 2, 7, 8, 15, 16) (B).



Supplemental Figure S2. Related to Figure 3C. CD45^{bright} are mature, CD45^{dim} are immature human thymocytes and the intensities of CD45 and CD44 expression correlate. Human thymocytes were thawed and stained directly (A) or after CD3, CD4, and CD8 (partial) pre-deletion of ~95% of all thymocytes (B) with various antibodies as indicated.



Supplemental Figure S3. Related to Figure 4. Bright field microscopic images of sorted and co-cultured human thymocyte populations I-III. Images (200x) were generated after 7 days of co-culture. OP9-DL1 stromal cells support T- and other lineage differentiation, whereas OP9-GFP control cells do not support T-lineage differentiation (the red outlines indicate images with only OP9 stromal cells). The middle column represents co-cultures on OP9-DL1:GFP 1:1 mixed cells.

Supplemental Tables

Table S1: Gene Symbols in each of the 16 Gene Signatures from Figure 1C-E

| Gene Signature | Gene Symbol | Gene Signature | Gene Symbol |
|----------------|------------------------------------|----------------|--------------------------|
| 1 | JHDM1D | 9 | S100A8 |
| 1 | BTG2 | 9 | S100A9 |
| 1 | NBEA | 9 | DEFA1///DEFA1B///DEFA3 |
| 1 | PFKFB2 | 9 | TACSTD2 |
| 1 | HIST1H2AC | 9 | CHI3L1 |
| 1 | ARID5B | 9 | CD24 |
| 1 | PELI2 | 9 | CSGALNACT1 |
| 1 | ERN1 | 9 | MS4A3 |
| 1 | BTBD11 | 9 | CSTA |
| 1 | ZNF704 | 9 | ACTN1 |
| 1 | RORA | 9 | PRG2 |
| 1 | MAFB | 9 | CLC |
| 1 | LOC283070 | 9 | RNASE3 |
| 1 | CAMK1D///LOC283070 | 9 | S100P |
| 1 | CAMK1D | 9 | RNASE2 |
| 1 | SPOCK2 | 9 | CTSG |
| 1 | DDIT4 | 9 | ELANE |
| 1 | TP53INP1 | 9 | CEACAM6 |
| 1 | HIST2H2BE | 9 | CST7 |
| 1 | CREBRF | 9 | PRTN3 |
| 1 | AMY1A///AMY1B///AMY1C///AMY2A | 9 | AZU1 |
| 1 | LOC728392 | 9 | CEACAM8 |
| 1 | DCHS1 | 9 | BPI |
| 1 | PCMTD1 | 9 | DEFA4 |
| 1 | BRWD1 | 9 | SLPI |
| 2 | TRERF1 | 9 | ANXA3 |
| 2 | DNTT | 9 | ETV5 |
| 2 | ITGB2-AS1 | 9 | TNFRSF4 |
| 2 | LINC00340 | 9 | TNFRSF18 |
| 2 | GAL3ST4 | 9 | PRF1 |
| 2 | OTTHUMG00000166495///RP11-728F11.1 | 9 | VPREB1 |
| 2 | MBD5 | 9 | CTSW |
| 2 | LBH | 9 | SVOP |
| 2 | SCN3A | 9 | GPA33 |
| 2 | IL10RA | 9 | PPIF |
| 2 | FAIM3 | 9 | ELOVL6 |
| 2 | CRIP1 | 9 | CTHRC1 |
| 2 | OTTHUMG00000178878///RP11-214C8.5 | 9 | CSRP2 |
| 2 | PCDH9 | 9 | TNFSF4 |
| 2 | NEGR1 | 9 | PRR5L |
| 2 | ETS1 | 9 | GCNT1 |
| 2 | BACH2 | 9 | RAB27B |
| 2 | PGM2L1 | 9 | GZMA |
| 2 | DUSP16 | 9 | NKG7 |
| 2 | TANC1 | 9 | DNMT3B |
| 2 | ITGA6 | 9 | TESC |
| 2 | ARRDC4 | 9 | ARRB1 |
| 2 | AGBL3 | 9 | APOBEC3B |
| 2 | PRSS2///PRSS3 | 9 | IGLL1 |
| 2 | SCN7A | 9 | GUSBP11///IGLL1///IGLL3P |
| 2 | PDE4D | 9 | COL5A2 |
| 2 | TUBB2A | 9 | SLC35F3 |
| 2 | ARHGAP29 | 9 | CISH |
| 2 | SIPA1L2 | 9 | CFH///CFHR1 |
| 2 | ID2///ID2B | 9 | TFEC |
| 2 | CDKN2D | 9 | SLC39A14 |
| 2 | S1PR1 | 9 | NLN |
| 2 | IKZF3 | 9 | FKBP4 |
| 2 | RBMS3 | 9 | TTLL12 |
| 2 | NSG1 | 9 | PUS7 |
| 2 | NLGN4X | 9 | OSM |
| 2 | LRRN3 | 9 | H19///MIR675 |
| 2 | CD8A | 9 | LIN28B |
| 2 | NINL | 10 | AGR2 |
| 2 | CD8B///LOC100996919 | 10 | MS4A1 |
| 2 | EFNB2 | 10 | KLRD1 |

| | | | |
|---|------------------------------------|----|------------------------------------|
| 2 | VCAN | 10 | IL18RAP |
| 2 | LOC100287598 | 10 | GNLY |
| 2 | ID1 | 10 | XCL1 |
| 2 | RASSF6 | 10 | ITGAM |
| 2 | SEC31B | 10 | SAMD3 |
| 2 | COX6A1 | 10 | KLRB1 |
| 2 | GBP1 | 10 | KLRC4-KLRK1///KLRK1 |
| 2 | KANSL1-AS1 | 10 | CRTAM |
| 2 | GAS2 | 10 | PTGDR |
| 2 | PALLD | 10 | PDE7B |
| 2 | HNRNPLL | 10 | EPHA4 |
| 2 | ANKRD36BP2///LOC101060554 | 10 | KLRC3 |
| 2 | PLEKHG1 | 10 | MAF |
| 2 | NCALD | 10 | KLRC1///KLRC2 |
| 2 | SYNM | 10 | IL2RB |
| 2 | GRK5 | 11 | FCRL4 |
| 2 | RDH10 | 11 | SIGLEC6 |
| 2 | ALOX5 | 11 | UGCG |
| 2 | FHIT | 11 | CTA-250D10.23///OTTHUMG00000172730 |
| 2 | TRERF1 | 11 | C1orf186///LOC100505650 |
| 2 | DNTT | 11 | TBC1D9 |
| 2 | ITGB2-AS1 | 11 | CDH1 |
| 2 | LINC00340 | 11 | FGL2 |
| 2 | GAL3ST4 | 11 | ENTPD1 |
| 2 | OTTHUMG00000166495///RP11-728F11.1 | 11 | MPEG1 |
| 2 | MBD5 | 11 | FGD2 |
| 2 | LBH | 11 | LGMN |
| 2 | SCN3A | 11 | CACNA2D3 |
| 2 | IL10RA | 11 | SLAMF7 |
| 2 | FAIM3 | 11 | LY9 |
| 2 | CRIP1 | 11 | PTPRJ |
| 2 | OTTHUMG00000178878///RP11-214C8.5 | 11 | MCOLN2 |
| 2 | PCDH9 | 11 | CCR2 |
| 2 | NEGR1 | 11 | TGFB1 |
| 2 | ETS1 | 11 | NAPSB |
| 2 | BACH2 | 11 | MYCL1 |
| 2 | PGM2L1 | 11 | TLR10 |
| 2 | DUSP16 | 11 | C5orf20///TIFAB |
| 2 | TANC1 | 11 | TIFAB |
| 2 | ITGA6 | 11 | LYZ |
| 2 | ARRDC4 | 11 | IL6R |
| 2 | AGBL3 | 11 | MRC1 |
| 2 | PRSS2///PRSS3 | 11 | MNDA |
| 2 | SCN7A | 11 | P2RY14 |
| 2 | PDE4D | 11 | TYROBP |
| 2 | TUBB2A | 11 | FCER1G |
| 2 | ARHGAP29 | 11 | CCR1 |
| 2 | SIPA1L2 | 11 | CD36 |
| 2 | ID2///ID2B | 11 | CTSZ |
| 2 | CDKN2D | 11 | MS4A7 |
| 2 | S1PR1 | 11 | CD86 |
| 2 | IKZF3 | 11 | NCF2 |
| 2 | RBMS3 | 11 | GSN |
| 2 | NSG1 | 11 | AMICA1 |
| 2 | NLGN4X | 11 | SLC41A2 |
| 2 | LRRN3 | 11 | DNASE1L3 |
| 2 | CD8A | 11 | TICAM2///TMED7-TICAM2 |
| 2 | NINL | 11 | GAS2L3 |
| 2 | CD8B///LOC100996919 | 11 | RGS18 |
| 2 | EFNB2 | 11 | IDH3A |
| 2 | VCAN | 11 | CD48 |
| 2 | LOC100287598 | 11 | IGJ |
| 2 | ID1 | 11 | TRIB2 |
| 2 | RASSF6 | 11 | SMIM14 |
| 2 | SEC31B | 11 | RNASE6 |
| 2 | COX6A1 | 11 | CLCN5 |
| 2 | GBP1 | 11 | GPR18 |
| 2 | KANSL1-AS1 | 11 | LOC158402 |
| 2 | GAS2 | 11 | FGR |
| 2 | PALLD | 11 | EPHB1 |
| 2 | HNRNPLL | 11 | OTTHUMG00000175583///RP11-4O1.2 |
| 2 | ANKRD36BP2///LOC101060554 | 11 | ARL4C |

| | | | |
|---|------------------------------------|----|--------------------------|
| 2 | PLEKHG1 | 11 | GPR183 |
| 2 | NCALD | 11 | CLECL1 |
| 2 | SYNM | 11 | FAM105A |
| 2 | GRK5 | 11 | SAMHD1 |
| 2 | RDH10 | 11 | CLEC4A |
| 2 | ALOX5 | 11 | ITGB7 |
| 2 | FHIT | 11 | BLNK |
| 3 | DENND1B | 11 | STX7 |
| 3 | PAPD4 | 11 | IFI44L |
| 3 | U2SURP | 11 | OAS1 |
| 3 | SF1 | 11 | CRYBG3 |
| 3 | SREK1 | 11 | ABHD6 |
| 3 | LOC100996467 | 11 | KYNU |
| 3 | MALAT1 | 11 | CCR6 |
| 3 | CLK4 | 11 | HCK |
| 3 | OTTHUMG00000168768///RP11-463O12.5 | 11 | GAS6 |
| 3 | KMT2A | 11 | IGK///IGKC |
| 3 | CAPZA1 | 11 | DAB2 |
| 3 | ZMYM2 | 11 | MS4A4A |
| 3 | CEP95 | 11 | HDAC9 |
| 3 | B2M | 11 | CLEC12A |
| 3 | IFNGR1 | 11 | KMO |
| 3 | EIF5///SNORA28 | 11 | IRF8 |
| 3 | MORF4L2 | 11 | CYBB |
| 3 | NFAT5 | 11 | DTX4 |
| 3 | C10orf118 | 11 | ADRB2 |
| 3 | PCNX | 11 | THBD |
| 3 | AC092620.2///OTTHUMG00000153633 | 11 | RBM47 |
| 3 | TGFBR1 | 11 | ATP10D |
| 3 | INPP5D | 11 | CD180 |
| 3 | MIR612///NEAT1 | 11 | GPR114 |
| 3 | WDR11 | 11 | GAPT |
| 3 | OTTHUMG00000168867///RP11-305O6.3 | 11 | PAX5 |
| 3 | LRRFIP1 | 11 | ARHGAP18 |
| 3 | CELF2 | 11 | KIAA0930 |
| 3 | LONRF1 | 11 | SUCNR1 |
| 3 | ANKRD28 | 11 | LY86 |
| 3 | ZNF274 | 12 | GZMB |
| 3 | LOC100506312 | 12 | NME8 |
| 3 | SEC61B | 12 | TCL1A |
| 3 | NKTR | 12 | CMKLR1 |
| 3 | PIK3C2A | 12 | PTPRS |
| 3 | RAB18 | 12 | IL18R1 |
| 3 | GNAS | 12 | SOGA2 |
| 3 | PELI1 | 12 | TPM2 |
| 3 | UBE2B | 12 | CYAT1///IGLC1///IGLV1-44 |
| 3 | TBRG1 | 12 | IGLC1 |
| 3 | LINC00597 | 12 | LAMP5 |
| 3 | SAMD4B | 12 | LOC285972 |
| 3 | DAPK1-IT1 | 12 | SEL1L3 |
| 3 | PGAP1 | 12 | IL3RA |
| 3 | MLLT3 | 12 | GNG7 |
| 3 | SPATA13 | 12 | LHFPL2 |
| 3 | RGCC | 12 | ZDHHC14 |
| 3 | BCL6 | 12 | BANK1 |
| 3 | KLF4 | 12 | CCDC50 |
| 3 | RGS1 | 12 | FZD3 |
| 3 | LOC100507577///LONP2 | 12 | FCRLA |
| 3 | ZBTB20 | 12 | IFNLRL1 |
| 3 | OTTHUMG00000175755///RP11-553L6.5 | 12 | SPIB |
| 3 | LOC100131564 | 12 | IRF7 |
| 3 | PPP2R5C | 12 | FAM129C |
| 3 | HECA | 12 | SULF2 |
| 3 | LOC100996653 | 12 | LILRA4 |
| 3 | APBB1IP | 12 | CLEC4C |
| 3 | TNFAIP3 | 12 | PAPLN |
| 3 | STK4 | 13 | CDH17 |
| 3 | PRO2852 | 13 | IL1R2 |
| 3 | ZFAND6 | 13 | IGSF6 |
| 3 | ZNF638///ZNF638-IT1 | 13 | CLEC10A |
| 3 | RBM5 | 13 | LGALS2 |
| 3 | RUNX1-IT1 | 13 | HLA-DQB1 |

| | | | |
|---|--|----|-----------------------------------|
| 3 | CCAR1 | 13 | HLA-DQA1 |
| 3 | GGNBP2 | 13 | AGPAT9 |
| 3 | PDCD6 | 13 | CLIC2 |
| 3 | RPS27 | 13 | CLEC7A |
| 3 | RPS16P5 | 13 | IL13RA1 |
| 3 | TLE4 | 13 | CST3 |
| 3 | LOC100996866 | 13 | AXL |
| 3 | IKZF1 | 13 | OTTHUMG00000173465///RP11-389C8.2 |
| 3 | GUSBP3///GUSBP9///LOC100653061///LOC101060519 | 13 | MS4A6A |
| 3 | MSI2 | 13 | GEM |
| 3 | RHOH | 13 | PKIB |
| 3 | LOC100190986///LOC101060564 | 13 | S100B |
| 3 | LINC00342 | 13 | TLR7 |
| 3 | FLJ31306 | 13 | SERPINF1 |
| 3 | FLJ38379 | 13 | CCR5 |
| 3 | TCF12 | 13 | JAZF1 |
| 3 | RAB11FIP3 | 13 | MARCH1 |
| 3 | ZNF124 | 13 | CPVL |
| 4 | C1orf54 | 13 | IDO1 |
| 4 | GGTA1P | 13 | FCGR2B |
| 4 | FCER1A | 14 | TASP1 |
| 4 | LGALS3 | 14 | GXYLT2 |
| 4 | HLA-DRB1///HLA-DRB3///HLA-DRB4///HLA-DRB5///LOC100507709///LOC100507714 | 14 | TRDC |
| 4 | SERPINB2 | 14 | YME1L1 |
| 4 | THEMIS2 | 14 | CD3D |
| 4 | SETBP1 | 14 | TRAT1 |
| 4 | SERPINB9 | 14 | GPR171 |
| 4 | EMP1 | 14 | GPR65 |
| 4 | HLA-DPA1 | 14 | LPXN |
| 4 | CD74 | 14 | CDK2AP2 |
| 4 | CSF2RB | 14 | VANGL1 |
| 4 | HLA-DOA | 14 | VASH2 |
| 4 | AHNAK | 14 | CD7 |
| 4 | HLA-DRB1///HLA-DRB3///HLA-DRB4///HLA-DRB5///LOC100507709///LOC100507714///LOC101060779 | 14 | HES4 |
| 4 | DUSP1 | 14 | ICOS |
| 4 | RGS2 | 14 | MYO1B |
| 4 | KLF6 | 14 | CEP41 |
| 4 | SIK1 | 14 | SLC35B3 |
| 4 | ARRDC3 | 14 | TMEM251 |
| 4 | ZBTB24 | 14 | ATG4C |
| 4 | PLK2 | 14 | VPS41 |
| 4 | KLF2 | 14 | ATP6V1D |
| 4 | PPP1R15A | 14 | BPNT1 |
| 4 | DUSP4 | 14 | KIAA1841 |
| 4 | ETNK1 | 14 | PPHLN1 |
| 4 | EGR3 | 14 | C17orf62 |
| 4 | PTGS2 | 14 | PWWP2A |
| 4 | NR4A2 | 14 | CBR4 |
| 4 | FOSB | 14 | WDR67 |
| 4 | FOS | 14 | DCAF12 |
| 4 | VIM | 14 | LRIF1 |
| 4 | SGK1 | 14 | ST8SIA4 |
| 4 | RASGEF1B | 14 | DNA2 |
| 4 | LOC284454 | 14 | LOC389831 |
| 4 | ANXA1 | 14 | HES1 |
| 4 | ITPKB | 14 | GIMAP7 |
| 4 | NASP | 14 | IL17RA |
| 4 | EGR2 | 14 | DPP4 |
| 4 | LAIR2 | 14 | IFNAR1 |
| 5 | VWF | 14 | PAQR8 |
| 5 | EFEMP1 | 14 | TRIM13 |
| 5 | MGP | 14 | ACSL3 |
| 5 | MAFIP | 14 | SHCBP1 |
| 5 | ELL2 | 14 | UMODL1 |
| 5 | FN1 | 14 | PHF20L1 |
| 5 | HBEGF | 14 | STK38L |
| 5 | ABLIM1 | 14 | CEP19 |
| 5 | NRIP1 | 14 | C9orf64 |
| 5 | PIK3R1 | 14 | HYLS1 |
| 5 | MYH10 | 14 | ATP6V1C1 |
| 5 | OTTHUMG00000182869///RP11-420K14.2 | 14 | ERCC4 |
| 5 | FLJ13197 | 14 | ARL13B |

| | | | |
|---|---------------------------------------|----|-----------|
| 5 | ZBTB16 | 14 | KBTBD7 |
| 5 | CDKN1C | 14 | HDHD1 |
| 5 | PXDN | 14 | FBXO5 |
| 5 | HIST3H2A | 14 | NCAPG2 |
| 5 | OCLN | 14 | SKA3 |
| 5 | HLA-DRB4 | 14 | CCNF |
| 5 | XIST | 14 | TROAP |
| 5 | FOXO3///FOXO3B | 14 | AURKB |
| 5 | MXI1 | 14 | CASC5 |
| 5 | DGKE | 14 | NUSAP1 |
| 5 | PPP1R3B | 14 | TMPO |
| 5 | ABCG1 | 14 | CCDC15 |
| 5 | RBM38 | 14 | SGOL2 |
| 5 | SHISA2 | 14 | GTSE1 |
| 5 | GNL1 | 14 | CDCA2 |
| 5 | BAG3 | 14 | E2F7 |
| 5 | SPTBN1 | 14 | E2F8 |
| 5 | CREM | 14 | KIF18B |
| 5 | KBTBD2 | 14 | MKI67 |
| 5 | IL6ST | 14 | KIF23 |
| 5 | SKIL | 14 | ECT2 |
| 5 | RIT1 | 14 | CENPF |
| 5 | HIST2H2AA3///HIST2H2AA4 | 14 | NUF2 |
| 5 | PIK3IP1 | 14 | FANCI |
| 5 | ANKDD1A | 14 | KIF15 |
| 5 | CCNL1 | 14 | KIF18A |
| 5 | JUN | 14 | ESPL1 |
| 5 | MEX3C | 14 | KIFC1 |
| 5 | CSRNP1 | 14 | BUB1B |
| 5 | BCL10 | 14 | STIL |
| 5 | EPC1 | 14 | KIF14 |
| 5 | BTG1 | 14 | CENPE |
| 5 | KRAS | 14 | DEPDC1 |
| 5 | CHMP3///RNF103///RNF103-CHMP3 | 14 | KIF2C |
| 5 | KLHL24 | 14 | HJURP |
| 5 | MIR4800///MXD4 | 14 | CDCA8 |
| 5 | OTUD1 | 14 | CCNB2 |
| 5 | CA2 | 14 | CENPA |
| 5 | RIN2 | 14 | NEK2 |
| 5 | LILRB2 | 14 | KIF11 |
| 5 | CPQ | 14 | KIF4A |
| 5 | ADAM8 | 14 | BUB1 |
| 5 | SQSTM1 | 14 | RRM2 |
| 5 | IL8 | 14 | TPX2 |
| 5 | TRIB1 | 14 | RACGAP1 |
| 5 | CCL3///CCL3L1///CCL3L3///LOC101060267 | 14 | MELK |
| 5 | ZFP36 | 14 | SPAG5 |
| 5 | TSC22D3 | 14 | NCAPG |
| 5 | METRNL | 14 | SRSF1 |
| 5 | ATF3 | 14 | NDC80 |
| 5 | NAMPT | 14 | SPDL1 |
| 5 | RC3H2 | 14 | CENPN |
| 5 | MAFF | 14 | C9orf40 |
| 5 | POLB | 14 | CKAP2L |
| 5 | CUL1 | 14 | ARHGAP11A |
| 5 | HEMGN | 14 | FOXM1 |
| 5 | PER1 | 14 | NCAPH |
| 5 | PCIF1 | 14 | RAD54L |
| 5 | CSNK1E///LOC100996460 | 14 | BIRC5 |
| 5 | CUL4A | 14 | CCDC167 |
| 5 | STAT3 | 14 | CCNE2 |
| 5 | FAM3C | 14 | ERCC6L |
| 5 | RAPGEF2 | 14 | PARPBP |
| 5 | RNF130 | 14 | C11orf82 |
| 5 | MECP2 | 14 | KNSTRN |
| 5 | SRSF11 | 14 | CDCA5 |
| 5 | NLRP1 | 14 | ASF1B |
| 5 | MED23 | 14 | AURKA |
| 5 | ST3GAL1 | 14 | CDKN3 |
| 5 | ZNF638 | 14 | MND1 |
| 5 | IREB2 | 14 | NIF3L1 |
| 5 | MIR22///MIR22HG | 14 | MTFR2 |

| | | | |
|---|--------------------|----|---|
| 5 | PPP1R16B | 14 | TCF19 |
| 5 | FLJ43663 | 14 | HIRIP3 |
| 5 | ISG20 | 14 | SPC24 |
| 5 | PIM2 | 14 | RTP4 |
| 5 | PDE4B | 14 | RBL1 |
| 5 | SOCS3 | 14 | TOP2A |
| 5 | LPGAT1 | 14 | FAM72A///FAM72B///FAM72C///LOC101060656 |
| 5 | HLA-E | 14 | UHRF1 |
| 5 | SMAD3 | 14 | WDR76 |
| 5 | RNF144B | 14 | INTS7 |
| 5 | IDS | 14 | CDK1 |
| 5 | ERO1LB | 14 | DLGAP5 |
| 5 | ACVR2A | 14 | TTK |
| 5 | TMEM200A | 14 | UBE2C |
| 5 | IRAK2 | 14 | CDCA3 |
| 5 | GABARAPL1 | 14 | KIF20A |
| 5 | SLC2A3 | 14 | CEP55 |
| 5 | SYTL3 | 14 | HMMR |
| 6 | HAS1 | 14 | ANLN |
| 6 | CTSL1 | 14 | BRIP1 |
| 6 | NINJ1 | 14 | ERI2 |
| 6 | CXCL3 | 14 | DIAPH3 |
| 6 | TM4SF1 | 14 | C1orf112 |
| 6 | INHBA | 14 | SMC2 |
| 6 | CXCL1 | 14 | DEPDC1B |
| 6 | AQP9 | 14 | UBE2T |
| 6 | CCL2 | 14 | CCNA2 |
| 6 | ICAM1 | 14 | CDC20 |
| 6 | PHLDA1 | 14 | CCNB1 |
| 6 | TNFRSF9 | 14 | OIP5 |
| 6 | IL6 | 14 | SPC25 |
| 6 | ZNF709 | 14 | PKP4 |
| 6 | C10orf10 | 14 | BRCC3 |
| 6 | RHOB | 14 | CDC25A |
| 6 | IER3 | 14 | FBXO22 |
| 6 | C15orf48 | 14 | SLC25A19 |
| 6 | BCL2A1 | 14 | POLR1B |
| 6 | TLN1 | 14 | IDH1 |
| 6 | HIST1H2BC | 14 | HK2 |
| 6 | TPP2 | 14 | GIMAP6 |
| 6 | IL1RL1 | 14 | GIMAP8 |
| 6 | CNST | 14 | GIMAP1 |
| 6 | USP36 | 14 | NABP1 |
| 6 | MAX | 14 | NRARP |
| 6 | FTH1 | 14 | GAS1 |
| 6 | USP15 | 14 | DCPS |
| 6 | GFOD1 | 14 | ARL11 |
| 6 | DOCK7 | 14 | SGK223 |
| 6 | AREG///AREGB | 14 | STAMBPL1 |
| 6 | C7orf73///SLC13A4 | 14 | OPN3 |
| 6 | SLC8A3 | 14 | NLRP2 |
| 6 | TPM3 | 14 | PSPH |
| 6 | SOD2 | 14 | FAM173B |
| 6 | GABPB1 | 14 | CHEK2 |
| 6 | STX11 | 14 | ICMT |
| 6 | CXCL2 | 14 | POLR3K |
| 6 | FAM198B | 14 | C2orf43 |
| 6 | CYLD | 14 | RAD51C |
| 6 | RELB | 14 | GART |
| 6 | BCL3 | 14 | CSE1L |
| 6 | NBL1 | 14 | OTTHUMG00000159786///RP11-85F14.5 |
| 6 | SLCO3A1 | 14 | TRMT1L |
| 6 | SMURF1 | 14 | METTL13 |
| 6 | SEMA6A | 14 | UBE2L3 |
| 6 | SDK2 | 14 | EPDR1 |
| 6 | MINOS1-NBL1///NBL1 | 14 | RPE |
| 6 | SERAC1 | 14 | BRI3BP |
| 6 | EIF4G3 | 14 | RM11 |
| 6 | NFKB2 | 14 | KIAA1524 |
| 6 | USP53 | 14 | GINS1 |
| 6 | ERMN | 14 | ARFIP1 |
| 6 | TCF25 | 14 | SHQ1 |

| | | | |
|---|-----------------------------------|----|-----------------------------------|
| 6 | GNG4 | 14 | TARP///TRGC2 |
| 6 | KSR1 | 14 | RAD51 |
| 6 | SKI | 14 | DHFR |
| 6 | MEG3 | 14 | BRCA1 |
| 6 | AVP | 14 | MCM8 |
| 6 | THBS1 | 14 | TK1 |
| 6 | HBA1///HBA2 | 14 | GUCY1B3 |
| 6 | NR4A3 | 14 | TAF5 |
| 6 | BMP6 | 14 | TDP1 |
| 6 | CDC42 | 14 | TUBG1 |
| 6 | LOC100507672 | 14 | PAGR1 |
| 6 | BEX5 | 14 | XRCC4 |
| 6 | CDC14B | 14 | SRSF10 |
| 6 | NEDD4L | 14 | CENPM |
| 6 | HBG1///HBG2 | 14 | WDHD1 |
| 6 | DLK1 | 14 | PANK1 |
| 6 | PPBP | 14 | LCMT2 |
| 6 | GOS2 | 14 | PSAT1 |
| 6 | TMEM163 | 14 | EXO1 |
| 6 | TNIK | 14 | C12orf4 |
| 6 | THSD7A | 14 | C18orf54 |
| 6 | LIMS3///LIMS3L | 14 | MED11 |
| 6 | ARHGEF12 | 14 | EXOSC3 |
| 6 | FOXO1 | 14 | GEMIN6 |
| 6 | FAM65C | 14 | UMPS |
| 6 | PDZD2 | 14 | TRIP13 |
| 6 | EREG | 14 | CDC6 |
| 6 | HBB | 14 | MCM10 |
| 6 | LOC254057 | 14 | CHAC2 |
| 6 | PEAR1 | 14 | CLUAP1 |
| 6 | FCHO2 | 14 | RAD54B |
| 6 | CRHBP | 14 | SAPCD2 |
| 6 | MYCT1 | 14 | GIN54 |
| 6 | DNAJC6 | 15 | CD1C |
| 6 | BEX2 | 15 | SMPD3 |
| 6 | ACSL1 | 15 | CD1E |
| 6 | BEX1 | 15 | HHIP-AS1 |
| 6 | H1FO | 15 | CHI3L2 |
| 6 | KLF3 | 15 | SYNE2 |
| 6 | LRPSL | 15 | TNRC6C |
| 6 | PRKACB | 15 | ARPP21 |
| 6 | NLK | 15 | LOC100130503 |
| 6 | C7orf41 | 15 | DOCK9 |
| 6 | MAGI2-AS3 | 15 | OTTHUMG00000176181///RP11-119F7.5 |
| 6 | OTTHUMG00000176268///RP11-18F14.2 | 15 | CCR9 |
| 6 | SLC35D2 | 15 | GSDMB |
| 6 | LOC285812 | 15 | CTC1 |
| 6 | PTX3 | 15 | HS3ST3B1 |
| 6 | GBP2 | 15 | ZBTB1 |
| 6 | ABCB1 | 15 | FAM69A |
| 6 | CEBPB | 15 | RUFY3 |
| 6 | FGD5 | 15 | CCNG2 |
| 6 | ABHD17C | 15 | PVRIG |
| 6 | MPP7 | 15 | NOTCH3 |
| 6 | IRS2 | 15 | HDAC4 |
| 6 | PPM1H | 15 | LRRC28 |
| 6 | PLOD2 | 15 | ACPL2 |
| 6 | RBPMS | 15 | LINC00226 |
| 6 | HLF | 15 | CR2 |
| 6 | PBX1 | 15 | ZC3H6 |
| 6 | CALN1 | 15 | NKD2 |
| 6 | GPR126 | 15 | GLRB |
| 6 | GNAI1 | 15 | OTTHUMG0000017520///RP11-144L1.4 |
| 6 | ALDH1A1 | 15 | CX3CR1 |
| 6 | TNFRSF10A | 15 | LOC101060391 |
| 6 | GATA2 | 15 | CCR8 |
| 6 | TNFRSF10D | 15 | TMSB15A///TMSB15B |
| 6 | CNRIP1 | 15 | SLC16A10 |
| 6 | ITGA9 | 15 | CD1B |
| 6 | TAL1 | 15 | RAG2 |
| 6 | TCF7L2 | 15 | AQP3 |
| 6 | DACH1 | 15 | CD1A |

| | | | |
|---|---|----|---------------------|
| 6 | LRP6 | 15 | GRAP2 |
| 6 | ZNF521 | 15 | CD3E |
| 6 | NPR3 | 15 | SLC7A3 |
| 7 | SOCS2 | 15 | RAPGEF5 |
| 7 | ARMCX1 | 15 | ATP6AP1L///FLJ41309 |
| 7 | MEST | 15 | SRPK2 |
| 7 | CXCR7 | 15 | TRBC1 |
| 7 | HSPA1A///HSPA1B | 15 | TRBC2 |
| 7 | C11orf96 | 15 | MAL |
| 7 | GSTM3 | 15 | ID3 |
| 7 | FHL1 | 15 | UBASH3A |
| 7 | SNTB1 | 15 | TLR5 |
| 7 | BSPRY | 15 | TSHR |
| 7 | EPCAM | 15 | IRS1 |
| 7 | HBD | 15 | HHIP |
| 7 | UCHL1 | 15 | PTCRA |
| 7 | NRIP3 | 15 | SLC8A1-AS1 |
| 7 | PDE4DIP | 15 | PAFAH2 |
| 7 | MATR3///SNHG4 | 15 | RAG1 |
| 7 | AC005682.5///OTTHUMG00000152563 | 15 | CD1D |
| 7 | TMEM220 | 15 | MME |
| 7 | GYPC | 15 | PLCL1 |
| 7 | FAM171B | 15 | SH3TC1 |
| 7 | ADM | 15 | SPRED2 |
| 7 | SMIM3 | 15 | BCL2L11 |
| 7 | CRYGD | 15 | FOXO6///FOXO6 |
| 7 | RIPK2 | 15 | CYP2U1 |
| 7 | ZBTB8A | 15 | PPP1R1C |
| 7 | GBP4 | 15 | LOC79015 |
| 7 | DRAM1 | 15 | STAG3 |
| 7 | MPL | 15 | CRNDE |
| 7 | ICA1 | 15 | ADAMTS1 |
| 7 | CMBL | 15 | WT1 |
| 7 | SYNGR1 | 15 | LOC283788 |
| 7 | FAM69B | 15 | LOC728613 |
| 7 | TCTEX1D1 | 15 | SLC16A6 |
| 7 | SDPR | 15 | IPP |
| 7 | OTTHUMG00000020937///RP11-251M1.1 | 15 | TFDP2 |
| 7 | CPA3 | 15 | SBK1 |
| 7 | KBTBD11 | 15 | POU2AF1 |
| 7 | SLC22A15 | 15 | ITK |
| 7 | RAB13 | 15 | RASGRP1 |
| 7 | GRB10 | 15 | CAMK4 |
| 7 | SLC2A5 | 15 | TCF7 |
| 7 | ELK3 | 15 | PCDH10 |
| 7 | PVRL2 | 15 | PTPRK |
| 7 | SORL1 | 15 | CD96 |
| 7 | PROK2 | 15 | GATA3 |
| 7 | GPR27 | 15 | IL32 |
| 7 | SNORD3A///SNORD3B-1///SNORD3B-2///SNORD3C///SNORD3D | 15 | TMOD2 |
| 7 | PRSS2 | 15 | PLCG1 |
| 7 | MSRB3 | 15 | NMT2 |
| 7 | NAP1L3 | 15 | P2RX5 |
| 7 | TPM1 | 15 | IKZF2 |
| 7 | HMGA2 | 15 | FAM63B |
| 7 | NUDT11 | 15 | DGKA |
| 7 | CABLES1 | 15 | RGPD1///RGPD2 |
| 7 | CKAP4 | 15 | LAT |
| 7 | RAMP1 | 15 | BCL11B |
| 7 | TFPI | 15 | ZAP70 |
| 7 | IL1RAP | 15 | THEMIS |
| 7 | IRAK3 | 15 | CD247 |
| 7 | GBP5 | 15 | SLFN5 |
| 7 | FOSL2 | 15 | LEF1 |
| 7 | SLC22A4 | 15 | CD3G |
| 7 | TRH | 15 | LDLRAD4 |
| 7 | TTC7B | 15 | AEBP1 |
| 7 | HOXB2 | 15 | SLC4A4 |
| 7 | ME3 | 15 | GALNT2 |
| 7 | CHRM3 | 15 | KIF3A |
| 7 | MECOM | 15 | PARP8 |
| 7 | MEIS1 | 15 | RIMS3 |

| | | | |
|---|-----------------------------------|----|-----------------------------------|
| 7 | PHTF1 | 15 | FYB |
| 7 | LOC100505573 | 15 | YPEL1 |
| 7 | LATS2 | 15 | KIAA0226L |
| 7 | MBOAT7 | 15 | NEDD9 |
| 7 | LPCAT2 | 15 | CD72 |
| 7 | AAED1 | 15 | ATM |
| 7 | WHAMMP2///WHAMMP3 | 15 | IPCEF1 |
| 7 | SEMA4C | 15 | APBA2 |
| 7 | KIT | 15 | SERHL2 |
| 7 | HOPX | 15 | CD79A |
| 7 | HOXA3 | 15 | EBF1 |
| 7 | HOXA5 | 15 | CDH2 |
| 8 | C11orf21 | 16 | KBTBD6 |
| 8 | GSAP | 16 | FRMD4A |
| 8 | GATM | 16 | DLEU2 |
| 8 | STAP1 | 16 | NAPEPLD |
| 8 | LYN | 16 | EVL |
| 8 | ALCAM | 16 | LOC100506776 |
| 8 | VCL | 16 | SLC4A7 |
| 8 | MEF2C | 16 | NREP |
| 8 | BTK | 16 | KATNB1 |
| 8 | MGLL | 16 | AHRR |
| 8 | B3GNT7 | 16 | FANCA |
| 8 | FAM49A | 16 | POLQ |
| 8 | PLXNB2 | 16 | ZNF519 |
| 8 | PHACTR1 | 16 | LOC100507312 |
| 8 | PLEK | 16 | ASPM |
| 8 | BASP1 | 16 | ESCO2 |
| 8 | IRF5 | 16 | FAM111B |
| 8 | DUSP6 | 16 | ATAD2 |
| 8 | CD300LF | 16 | RHOU |
| 8 | RAB31 | 16 | LYST |
| 8 | PIK3AP1 | 16 | LOC284757 |
| 8 | MYO1F | 16 | GTf2H2B |
| 8 | CTBP2 | 16 | TXLNG2P |
| 8 | CDC42BPA | 16 | DSERG1 |
| 8 | IPO11//LRRC70 | 16 | PLGLB1///PLGLB2 |
| 8 | DUSP10 | 16 | SEPT1 |
| 8 | C1orf186 | 16 | HOMER1 |
| 8 | RPS11 | 16 | RRN3P3 |
| 8 | BHLHE40 | 16 | UGP2 |
| 8 | SPRY1 | 16 | LOC100996511 |
| 8 | NFIL3 | 16 | MIR181A2HG |
| 8 | KLF11 | 16 | GOLGA2P5 |
| 8 | SPRY2 | 16 | OTTHUMG00000183913//RP11-93209.10 |
| 8 | CD69 | 16 | EPM2AIP1 |
| 8 | EGR1 | 16 | C1orf132 |
| 8 | DUSP5 | 16 | CD28 |
| 8 | FXYS | 16 | DHFR1 |
| 8 | CD44 | 16 | SUV420H1 |
| 8 | OTTHUMG00000176545//RP11-510116.3 | 16 | AKAP2///PALM2-AKAP2 |
| 8 | MAML3 | 16 | TRAF3IP3 |
| 8 | LAPTM5 | 16 | SATB1 |
| 8 | CRIM1 | 16 | INPP4A |
| 8 | IGHM | 16 | ITGAL |
| 8 | SNX9 | 16 | CDC25B |
| 8 | FLT3 | 16 | SH3KBP1 |
| 8 | HLX | 16 | NDST3 |
| 8 | TCF4 | 16 | HNRNPR |
| 8 | NFKBIZ | 16 | ARHGAP19 |
| 8 | CAPN2 | 16 | GRSF1 |
| 8 | MAP3K8 | 16 | CEP70 |
| 8 | ADAM28 | 16 | WDR7 |
| 8 | MYLIP | 16 | CEP78 |
| 8 | KLF9 | 16 | STAU2 |
| 8 | LILRA2 | 16 | TBCD |
| 8 | HPGDS | 16 | LRRC1 |
| 8 | CD302//LY75-CD302 | 16 | WBP1L |
| 8 | HHEX | 16 | SELPLG |
| 8 | TNS3 | 16 | JPH1 |
| 8 | SH3RF1 | 16 | PRKCA |
| 8 | ANKRD33B | 16 | BIK |

8 NLRP3
 8 GNG11
 8 ARSD
 8 FNDC3B
 8 SLC25A13
 8 GOLIM4
 8 TIMP1
 8 NEK3
 8 C10orf54
 8 DUSP3
 8 TNFRSF1B
 8 IL18
 8 ITPRIPL2
 8 CEBPD
 8 FAM46A
 8 CTNNA1
 8 LMO2
 8 TNFRSF10B
 8 CFD
 8 SLC7A5
 8 BATF
 8 LOC729680
 8 EIF4EBP1
 8 TRIM6
 8 MGST1
 8 MPO
 8 SLC22A16
 8 FAH
 8 BCAT1
 8 MTHFD1L
 8 TNFAIP2
 8 TNFSF13B
 8 CD33
 8 PLA2G4A
 8 MOB3B
 8 KCNK17
 8 GCSAML
 8 NFE2
 8 HSH2D
 8 LAT2
 8 P2RY2
 8 SIGLEC17P
 8 ACY3
 8 S100Z
 8 PAM
 8 CD34
 8 C11orf74
 8 PYGL
 8 CLDN10
 8 CYB561
 8 CLEC11A
 8 MYCN
 8 SPON1
 8 ANGPT1
 8 PTPRD
 8 SPINK2
 8 KIAA0125
 8 EFHC2
 8 MCTP2
 8 AP001171.1.1///OTTHUMG00000074606
 8 ERG
 8 IGFBP7
 8 ARHGEF40
 8 MYLK
 8 LYL1
 8 PRAM1
 8 SERPINB1
 8 ATP8B4
 8 XBP1
 8 LAMC1
 8 C19orf77
 8 MAP7

16 OTTHUMG00000178927///RP11-196G18.23
 16 CHST2
 16 STS
 16 RHBDD1
 16 OXNAD1
 16 MNS1
 16 PSRC1
 16 CDC25C
 16 LOC100288637
 16 FBXO43
 16 MYBL1
 16 LEF1-AS1
 16 NEIL3
 16 LOC100506100
 16 KIAA0922
 16 MTA3
 16 CEP128
 16 ITPR2
 16 TCF5
 16 STK11IP
 16 ZNF280D
 16 APOLD1
 16 ZEB1-AS1
 16 CD2
 16 LOC100507600
 16 CAPSL
 16 MGAT4A
 16 NDFIP2
 16 HIVEP3
 16 PCSK5
 16 IGHG1
 16 IGH///IGHA1///IGHA2///IGHD///IGHG1///IGHG3///IGHG4///IGHM///IGHV3-23///IGHV4-31
 16 MAP1A
 16 RCAN1
 16 SH2D1A
 16 PBK
 16 IL7R
 16 ADA
 16 LCK
 16 LIG4
 16 CDKN2C
 16 E2F2
 16 NETO2
 16 GAS7
 16 FXYD2
 16 IGF2R

8 PROSER2
8 HTR1F
8 SPARC
8 SHANK3
8 F2RL1
8 ZC3H12C
8 CNKSR3
8 MN1
8 TUSC1
8 C1QTNF4
8 C9orf43
8 KIF13A
8 BAALC
8 PROM1
8 IQCJ-SCHIP1///SCHIP1
8 IL1B
8 B3GNT5
8 NIPAL2
8 DEPTOR
8 STOM
8 SERPINB6
8 CREG1
8 CD63
8 C17orf58
8 MYO5C
8 RAB3D
8 ASB9
8 SETD9
8 CYTL1
8 DPPA4
8 KCTD15
8 PDGFC
8 MATR3
8 CCND2
8 CPXM1
8 SLC39A8
8 SERPINE2
8 LAPTM4B
8 KIAA1211
8 STON2
8 LDLRAD3
8 KCNQ5
8 SLC27A2
8 KIF9
8 SLC35F2

Table S2: Gene set enrichment on gene expression of murine T-cell development, Figure 2B

Gene Signatures #7 and #8

| Gene Symbol | Rank in gene list | Running ES | Core enrichment |
|-------------|-------------------|------------|-----------------|
| RAB31 | 0 | 0.023 | Yes |
| PLEK | 2 | 0.043 | Yes |
| HHEX | 4 | 0.061 | Yes |
| KIT | 9 | 0.079 | Yes |
| LYN | 10 | 0.096 | Yes |
| MEF2C | 12 | 0.113 | Yes |
| BTK | 14 | 0.131 | Yes |
| LYL1 | 15 | 0.147 | Yes |
| ANGPT1 | 19 | 0.164 | Yes |
| CD34 | 28 | 0.179 | Yes |
| PLA2G4A | 32 | 0.194 | Yes |
| KIF13A | 45 | 0.207 | Yes |
| PLXNB2 | 54 | 0.220 | Yes |
| ERG | 63 | 0.233 | Yes |
| MYO1F | 69 | 0.245 | Yes |
| PYGL | 72 | 0.257 | Yes |
| CTBP2 | 82 | 0.269 | Yes |
| B3GNT5 | 99 | 0.280 | Yes |
| CABLES1 | 101 | 0.291 | Yes |
| IRF5 | 103 | 0.303 | Yes |
| CD44 | 108 | 0.314 | Yes |
| PIK3AP1 | 114 | 0.325 | Yes |
| VCL | 126 | 0.336 | Yes |
| MYCN | 133 | 0.346 | Yes |
| MPO | 142 | 0.357 | Yes |
| ALCAM | 148 | 0.367 | Yes |
| NFE2 | 171 | 0.377 | Yes |
| TCF4 | 284 | 0.381 | Yes |
| SNX9 | 290 | 0.389 | Yes |
| ZC3H12C | 330 | 0.396 | Yes |
| IL18 | 335 | 0.404 | Yes |
| HOXA5 | 344 | 0.412 | Yes |
| SLC25A13 | 354 | 0.420 | Yes |
| FAM69B | 355 | 0.428 | Yes |
| CTNNA1 | 360 | 0.436 | Yes |
| FAM49A | 370 | 0.444 | Yes |
| TRIM6 | 386 | 0.451 | Yes |
| CNKSR3 | 392 | 0.459 | Yes |
| LMO2 | 394 | 0.467 | Yes |
| LAT2 | 452 | 0.472 | Yes |
| HLX | 467 | 0.479 | Yes |
| P2RY2 | 501 | 0.485 | Yes |
| IL1RAP | 519 | 0.491 | Yes |
| EIF4EBP1 | 523 | 0.498 | Yes |
| ICA1 | 539 | 0.505 | Yes |
| MEIS1 | 573 | 0.510 | Yes |
| LPCAT2 | 595 | 0.516 | Yes |
| ITPR1PL2 | 613 | 0.523 | Yes |
| HPGDS | 628 | 0.529 | Yes |
| FLT3 | 631 | 0.536 | Yes |
| CKAP4 | 636 | 0.542 | Yes |
| PAM | 649 | 0.549 | Yes |
| GATM | 669 | 0.555 | Yes |
| ARMCX1 | 717 | 0.559 | Yes |
| HMGA2 | 755 | 0.564 | Yes |
| GSTM3 | 763 | 0.570 | Yes |
| GOLIM4 | 779 | 0.576 | Yes |
| CD33 | 892 | 0.577 | Yes |
| NAP1L3 | 901 | 0.582 | Yes |
| MAML3 | 963 | 0.585 | Yes |
| MGST1 | 1076 | 0.586 | Yes |
| GBP5 | 1077 | 0.592 | Yes |
| SEMA4C | 1117 | 0.595 | Yes |
| DUSP3 | 1147 | 0.599 | Yes |
| MSRB3 | 1181 | 0.603 | Yes |
| IRAK3 | 1227 | 0.607 | Yes |
| RAB3D | 1254 | 0.611 | Yes |

Gene Signatures #2, #15 and #16

| Gene Symbol | Rank in gene list | Running ES | Core enrichment |
|-------------|-------------------|------------|-----------------|
| CD72 | 21 | 0.021 | No |
| PARP8 | 369 | 0.016 | No |
| TFDP2 | 399 | 0.026 | No |
| ACPL2 | 587 | 0.027 | No |
| SLC4A7 | 818 | 0.025 | No |
| PAFAH2 | 851 | 0.032 | No |
| STAU2 | 1238 | 0.022 | No |
| KIF3A | 1913 | -0.002 | No |
| ARRDC4 | 1914 | 0.003 | No |
| CAPSL | 2025 | 0.004 | No |
| FAM69A | 2152 | 0.004 | No |
| SH3TC1 | 2312 | 0.002 | No |
| FRMD4A | 2418 | 0.002 | No |
| SLC16A6 | 2503 | 0.003 | No |
| SYNE2 | 2504 | 0.007 | No |
| EBF1 | 3291 | -0.023 | No |
| STK11IP | 3359 | -0.023 | No |
| ALOX5 | 3451 | -0.023 | No |
| RASSF6 | 3470 | -0.021 | No |
| NMT2 | 3613 | -0.024 | No |
| IL10RA | 3730 | -0.026 | No |
| CEP70 | 3799 | -0.025 | No |
| PGM2L1 | 3814 | -0.023 | No |
| UGP2 | 4240 | -0.039 | No |
| GRK5 | 4308 | -0.039 | No |
| ATM | 4523 | -0.046 | No |
| RUFY3 | 4632 | -0.049 | No |
| DUSP16 | 4635 | -0.047 | No |
| CD79A | 4710 | -0.048 | No |
| GRSF1 | 4822 | -0.051 | No |
| DOCK9 | 4974 | -0.055 | No |
| HNRNP9 | 4991 | -0.054 | No |
| PLEKHG1 | 5034 | -0.054 | No |
| CEP78 | 5175 | -0.058 | No |
| MTA3 | 5189 | -0.057 | No |
| INPP4A | 5287 | -0.060 | No |
| ARHGAP29 | 5378 | -0.062 | No |
| PALLD | 5420 | -0.062 | No |
| RHBDD1 | 5450 | -0.062 | No |
| LRRN3 | 5481 | -0.062 | No |
| NINL | 5490 | -0.061 | No |
| CR2 | 5603 | -0.064 | No |
| RAPGEF5 | 5719 | -0.068 | No |
| OXNAD1 | 5783 | -0.070 | No |
| EFNB2 | 5905 | -0.074 | No |
| ITGA6 | 6036 | -0.078 | No |
| TBCD | 6251 | -0.087 | No |
| RDH10 | 6597 | -0.101 | No |
| JPH1 | 6862 | -0.113 | No |
| ZC3H6 | 6869 | -0.112 | No |
| SLFN5 | 6938 | -0.115 | No |
| ID1 | 6951 | -0.115 | No |
| CRIP1 | 7203 | -0.126 | No |
| FHIT | 7332 | -0.131 | No |
| FAIM3 | 7477 | -0.138 | No |
| PSRC1 | 7753 | -0.150 | No |
| BCL2L11 | 7796 | -0.151 | No |
| CHST2 | 8127 | -0.166 | No |
| YPEL1 | 8448 | -0.179 | No |
| SCN7A | 8951 | -0.201 | No |
| AGBL3 | 9194 | -0.210 | No |
| STAG3 | 9339 | -0.216 | No |
| AHRR | 9658 | -0.229 | No |
| NDFIP2 | 9800 | -0.234 | No |
| RCAN1 | 9909 | -0.238 | No |
| NDST3 | 10158 | -0.247 | No |
| NSG1 | 10250 | -0.250 | No |

| | | | | | | | |
|-----------|------|-------|-----|----------|-------|--------|-----|
| KIF9 | 1313 | 0.613 | Yes | POU2AF1 | 10391 | -0.255 | No |
| CD63 | 1317 | 0.618 | Yes | APBA2 | 10449 | -0.256 | No |
| FNDC3B | 1320 | 0.623 | Yes | GAL3ST4 | 10611 | -0.262 | No |
| FAM46A | 1328 | 0.628 | Yes | GAS7 | 10758 | -0.267 | No |
| TNFAIP2 | 1332 | 0.633 | Yes | CX3CR1 | 10866 | -0.271 | No |
| SPRY2 | 1338 | 0.638 | Yes | S1PR1 | 11101 | -0.279 | No |
| TFPI | 1404 | 0.640 | Yes | ADAMTS1 | 11356 | -0.289 | No |
| MPL | 1418 | 0.645 | Yes | EVL | 11375 | -0.288 | No |
| MECOM | 1498 | 0.646 | Yes | IKZF2 | 11638 | -0.298 | No |
| KLF11 | 1547 | 0.648 | Yes | TSHR | 11754 | -0.302 | No |
| CDC42BPA | 1691 | 0.647 | Yes | PCSK5 | 11874 | -0.305 | No |
| GRB10 | 1809 | 0.646 | Yes | SLC7A3 | 12006 | -0.309 | No |
| LATS2 | 1823 | 0.650 | Yes | NETO2 | 12466 | -0.328 | No |
| TNS3 | 1912 | 0.650 | Yes | TMOD2 | 12502 | -0.328 | No |
| SHANK3 | 1915 | 0.654 | Yes | GLRB | 12787 | -0.338 | No |
| MN1 | 2041 | 0.653 | Yes | SUV420H1 | 12849 | -0.339 | No |
| MCTP2 | 2067 | 0.656 | Yes | PLCL1 | 13625 | -0.371 | No |
| MAP3K8 | 2115 | 0.657 | Yes | NEIL3 | 13626 | -0.369 | No |
| MYO5C | 2194 | 0.658 | Yes | SYNM | 13747 | -0.372 | No |
| SNTB1 | 2261 | 0.659 | Yes | WT1 | 14113 | -0.386 | No |
| NFIL3 | 2317 | 0.660 | Yes | MBD5 | 14171 | -0.387 | No |
| FXYD5 | 2391 | 0.660 | Yes | ID2 | 14366 | -0.393 | No |
| CPXM1 | 2433 | 0.662 | Yes | PCDH9 | 14484 | -0.396 | No |
| DUSP6 | 2633 | 0.657 | Yes | ATAD2 | 14940 | -0.414 | No |
| CRIM1 | 2660 | 0.659 | Yes | SCN3A | 15277 | -0.426 | No |
| NFKBIZ | 2740 | 0.659 | Yes | PCDH10 | 15373 | -0.428 | No |
| RPS11 | 2789 | 0.660 | Yes | DLEU2 | 15620 | -0.436 | No |
| GYPC | 2795 | 0.663 | Yes | HHIP | 15683 | -0.436 | No |
| FOSL2 | 2804 | 0.666 | Yes | MME | 15763 | -0.437 | No |
| LAPTM4B | 2869 | 0.666 | Yes | NAPEPLD | 15833 | -0.438 | No |
| GNG11 | 2870 | 0.669 | Yes | CDH2 | 15903 | -0.438 | No |
| MYLK | 2886 | 0.672 | Yes | CYP2U1 | 15979 | -0.439 | No |
| TNFRSF10B | 3083 | 0.666 | Yes | FANCA | 16117 | -0.442 | No |
| BATF | 3145 | 0.666 | Yes | VCAN | 16153 | -0.441 | No |
| SDPR | 3180 | 0.668 | Yes | PBK | 16185 | -0.440 | No |
| UCHL1 | 3306 | 0.665 | Yes | IL7R | 16226 | -0.439 | No |
| SOC52 | 3468 | 0.660 | Yes | FXYD2 | 16538 | -0.450 | No |
| HSH2D | 3496 | 0.662 | Yes | IPCEF1 | 16750 | -0.457 | No |
| CPA3 | 3506 | 0.664 | Yes | AEBP1 | 16869 | -0.459 | No |
| FHL1 | 3516 | 0.666 | Yes | SLC4A4 | 16978 | -0.461 | No |
| CD69 | 3519 | 0.669 | Yes | POLQ | 17262 | -0.471 | No |
| TPM1 | 3521 | 0.671 | Yes | HOMER1 | 17373 | -0.472 | No |
| CAPN2 | 3524 | 0.674 | Yes | ASPM | 17379 | -0.470 | No |
| BHLHE40 | 3525 | 0.676 | Yes | PTPRK | 17490 | -0.472 | No |
| DUSP5 | 3586 | 0.676 | No | PPP1R1C | 17543 | -0.471 | No |
| EFHC2 | 3800 | 0.669 | No | RHOA | 17688 | -0.474 | No |
| NLRP3 | 3826 | 0.670 | No | AQP3 | 17833 | -0.478 | No |
| MATR3 | 3861 | 0.671 | No | BIK | 17936 | -0.479 | No |
| BASP1 | 3921 | 0.671 | No | NKD2 | 17946 | -0.476 | No |
| ATP8B4 | 4068 | 0.666 | No | LRRC1 | 17995 | -0.476 | No |
| NUDT11 | 4182 | 0.663 | No | MAL | 18144 | -0.479 | No |
| SLC22A15 | 4250 | 0.662 | No | MYBL1 | 18292 | -0.482 | No |
| MBOAT7 | 4381 | 0.658 | No | FBXO43 | 18398 | -0.484 | No |
| DRAM1 | 4441 | 0.658 | No | GALNT2 | 18506 | -0.485 | No |
| SLC27A2 | 4640 | 0.651 | No | NEGR1 | 18581 | -0.485 | No |
| FAM171B | 4755 | 0.647 | No | LRRC28 | 18595 | -0.482 | No |
| PDGFC | 5119 | 0.633 | No | P2RX5 | 18891 | -0.492 | No |
| RIPK2 | 5382 | 0.622 | No | SMPD3 | 18946 | -0.491 | No |
| FAH | 5468 | 0.620 | No | MGAT4A | 19097 | -0.494 | No |
| CCND2 | 5470 | 0.621 | No | CCNG2 | 19220 | -0.496 | No |
| EGR1 | 5727 | 0.611 | No | CCR8 | 19280 | -0.495 | No |
| TMEM220 | 5735 | 0.611 | No | ESCO2 | 19389 | -0.497 | No |
| SLC39A8 | 5887 | 0.605 | No | RBMS3 | 19530 | -0.499 | No |
| HOXB2 | 5890 | 0.606 | No | BACH2 | 19576 | -0.498 | No |
| ACY3 | 5923 | 0.606 | No | NEDD9 | 19624 | -0.496 | No |
| CEBPD | 6050 | 0.601 | No | DNTT | 19641 | -0.493 | No |
| TUSC1 | 6088 | 0.600 | No | CDKN2C | 19717 | -0.493 | No |
| CYB561 | 6104 | 0.600 | No | CDC25C | 19833 | -0.494 | No |
| STOM | 6658 | 0.576 | No | ZBTB1 | 19978 | -0.497 | No |
| STAP1 | 6893 | 0.566 | No | PDE4D | 20213 | -0.503 | Yes |
| KCNQ5 | 7180 | 0.554 | No | EPM2AIP1 | 20231 | -0.500 | Yes |
| HOPX | 7376 | 0.545 | No | TANC1 | 20370 | -0.502 | Yes |

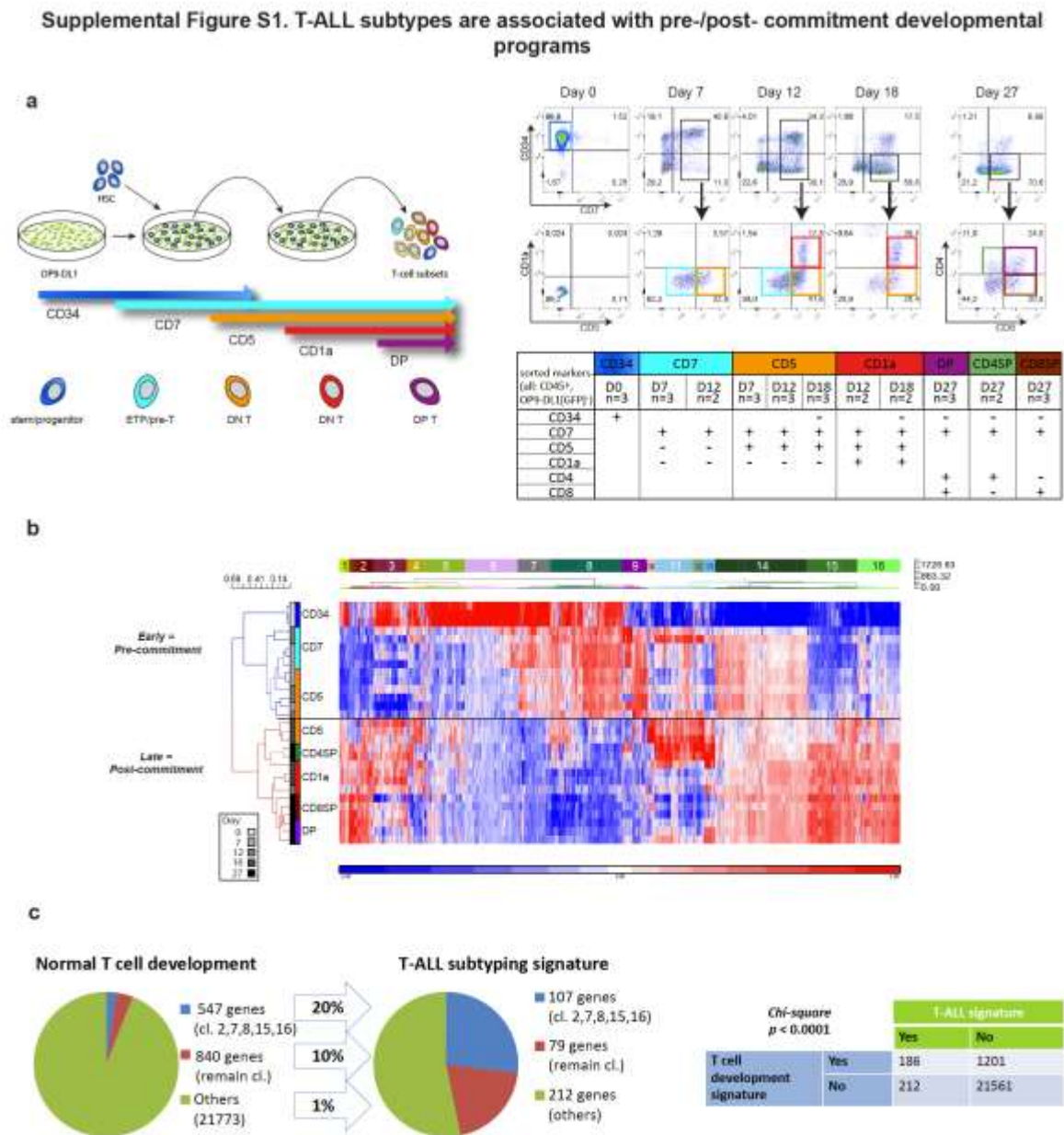
| | | | | | | | |
|-----------|-------|--------|----|----------|-------|--------|-----|
| ANKRD33B | 7426 | 0.543 | No | WDR7 | 20405 | -0.499 | Yes |
| BSPRY | 7520 | 0.539 | No | FAM63B | 20481 | -0.499 | Yes |
| MTHFD1L | 7705 | 0.531 | No | IRS1 | 20511 | -0.496 | Yes |
| PHACTR1 | 7887 | 0.523 | No | TCFL5 | 20559 | -0.494 | Yes |
| LDLRAD3 | 7926 | 0.522 | No | NOTCH3 | 20582 | -0.491 | Yes |
| B3GNT7 | 8251 | 0.508 | No | RIMS3 | 20758 | -0.494 | Yes |
| LAMC1 | 8390 | 0.502 | No | COX6A1 | 20841 | -0.493 | Yes |
| SPRY1 | 8581 | 0.494 | No | HDAC4 | 20879 | -0.491 | Yes |
| SLC22A4 | 8690 | 0.490 | No | SEC31B | 21056 | -0.494 | Yes |
| RAB13 | 8804 | 0.486 | No | APOLD1 | 21084 | -0.490 | Yes |
| TNFRSF1B | 9005 | 0.477 | No | CDC25B | 21224 | -0.492 | Yes |
| SPARC | 9169 | 0.471 | No | GAS2 | 21363 | -0.493 | Yes |
| RAMP1 | 9170 | 0.471 | No | ARHGAP19 | 21464 | -0.493 | Yes |
| XBP1 | 9206 | 0.470 | No | NCALD | 21545 | -0.491 | Yes |
| MYLIP | 9290 | 0.467 | No | HS3ST3B1 | 21595 | -0.488 | Yes |
| CHRM3 | 9431 | 0.462 | No | ITGAL | 21773 | -0.490 | Yes |
| CD300LF | 9453 | 0.462 | No | PTCRA | 21814 | -0.487 | Yes |
| CMBL | 9680 | 0.452 | No | SEPT9 | 21820 | -0.481 | Yes |
| TNFRSF13B | 9730 | 0.451 | No | LYST | 21843 | -0.477 | Yes |
| NIPAL2 | 9768 | 0.450 | No | MNS1 | 21880 | -0.473 | Yes |
| IL1B | 9901 | 0.445 | No | PRKCA | 21923 | -0.469 | Yes |
| ZBTB8A | 10028 | 0.441 | No | SLC16A10 | 21982 | -0.465 | Yes |
| SORL1 | 10148 | 0.436 | No | SPRED2 | 22073 | -0.463 | Yes |
| IGFBP7 | 10154 | 0.437 | No | LBH | 22103 | -0.459 | Yes |
| LAPTM5 | 10199 | 0.436 | No | TUBB2A | 22278 | -0.460 | Yes |
| SLC35F2 | 10335 | 0.431 | No | RAG2 | 22294 | -0.454 | Yes |
| SERPINE2 | 10478 | 0.425 | No | CD2 | 22351 | -0.449 | Yes |
| C1QTNF4 | 11047 | 0.402 | No | ADA | 22414 | -0.445 | Yes |
| SLC7A5 | 11283 | 0.392 | No | E2F2 | 22506 | -0.441 | Yes |
| S100Z | 11421 | 0.387 | No | ARPP21 | 22519 | -0.434 | Yes |
| PVRL2 | 11712 | 0.376 | No | SBK1 | 22543 | -0.427 | Yes |
| NEK3 | 12156 | 0.358 | No | PLCG1 | 22551 | -0.420 | Yes |
| PRSS2 | 12402 | 0.348 | No | CD96 | 22585 | -0.413 | Yes |
| STON2 | 12861 | 0.329 | No | TNRC6C | 22588 | -0.405 | Yes |
| SPINK2 | 13304 | 0.311 | No | SEPLG | 22617 | -0.398 | Yes |
| ADM | 13365 | 0.310 | No | CCR9 | 22628 | -0.390 | Yes |
| CXCR7 | 13385 | 0.311 | No | UBASH3A | 22639 | -0.383 | Yes |
| PROM1 | 13415 | 0.311 | No | CD8A | 22641 | -0.374 | Yes |
| CLEC11A | 13474 | 0.310 | No | TRERF1 | 22661 | -0.367 | Yes |
| ADAM28 | 13623 | 0.305 | No | SIPA1L2 | 22711 | -0.360 | Yes |
| ASB9 | 13845 | 0.297 | No | DGKA | 22740 | -0.352 | Yes |
| CFD | 14062 | 0.289 | No | ID3 | 22806 | -0.345 | Yes |
| TTCT7B | 14669 | 0.264 | No | CD28 | 22808 | -0.335 | Yes |
| CRYGD | 14848 | 0.258 | No | SATB1 | 22813 | -0.325 | Yes |
| PTPRD | 15343 | 0.238 | No | SH2D1A | 22825 | -0.316 | Yes |
| SLC22A16 | 15517 | 0.232 | No | LIG4 | 22828 | -0.305 | Yes |
| BAALC | 17146 | 0.163 | No | SH3KBP1 | 22829 | -0.295 | Yes |
| NRIP3 | 17299 | 0.158 | No | HIVEP3 | 22835 | -0.285 | Yes |
| MGLL | 17390 | 0.157 | No | IKZF3 | 22863 | -0.275 | Yes |
| TRH | 17667 | 0.147 | No | CAMK4 | 22868 | -0.264 | Yes |
| KBTBD11 | 18498 | 0.113 | No | GRAP2 | 22873 | -0.253 | Yes |
| CYTL1 | 18655 | 0.108 | No | THEMIS | 22876 | -0.241 | Yes |
| PDE4DIP | 18952 | 0.098 | No | GATA3 | 22878 | -0.230 | Yes |
| F2RL1 | 19071 | 0.095 | No | ITK | 22887 | -0.218 | Yes |
| PROK2 | 19095 | 0.097 | No | RASGRP1 | 22899 | -0.206 | Yes |
| KCTD15 | 19117 | 0.098 | No | FYB | 22904 | -0.193 | Yes |
| HTR1F | 19175 | 0.098 | No | TCF7 | 22907 | -0.180 | Yes |
| SPON1 | 19237 | 0.098 | No | IGF2R | 22910 | -0.166 | Yes |
| TIMP1 | 19830 | 0.075 | No | LCK | 22911 | -0.153 | Yes |
| SLC2A5 | 20039 | 0.069 | No | ITPR2 | 22915 | -0.139 | Yes |
| PRAM1 | 20423 | 0.055 | No | ETS1 | 22918 | -0.125 | Yes |
| SYNGR1 | 20540 | 0.053 | No | CD247 | 22923 | -0.110 | Yes |
| DPPA4 | 20676 | 0.050 | No | ZAP70 | 22928 | -0.095 | Yes |
| ME3 | 20966 | 0.041 | No | BCL11B | 22932 | -0.080 | Yes |
| EPCAM | 20977 | 0.043 | No | LEF1 | 22933 | -0.064 | Yes |
| GPR27 | 21492 | 0.025 | No | LAT | 22938 | -0.046 | Yes |
| TCTEX1D1 | 21586 | 0.024 | No | CD3E | 22939 | -0.026 | Yes |
| PHTF1 | 21786 | 0.020 | No | CD3G | 22941 | 0.000 | Yes |
| BCAT1 | 22231 | 0.005 | No | | | | |
| SH3RF1 | 22848 | -0.014 | No | | | | |
| DUSP10 | 22869 | -0.007 | No | | | | |
| GBP4 | 22905 | 0.002 | No | | | | |

Table S3: Gene Symbols in Gene Clusters I-V from Figure 2C

| Gene Cluster I | Gene Cluster II | Gene Cluster III | Gene Cluster IV | Gene Cluster V |
|-------------------|----------------------|-----------------------------------|-----------------|---------------------|
| SNTB1 | CRIM1 | KATNB1 | RCAN1 | IPP |
| DRAM1 | CYTL1 | OTTHUMG00000176181///RP11-119F7.5 | PAFAH2 | SLC16A10 |
| IL1RAP | TNS3 | GRSF1 | SPRED2 | STAG3 |
| SDPR | PROM1 | STAU2 | GAS2 | GSDMB |
| ARHGAP29 | BAALC | TXLNG2P | CX3CR1 | CD8A |
| HLX | LMO2 | POU2AF1 | NLGN4X | CD8B///LOC100996919 |
| MEIS1 | LAPTM4B | GATA3 | NINL | LYST |
| NFIL3 | BASP1 | BATF | SMPD3 | IRS1 |
| KIT | FNDC3B | SPON1 | SH3TC1 | AQP3 |
| IRAK3 | FAM46A | DUSP6 | CD1E | NCALD |
| TFPI | CREG1 | SCN3A | CD1A | GAL3ST4 |
| SLC22A4 | BSPRY | HSPA1A///HSPA1B | CD1B | CCR9 |
| CPA3 | HOXA5 | ADAMT51 | CD1D | CTC1 |
| RIPK2 | FLT3 | CDH2 | CD1C | PTPRK |
| PVRL2 | RAB13 | NDST3 | AEBP1 | SATB1 |
| BHLHE40 | CD302///LY75-CD302 | RIMS3 | E2F2 | CD79A |
| PHACTR1 | IL1B | TBCD | GT2H2B | SEC31B |
| KLF9 | EFHC2 | APBA2 | DOCK9 | PARP8 |
| FOSL2 | SLC35F2 | NSG1 | DSERG1 | TLR5 |
| ARHGEF40 | MCTP2 | CR2 | CDKN2D | PLCG1 |
| CEBPD | ADAM28 | SLC4A7 | CD72 | IKZF2 |
| HBD | PTF1 | ZBTB1 | PCDH9 | IKZF3 |
| CHRM3 | IL18 | WDR7 | EFNB2 | INPP4A |
| RAMP1 | GNG11 | NMT2 | CYP2U1 | ETS1 |
| CRYGD | MN1 | SUV420H1 | GPR27 | CD28 |
| SYNGR1 | HMGA2 | WBP1L | MAML3 | IL10RA |
| WHAMMP2///WHAMMP3 | PDGFC | KIF3A | FRMD4A | PLCL1 |
| MECOM | SLC27A2 | SELPLG | RPS11 | GRK5 |
| CDC42BPA | KLF11 | NREP | MAP1A | CRIP1 |
| CFD | ATP8B4 | UGP2 | CCR8 | NETO2 |
| C11orf21 | MAP3K8 | MATR3///SNHG4 | HIVEP3 | MBD5 |
| CD69 | SLC25A13 | HDAC4 | BCL2L11 | AGBL3 |
| TUBB2A | ARMCX1 | SRPK2 | SLC16A6 | PRKCA |
| S1PR1 | PAM | EPM2AIP1 | BIK | TMOD2 |
| SERPINE2 | F2RL1 | ITPR2 | DUSP5 | CD3E |
| ELK3 | GSAP | ZNF280D | DUSP3 | DGKA |
| PDE4D | DPPA4 | ITGA6 | SCN7A | TRBC1 |
| EGR1 | CXCR7 | CCNG2 | HOXA3 | GRAP2 |
| ID2 | NAP1L3 | FANCA | RAB3D | FAM63B |
| MYLIP | GATM | GOLGA2P5 | MYCN | FYB |
| SPRY2 | ASB9 | ARPP21 | TNFRSF10B | LEF1 |
| LAPTM5 | PLA2G4A | POLQ | ICA1 | SEPT9 |
| CD44 | IQCJ-SCHIP1///SCHIP1 | STS | NRIP3 | LIG4 |
| TNFRSF1B | RAB31 | CDKN2C | MOB3B | SYNE2 |
| KBTBD11 | NUDT11 | PSRC1 | ARSD | TCF7 |
| HOXB2 | XBP1 | DLEU2 | PDE4DIP | SH2D1A |
| SEMA4C | EPCAM | NEIL3 | UCHL1 | LCK |
| CCND2 | SLC39A8 | ASPM | VCAN | PVRIG |
| FXYD5 | MEST | PBK | ADM | BCL11B |
| GYPC | CD63 | ATAD2 | HS3ST3B1 | EVL |
| FHIT | SPRY1 | APOLD1 | ALOX5 | CD96 |
| DUSP10 | STOM | PTCRA | RBMS3 | RASGRP1 |
| CAPN2 | SORL1 | TSHR | KIF13A | IL7R |
| HOPX | SERPINB1 | RAPGEF5 | KCTD15 | CD3G |
| TIMP1 | SPINK2 | KIAA0226L | IGHG1 | ITK |
| GRB10 | MYO1F | RAG2 | CDC25C | CHST2 |
| CKAP4 | LYL1 | RAG1 | NIPAL2 | CD2 |
| CTNNA1 | CD34 | NOTCH3 | GLRB | ZAP70 |
| TPM1 | MPO | LRRC1 | | CD247 |
| | SERPINB6 | IGF2R | | IL32 |
| | BTX | KIAA0922 | | PCSK5 |
| | CLDN10 | LOC79015 | | HOMER1 |
| | MYO5C | MAL | | NEDD9 |
| | HPGDS | ARHGAP19 | | CHI3L2 |
| | MGLL | MYBL1 | | ITGAL |
| | NFE2 | COX6A1 | | MGAT4A |

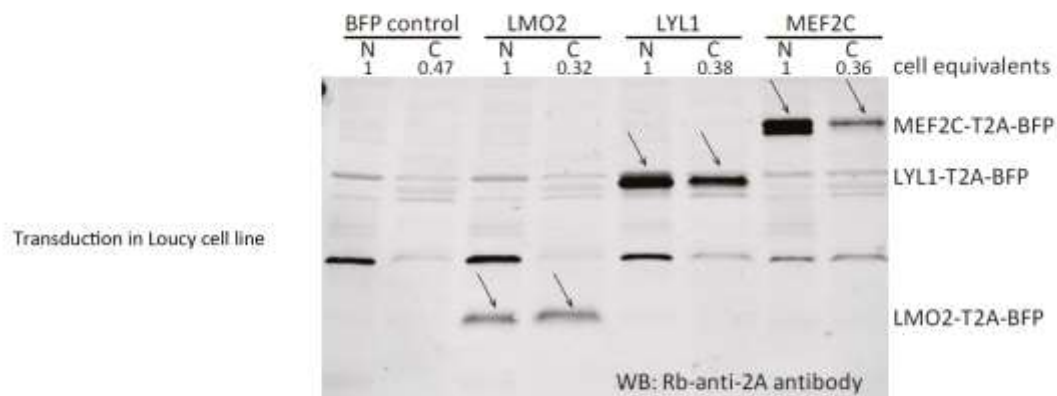
| | | |
|---------------|---------------------|----------|
| FAH | CDC25B | LAT |
| ANGPT1 | YPEL1 | LBH |
| SLC2A5 | MME | ID3 |
| PYGL | AKAP2///PALM2-AKAP2 | FAM69A |
| CTBP2 | GALNT2 | CAMK4 |
| VCL | LDLRAD4 | UBASH3A |
| IGFBP7 | CEP70 | ATM |
| SPARC | DNTT | IPCEF1 |
| DEPTOR | ADA | BACH2 |
| KIAA0125 | MNS1 | FAIM3 |
| PRSS2///PRSS3 | TFDP2 | LRRN3 |
| MYLK | FXVD2 | TRAF3IP3 |
| PRSS2 | SLC4A4 | P2RX5 |
| LILRA2 | RUFY3 | GBP1 |
| LAMC1 | GAS7 | |
| HTR1F | HNRNPR | |
| MPL | IGHM | |
| NLRP3 | PALLD | |
| MBOAT7 | SERHL2 | |
| CYB561 | | |
| TRH | | |
| WT1 | | |
| GSTM3 | | |
| SYNM | | |
| LAT2 | | |
| PTPRD | | |
| MAP7 | | |
| P2RY2 | | |
| ERG | | |
| CD33 | | |
| GOLIM4 | | |
| IRF5 | | |
| TNFAIP2 | | |
| EIF4EBP1 | | |
| ID1 | | |
| FHL1 | | |
| BCAT1 | | |
| SLC7A5 | | |
| FAM49A | | |
| TCF4 | | |
| PLEK | | |
| PLXNB2 | | |
| LYN | | |
| MEF2C | | |
| ALCAM | | |
| HHEX | | |
| STAP1 | | |
| NEK3 | | |
| ME3 | | |
| SOCS2 | | |
| CLEC11A | | |

Chapter 6
Supplementary Figures



Supplemental Figure S1. (a) Schematic representation of the OP9-DL1 co-culture with the sorting strategy of consecutive T-cell differentiation stages. The table displays the sorted populations based on surface markers (CD45, CD34, CD7, CD5, CD1a, CD4, CD8) and number of days in co-culture. All populations were also gated for CD45+GFP- to exclude OP9-DL1 cells that are GFP positive. (b) Hierarchical clustering analysis using 2179 probesets with high variance (top 5%) and log2 expression values (>8). These probesets were summarized into 1387 genes by taking the median of all probesets across a gene. Pearson dissimilarity measure and average linkage were applied to cluster the samples, and Euclidean distance and Ward's method were applied to define 16 distinct gene clusters with similar expression patterns. (c) The left Pie Chart represents all the genes analyzed in the affymetrix gene expression array. In blue is the portion of genes that represent the clusters 2,7,8,15 and 16 (547 genes), in red are 840 genes from the remaining 11 clusters (1387 genes minus the 547 genes) that were filtered based on high variance (>0.8) and log2 expression values (>8), and in green are all the remaining genes in the array (others). The right Pie Chart represents the 435 probeset gene signature used to cluster T-ALL patient samples in Homminga et al, (2011) Cancer Cell. It contains 107 genes (20%) that overlap with the 547 gene signature, 79 genes (10%) that overlap with the remaining 11 clusters, and 212 genes (1%) overlapping the remaining genes in the array (others). At the right is a table used to analyse the similarity between the pre-post commitment signature and the T-ALL gene signature using the Chi-square test.

Supplemental Figure S2. Validation of lentiviral vectors expressing MEF2C, LYL1, or LMO2 in Loucy cells.



Supplemental Figure S2. (a) Cellular distribution of overexpressed MEF2C,LYL1, and LMO2 in LOUCY cells. Western blot of nuclear (N) and cytoplasmic (C) lysates of LOUCY cells transduced with SFFV-promoter-driven fusion transcripts of MEF2C, LYL1, or LMO2 with a blue fluorescence protein (BFP) reporter, separated by the *Thosea asigna* virus 2A (T2A) cleavage recognition peptide. After cleavage of the proteins, the MEF2C, LYL1, or LMO2 retain the majority of the T2A peptide that can serve as an epitope tag and can be recognized by the anti-2A antibody. The BFP reporter that follows the T2A sequence is separated and not recognized by anti-2A.

[illegible]

178

Supplemental Tables

Table S1: Gene Symbols in each of the 4 gene groups (IV) from Figure 1d

Gene Group // Gene Symbol

| | | | | | |
|---|-----------|---|-------------------------------------|---|-------------------|
| 1 | BIK | 2 | NSG1 | 4 | RASGRP1 |
| 1 | CCR9 | 2 | MSRB3 | 4 | PBK |
| 1 | KIAA0226L | 2 | ARRDC4 | 4 | LEF1 |
| 1 | TFPI | 2 | MME | 4 | MAL |
| 1 | LATS2 | 2 | EGR1 | 4 | THEMIS |
| 1 | LMO2 | 2 | NCALD | 4 | CD8A |
| 1 | B3GNT5 | 2 | RPS11 | 4 | AQP3 |
| 1 | MAP3K8 | 2 | CAPN2 | 4 | IL32 |
| 1 | FOSL2 | 2 | LAPTM4B | 4 | LIG4 |
| 1 | BHLHE40 | 2 | PRKCA | 4 | SH2D1A |
| 1 | ID2 | 2 | GALNT2 | 4 | ITGAL |
| 1 | SPINK2 | 2 | NOTCH3 | 4 | CD2 |
| 1 | HOPX | 2 | CR2 | 4 | HHIP |
| 1 | DUSP5 | 2 | RHOU | 4 | SCN3A |
| 1 | HHEX | 2 | NETO2 | 4 | NAP1L3 |
| 1 | CD69 | 2 | SPRY2 | 4 | CD28 |
| 1 | GATA3 | 2 | RAB13 | 4 | SPRED2 |
| 1 | MYLIP | 2 | GAS2 | 4 | TRBC1 |
| 2 | NFIL3 | 2 | MEST | 4 | CRIP1 |
| 2 | IGFBP7 | 2 | FRMD4A | 4 | LCK |
| 2 | LAT2 | 2 | LOC101060391 | 4 | CHI3L2 |
| 2 | P2RX5 | 2 | ID3 | 4 | FHL1 |
| 2 | ZBTB8A | 2 | ID1 | 4 | FXVD2 |
| 2 | GBP1 | 2 | IGHG1 | 4 | LOC728613 |
| 2 | HIVEP3 | 2 | SH3TC1 | 4 | CRNDE |
| 2 | NFKBIZ | 3 | CD1A | 4 | CLEC11A |
| 2 | CCND2 | 3 | RCAN1 | 4 | C11orf96 |
| 2 | SOCS2 | 3 | CD1E | 4 | NUDT11 |
| 2 | NDFIP2 | 3 | CD1B | 4 | TUBB2A |
| 2 | DUSP6 | 3 | OTTHUMG00000017520 /// RP11-144L1.4 | 4 | HSPA1A /// HSPA1B |
| 2 | TNFRSF10B | 3 | DNTT | 4 | HSPA1A /// HSPA1B |
| 2 | TIMP1 | 3 | CEP128 | 4 | PGM2L1 |
| 2 | BCAT1 | 3 | TSHR | 4 | BCL2L11 |
| 2 | NREP | 3 | FYB | 4 | NEDD9 |
| 2 | TCF4 | 3 | TXLNG2P | 4 | PCDH10 |
| 2 | FHIT | 3 | TXLNG2P | | |
| 2 | KCNK17 | 4 | ARHGAP29 | | |

Curriculum Vitae

Rui Daniel Martins Nunes Mendes was born on July 17th, 1984 in Lisbon, Portugal. After finishing high school in Torres Vedras at Escola Secundaria Madeira de Torres, he studied Molecular Biology and Genetics at the University of Lisbon, where he obtained his Bachelor of Science (BSc.) degree in 2007, and Master of Science (MSc.) degree in 2009. He did his master thesis internship “Non-viral gene delivery to human mesenchymal stem cells using cationic liposomes” at the Stem Cell Bioengineering Laboratory at Instituto Superior Tecnico, under the supervision of Dr. Claudia Lobato da Silva.



In 2009, he moved to the Netherlands and started working as a researcher for the Orthopedics department at Erasmus Medical Center, Rotterdam. In 2011 he started as a PhD student at the department of Pediatric Oncology/Hematology at Erasmus Medical Center under the supervision of Dr. Jules Meijerink, focusing on the molecular mechanisms that lead to T-cell acute lymphoblastic leukemia.

In 2016, he started working as a Scientist for a Biopharmaceutical company in Leiden, ProQR Therapeutics N.V.

List of publications

Mendes RD, Canté-Barrett K, Pieters R, Meijerink JP. The relevance of PTEN-AKT in relation to NOTCH1-directed treatment strategies in T-cell acute lymphoblastic leukemia. *Haematologica*. 2016 Sep;101(9):1010-7.

Canté-Barrett K, **Mendes RD**, Smits WK, van Helsdingen-van Wijk YM, Pieters R, Meijerink JP. Lentiviral gene transfer into human and murine hematopoietic stem cells: size matters. *BMC Res Notes*. 2016 Jun 16;9:312.

Mendes RD, Sarmiento LM, Canté-Barrett K, Zuurbier L, Buijs-Gladdines JG, Póvoa V, Smits WK, Abecasis M, Yunes JA, Sonneveld E, Horstmann MA, Pieters R, Barata JT, Meijerink JP. PTEN micro-deletions in T-cell acute lymphoblastic leukemia are caused by illegitimate RAG-mediated recombination events. *Blood*, 2014. 124(4): p. 567-78

Madeira C, **Mendes RD**, Ribeiro SC, Boura JS, Aires-Barros MR, da Silva CL, Cabral JM. Nonviral gene delivery to mesenchymal stem cells using cationic liposomes for gene and cell therapy. *J Biom Biotech*, 2010; 2010:735349

Ribeiro SC, **Mendes RD**, Madeira C, Monteiro GA, da Silva CL, Cabral JM. A quantitative method to evaluate mesenchymal stem cell lipofection using real-time PCR. *Biotechnol Prog*. 2010 Sep-Oct;26(5):1501-4

PhD Portfolio

Name PhD student: Rui Daniel Martins Nunes Mendes

Erasmus MC department: Pediatric Oncology

PhD period: 2011-2015

Research School: Postgraduate School Molecular Medicine

Promoter: Prof.dr. R. Pieters

Co-promoter: Dr. J.P.P. Meijerink

Seminars/Courses/Workshops

- Course on Laboratory Animal Science (Article 9)
- Basic and Translational Oncology
- Analysis of Microarray Gene Expression Data
- Basic course on R
- Biomedical Scientific English Writing
- Indesign workshop
- Photoshop & Illustrator workshop
- Presenting skills for Scientists
- Research Management for PhDs/Postdocs
- SNPs and Human Diseases
- MolMed Days

Presentations at national and international conferences

- Sophia Research Day, Rotterdam, the Netherlands. June 26, 2015, oral presentation
- Course on molecular aspects of Hematological Disorders, Rotterdam, the Netherlands. June 9-10, 2015, oral presentation
- 1st International Venice Thymus Meeting, San Servolo Island, Italy. April 9-13, 2015, poster presentation
- 3rd Daniel den Hoed day, Rotterdam, the Netherlands. February 18, 2015, oral presentation
- 19th Congress of the European Hematology Association, Milan, Italy. June 12-15, 2014, poster presentation
- Course on molecular aspects of Hematological Disorders, Rotterdam, the Netherlands. June 10-11, 2014, oral presentation
- ESH-EHA Scientific Workshop on T-Cell Acute Lymphoblastic Leukemia, Lisbon, Portugal. Mar 22-24, 2013, poster presentation
- Erasmus Postgraduate School Molecular Medicine Day, Rotterdam, The Netherlands. Feb 13, 2013, poster presentation
- Dutch Hematology Congress, Arnhem, The Netherlands. Jan 23-25, 2013, oral presentation

Acknowledgements

This is the end of a very important chapter in my life, the last days of my life as a PhD student. So far, the biggest adventure in my life, taking into account the challenge of enrolling a PhD program and living abroad. I could not anticipate what was about to come when I decided to embark in this ship and sail through the unknown, in a journey full of ups-and-downs. At this moment, I can say that every single moment was extremely special, not only for all the people that I met along the way, but also because I had the opportunity to get to know myself even better. Some people were crucial to this achievement and I will never be able to express my words of gratitude your friendship, guidance and support.

My appreciation to my promotor Prof. Dr. Rob Pieters for giving me the opportunity to join the Department of Pediatric Oncology and for all the critical thinking and scientific discussions, especially in guiding us into a patient-oriented research.

To Dr. Jules Meijerink, you were such a great supervisor, and the most important person during the last 5 years of my PhD track. More than a supervisor, you were there every moment that I needed your special support. You always heard my complaints and told me the wisest words of encouragement in the most challenging moments. I truly admire you as an outstanding investigator for your extreme passion and dedication to science, and because of your will and drive to find the answers to the most difficult research questions.

To my colleague (and mentor) Dr. Kirsten Canté-Barrett, you were also extremely important during all these years. You were the one that helped me to start working in most of the projects and whether I needed support in the lab, to discuss ideas or interpret new data, you were you were always eager to help me. I learnt a lot from you and I admire the way you are always open to new ideas and promptly ready to apply them if necessary. Besides that, I truly appreciate your scientific input during our meetings and your management skills, especially how you can efficiently manage so many projects at the same time.

To the whole T-ALL research team, current and former members, you guys made our group a very cohesive team with a great atmosphere to work. Thank you all for you scientific input and for your exceptional support during my PhD track. A very special word to Wilco, Yunlei, Jessica, Elnaz and Eric. Wilco, you were such a great colleague, always eager to help everyone in the

lab, enthusiastic to try new ideas and to have a nice scientific discussion. Yunlei, I really had a great time during our collaboration, I just learnt a lot from you and I appreciate all your dedication and patience. Jessica, although we did not have much overlap during our lab activities, I recognize that you are a key in the group and that most of the work and results that I achieved could not be possible without all your commitment since you joined the group. Elnaz, it was very pleasant to share the same office with you, I carry all your words of wisdom with me. Eric, I do appreciate all your advices and that you are so open about your opinions. I will never forget our little adventure when we tried to find something to eat, late in Venice.

To Monique and Ronald, thank you for your great input in my work either directly during our meetings or indirectly through the knowledge and expertise that you created over the years in the department.

To Patricia, I am so happy that I had the opportunity to share the office and have you as a friend. You gave me good advices for my research and for my own career. Besides that it was my pleasure to nurture our friendship along the way. You are a trustworthy friend and a very talented scientist.

To Sandra, you know how your presence was really important to me. Thank you for all your precious support and for being such a great friend.

To all the members of the Pediatric Oncology, my words of gratitude for the scientific discussions, for the support that we always gave to each other, and for all the good moments passed together. In particular, I would like to special thank João, Isabel, Roel, Bob, Mark and Fhrad.

To all the members of the Pediatrics department I want to thank you all for the great environment and for all the events organized. It was a great pleasure to share the labs and having a good time at lunch. My special thanks to Celia and Silvia.

To my new family in Rotterdam: all the Portuguese and international friends that I made during the last years. All the moments we passed together were so important to make me feel at home and learn to be in love with Rotterdam. A special hug to Rodrigo and Leti, Miguel and Lilia, Rui, João, Fernando and Henrique.

Finally, the last words go to my origins.

To all the friends and family that I have in Portugal and that were always ready to receive me with open arms, warm hugs and love. Those moments passed together were always so special and gave me all the strength to come back and continue my adventure.

To the most important people in my life: my parents and my brother. I honestly don't have words to say how much I love you! This is not my thesis, this is our thesis. The love that we feel for each other makes us a perfect team that makes impossible things to happen. To my parents, thank you so much for always support me in my decisions and to always encourage me in the most difficult moments, and celebrate with me every little achievement. To my brother, you are just an incredible person, I feel the luckiest guy to have you as my best friend and to have the opportunity to share my whole life with you, thanks for being who you are!

